UNIVERSITY OF CALIFORNIA

Los Angeles

A search for heavy resonances decaying to a pair of Higgs bosons with leptons and a bottom quark pair in the final state

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Physics

by

Brent Stone

2021

ABSTRACT OF THE DISSERTATION

A search for heavy resonances decaying to a pair of Higgs bosons with leptons and a bottom quark pair in the final state

by

Brent Stone
Doctor of Philosophy in Physics
University of California, Los Angeles, 2021
Professor David Saltzberg, Chair

This dissertation presents a search for new heavy resonances decaying to a pair of Higgs bosons in proton-proton collisions at a center-of-mass energy of 13 TeV. The data were collected with the CMS detector at the LHC in 2016–2018, corresponding to an integrated luminosity of 138 fb^{-1}. Resonances with a mass between 0.8 and 4.5 TeV are considered using events in which one Higgs boson decays into a bottom quark pair and the other into final states with either one or two charged leptons. Specifically, the single-lepton decay channel $\mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}\ell\nu\mathrm{q\bar{q}}$ and the dilepton decay channels $\mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}\ell\nu\ell\nu$ and $\mathrm{HH} \to \mathrm{b\bar{b}}\tau\tau \to \mathrm{b\bar{b}}\ell\nu\nu\ell\nu\nu$ are examined. The signal is extracted using a two-dimensional maximum likelihood fit of the $\mathrm{H} \to \mathrm{b\bar{b}}$ jet mass and HH invariant mass distributions. No significant excess above the standard model expectation is observed in data. Model-independent exclusion limits are placed on the product of the cross section and branching fraction for narrow spin-0 and spin-2 massive bosons decaying to HH. The results provide the most stringent limits to date for $\mathrm{X} \to \mathrm{HH}$ signatures with final-state leptons and at some masses provide the most sensitive limits of all $\mathrm{X} \to \mathrm{HH}$ searches.

The dissertation of Brent Stone is approved.

Zvi Bern

Jay Hauser

Michail Bachtis

David Saltzberg, Committee Chair

University of California, Los Angeles

2021

To my parents, Ira and Lori, and my brother Drew for all their support and inspiration

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I first learned of CERN and particle physics when I read Dan Brown's Angels and Demons as a kid. I developed this notion then that those scientists described in the book were the smartest people on the planet. In the beginning of my high school years, my classmates and I were asked on a get-to-know-you worksheet in a science class what our dream career was. At this point, I was resigned to the fact that I would not be a professional hockey player. I ended up writing that I wanted to be a scientist at CERN, inspired by the book I had read a couple years earlier, and unsure of what else I would've wanted. I didn't seriously entertain a path as a scientist until sometime during my junior year of college, when I first considered graduate school in physics. It is somewhat surreal that I ended up with the opportunity to work as a scientist in the CERN community and realize those aspirations from many years before. However, my accomplishments with this dissertation would not have been possible without many others in my life.

I'd like to begin by thanking my advisor, David Saltzberg, for the wisdom and support he has provided me during my graduate career. David's enthusiasm for physics is contagious, and I consider him to be among the kindest and most genuine figures in the physics community whom I have encountered. I appreciate the energy he has invested into me while simultaneously serving as Department Chair, and I feel very lucky to have found him as an advisor the summer after my first year of graduate studies.

I would like to thank Nick McColl, David's postdoc when I joined the group, who has had the largest impact on my development as a scientist. I feel quite lucky to have been able to work with and learn from Nick intimately for the first several years of my research studies. Nick showed me how to lead a project within the CERN community, and he is certainly among the most brilliant scientists I have encountered during my studies. When Nick first left the field for Google, I was intimidated to have to lead this project on my own, but I ultimately gained the confidence to do so because of the foundation he laid for me. I

will always look up to him as a scientist and a thinker.

There are several other individuals I would like to thank for providing me with valuable discussion and expert advice throughout my studies. Thank you to Michalis Bachtis, with whom I've developed a nice relationship despite him not being my advisor, who has been both an entertaining and extremely helpful voice when I've needed it. Thank you to Bob Cousins, who taught me quite a bit about particle physics and statistics in the years leading up to his retirement. Thank you to Andrew Peck, with whom I shared an office for several years and was a valuable and intelligent voice when working on the GEM hardware project. Thank you to Brian Dorney, who was my primary mentor for GEM hardware studies when I spent time at CERN.

I'd also like to thank several of my peers for providing valuable discussions and making life in the office at UCLA enjoyable. Thanks to Chris Schnaible, whom I admired as a senior graduate student when I was a younger one and whom I would annoy with questions that I felt were not worth Nick McColl's time. I appreciate the friendship and advice he provided when he moved back to Los Angeles from Geneva. Thanks to Tyler Lam, David Hamilton, and Maxx Tepper, who all served as graduate students or junior researchers with me for the majority of my time at UCLA, and who made the office an enjoyable place to be. Thanks particularly to Will Nash, who started grad school with me and who has been on this same journey with me during our entire time at UCLA. Will has been my primary confidant for all the ups and downs of our graduate career, and I am thankful for the friendship we have developed.

There are many other friends whom I've met and with whom I've shared lovely times with in Los Angeles, but I'd like to particularly thank my college friend Jordan Rodnizki, who has been a constant throughout my life in Los Angeles. We moved together from Philly to LA at the same time (albeit for very different reasons), and our friendship has grown dramatically since then. I know that will persist for the rest of our lives.

A special thanks is warranted for Grace Schnaible, who created the beautiful standard

model diagram in Figure 2.1.

Finally, thank you to my parents, Ira and Lori, whose support in my life is immeasurable. They have stopped at nothing to provide for me in every imaginable way through my entire life, but especially during my time far away in California. Thank you to my brother Drew as well. Despite being my younger brother, I look up to him and am amazed and inspired by his accomplishments in life and in the field of blockchains. His example pushes me to keep moving forward intellectually.

This dissertation includes material that is based upon work supported by the U.S. Department of Energy under Award Number DE-SC0009937. Parts of Chapters 4, 5, and 6 have been adapted from research works for which I am the primary author but have also had contributions from both Nick McColl and David Saltzberg. Nick contributed significantly to the development of the code used in this analysis, and both Nick and David further contributed to the analysis procedures and writing as well. These research works include a CERN CMS analysis note (AN-2019-187), a publicly available physics analysis summary [13], and a paper that has recently been submitted to the Journal of High Energy Physics [14] at the time of this writing.

VITA

2016	B.A. (Physics), University of Pennsylvania
2016-2017	Teaching Assistant, Department of Physics and Astronomy, UCLA, Los Angeles, California
2017	M.S. (Physics), UCLA
2017-2021	Graduate Research Assistant, Department of Physics and Astronomy, UCLA

CHAPTER 1

Introduction

For thousands of years, humans have explored the notion that matter — all the "stuff" in the universe — can be broken down repeatedly into smaller units and that at some tiny length scale, there must exist some unit of matter that is no longer composed of smaller parts. These units, which have no substructure, are called elementary or fundamental particles, and the field of particle physics is humanity's best attempt at explaining their nature and thus the fundamental nature of matter in the universe.

Particle physics was essentially born with the first observation of an elementary particle, J. J. Thomson's discovery of the electron [15]. By 1932, it was clear that electromagnetic (EM) radiation as described classically by Maxwell was actually quantized into particles called photons, and a complete atomic picture of matter had been formulated. It seemed that atoms were fundamentally composed of negatively charged electrons bound by Coulombic forces to a positively charged nucleus of protons and neutrons.

From the 1930s to the 1970s, deeper questions about the particle nature of matter would be addressed. Until this point, the question of what held together the protons and neutrons in the nucleus remained a mystery. This would lead physicists to explore a *strong* nuclear force that holds together the atomic nucleus, a pursuit that would reveal a complex landscape of interactions and new particles. By the end of the 1970s, it was understood that protons and neutrons were two of many hadrons, composite particles composed of fundamental particles called quarks. Still there remained puzzling observations that suggested the existence of yet another force of nature. For example, results from nuclear β-decay experiments seemed to suggest the production of an undetectable particle [16]. Furthermore, the observed lifetimes of some hadrons were roughly a dozen orders of magnitude larger than expected for strongly interacting particles, suggesting that a weaker nuclear force may be at play. Indeed, such a theory of weak nuclear interactions was developed from the 1930s through the 1970s. It revealed that the undetectable particle was the neutrino, which does not interact strongly or electromagnetically, and explained the lifetime discrepancies observed in hadrons.

In order to understand these fundamental interactions between elementary particles, a mathematical framework was needed. In the early 20th century, the introduction of quantum mechanics and special relativity provided more precise descriptions of nature at the smallest of length scales and the highest of energies, respectively. To best describe subatomic phenomena, a union of both is necessary, and this became accomplished with quantum field theory (QFT). The quantum field theory of electrodynamics (QED) was the first to be successfully developed, beginning in the 1930s. QED models EM interactions — like the Coulomb force binding the electrons and nucleus in the atom — as the exchange of massless photons between charged particles. Similarly, a quantum field theory of strong nuclear interactions, called quantum chromodynamics (QCD), predicts the existence of massless force-carrying bosons called gluons and a "charge" for strong interactions called color. Like in QED, the strong force between colored particles, which include quarks and gluons, is mediated by the gluon.

The structure of QED and QCD might suggest that a quantum field theory of weak nuclear interactions could follow that predicts some weak "charge" and new massless weak bosons to mediate the force. However, it later became clear that the weak force was in essence the same as the EM force, except that the force-mediating bosons were massive, unlike the photon in QED. Physicists successfully unified the two forces into one electroweak force at higher energies and showed that a mechanism called spontaneous symmetry breaking (SSB) was responsible for the masses of the force-carrying bosons; this is in essence why the weak couplings are dwarfed in strength by the EM couplings. A fourth force of nature exists as well and is even about a dozen orders of magnitude weaker than the weak nuclear force: gravity. However, to date there is however no successful quantum field theory of gravitation.

The quantum field theory that collectively describes the electroweak and strong interactions is today known as the standard model (SM). The SM is regarded as a wildly successful theory; no measurements performed to date at the GeV and TeV scales have been able to strongly challenge the predictions of the SM. The prediction of the force-carrying bosons of the electroweak and strong interactions is also not the full picture. The existence of a special scalar boson is predicted as a result of the SSB of the electroweak interactions — the Higgs boson (H), with a mass m_{H} that is a free parameter in the SM. For decades the Higgs boson eluded experimental observation after all other SM particles had been found, and so the Large Hadron Collider (LHC) was constructed at the European Center for Nuclear Research (CERN) with the discovery of the Higgs boson as a top priority. In 2012 that pursuit, the final piece to the SM puzzle, was realized [17–19].

Despite the profound success of the SM, there are several important questions and observations in nature that are not addressed by the theory. For example, the SM only provides a description of normal baryonic matter, which accounts for about 5% of the composition of the mass-energy of the observed universe. The nature of dark matter, accounting for 27%, and dark energy, accounting for 68%, are not addressed in the SM. The SM also does not provide insight to the matter-antimatter asymmetry in the universe - that is, why are there only stable matter particles and no stable antimatter particles in the universe? The SM does not address gravity since no successful quantum theory exists for it yet, and the SM offers no insight into many observations in neutrino physics, such as neutrino oscillations [20–22], which indicate that neutrinos do in fact have mass. It is merely logical to deduce that the SM is incomplete, and there must be new physical phenomena beyond the SM (BSM) waiting to be discovered.

This dissertation describes a general, model-independent search for the production of massive, unstable BSM particles that decay to a pair of Higgs bosons in proton-proton

(pp) collisions collected by the Compact Muon Solenoid (CMS) detector at the CERN LHC at a center-of-mass beam energy $\sqrt{s} = 13$ TeV. A number of searches have been conducted for such diboson resonances by the ATLAS [23–46] and CMS [47–72] Collaborations, including resonances decaying to WW, WZ, ZZ, WH, ZH, and HH. Existing searches for HH resonances in both experiments include the final states of $\mathrm{b\bar{b}b\bar{b}}$, $\mathrm{b\bar{b}}\tau\tau$ (with at least one hadronic τ lepton), and $\mathrm{b\bar{b}}\gamma\gamma$. In this search, we consider a final state of the HH decay where one of the Higgs bosons decays to $\mathrm{b\bar{b}}$ and the other Higgs decays into either a single-lepton or dilepton final state. The single-lepton final state is due to $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}\ell\nu\mathrm{q\bar{q}}$, and the dilepton final states are due to both $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}\ell\nu\ell\nu$ and $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}}\tau\tau \to \mathrm{b\bar{b}}\ell\nu\nu\ell\nu\nu$, where ℓ in the final state is either a muon or an electron. In the $\mathrm{b\bar{b}WW^*}$ decays, this search is sensitive to leptonically decaying τ leptons from $\mathrm{W} \to \tau\nu$.

The search considers BSM resonances X in the mass range $0.8 < m_{\mathrm{X}} < 4.5$ TeV. Because we consider X bosons that are much heavier than the Higgs boson, the Higgs bosons in signal events are boosted, and their decay products are typically detected in a collimated region. The two b quarks from the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay that is characteristic of both channels are often produced very close together, and so this decay is reconstructed as a single, large jet (the $\mathrm{b\bar{b}}$ jet) that has substructure consistent with a two-prong decay. The kinematics of the boosted system also tend to result in the decay products of the X boson being produced opposite each other in the transverse plane. Thus, we aim to reconstruct the other Higgs boson from a collimated collection of other particles that include either one or two leptons and that is far away from the $\mathrm{b\bar{b}}$ jet. The search is executed in a two-dimensional (2D) plane constructed from the mass of the $\mathrm{b\bar{b}}$ jet (called $m_{\mathrm{b\bar{b}}}$) and the invariant mass of the two Higgs boson momenta (called m_{HH}). If it exists, the signal is expected to appear as a bump over a smoothly falling spectrum, peaking around $m_{\mathrm{H}} = 125$ GeV in the $m_{\mathrm{b\bar{b}}}$ dimension and at some unknown mass *a priori* in the m_{HH} dimension.

In the single-lepton (1ℓ) channel, the quarks in the $\mathrm{H} \to \mathrm{WW^*} \to \ell\nu\mathrm{q\bar{q}}$ decay are reconstructed as a single large jet, referred to as the $\mathrm{q\bar{q}}$ jet. The $\mathrm{q\bar{q}}$ jet is required to have substructure arising from a two-prong decay and be accompanied by a nearby electron or muon. The $\mathrm{H} \to \mathrm{WW}^*$ decay chain is reconstructed from the momenta of the $\mathrm{q\bar{q}}$ jet, the lepton, and the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ that is largely due to the undetectable neutrino. Because the proton beams are directed along $\hat{z}$ in the event frame of reference, the missing momentum in the transverse plane can be inferred from momentum conservation. However, the incoming momenta of the partons that create the hard collision cannot be known because of the parton distribution functions of the proton, and so the missing longitudinal momentum cannot be directly inferred. In the dilepton (2ℓ) channel, two leptons are reconstructed in close proximity to each other and with $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ nearby in the same azimuthal direction, consistent with the multiple neutrinos.

The main SM background in this search arises from top quark pair production ($\mathrm{t\bar{t}}$). The top quarks in this analysis have collimated decay products because of large Lorentz boosts. In the 1ℓ channel, the largest background comes from semileptonic $\mathrm{t\bar{t}}$ decays: one top quark decays to a charged lepton and a neutrino ($\mathrm{t} \to \mathrm{Wb} \to \ell\nu\mathrm{b}$), and the other decays exclusively to quarks ($\mathrm{t} \to \mathrm{Wb} \to \mathrm{q\bar{q}b}$), which can be mistakenly reconstructed as the $\mathrm{b\bar{b}}$ jet candidate. Other significant backgrounds in this channel are production of a W boson in association with jets (with $\mathrm{W} \to \ell\nu$, hereafter referred to as $\mathrm{W}+\mathrm{jets}$) and QCD multijet events with either a non-prompt lepton or a hadron misidentified as a lepton. In the 2ℓ channel the background yield is smaller than that of the 1ℓ channel by a factor of $\approx$60. Top quark pair production is also the dominant background with approximately equal contributions from events where both top quarks decay leptonically and events where one decays hadronically but a lepton is misidentified. The other significant background in this channel is production of Z/γ^* bosons in association with jets ($\mathrm{Z}/\gamma^* + \mathrm{jets}$). These backgrounds are distinguished in data with their unique $m_{\mathrm{b\bar{b}}}$ spectra.

In this search, the events are divided into 12 exclusive categories with similar ratios of signal to background, a process that helps to maximize our sensitivity to signal. The signal and SM background yields are estimated using a simultaneous 2D maximum likelihood fit to the $m_{\mathrm{b\bar{b}}}$, m_{HH} mass distributions in all 12 categories.

The content of the search presented in this dissertation is the focus of a paper that has be submitted to the Journal of High Energy Physics [14]. This information is currently summarized in a Physics Analysis Summary (PAS) published on the CERN Document Server [13]. The analysis has been produced and reviewed within the Beyond 2 Generations (B2G) physics analysis group of CMS. The full results of this dissertation also serve as an expansion of a published paper searching for $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}\ell\nu\mathrm{q\bar{q}}$ in just the 2016 data set [72].

This dissertation is organized as follows. Chapter 2 provides the theoretical and phenomenological motivation to the search, which includes descriptions of the SM and important BSM ideas that are relevant for HH production. Chapter 3 presents the experimental apparatus used to perform the search. This chapter includes descriptions of the LHC accelerator complex and the CMS detector, and it also reviews the basic reconstruction of a *pp* collision event using the CMS detector. Chapter 4 describes the global event selection strategy, which includes triggering events, reconstructing all states in the full $\mathrm{X} \to \mathrm{HH}$ decay chain, and then optimizing the signal purity of final sample. Chapter 5 presents the strategy for modeling the signal and background probability densities and associated systematic uncertainties. Finally, Chapter 6 details the statistical analysis of the data and the results of the search. In every chapter, we assume $c = 1$ such that masses and momenta are given in units of energy.

CHAPTER 2

Theoretical Considerations for the Resonant HH Search

In this chapter, theoretical descriptions of both the SM and BSM theories that motivate searches for X $\rightarrow$ HH are presented. The chapter concludes with a discussion on X $\rightarrow$ HH searches at the LHC up to this point in time.

2.1 The Standard Model

The SM is the quantum field theory that best describes the properties and interactions of all elementary particles known to exist. The known interactions in the universe correspond to four fundamental forces: gravity, electromagnetism, and the weak and strong nuclear forces. The SM provides a description of three of these forces, while a suitable quantum field theory of gravitational interactions still eludes physicists to this day.

In general, the SM describes matter with fermion fields of spin-$\frac{1}{2}$ and interactions mediated by boson fields of spin-1. Such spin-1 fields in the SM are also known as vector fields or gauge fields. Experimental observations show that there are six pairs of fermions: three generations of quarks and three generations of leptons. Fermions in one generation differ from the other generations only by their masses, a consequence of their couplings to the scalar field in the SM, described in Section 2.1.5. Each fermion also possesses a corresponding antiparticle that has identical properties but opposite quantum numbers.

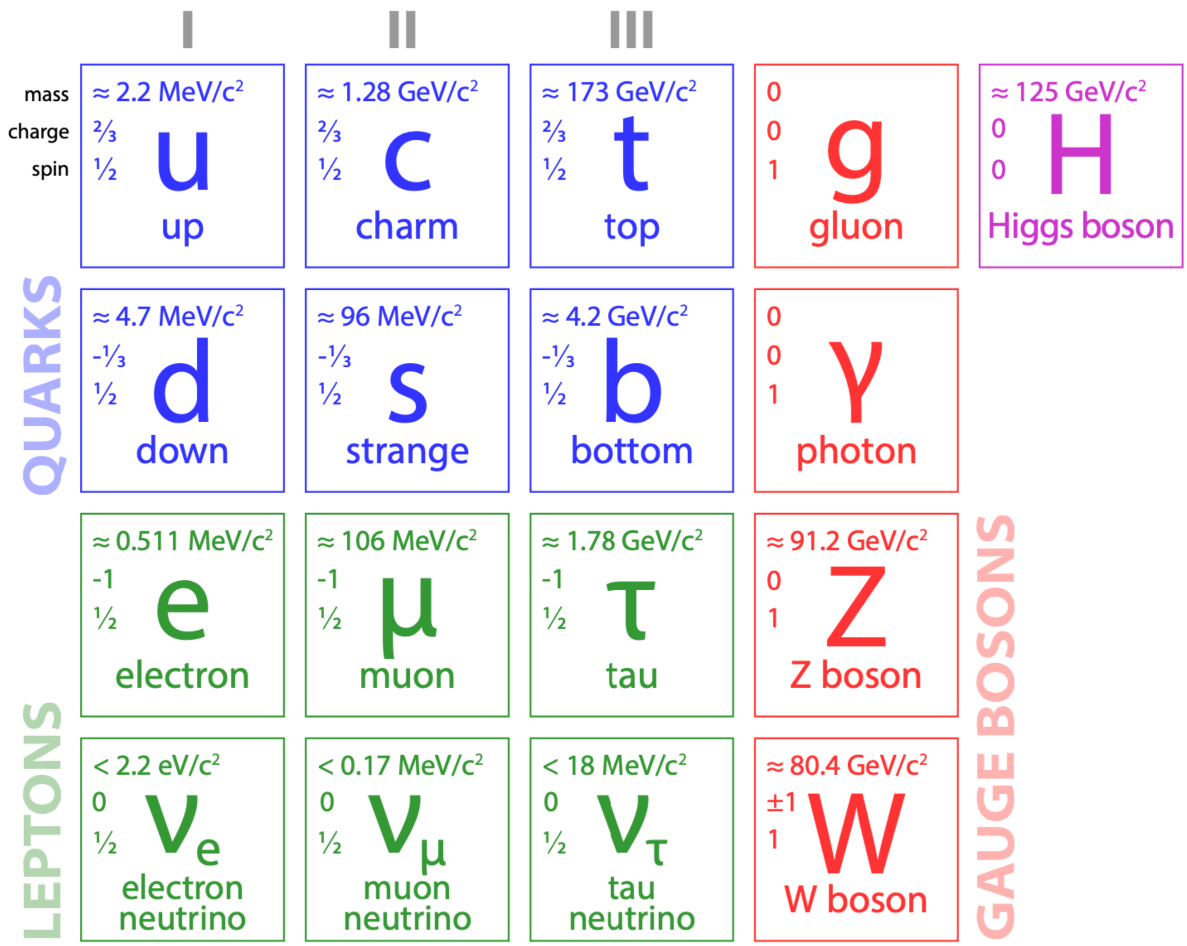

Figure 2.1: Table showing all SM particles, with blue for quarks, green for leptons, red for gauge bosons, and purple for scalar bosons. The columns containing the fermions (quarks and leptons) are grouped by generation and type, with increasing mass from left to right. The mass, charge, and spin are shown for all particles.

2.1.1 Elementary particles

All ordinary matter in the universe is composed of quarks and leptons, each organized into three generations, which themselves are each composed of a particle doublet. In each quark doublet, the "up"-type quarks (u, c, t) have electric charge of $+2/3$, while the "down"-type quarks (d, s, b) have electric charge of $-1/3$. Each doublet in a lepton generation contains a charged lepton with electric charge of -1 and a corresponding electrically neutral lepton called a neutrino that is much less massive than any other massive SM fermion. The gauge bosons act as force carriers (mediators), allowing fermions to interact with each other via the gauge fields. The photon is the mediator of the EM force, the gluon of the strong nuclear force, and the W and Z bosons of the weak nuclear force. While the number of fermion generations is known only from experiment and not predicted by the SM, the different kinds of gauge bosons are exactly a consequence of the theory. Furthermore, the theory predicts the existence of at least one special scalar boson, the Higgs boson. A table summarizing the fundamental particles is shown in Fig. 2.1, and the mechanics of this hierarchy are explained in the rest of this section.

2.1.2 Gauge theory and local gauge invariance

The success of the SM relies on the local gauge invariance of the fermion fields under symmetry transformations of the Lagrangian $\mathcal{L}$, a function from which the equations of motion of a physical system can be derived [15,73]. The consequences of enforcing local gauge invariance are outlined now.

The Lagrangian for a free massive spin-$\frac{1}{2}$ fermion field ψ with no interactions is given by the Dirac Equation:

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi, \tag{2.1}$$

where m is the fermion mass. The first term corresponds to the kinetic energy of the field, and the second term corresponds to the fermion mass. For a physical system to possess gauge invariance, the equations of motion derived from the Lagrangian must be the same under the gauge transformations that affect the phase of the fields. Local gauge invariance is the most general type of phase invariance, valid under transformations of the form:

$$\psi(x) \to U\psi(x) = e^{i\alpha_a(x)T_a}\psi(x) \tag{2.2}$$

where U is a unitary operator corresponding to some group symmetry, the T_a are the generators of the group (with a enumerating the generators), and $\alpha_a(x)$ is the phase, hereafter denoted simply as α_a but understood to be a function of spacetime x (3+1 dimensions).

Considering infinitesimal transformations of the fields, we can show that:

$$\psi(x) \to (1 + i\alpha_a T_a)\psi(x) \tag{2.3}$$

$$\partial_\mu\psi(x) \to (1 + i\alpha_a T_a)\partial_\mu\psi(x) + iT_a\psi(x)\partial_\mu\alpha_a, \tag{2.4}$$

where the subscript μ runs over the 3+1 spacetime indices. The presence of the last term in the transformation of $\partial_\mu\psi(x)$ spoils the local gauge invariance at this point. In order to preserve the gauge invariance, we can introduce new vector fields G^a_μ into the theory that transform under the gauge symmetry operations as:

$$G^a_\mu \to G^a_\mu - \frac{1}{g}\partial_\mu\alpha_a(x). \tag{2.5}$$

This amounts to forming a new covariant derivative that contains the gauge fields and does not destroy the invariance under infinitesimal gauge transformations of the fields:

$$\partial_\mu \to D_\mu = \partial_\mu + igT_a G^a_\mu, \tag{2.6}$$

where D_μ is known as the covariant derivative. Replacing ∂_μ with D_μ and adding in a gauge invariant kinetic term for the new fields, the Lagrangian of Eq. (2.1) then becomes:

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi - g(\bar{\psi}\gamma^\mu T_a\psi)G^a_\mu - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a. \tag{2.7}$$

Demanding that the Lagrangian be invariant under local gauge transformations thus produces new vector fields G_μ^a that couple to the fermion field ψ with coupling strength g and with kinetic energy described by the field strength tensor $F_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a$. There is no mass term ($\propto -G_\mu^a G_a^\mu$) for the gauge fields in this new form of the Lagrangian, and so these new gauge bosons must be massless to preserve the invariance.

For non-Abelian symmetries, wherein the group generators do not commute ($[T_a, T_b] \neq 0$), an extra term is needed in the transformation rule for the gauge fields

$$G_\mu^a \rightarrow G_\mu^a - \frac{1}{g}\partial_\mu \alpha_a - f_{abc}\alpha_b G_\mu^c, \tag{2.8}$$

where f_{abc} are the structure constants of the symmetry group.

It is a simple exercise to see how this works out for QED to reproduce the Maxwell Lagrangian for classical EM fields. QED obeys a $U(1)$ gauge symmetry, of which the only generator of the group is unity. With the replacements $T_a \rightarrow \mathbb{1}$ and $G_\mu^a \rightarrow A_\mu$, the Lagrangian of Eq. (2.7) then becomes:

$$\mathcal{L}_{\text{QED}} = i\bar{\psi}\gamma_\mu \partial^\mu \psi - m\bar{\psi}\psi - g(\bar{\psi}\gamma^\mu\psi)A_\mu - \frac{1}{4}F_{\mu\nu}F_{\mu\nu}, \tag{2.9}$$

where g can be identified as $-e$ with Maxwell's electrodynamics in mind. The field A_μ corresponds to the photon, which is massless and has field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

The SM as a whole, which describes the interactions due to three fundamental forces, is invariant under gauge transformations of the symmetry group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y, \tag{2.10}$$

where $\times$ denotes the direct product of each of the three symmetry groups. The subscripts next to the group names represent the conserved charges of the symmetry due to Noether's Theorem. In the following sections, each of the components of the full SM symmetry group are explained.

2.1.3 Strong interactions

In Eq. (2.10), $SU(3)_C$ is the symmetry group used to describe the strong interactions (QCD), where the subscript C represents "color," a quantum number and the conserved charge in these interactions. There are three color charges a particle can possess, known commonly as red (R), green (G), and blue (B), each of which also have an anticolor partner. Only colored particles can interact strongly. Only quarks and gluons carry color charge. There are eight gluons, each of which is bicolored, corresponding to the eight generators of the $SU(3)$ group.

The Lagrangian of QCD looks like the that for EM as seen in Eq. (2.9), except it has some extra terms due to the gauge transformation rule of Eq. (2.8) that is required because of the non-Abelian structure of $SU(3)$. Extra terms that appear and are proportional to G^3 and higher orders are responsible for the ability of the gluon to self-couple. This is an important distinction relative to QED, in which the $U(1)_{\text{em}}$ symmetry is Abelian and thus does not provide a mechanism for photon self-coupling.

2.1.4 Electroweak interactions

The symmetry group $SU(2)_L \times U(1)_Y$ in Eq. (2.10) describes the electroweak (EW) interactions, the unified and underlying symmetry of the existence of the weak nuclear and EM interactions. Experimental observations in the 1950s [16] showed that parity is violated in the weak sector, and EW theory accounts for this with fields that couple differently to fermions of opposite chiralities (left- and right-handedness). For massless particles, the chirality reduces to the helicity, which is the sign of the projection of the spin vector onto the momentum vector $\hat{s} \cdot \hat{p}$.

The subscript L in $SU(2)_L$ is used to indicate fields that only couple to left-handed fermions. The gauge field from the $U(1)_Y$ symmetry couples to both left- and right-handed fermions. The EW eigenstates for the fermions then consist of left-handed doublets and right-handed singlets:

$$\text{Leptons:} \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix}, e_R^i \qquad \text{Quarks:} \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}, u_R^i, d_R^i \tag{2.11}$$

where i runs over the three fermion generations, and where L and R denote left- and right-handedness, respectively. Neutrinos are assumed to be massless in the SM, and right-handed neutrinos have not yet been observed, so no ν_R state is assumed to exist in this formulation. Each field component in Eq. (2.11) is characterized by the two quantum numbers T_3 (weak isospin) and Y (weak hypercharge). Weak isospin T_3 is the eigenvalue of the third $SU(2)$ generator (the Pauli matrix σ_z), and weak hypercharge Y is the associated charge of the $U(1)_Y$ symmetry. The electric charge Q of these fields is related to the quantum numbers by $Q = T_3 + \frac{Y}{2}$. The values for these quantum numbers are summarized in Table 2.1.

Fermion field	T_3	Y	Q
ν_L^i	1/2	-1	0
e_L^i	-1/2	-1	-1
e_R^i	0	-2	-1
u_L^i	1/2	1/3	2/3
d_L^i	-1/2	1/3	-1/3
u_R^i	0	4/3	2/3
d_L^i	0	-2/3	-1/3

Table 2.1: Electroweak quantum numbers for each of the different fermion fields, which include the left- and right-handed leptons and quarks. Included are the weak isospin T_3, weak hypercharge Y, and the electric charge Q associated with $U(1)_{\text{em}}$.

The $SU(2)_L$ group gives rise to three gauge fields W_μ^a ($a = 1, 2, 3$) that only interact with the left-handed chiral fields, while the $U(1)_Y$ group gives rise to one gauge field B_μ. It is useful and revealing to recast these gauge fields as linear combinations of each other. The two fields W_μ^1 and W_μ^2 can be rewritten as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp W_\mu^2\right), \tag{2.12}$$

which are the fields that correspond to the physical charged $W^\pm$ bosons that mediate the weak interactions. The physical neutral Z boson that mediates weak interactions and the photon A of the EM interactions can be identified by mixing the W_μ^3 and B_μ fields [73]:

$$Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \tag{2.13}$$

$$A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}}, \tag{2.14}$$

where g is the fermion coupling to the fields W_μ^a and g' is the coupling to B_μ. After spontaneous symmetry breaking, reviewed next in Section 2.1.5, the fields $W^\pm$ and Z_μ acquire mass terms in the Lagrangian, while the photon field A_μ remains massless. It is this mechanism that breaks the EW symmetry down into separate weak and EM interactions.

2.1.5 Spontaneous symmetry breaking and the Brout-Englert-Higgs mechanism

As described in Section 2.1.2, local gauge invariance results in the existence of massless force-mediating bosons. This alone works well for the EM and strong forces, since the photon and gluon are massless. However, the W and Z weak vector bosons are massive, and so extra machinery is needed to produce mass terms in the Lagrangian. This is ameliorated with the concept of spontaneous symmetry-breaking and the Brout-Englert-Higgs (BEH) mechanism [74, 75].

We can best illustrate the mass generation of gauge bosons due to spontaneous symmetry-breaking by considering the case of a local $U(1)$ gauge symmetry. We start by introducing a complex scalar field

$$\phi = \sqrt{\frac{1}{2}}(\phi_1 + i\phi_2), \tag{2.15}$$

whose Lagrangian takes the form for a massive scalar field with a quartic potential:

$$\mathcal{L} = (\partial_\mu \phi)^\star (\partial^\mu \phi) - \mu^2 \phi^\star \phi - \lambda(\phi^\star \phi)^2, \tag{2.16}$$

The scalar potential, as seen in the Lagrangian, has the form:

$$V(\phi) = \mu^2 \phi^\star \phi + \lambda(\phi^\star \phi)^2. \tag{2.17}$$

For the case $\mu^2 > 0$, this scalar potential has a global minimum at $\phi = 0$. The $\mu^2 \phi^* \phi$ term represents the mass of the scalar field, but when working out the full Lagrangian with local gauge invariance, no mass terms appear for the gauge fields. However, for the case $\mu^2 < 0$, the potential is minimized along the manifold corresponding to

$$\phi^\star \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2) = -\frac{\mu^2}{\lambda} \equiv v^2, \tag{2.18}$$

which is a circle in the (ϕ_1, ϕ_2) plane, and where v, the ground state of this potential, is known as the vacuum expectation value. This is illustrated in Fig. 2.2.

This circle of points is also invariant under the $U(1)$ transformations. Perturbative calculations such as the ones we will perform on ϕ should involve expansions around a potential minimum, and we can choose any point we desire without loss of generality. It is convenient to choose a minimum point such as $\phi_1 = v$ and $\phi_2 = 0$. The ground state is then taken to be

$$\phi_0 = \frac{v}{\sqrt{2}}, \tag{2.19}$$

and fluctuations of the fields around this ground state are expressed as

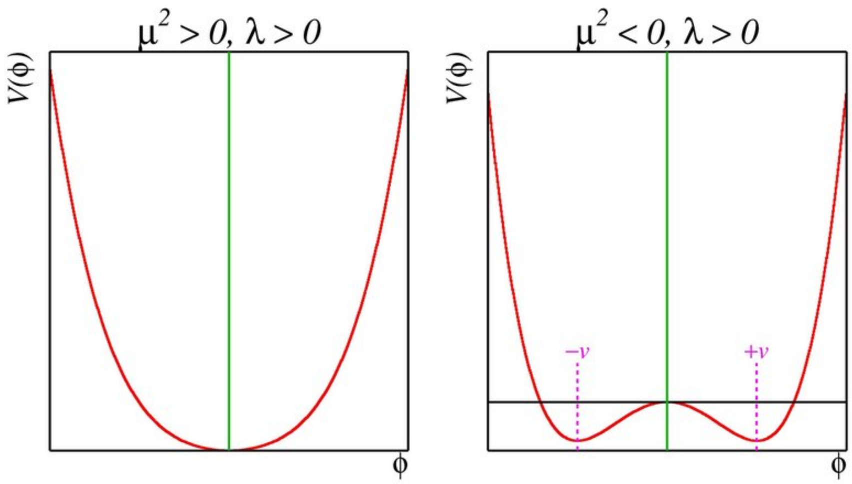

Figure 2.2: The shape of the Higgs potential $V(\phi)$ from Eq. (2.17) for $\lambda > 0$ and the cases $\mu^2 > 0$ (left) and $\mu^2 > 0$ (right) [1].

$$\phi(x) = \sqrt{\frac{1}{2}}\,(v + h(x) + i\xi(x))\,, \tag{2.20}$$

where $h(x)$ corresponds to fluctuations about $\phi_1 = v$ and $\xi(x)$ to fluctuations about $\phi_2 = 0$.

These scalar fields are inherently invariant under global $U(1)$ transformations, but to enforce local gauge invariance, we must replace the ∂_μ with the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + igA_\mu, \tag{2.21}$$

where A_μ is the newly introduced gauge field. The gauge invariant Lagrangian then looks like

$$\mathcal{L} = (\partial^\mu - igA^\mu)\phi^\star(\partial_\mu + igA_\mu)\phi - \mu^2\phi^\star\phi - \lambda(\phi^\star\phi)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.22}$$

Upon translating to the ground state and substituting (2.20) into (2.22), the Lagrangian then becomes

$$
\begin{aligned}
\mathcal{L} = &\frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 + \frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2} g^2 v^2 A_\mu A^\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\
&+ gvA_\mu(\partial^\mu \xi) + g[(\partial_\mu \xi) A^\mu h - (\partial_\mu h) A^\mu \xi] - \lambda v h^3 - \frac{1}{4}\lambda(h^4 + \xi^4) \\
&+ g^2 v h A_\mu A^\mu + \frac{1}{2} g^2 (h^2 + \xi^2) A_\mu A^\mu - \lambda v h \xi^2 - \frac{1}{2}\lambda h^2 \xi^2 - \frac{1}{4}\lambda v^4.
\end{aligned}
\tag{2.23}
$$

From the first line of the Lagrangian of Eq. (2.23), the particle spectrum includes a massive scalar boson h with mass $m_h = \sqrt{2\lambda v^2}$, a massive vector field A_μ with mass $m_A = gv$, and a massless boson ξ. This boson is known as a Goldstone boson; for a spontaneously broken group symmetry, one massless Goldstone boson is predicted to appear for each generator in the broken symmetry, according to Goldstone's Theorem [73]. Thus, in the case of the $U(1)$ symmetry here, we expect one Goldstone boson.

The terms in the last two lines of Eq. (2.23) represent various interactions among these fields. Some of these terms, such as $gvA_\mu(\partial^\mu \xi)$, do not seem to represent any kind of physical process, however. It turns out that there is another gauge transformation that can be made that will absorb the Goldstone boson ξ and remove it from the theory. Since the scalar fields are invariant under the transformation $\phi \to e^{i\theta(x)}\phi$, we can choose a gauge in which the fields are only real and $\phi_2 = \xi = 0$. Under such gauge transformations, all terms with ξ vanish from Eq. (2.23), and the Lagrangian reduces to (ignoring constant terms):

$$
\begin{aligned}
\mathcal{L} = &\frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 + \frac{1}{2} g^2 v^2 A_\mu A^\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \lambda v h^3 - \frac{1}{4}\lambda h^4 \\
&+ g^2 v h A_\mu A^\mu + \frac{1}{2} g^2 h^2 A_\mu A^\mu.
\end{aligned}
\tag{2.24}
$$

The theory now predicts both a massive gauge boson and a massive scalar boson as well as several interactions of these fields with each other and themselves. The interactions predicted by Eq. (2.24) are depicted graphically in the Feynman diagrams of Fig. 2.3.

The process of spontaneous symmetry breaking works similarly to this in the $SU(2)$ language of EW theory, wherein we consider a complex doublet scalar field

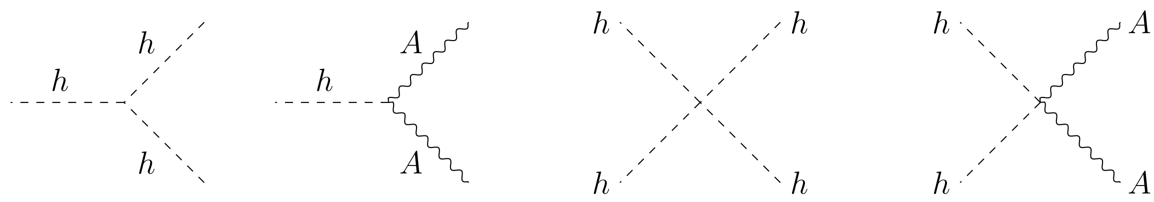

Figure 2.3: Feynman diagrams of the interactions between the scalar field h and the gauge field A_μ described in the Lagrangian of Eq. (2.24). From left to right, the diagrams correspond to the terms $\lambda v h^3$, $g^2 v h A_\mu A^\mu$, $\lambda h^4/4$, and $g^2 h^2 A_\mu A^\mu/2$.

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.25}$$

For the case of a local $U(1)$ symmetry and a scalar field with two degrees of freedom (ϕ_1 and ϕ_2 from Eq. (2.15), that one degree of freedom gave rise to the massive Higgs boson, while the other gave mass to the gauge field A_μ by absorbing the Goldstone boson. In the $SU(2)$ case now, there are four degrees of freedom in the scalar field. The Higgs boson still takes one of these, and the remaining three degrees of freedom, each corresponding to a Goldstone boson, are absorbed to become responsible for the three massive weak vector bosons $W^\pm$ and Z^0.

So far, we have seen that the BEH mechanism gives mass to the gauge bosons, but we have not seen how the fermions acquire mass. The fermion mass term $m\psi\bar{\psi}$ is excluded by gauge invariance under the local $SU(2)_L \times U(1)_Y$ symmetry, but this can be remedied by introducing the gauge invariant Yukawa Lagrangian term

$$\mathcal{L}_{\text{Yukawa}} = -y_f(\bar{f}_L \phi f_R + \bar{f}_R \phi^\dagger f_L + \text{ h.c.}), \tag{2.26}$$

where y_f is the Yukawa coupling, ϕ is the scalar doublet from Eq. (2.25), f_L and f_R are the chiral doublet and singlet fermion fields as in Eq. (2.11), and "h.c." denotes the hermitian conjugate terms. Consider the case for the first lepton generation (with the h.c. dropped):

$$\mathcal{L}_e = -y_e \left[\begin{pmatrix} \bar{\nu}_e & e \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R \begin{pmatrix} \phi^- & \bar{\phi}^0 \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right]. \tag{2.27}$$

Spontaneously breaking the symmetry and choosing a convenient gauge in which

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.28}$$

the electron Yukawa Lagrangian becomes

$$\mathcal{L}_e = -\frac{y_f}{\sqrt{2}} v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{y_f}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) h. \tag{2.29}$$

Identifying the Yukawa coupling such that $y_f = \sqrt{2} m_e / v$, this Lagrangian reduces to

$$\mathcal{L}_e = -m_e \bar{e} e - \frac{m_e}{v} \bar{e} e h, \tag{2.30}$$

which contains the mass term for the electron and an electron coupling to the Higgs field with strength proportional to its mass m_e. It is these Yukawa interactions that describe the Higgs coupling to the mass of a fermion.

2.2 Beyond the Standard Model

While the SM provides a framework compatible with all measurements made using particle colliders thus far, there still exist some glaring shortcomings of the model. In the description of EW interactions in Section 2.1.4, neutrinos are assumed to be massless (and thus have no right-handed field components), but experimental observations of neutrino flavor oscillations provide evidence that neutrinos must have some non-zero mass. Furthermore, there are a variety of astronomical observations that conflict with or are not addressed by

SM predictions. The SM provides no insight into the nature of dark matter, understood from the observed rotation curves of galaxies. There is also a constant term of $\lambda v^4/4$ in the EW Lagrangian seen in Eq. (2.23) that contributes to the vacuum energy density of the universe. Astronomical observations that constrain the upper bound of the cosmological constant are used to also interpret an upper bound on the vacuum energy density. However, the upper bound on the vacuum energy density interpreted from large scale cosmology is smaller by at least 40 orders of magnitude compared to that interpreted from QFT [76]. Some additional interactions or other mechanisms could reduce the vacuum energy density due to Higgs potential.

2.2.1 Hierarchy Problems

Of particular interest in this dissertation is a peculiarity in the relationship between the SM and gravity. If we believe the SM to be part of a higher-energy unified field theory that includes gravity, the energy scale of gravitational interactions must be considered. With the use of fundamental constants, we may estimate the energy scale at which quantum gravitational effects become apparent and the gravitational couplings become comparable to the EW couplings. This energy scale corresponds to the Planck mass:

$$M_{\mathrm{Pl}} = \sqrt{\frac{\hbar c}{G}} \approx 1.5 \times 10^{19}\ \mathrm{GeV}. \tag{2.31}$$

Relative to the scale of EW symmetry breaking, set by the value of the Higgs field vacuum expectation value of $v \approx 246$ GeV, the Planck scale at $\sim 10^{19}$ GeV represents a discrepancy of 17 orders of magnitude; this is a hierarchy problem. This mass scale is 15 orders of magnitude larger than what is even accessible with the LHC. Exploring this scale directly in collider experiments is not feasible.

The situation becomes more troublesome when considering the radiative loop corrections to the mass term of the Higgs field. Quantum corrections to the Higgs boson mass squared

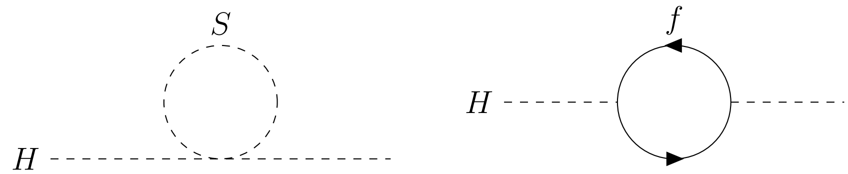

Figure 2.4: One-loop Feynman diagrams that represent higher-order corrections to the mass of the Higgs field due to a virtual loop of scalar bosons (left) and fermions (right).

due to loop diagrams such as those seen in Fig. 2.4 take the form:

$$\Delta m_{\mathrm{H}}^2 \sim \Lambda_{\mathrm{UV}}^2 + ..., \tag{2.32}$$

where Λ_{UV} is a momentum cut-off needed to evaluate the Feynman loop diagrams [77]. Taking $\Lambda_{\mathrm{UV}} \to M_{\mathrm{Pl}}$ would cause the Higgs boson mass to blow up well beyond its observed value, but taking Λ_{UV} only up to some energy scale in between the EW and Planck scales requires new physics at that scale, which is not known to exist yet. To prevent such divergences in Δm_{H}^2 without the introduction of new terms to the Lagrangian, extreme fine-tuning of parameters in the SM on the order of $M_{\mathrm{Pl}}^2/v^2 \approx 10^{34}$ is necessary. The degree of fine-tuning needed in the SM is considered to be "unnatural," unlike the elegance seen in other physical models of nature [78]. Basing the hierarchy problem as a problem with the SM on arguments of naturalness and fine-tuning should be accepted with some caution, however, as this argument relies on unphysical parameters that cannot be measured [79].

A variety of models of new physics attempt to solve the hierarchy problem by either introducing new terms that may provide delicate cancellations of the Higgs boson mass squared corrections or by providing an explanation as to why M_{Pl} is so far from v of the Higgs field. The phenomenology of these models includes new heavy particles that could be detected at the LHC and could provide smoking-gun evidence of new physics perhaps at the TeV scale. Some relevant paradigms of new physics that predict new particles that can couple to Higgs boson pairs are described in the next two sections.

2.2.2 Extended Higgs sector

The choice to include a complex doublet of scalar fields in the SM in order to break the EW symmetry is the minimal one: the four degrees of freedom are enough to give mass to the three gauge bosons and produce one scalar particle. The Higgs sector could very well be more complex, with more degrees of freedom, which would lead to multiple different Higgs fields (and thus particles). The most widely-known class of models that involve an extended Higgs sector come from the concept of Supersymmetry.

Supersymmetry (SUSY) is a local symmetry that can transform fermionic states into bosonic states and vice-versa [80–87]. In the paradigm of SUSY, all the known SM particles have supersymmetric partners that differ by half a unit of the spin quantum number. For example, in the theory, the top quark has a spin-0 supersymmetric partner, known as a "stop," while the gluon has a spin-$\frac{1}{2}$ supersymmetric partner called a "gluino." SUSY provides an elegant and attractive solution to the hierarchy problem, as the new radiative corrections to the Higgs mass due to the inclusion of the SUSY particles in the loops cancel out the typical quadratically divergent corrections. Thus, SUSY provides a mechanism to naturally place the Higgs mass at the EW scale, and it predicts a large spectrum of new particles (the superpartners) that could be produced in high energy proton particle collisions. No SUSY particles have been observed yet with the LHC, though they could be lurking at the TeV scale, a possibility with the circumstance that SUSY is a spontaneously broken symmetry.

Of particular interest in this dissertation is the fact the SUSY requires the presence of two complex scalar doublets rather than the lone one needed in the SM. The extra four degrees of freedom from the second complex scalar doublet relative to what was seen in Section 2.1.5 each give rise to a different scalar field (other Higgs bosons). Thus, in total SUSY requires five physical spin-0 particles: two charged Higgs bosons ($H^{\pm}$), one neutral CP-odd Higgs boson (A), and two CP-even neutral Higgs bosons (H and h). The discovery of the 125 GeV scalar boson consistent with the SM Higgs in 2012 suggests that this state

H be identified with the lightest CP-even Higgs boson h. In these models, there are Hhh couplings, allowing for possible decays of $H \to hh$ in the case that $m_H > 2m_h$. Models with an expanded Higgs sector that are not supersymmetric exist as well. SUSY itself is a specific case of a Two-Higgs-Double Model (2HDM), which plays an important role in some axion models and models that provide the baryon asymmetry observed in the universe [88]. There is an even simpler extension of the scalar sector that contains an extra Higgs singlet [89], giving one extra Higgs boson, which can couple to the lighter one as in the 2HDM.

2.2.3 Extra dimensions

Earlier attempts to unify the forces of nature involved modeling the universe with more than three spatial dimensions, first undertaken by Kaluza and Klein in the 1920s [90, 91]. Kaluza and Klein sought a unified field theory that incorporated both gravity and electromagnetism with the addition of a fourth spatial dimension that is compactified into a miniscule circle with some radius r_C that is much smaller than the length scales that can be detected in an experiment. In this case, the $5D$ Einstein field equations yield the nominal field equations for gravitation (Einstein), electromagnetism (Maxwell), and a new scalar field.

More recently, the notion of extra spatial dimensions has provided a possible solution to the hierarchy problem. In models containing extra spatial dimensions, the extreme weakness of gravity can be explained by allowing the gravitational field to propagate through the extra dimensions. Of particular interest to this dissertation are models with one very small extra dimension that has a warped geometry, as proposed by Randall and Sundrum [2, 92]. Influenced by Kaluza-Klein (KK) theory, the Randall-Sundrum (RS) models envision the (4+1)-dimensional spacetime as consisting of two (3+1)-dimensional subspaces, known as "branes," connected via the (4+1)-dimensional subspace known as the "bulk." Elementary fields are localized at one of the branes, which correspond to different energy scales: the electroweak fields are localized at an infrared (IR) brane, while gravity is localized at an ultraviolet (UV) brane. The UV brane is also known as the Planck brane, while the IR brane is commonly called the TeV brane (as this is where TeV scale resonances may appear). The warped geometry of the bulk diminishes the strength of the gravitational field as it propagates through the bulk from the Planck brane to the IR brane, where it appears to be much weaker than the SM forces. There are two notable scenarios of the RS models: the RS1 scenario confines all SM particles to the IR brane, while the bulk scenario allows them to propagate through the extra dimension.

In the RS models, the geometry of the bulk space is characterized by a spacetime metric with a warp factor that is rapidly decreasing as a function of the extra dimension:

$$ds^2 = e^{-2kr_c\phi}\eta_{\mu\nu}dx^\mu dx^\nu + r_c^2 d\phi^2, \tag{2.33}$$

where k is a factor on order of $M_{\rm Pl}$, r_C is the compactification radius (with πr_C corresponding to the size of the extra dimension), $x_{\mu\nu}$ are the familiar (3+1) spacetime dimensions, $\eta_{\mu\nu}$ is the Minkowski metric, and $0 \leq \phi \leq \pi$ is the coordinate of the extra dimension. The two 3-branes are located at $\phi = 0$ and π. The exponential warp factor is responsible for the hierarchy between the Planck and EW scale and has an important implication for the physical mass of fields located on the the 3-brane where we live: for a mass parameter m_0 in the IR 3-brane, the physical mass is given by

$$m = e^{-kr_c\pi}m_0. \tag{2.34}$$

With $kr_C \sim 10$, and without the need for fine-tuning parameters, Planck brane particles with Planck scale mass can appear with TeV-scale mass on the IR brane, possibly within reach of the LHC. So as not to deal with large numbers, the fundamental parameters of the theory are typically taken to be kr_C and $\tilde{k} = k/\bar{M}_{\rm Pl}$.

Because of the compactified geometry, the RS theory predicts the existence of a spin-2 [93–95] gravitational field along with a spin-0 (scalar) [96–99] field. The scalar field is needed to stabilize the value of r_C, which itself is the vacuum expectation value of the potential of ϕ. The graviton G serves as the mediator of the gravitational field, while the radion R corresponds to excitations of the scalar field. As in KK theory, for both gravitons and radions, these fields correspond not only to one massive state, but to a tower of states with mass m_n proportional to a quantum number n. The zero modes of these states are massless, while the excited states are massive, with the lowest lying ones expected to appear at TeV scales.

In both RS model scenarios, the spin-2 gravitons and spin-0 radions are expected to have significant HH couplings. As seen in Fig. 2.5, in both scenarios, $R \to$ HH decays, with a branching fraction of about 25%, are expected to be second only to $R \to$ WW, constant as a function of m_R above about 500 GeV. As seen in Fig. 2.6, $G \to$ HH is expected to be significant in the bulk scenario, with a branching fraction of about a constant 10% for $m_G \gtrsim 800$ GeV.

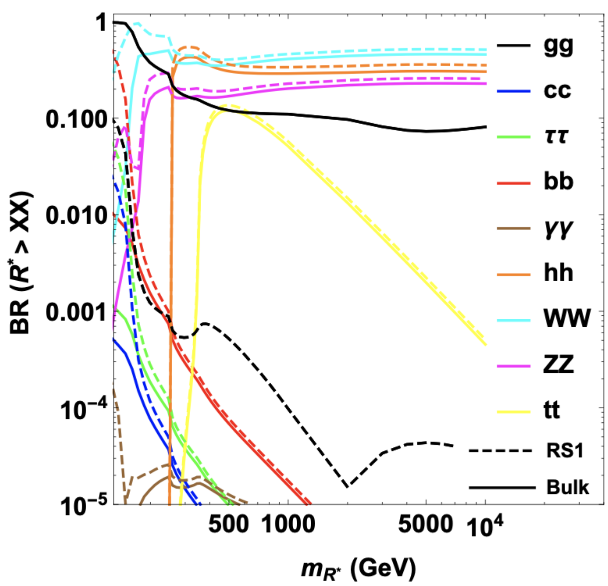

Figure 2.5: Branching fractions for a radion R decaying to various SM particles. Solid lines are used to refer to the bulk RS scenario, while dashed lines refer to the RS1 scenario [2].

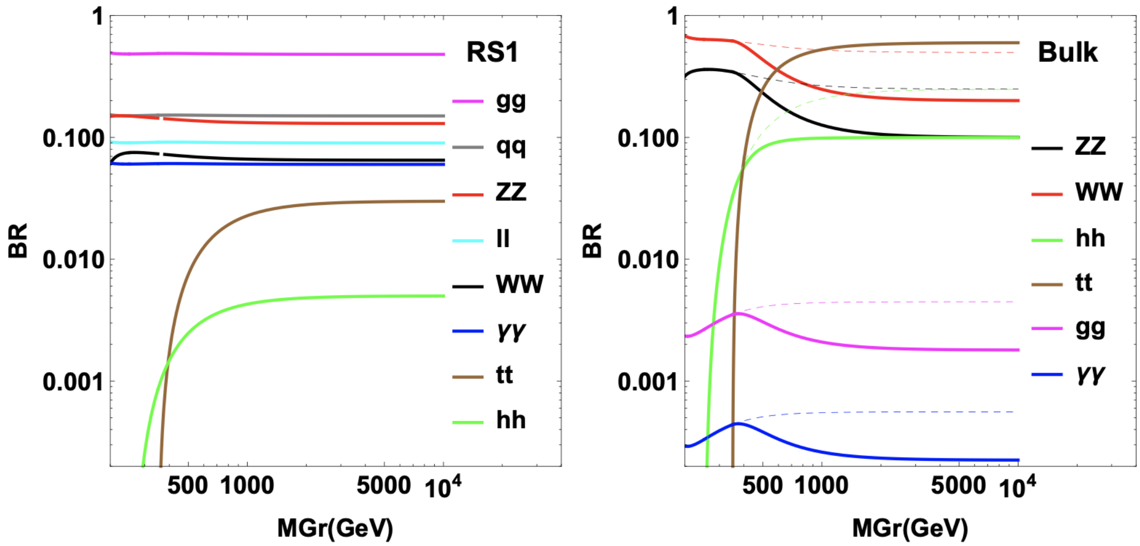

Figure 2.6: Branching fractions for a graviton G decaying to various SM particles for the RS1 (left) and bulk (right) scenarios [2].

2.2.4 HH searches at the LHC

Searches for X → HH in hadron colliders consider spin-0 or spin-2 resonances X (spin-1 X resonances decaying to HH are forbidden). In order to find such resonances decaying to HH, we must consider the optimal decay channels to explore. Typically for each decay channel, there will be unique considerations such as the reconstruction of the final state, the expected amount of signal, and the different dominant backgrounds. As these things vary depending on the decay channel of interest, only certain channels are worthwhile for searching for signal. The most important ones are discussed here.

The branching fractions for decays of HH are shown in Fig. 2.7 assuming the SM mass of 125 GeV. The most common Higgs decay is H → $b\bar{b}$, which occurs with about a 58% probability. Because of this large branching fraction and the ability of CMS to efficiently identify b hadron decays, almost every search for HH involves at least one H → $b\bar{b}$. The HH → $b\bar{b}b\bar{b}$ channel, constituting almost 34% of HH decays, is thus the most common search mode in CMS, explored for both resonant and nonresonant topologies. This search channel suffers from a large multijet background, more severe at lower values of m_{HH}, but at larger values of m_{HH} where the background levels decrease, it is the most sensitive due to the expected signal yield.

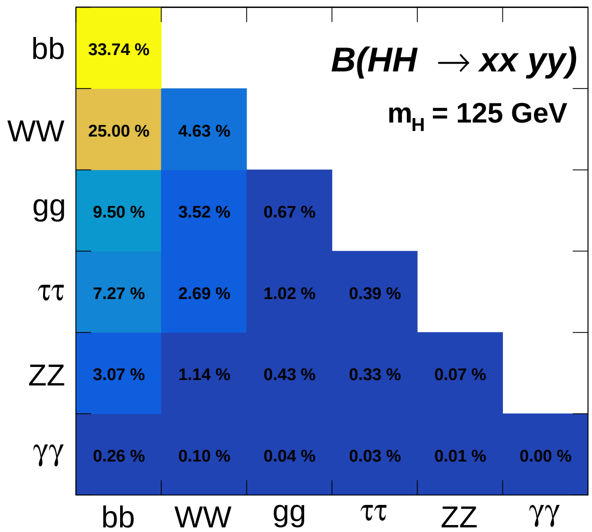

Figure 2.7: Branching fractions for HH decaying into its most notable final states. The $\mathrm{b\bar{b}b\bar{b}}$ and $\mathrm{b\bar{b}WW^*}$ modes together constitute over half of all HH decays.

The next most common decay channel is $\mathrm{HH} \to \mathrm{b\bar{b}WW^*}$, with a total branching fraction of about 25%. The most useful $\mathrm{b\bar{b}WW^*}$ modes include at least one lepton in the final state, which accounts for about 5/9 of the total $\mathrm{b\bar{b}WW^*}$ yield and significantly reduces the background contributions, which are dominated by $\mathrm{t\bar{t}} \to \mathrm{b\bar{b}WW}$. Similar background discrimination can be achieved with $\mathrm{HH} \to \mathrm{b\bar{b}ZZ^*}$, but this suffers from almost 90% fewer expected signal events, so this is not usually analyzed. The next most commonly analyzed channel is $\mathrm{HH} \to \mathrm{b\bar{b}}\tau\tau$, which has less background contamination than the two already mentioned but suffers from a reduced signal yield relative to $\mathrm{b\bar{b}WW^*}$. Significant background discrimination can be achieved with $\mathrm{HH} \to \mathrm{b\bar{b}}\gamma\gamma$, in which two high energy photons provide a very clean signature, but the effectiveness of this channel is limited by the small branching fraction. While the third most probable HH decay channel, $\mathrm{HH} \to \mathrm{b\bar{b}}gg$ suffers from an even larger multijet background than HH $\rightarrow$ $b\bar{b}b\bar{b}$ and only about 30% of the expected signal, so it is not typically analyzed.

In this dissertation, both the $b\bar{b}WW^*$ and $b\bar{b}\tau\tau$ modes with leptons in the final state are analyzed. By doing this, we maintain high acceptance to signal events from HH decays while suppressing large backgrounds from multijet events. Three different final states are used, which are shown with Feynman diagrams in Fig. 2.8. The dilepton final states of the $b\bar{b}WW^*$ and $b\bar{b}\tau\tau$ modes can be reconstructed with a common set of selections, as they only have small differences in the lepton p_T and p_T^{miss} spectra. However, reconstructing the single-lepton final states of these two modes requires different set of selections, and so the single-lepton final states of HH $\rightarrow$ $b\bar{b}\tau\tau$ are not considered.

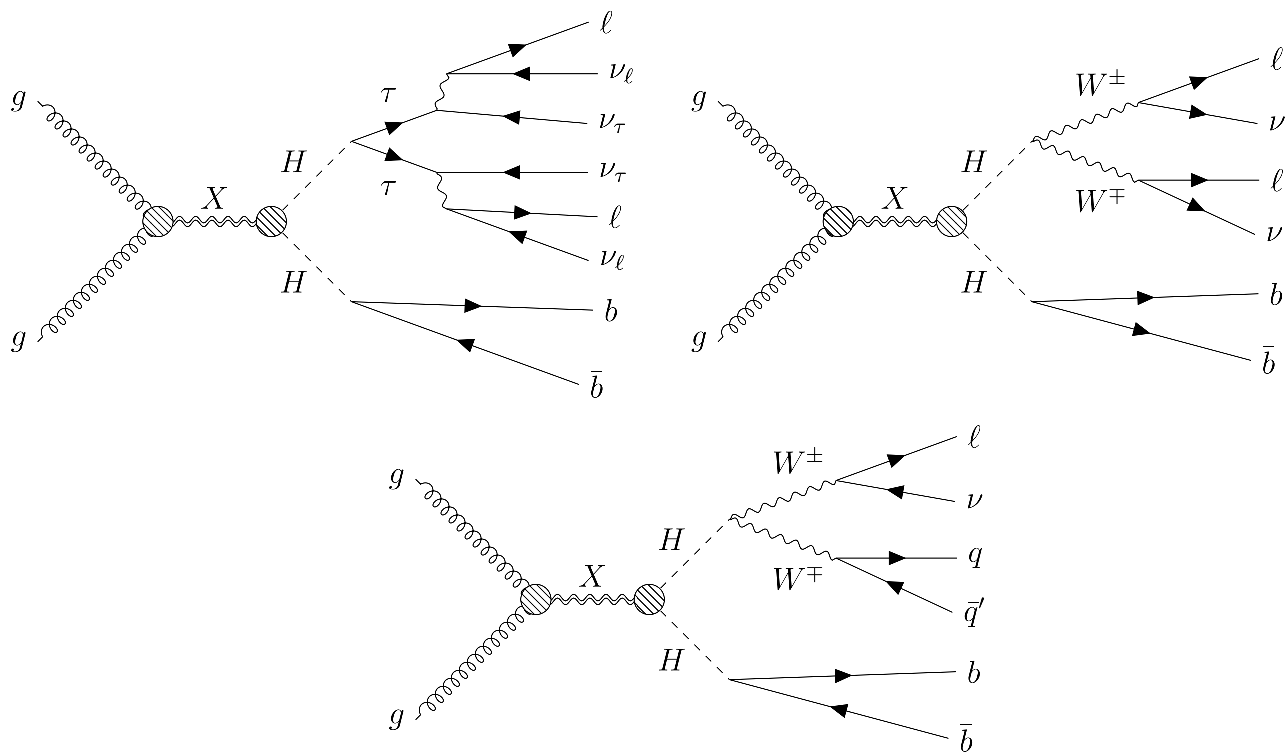

Figure 2.8: Feynman diagrams for the signal processes considered in the analysis. The dilepton processes are depicted on top, with HH $\rightarrow$ $b\bar{b}\tau\tau$ on the left and HH $\rightarrow$ $b\bar{b}WW^*$ on the right. The single-lepton process from HH $\rightarrow$ $b\bar{b}WW^*$ is shown on the bottom diagram.

CHAPTER 3

Experimental Apparatus

This chapter describes the experimental apparatus used to search for BSM phenomena in this dissertation. To search for heavy resonances that push the boundaries of the highest energy scales accessible, it is most advantageous to collide massive particles in a circular geometry using intense and high energy counter-rotating beams. In Section 3.1, we describe the Large Hadron Collider (LHC), the machine used to accelerate and collide high energy protons in this search for BSM physics. In Section 3.2, we describe the Compact Muon Solenoid (CMS) experiment, the detector used to image the LHC beam collisions and reconstruct hundreds of particles in an event. In Section 3.3, we describe the reconstruction of basic physics objects using signals from the various CMS subdetectors reviewed in Section 3.2. These basic physics objects serve as inputs for the global analysis selection that is described in Chapter 4.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator, primarily operating with beams of protons but also capable of producing beams of heavy ions. Protons and heavy ions are well suited for the LHC's high energy physics objectives because they do not lose as much energy due to synchrotron radiation as electrons. The LHC, a 27 km ring in circumference and located about 100 m below the ground, straddles the Swiss-French border just outside of Geneva, Switzerland and is located at the European Organization for Nuclear Research (CERN).

The LHC tunnels contain two counter-rotating beams of particles, each contained within their own high vacuum pipe at energies of 6.5 TeV (during Run-II), that merge and collide at four different interaction points (IPs). Beams are organized into localized bunches of particles; radiofrequency (RF) cavity electric fields are used to accelerate the bunches, and a variety of magnetic fields are used to steer (bend) and focus them. These fields are produced with a system of 9593 magnets, 1232 of which are superconducting dipole magnets for bending, and 392 of which are quadrupole magnets for focusing. The magnets are cooled using superfluid helium that is circulated in the LHC [100].

At each of the four IPs on the LHC, there exists a large-scale detector. There are two low luminosity experiments, ALICE and LHCb, each focusing on specialized physics programs: ALICE is primarily used to study quark-gluon plasmas in heavy ion collisions, and LHCb focuses on b physics and tests of CP violation. The two high luminosity experiments, ATLAS and CMS, focus on similar general-purpose physics programs, which include the discovery and study of the Higgs boson, precision measurements, and searches for BSM physical phenomena. This dissertation analyzes pp collision data collected by the CMS detector, which is described in Section 3.2.

In order to reach high energies of 6.5 TeV, protons must be accelerated in several stages, accomplished with CERN's complex of accelerators, displayed in Fig. 3.1. Protons used for LHC collisions are first obtained at the linear accelerator facility Linac2 from bottled hydrogen gas, which is released into electric fields that strip away the electrons from the protons. These protons are accelerated to energies of 50 MeV before entering the Proton Synchrotron Booster (PSB), which further accelerates the protons to 1.4 GeV. Protons are then injected into the Proton Synchrotron (PS), accelerated to 25 GeV, and then fed into the Super Proton Synchrotron (SPS) to be accelerated to 450 GeV. At this energy, the SPS finally injects the protons into the LHC, which itself then accelerates them to the operating beam energy of 6.5 TeV.

The crucial task of this machine is to produce as many proton collisions as possible, since

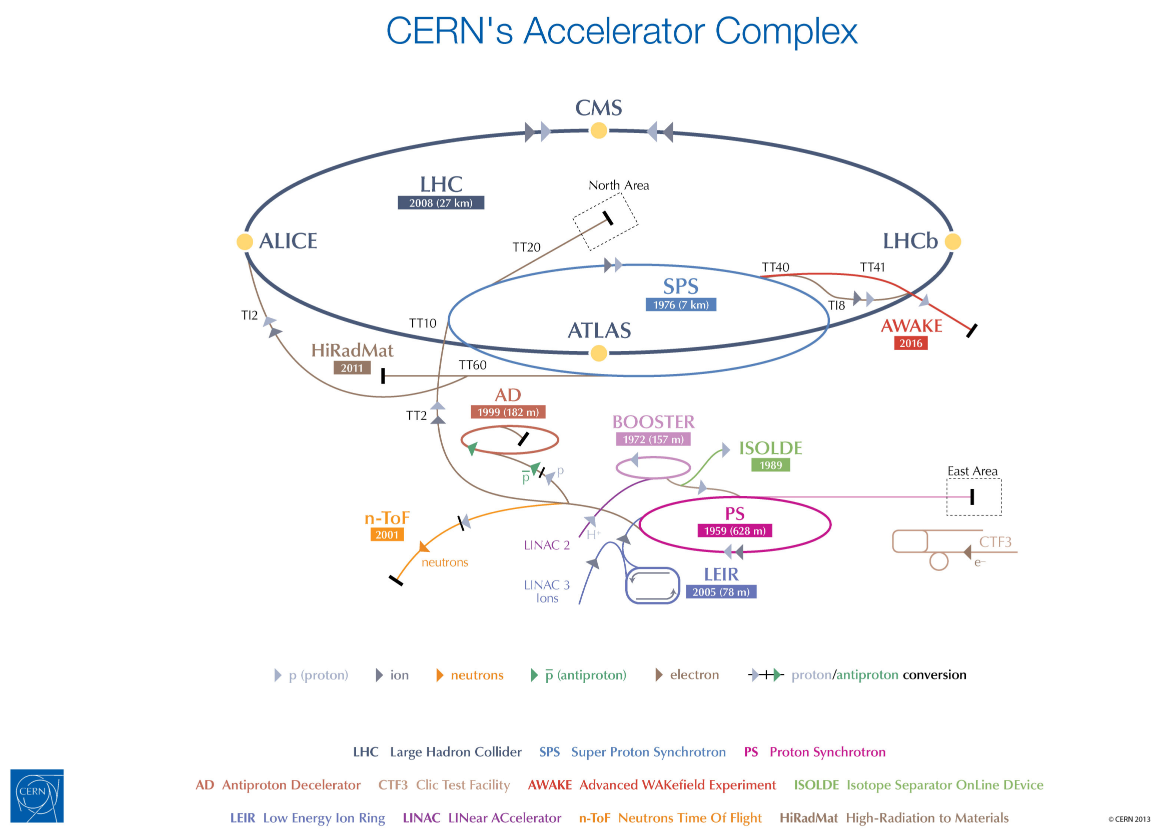

Figure 3.1: Schematic overview of CERN's accelerator complex, currently used to accelerate protons to 6.5 TeV [3].

we are now probing BSM phenomena with increasingly smaller interaction cross sections. The number of events N we can expect to produce is given by

$$N = \sigma \int \mathcal{L}(t)dt = \sigma \mathcal{L}_{int}, \tag{3.1}$$

where σ is the interaction cross section of the process of interest, $\mathcal{L}(t)$ is the instantaneous luminosity, and $\mathcal{L}_{int}$ is the integrated luminosity over some fixed amount of time. The luminosity is completely dependent on the design and operation of the collider, and for a general collider with bunched beams it can be expressed as

$$\mathcal{L} = f_{coll} \frac{n_1 n_2}{4\pi\sigma_x^* \sigma_y^*} \mathcal{F}, \tag{3.2}$$

where f_{coll} is the collision frequency, n_1 and n_2 are the number of particles in each colliding bunch, σ_x^* and σ_y^* are the rms transverse beam sizes in the horizontal (bend) and vertical directions at the IP, and $\mathcal{F}$ is a factor of order 1 that accounts for geometric effects such as the beam crossing angle and finite bunch size. Thus, the magnet and the RF systems are designed such that we can pack bunches as densely as possible with protons at the IP and collide them as frequently as possible.

During Run-II, a maximum instantaneous luminosity of 2.1×10^{34} cm^{-2}s^{-1}, double that of the LHC design, was achieved, and the LHC delivered 163.5 fb^{-1} of total integrated luminosity. Both the peak instantaneous luminosity and the total integrated luminosity for Run-II are shown in Fig. 3.2. A maximum of 2556 bunches, each initially containing around 1.1×10^{11} protons, could populate the LHC beams. The bunch spacing was 25 ns, corresponding to a 40 MHz collision rate [101].

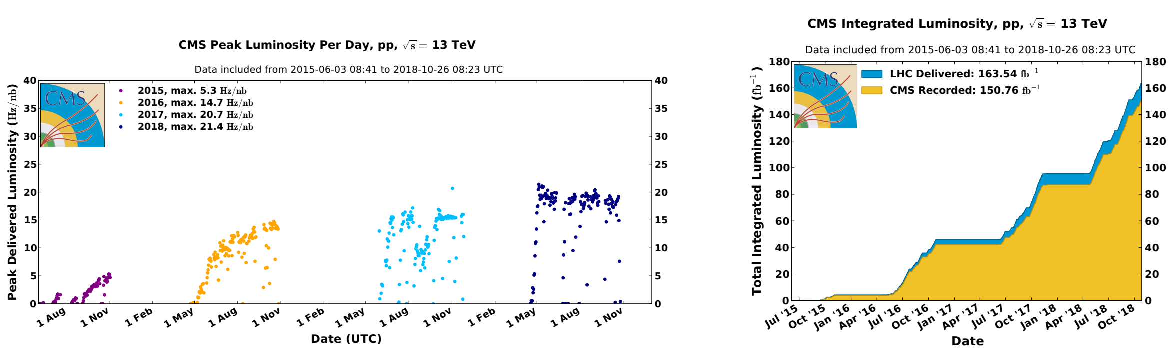

Figure 3.2: The LHC performance during Run-2 (2015 to 2018). Left: peak instantaneous luminosity for each day of operation. The maximum instantaneous luminosity was achieved early in the 2018 run. Right: the running total integrated luminosity as a function of time, delivered by the LHC and recorded by the CMS detector.

3.2 The CMS Detector

The Compact Muon Solenoid (CMS) experiment is a cylindrical general-purpose detector capable of reconstructing physics events at the TeV scale and consisting of a barrel and two endcap sections of detectors. The detector has a diameter of 15 m and a length of 28.7 m. Interesting physics events, such as the ones concerned in this dissertation, typically include highly unstable particles that promptly decay to other more stable particles before they can interact with the detector directly. The goal of CMS is to precisely measure the momenta and trajectory of all these more stable particles, which include muons, electrons, photons, and hadrons such as pions. This information can be used to reconstruct the intermediate and initial states in the full event.

The namesake feature of the CMS detector is its superconducting solenoid magnet, which achieves a roughly uniform axial field of 3.8 T that is crucial for bending particles in order to measure their momenta precisely. The magnet itself has a length of 12.5 m and a diameter of 6 m. The solenoid magnet is large enough to enclose the silicon tracking system, the electromagnetic calorimeter (ECAL), and the hadronic calorimeter (HCAL), in the order of increasing distance from the beamline. Outside the magnet, there is a final thin layer of HCAL surrounded by a system of muon detectors. The muon system is interspersed with steel yoke that facilitates the return of the magnetic field lines through the solenoid. Figure 3.3 shows a cross-sectional view of CMS and its major detector subsystems.

3.2.1 CMS coordinate system

The coordinate system of CMS is defined with its origin at the IP of the crossing proton beams. The $\hat{z}$ axis is taken to be along the direction of the beamline. The $+\hat{x}$ direction is the bend axis, pointed toward the center of the LHC ring, and the $+\hat{y}$ direction is pointed upwards toward the surface of the earth.

Since CMS has cylindrical geometry, a polar coordinate system is most commonly used to

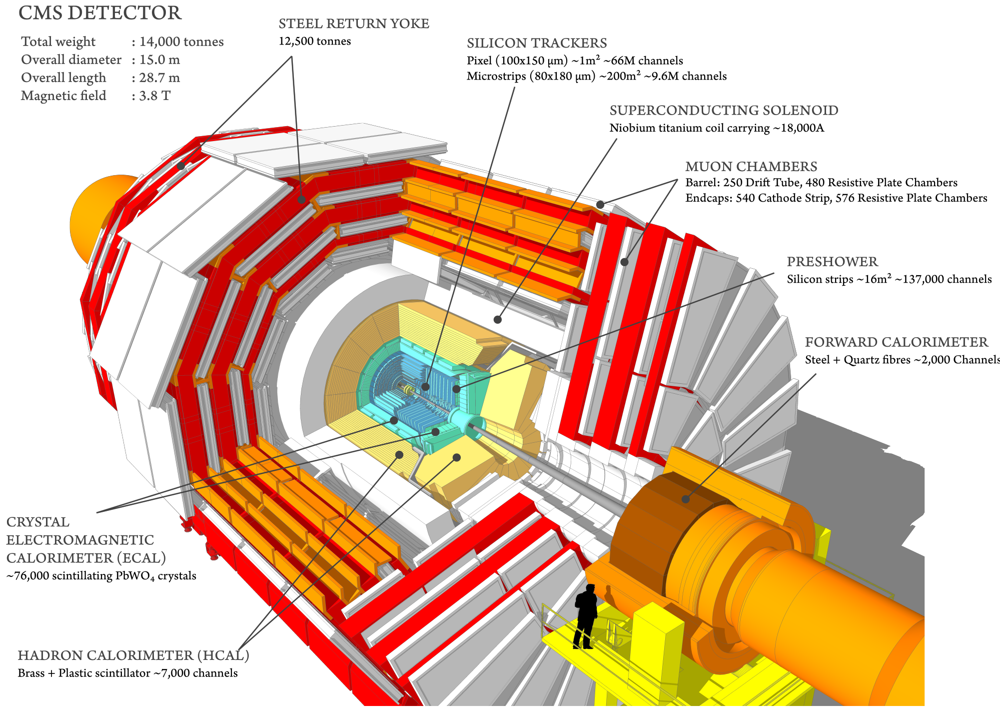

Figure 3.3: Layout of the CMS detector, with the major subsystems labelled and a black figure for scale [4].

describe physics in the transverse plane. Since the beams are directed along $\hat{z}$, there is zero total momentum in the transverse plane, and particle quantities in this plane are unaffected by the incoming states. In the cylindrical coordinate system, r is defined as the distance from the beamline in the transverse plane, ϕ is the azimuthal angle measured with respect to the $+\hat{x}$-axis, and θ is the polar angle measured with respect to the $+\hat{z}$-axis. In hadron collider physics, another variable is used instead of θ: the pseudorapidity η, defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) = \frac{1}{2}\ln\left(\frac{p+p_z}{p-p_z}\right), \tag{3.3}$$

where p and p_z are the magnitude and z-component of a particle's three-momentum, respectively. Differences in η are invariant under Lorentz boosts along the beamline direction. The spatial distance between two particles can then be expressed in terms of their angular separation as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The cylindrical geometry of CMS and the physics of hadron colliders also create an environment in which measuring the transverse momentum p_{T} is important. By measuring the p_{T}, η, and ϕ of particles using a variety of detector subsystems, CMS can reconstruct the full event.

3.2.2 Inner tracking system

The tracking system is the innermost part of the CMS detector and is crucial for accurately and precisely measuring final state particles. Due to the high magnetic field within CMS, charged particles produced in the collisions will bend, and the tracks left behind in the detector can be used to measure their momenta. The tracker is capable of reconstructing the trajectories of hundreds of particles and the displaced secondary vertices from b hadron decays (particularly important in this analysis where we search for $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$ decays) with the use of silicon (Si) tracking modules.

A Si tracking module essentially acts as a solid-state ionization chamber. When a charged particle moves through a silicon detector operated under an applied electric field, the absorbed energy creates electron-hole pairs, which drift toward their respective electrodes, after which an electronic signal is collected, amplified, and then read out. These signals may be identified as "hits" in the tracker and are then used to reconstruct particle trajectories, which are used the measure quantities like p_{T}, η, ϕ, and the impact parameter. Silicon is also useful for its small bandgap energy: electron-hole pairs can be created with only a small amount of absorbed energy. Considering this, the silicon modules are also constructed with low amounts of material so that particles do not lose significant energy before their energies are ultimately measured by the calorimeters or muon systems. The tracking system is composed of an inner layer of Si pixels and an outer layer of Si strips, both of which have a cylindrical barrel and two endcap disk sections, giving a detection coverage of $|\eta| < 2.5$ [102].

3.2.2.1 Pixel Detector

The highest granularity is in the pixel layer, where there is the largest density of tracks in an event due to the proximity to the IP. The layer is composed of approximately 123 million Si pixels, each with an area of 100 μm $\times$ 150 μm and a thickness of about 300 μm [5]. At the beginning of Run-II, the pixel detector consisted of three layers in the barrel region (BPIX) and two layers in the endcap region (FPIX), but another layer was added to each during the shutdown between the 2016 and 2017 data-taking periods. This development particularly improved the efficiency for measuring displaced secondary vertices in b decays, which noticeably improved this analysis. The current four BPIX layers are 54 cm in length and are located at mean radii of 2.9 cm, 6.8 cm, 10.9 cm, and 16 cm with respect to the beamline. The current three FPIX layers are located at $z = \pm 29.1$ cm, ± 39.6 cm, and ± 51.6 cm with respect to the IP, with radii ranging from 4.5 to 16.1 cm. The spatial resolution of charged particle hits for track measurement is 10 μm for $r - \phi$ position measurements and 20 μm for z position measurements. The composition of the pixel detector before (2016) and after (2017 and 2018) the Phase-1 upgrade is shown in Fig. 3.4.

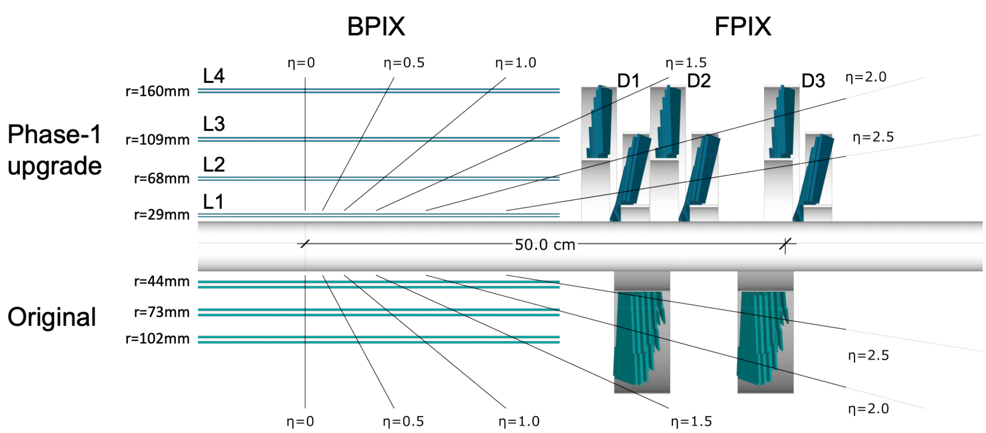

Figure 3.4: The CMS pixel detector, the innermost section of the Si tracking system. The image compares the pixel detector configuration during 2016 (original) and during 2017 and 2018 (Phase-1 upgrade). The barrel (forward) region is denoted as BPIX (FPIX) [5].

3.2.2.2 Strip detector

The density of particle tracks rapidly falls further away from the IP, and so the Si strip tracker outside of the pixel detector does not require as many small modules as the collection of pixels. The strip tracker is divided into two different parts, each with differing strip sizes, in both the barrel and endcap regions: strips of 10 cm in length compose the inner regions, while strips of 25 cm are in the outer regions.

The tracker inner barrel and disks (TIB and TID) cover the radial region between 20 and 55 cm outside of the pixel detector. In the TIB, there are four layers of strips, with each layer about 140 cm in length. In the TID, there are three disks at each endcap located between $z = \pm 80$ cm and $z = \pm 90$ cm. The tracker outer barrel (TOB) region is set outside of the TIB and TID, extending to a radius of 116 cm and containing six layers spanned between $z = \pm 118$ cm. The tracker endcap modules (TEC) contain nine discs, located between ± 124 and 280 cm and extending from 22.5 to 113.5 cm in radial distance from the beamline.

3.2.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL), surrounding the Si tracker, is primarily designed to measure the energies of electrons and photons [103]. It is also the location for the initiation of many hadronic showers. The ECAL is a homogeneous calorimeter with high granularity, constructed with high-Z lead-tungstate ($PbWO_4$) crystals. When an electron or photon is incident upon such a crystal, the particle decays into an EM shower, and the showering particles further interact with the ECAL material to produce scintillation light. This scintillation light can be collected by photodetectors attached to the crystal cells, after which the signal can be amplified and read out with electronics. The $PbWO_4$ crystals are chosen and designed such that the energies of incident electrons and photons can be fully absorbed within a small crystal volume, necessary to achieve a strong energy resolution. In particular, the $PbWO_4$ crystals have high density (8.3 g/cm^3), a small radiation length ($X_0 = 0.89$ cm), a small

Molière radius (2.2 cm), and a scintillation decay time similar to the LHC bunch spacing of 25 ns, enabling fast signal readout that minimizes interference from adjacent bunch crossings. The high density allows the whole detector to be compact and fit well within the CMS solenoid. The energy resolution of the ECAL, for particle energies below 500 GeV, improves with higher energies and can be parameterized as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{3.4}$$

where S is a stochastic factor, N is a noise factor, and C is a constant term. The stochastic term in the energy resolution arises from inherent effects such as intrinsic shower fluctuations and photo-electron statistics, while the noise term is due to both the electronics noise summed over all the readout channels and the effects of pileup. The constant term arises from detector non-uniformity, calibration uncertainties, and radiation damage of the active media. The ECAL design is also optimized to include radiation hard media and electronics components to reduce this effect in the uncertainty.

The ECAL, shown in Fig. 3.5, is constructed with a barrel section consisting of 61200 $PbWO_4$ crystals and two endcap sections, each containing 7324 crystals. In the barrel (EB) section, the crystals are arranged in an $\eta - \phi$ grid, and they each are 23 cm long (25.8 X_0) with cross sectional area of 2.2×2.2 cm^2, such that about a whole Molière radius encapsulates the area of a 2×2 collection of crystals. In the endcap (EE) section, the crystals are arranged in an $x - y$ grid, are 22 cm long (24.7 X_0), and have a cross sectional area of 2.86×2.86 cm^2. The EB region of the detector provides a detector coverage of $|\eta| < 1.479$, while the EE region covers $1.479 < |\eta| < 3.0$.

A sampling preshower detector (ES) is located right in front of the two endcaps, providing coverage for $1.654 < |\eta| < 2.6$, and is primarily used to improve the identification of individual photons from $\pi^0 \to \gamma\gamma$ decays. The preshower detector contains two sections: one of lead absorber to induce EM showers from electrons or photons, and then behind this,

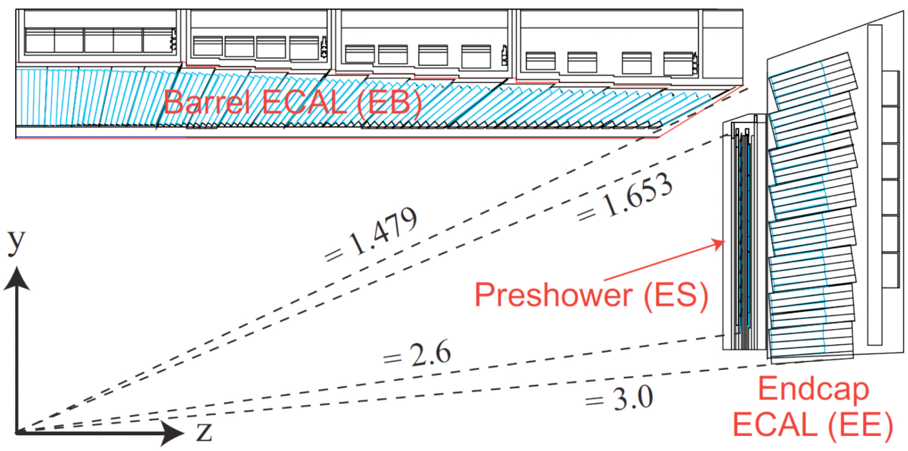

Figure 3.5: Longitudinal view of the ECAL system, with indications of the acceptance in $|\eta|$ of each of the detector regions. The barrel region (EB), endcap region (EE), and preshower detector (ES) are depicted [6].

Si strip sensors to measure the energy and spatial profile of showers.

3.2.4 Hadronic calorimeter

Hadrons typically begin showering in the ECAL, but the nuclear interaction length of the ECAL material is insufficient to contain their energies. Surrounding the ECAL, the hadronic calorimeter (HCAL) is a hermetic sampling calorimeter, made of alternating layers of brass absorber and plastic scintillator, designed to fully absorb and measure the energies of these hadrons [104]. A hadron incident on the brass absorber will begin showering, and the secondary particles produced can interact further with successive layers or brass. When the shower particles pass through the plastic scintillator tiles, scintillation light is emitted, collected by wavelength-shifting fibers, and used to measure the energies.

The HCAL is divided into four distinct regions: the barrel (HB), the endcap (HE), the outer calorimeter (HO), and the forward calorimeter (HF), all shown in Fig. 3.6. The primary parts of the detector, HB and HE, cover the regions $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. The detectors are segmented into "towers" constructed of alternating layers of absorber and scintillating tiles. Each of the towers covers a $\Delta\eta \times \Delta\phi$ area of 0.087×0.087 in the HB and of 0.17×0.17 in the HE. While the HB section encompasses the radial volume between the ECAL and the solenoid, the EB and HB collectively do not provide enough material to fully contain hadronic showers. The tails of these showers are detected with the HO sections, located just outside of the solenoid and covering $|\eta| < 1.3$. The HO detectors also partially rely on the solenoid to induce further hadronic showering for energy measurement. In order to detect hadronic particles closer to the beamline, the HF detector is placed 11.2 m from the IP, beyond the muon system, covering $3.0 < |\eta| < 5.2$. To withstand the high radiation in this forward region, the detector is outfitted with quartz fibers instead of plastic for the scintillating material and more steel than brass for the absorber material.

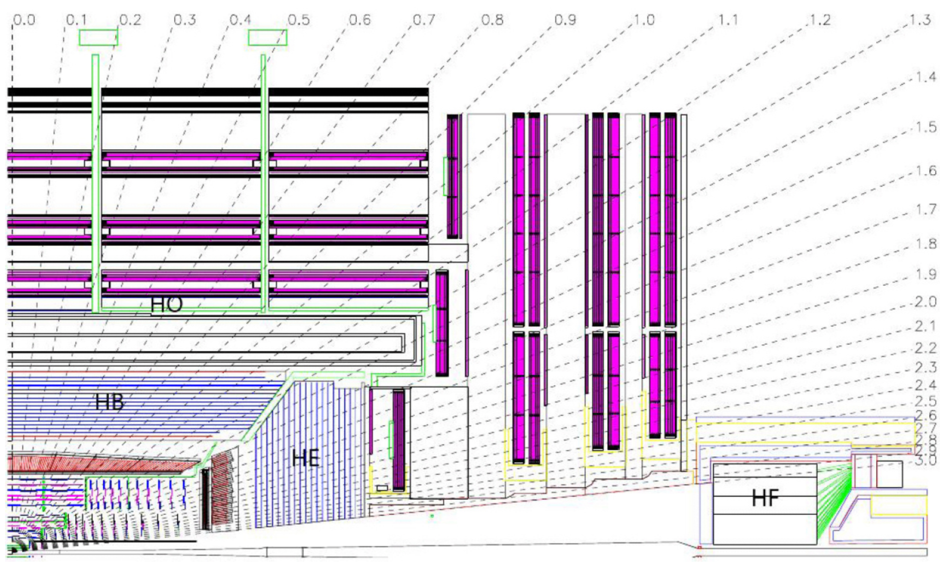

Figure 3.6: Longitudinal view of the CMS HCAL system [7]. The barrel (HB), endcap (HE), outer barrel (HO) and very forward (HF) regions are indicated in the figure.

3.2.5 Muon system

Due to their relatively long lifetime and energetic properties, muons are the only detectable particles that typically reach beyond the solenoid and outer HCAL system. A gaseous detector system designed specifically for muons is placed in the outermost region of the CMS detector [105, 106]. The muon chambers are interspersed between the iron return yoke for the magnetic fields, necessary to bend passing muons and measure their momenta. Three different types of muon detectors are used in the muon system, considering expected background rates and other factors related to performance and cost. Drift tube (DT) chambers are used only in the barrel region, cathode strip chambers (CSCs) are employed in the endcap region, and resistive plate chambers (RPCs) are employed in parts of both regions. A diagram showing the layout of these detectors is seen in Fig. 3.7.

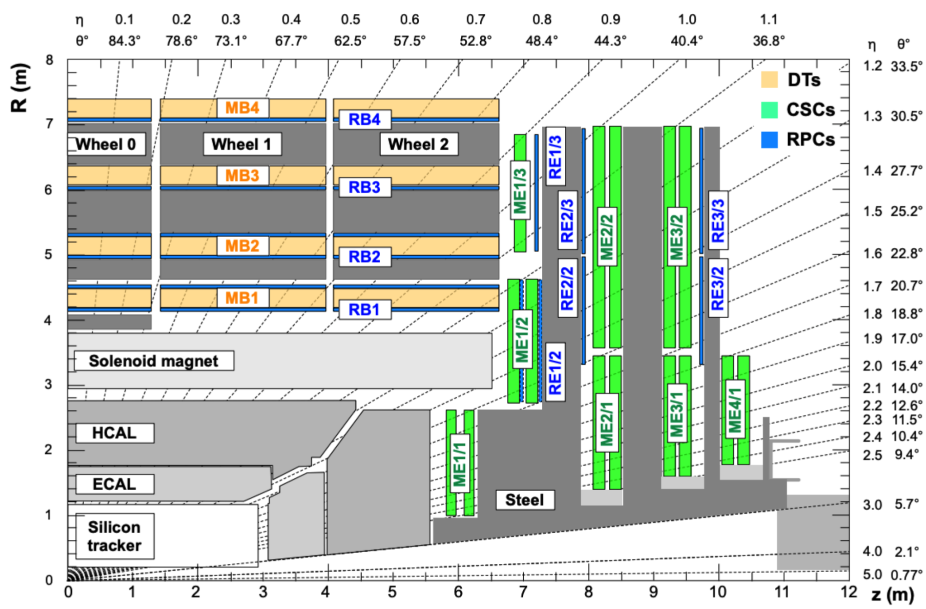

Figure 3.7: Longitudinal view of the CMS muon system [8], highlighted by the colored parts of the detector. Drift tubes (DTs) are shown in yellow, cathode strip chambers (CSCs) in green, and resistive plate chambers (RPCs) in blue.

In the barrel region, the DT system consists of five wheels, each with four radial stations, providing pseudorapidity coverage up to $|\eta| < 1.2$. In each station, DT chambers overlap along ϕ to eliminate gaps in the detection system. The DTs are segmented into long aluminum drift cells, each with a cross-sectional area of 4.2×1.3 cm^2. The DTs contain an Ar/CO_2 gas mixture and an anode wire with a 50 μm diameter running through the center of the gas volume. The walls of the cell are covered with electrodes above and below the wire and cathode strips on the sides. A muon that passes through the gas volume will produce ionization electrons that drift toward the anode wire with a uniform velocity of about 55 μm/s. By measuring the drift time, the position of the muon in the plane orthogonal to the wire can be determined, with better than a 250 μm resolution per cell. Each DT chamber contains three "superlayers" (SLs) grouped radially, each of which contains four staggered layers of drift cells. In the inner and outer SL, the anode wires are oriented parallel to the beamline to provide a measurement in the $r - \phi$ plane, while the middle SL measures the $r - z$ position with wires perpendicular to the beamline.

Located in the endcap region, the CSCs are multi-wire proportional chambers that are arranged into four disks at each end and provide a pseudorapidity coverage of $0.9 < |\eta| < 2.4$. The CSCs have fine segmentation, fast response, and are radiation hard, all necessary to operate in an environment with high particle rates and strong non-uniform magnetic fields. The chambers are trapezoidal, containing cathode strips running radially that precisely measure the muon $r - \phi$ position and anode wires orthogonal to them that provide η measurements. The chamber is organized into six layers of anode wires interspersed between seven layers of cathode strips, with six independent gas volumes containing an $Ar/CO_2/CF_4$ mixture that is ionized by a passing muon. The CSCs are capable of identifying the bunch crossing of a *pp* collision and can provide a typical time resolution of 3 ns and spatial resolution of $40 - 150$ μm per chamber.

The RPCs are double-gap chambers consisting of electrode plates enclosing a gas volume that is a mixture of $C_2H_2F_4$, C_4H_{10}, and SF_6. RPCs are placed in both the barrel and endcap near the DTs and CSCs, covering the pseudorapidity range $|\eta| < 1.9$. The detectors are operated in avalanche mode at a high electric field, providing the best time resolution of any of the muon detectors at about 1.5 ns. The RPCs are primarily used for accurate timing and fast triggering, as they have a the worst position resolution of any muon detectors at 0.8–1.3 cm.

3.3 Physics object reconstruction

Proton-proton (pp) collisions occur in the crossing of LHC bunches every 25 ns, corresponding to a 40 MHz rate. Not only is this far too high a rate at which event data can be stored, but most of the collisions are not interesting enough to search for new phenomena or perform measurements of standard model physics. Most of the collisions are low-energy glancing interactions between protons, yielding many events dominated by QCD processes and jets found near the beamline at high $|\eta|$. On the other hand, interesting events for searches and precision measurements often involve electroweak processes and usually arise from high-energy, head-on collisions of protons. Figure 3.8 shows the production cross sections of such interesting SM processes. Compared to the total inelastic pp collision cross section of $\sigma_{\mathrm{pp}} \approx 10^{11}$ pb [107], even the most common interesting processes occur six or more orders of magnitude less frequently.

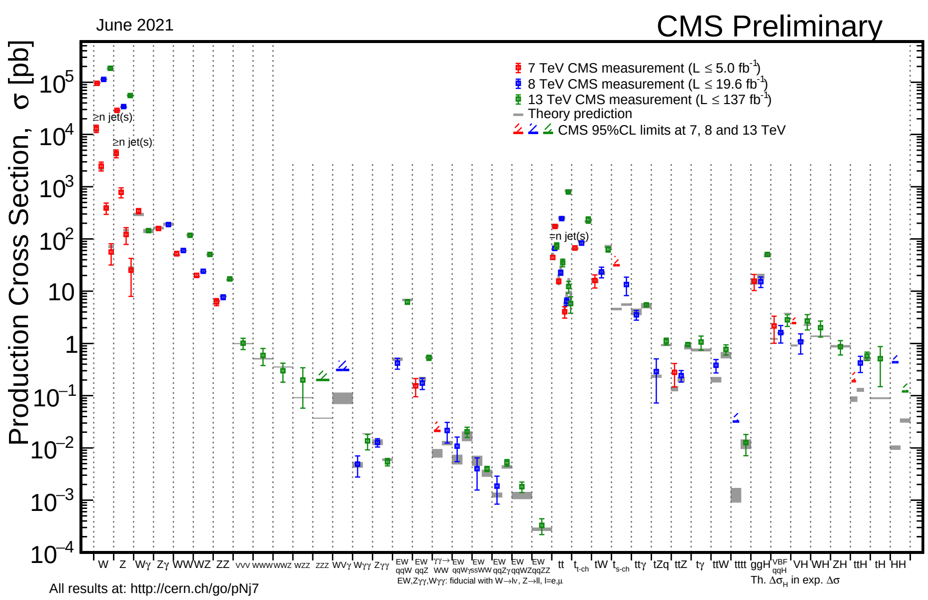

Figure 3.8: Interaction cross sections for various SM processes of interest in the LHC physics program [9]. Measurements from 2016–2018 at 13 TeV are shown in green, while predictions from theory are shown in gray filled bands.

In order to select the most interesting events and reduce the recorded event rate to some feasible amount, CMS employs a complex two-tier trigger system. This trigger system performs a low-level (online) reconstruction of the event using algorithms of varying complexity and makes a decision on whether to write the data to tape offline. Events that are stored offline are processed with more advanced, computationally intensive algorithms that aim to reconstruct all particles produced in the event. After this offline processing, data sets can be used for analysis.

3.3.1 Trigger system

The trigger system of CMS is the first step in the event selection process, and it is two-tiered. The first level of triggering, called Level 1 (L1), consists of on-detector custom hardware that processes and selects events at a rate of 100 kHz using information from the calorimeters and muon system [108]. The next level, referred to as the high-level trigger (HLT), is composed of a conventional farm of CPUs running a more detailed event reconstruction in software that reduces the event rate to about 1 kHz [109]. The L1 trigger is fast in order to quickly discard uninteresting events, while the HLT system contains more complex and time intensive algorithms.

The L1 trigger is hardware-based and operates with a fixed latency of about 4 μs; in this amount of time after a collision, the system must decide whether the event is tentatively accepted or rejected. This decision is made using coarsely segmented data, known as trigger primitives, received only from the calorimeters and the muon system. These trigger primitives are built from energy deposits in the calorimeters and track segments or hit patterns in the muon system. The trigger primitives are then grouped and combined to form trigger objects that correspond to candidate electrons, photons, muons, and jets. The trigger objects are then sorted by their transverse energies and momenta and some reconstruction quality criteria. The highest ranked objects, corresponding to high energy and reconstruction quality, are sent to the Global Trigger (GT), which employs a programmable menu of up to 128 selection algorithms that are used to determine if an event is desired. During this process, high-resolution data is stored in pipelined memory on the front-end electronics. If the GT decides to tentatively select an event, an L1 accept signal is generated, and the high-resolution data is passed on to the HLT more complex evaluation [110].

After the L1 accept signal is generated, the HLT receives the full precision data from all subsystems in the detector, now including the tracker in addition to the calorimeters and muon detectors. The HLT is software-based and processes data with a similar but faster version of the offline reconstruction software. Data is processed in a modular manner via trigger paths, a set of algorithmic processing steps runs in an order with increasing complexity that reconstructs physics objects in the full event and makes selections on these objects. Selections relying on calorimeter and muon system information are made first to reduce the rate before selections that rely on more CPU-expensive tracking reconstruction. Events that pass the full set of selections in an HLT path are written to tape. Overall, the HLT system operates with a latency of about 200 ms and reduces the rate functionally down to a few hundred kHz.

3.3.2 Global event reconstruction

Reconstruction of individual elementary particle objects is possible with the CMS detector because of its strong magnetic field, highly granular tracker and ECAL, hermetic HCAL, and excellent highly efficient muon tracking system. This endeavor relies on the Particle Flow (PF) algorithm [10], which correlates the information sent from all detector subsystems to identify individual candidate particles such as charged and neutral hadrons, electrons, photons, and muons. Each of these particles leaves a distinct characteristic signature, illustrated in Fig. 3.9, that is used by the PF algorithm. The trajectories of charged particles, known as tracks, are reconstructed using the Si tracker and are matched to other detector signatures for further identification. Electron tracks are matched to ECAL deposits, charged hadrons to both ECAL and HCAL deposits, and muons to additional tracks in the muon system outside the solenoid. A photon is identified as a cluster of energy deposited only in the ECAL without an accompanying track, while neutral hadrons also leave no track and mainly deposit a cluster of energy in the HCAL.

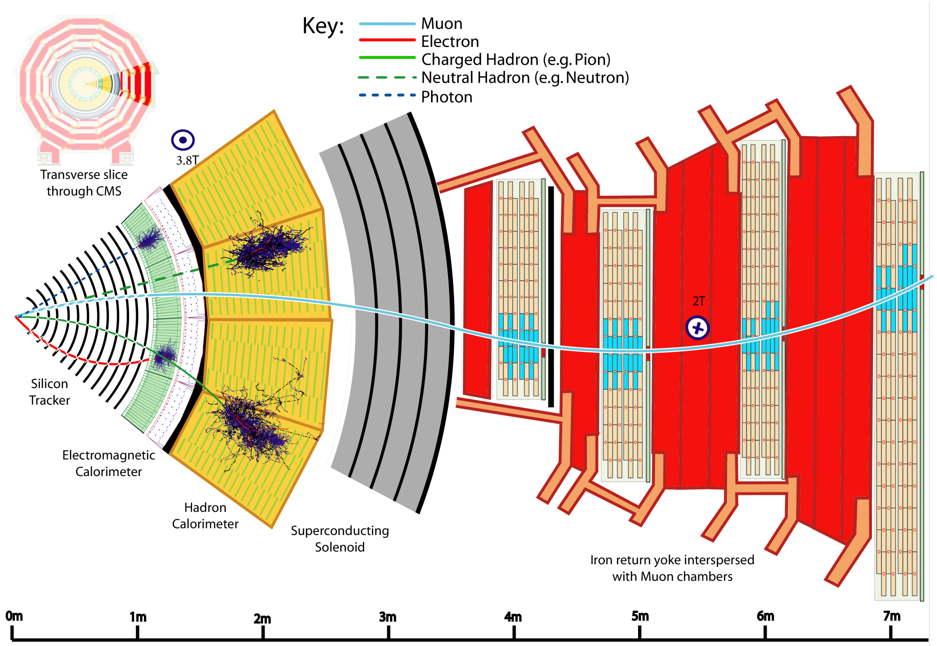

Figure 3.9: Transverse slice of the CMS detector with a depiction of specific particle interactions extending from the beamline through the muon system [10]. The typical detector signatures of photons, neutral and charged hadrons, electrons, and muons are shown. In the image, the muon and charged hadron are positively charged, while the electron is negatively charged.

The PF algorithm operates by first building basic elements for particle identification: tracks and clusters. Reconstruction of tracks occurs in three stages: initial seeds are built from a few hits compatible with a charged-particle trajectory, and then pattern recognition is used to to gather all the hits from other tracker layers along the trajectory. Finally, the track is refitted to obtain the charged particle properties, including the vertex, transverse momentum, and direction. An iterative tracking approach is employed to increase efficiency with a marginal change in misreconstruction rate. Reconstruction of PF clusters begins by identifying local maxima of energy (seeds) deposited into the calorimeter. Then, energy clusters are built around these seeds according to topological and energetic constraints before even smaller clusters are identified within each initial cluster with an expectation-maximization algorithm based on a Gaussian-mixture model.

Once constructed, PF elements from different subdetectors are associated with each other into "PF blocks" using a linking algorithm. The PF blocks are then fed into the PF algorithm as inputs for reconstructing individual particles; these output particles are known as the PF candidates. In the search for X $\rightarrow$ HH, reconstruction and identification of both the intermediate and final state particles relies heavily on the use of the PF candidates.

In each bunch crossing of the LHC beams, not just one but many pp collisions occur, as the bunches contain many protons. An interesting event typically involves a single hard-scatter interaction, referred to as the primary vertex (PV). All other vertices are referred to as pileup. The PF algorithm aims to reconstruct each one of these vertices and identify the PV, which is done by considering the one with the largest $\sum p_{\mathrm{T}}^2$ of tracks. The distribution of the number of pileup vertices in an event can be measured using a minimum bias trigger that has very loose criteria and aims to collect a large fraction of the total inelastic cross section [111]. As can be seen in the measured pileup profiles within CMS for each year of 2016–2018 in Fig. 3.10, the average number of pp interactions within a single bunch crossing during Run-II was approximately 30. When the PF algorithm associates charged hadron tracks to any of the pileup vertices, these tracks are discarded in the full event reconstruction.

3.3.3 Muon reconstruction

Muons tend to produce the cleanest particle signatures in the CMS detector due to the efficient muon chambers that are placed beyond the solenoid, where other particles typically do not reach. Muon track reconstruction is performed with dedicated algorithms that sometimes rely on separate muon reconstruction by the PF algorithm [8]. The track reconstruction relies on a Kalman filter (KF) method that accounts for muon energy loss in

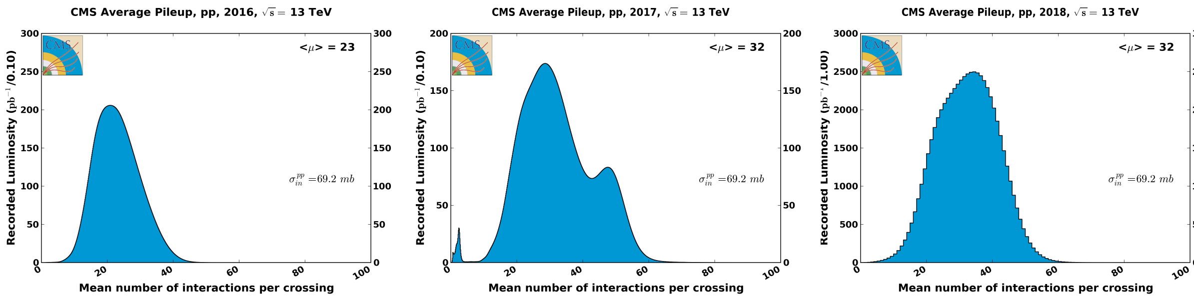

Figure 3.10: Pileup profiles for each year 2016–2018 data-taking (from left to right) [11]. These are the mean number of pp interactions per bunch crossing. To construct these distributions, the assumed minimum bias cross section is 69.2 mb.

the detector material, and the shape of the track is used to infer the muon momentum due to its bending in the magnetic field. There are three different approaches to muon track reconstruction:

- **Standalone muon** tracks are constructed solely from track segments in the muon system, combining information from the DT, CSC, and RPC detectors.
- **Tracker muon** tracks are constructed from hits in the Si tracking system. A tracker track with $p_{\mathrm{T}} > 0.5$ GeV and $p > 2.5$ GeV that can be successfully extrapolated to at least one track segment in the muon chambers at a compatible position qualifies as a tracker muon track.
- **Global muon** tracks are constructed by matching standalone muon tracks to tracker muon tracks. The two different tracks are propagated to a common surface and then a combined fit is performed using information from both the tracker and muon systems to construct the global track. Global muon reconstruction is designed to be highly efficient for muons that penetrate through more than one muon station.

Because the tracker track and muon segment reconstruction is highly efficient, about 99% of muons produced within the geometrical acceptance of the muon system are reconstructed as either a tracker or global muon track, and often as both. In this case, the tracks are merged into a single object. Tracker muons that are not global muons typically match only to segments in the innermost muon station, yielding an increased probability to misidentify "punch through," hadronic shower particles that can traverse the whole HCAL system into the muon system, as a muon.

All reconstructed muon tracks are fed into the PF algorithm, which itself employs different selection criteria for different muon candidate types using information from other CMS subdetectors in order to identify a particle as PF muon. All muons used in this dissertation work are required to be PF muons, and most are global muons. Standalone-only muons are not used in this analysis because they have worse momentum resolution and suffer from cosmic-ray background contamination.

3.3.4 Electron reconstruction

Electrons typically deposit all of their energy into the ECAL. Before electrons produced in pp collisions reach this detector, however, they will often interact with the material of the Si tracker, emitting bremsstrahlung photons, which themselves may convert into electron-positron pairs. Thus, the signature of an electron in the ECAL typically arises from a shower of multiple electrons and photons rather than a hit from a single object. As the bremsstrahlung photons are produced tangentially to the electron track, which bends along ϕ in the magnetic field, the energy of the electron will be more spread out in ϕ than in η. As charged particles, electrons also leave tracks in the tracker, important information that is primarily needed to distinguish electrons from photons. Thus, electron reconstruction requires dedicated algorithms for energy clustering and tracking, and then it aims to find the tracks that are compatible with the clusters. This reconstruction is fully integrated into the PF framework [112].

Electron energy clustering focuses on building "superclusters" (SCs) from ECAL clusters. First, seed clusters that contain local energy maxima are identified, and then other ECAL clusters within some geometric area around the seed are grouped along with the seed to form the SCs. The SC procedure is dependent on the cluster E_{T}, and the geometric area is more spread out in ϕ than in η to collect the bremsstrahlung photons.

Reconstruction of the electron tracks must consider that the curvature of the trajectory changes as the electron radiates bremsstrahlung photons and loses energy. To accomplish this, a track fitting procedure that models the radiated energy loss probability with a sum of Gaussian distributions, called Gaussian sum filtering (GSF), is used rather than the standard KF tracking. The GSF tracks can be seeded using either an ECAL-based or a tracker-based approach. The ECAL-based method estimates the track starting from an input SC, while the tracker-based method searches for a normal tracker track that is compatible with an SC and then uses the KF track as the seed.

The final SCs and GSF tracks are associated together to an electron candidate, provided that they satisfy some loose criteria relating to track topology, shower shape, and detector isolation. Both the GSF track and the SC are used to determine the electron charge and momentum, the latter of which is constrained using both the track momentum and the SC energy. Detector isolation involves criteria that serve to discriminate electrons from hadrons by placing an upper bound on the allowed HCAL energy within a cone of some fixed radius behind the energy-weighted center of an ECAL SC. This constraint produces severe inefficiencies for electrons produced near a jet and from a boosted object, like in the $\mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}\ell\nu\mathrm{q\bar{q}}$ search in this dissertation. This is discussed more in Section 4.3.1, as this effect has important implications for the electron channel efficiency in the search.

3.3.5 Jet reconstruction

Quarks and gluons produced in the hard interaction of a pp collision will shower into partons that will then hadronize before being absorbed in the calorimeters. While some of the energy is absorbed in the ECAL, most of the energy is absorbed in the HCAL that lies beyond it. The decay signature of such a quark or gluon in the CMS detector appears as a jet of collimated particles that include charged hadrons ($K^{\pm}$, $\pi^{\pm}$, or protons), neutral hadrons (K_L^0 or neutrons), nonisolated photons (from π^0 decays), and more rarely additional muons (from early decays of charged hadrons). Jets used in this analysis are constructed by clustering PF candidates together within some maximum distance R in $\eta - \phi$ space, and the jet momentum is determined as the vectorial sum of all particle momenta in the jet. The jet momentum is found to be, on average, within 5–10% of the true momentum over the entire p_{T} spectrum and detector acceptance. Jet energy corrections (JEC) are applied to calibrate the jet response using both simulation-driven and data-driven methods. The jet energy scale is corrected with the JEC by considering contributions from pileup, nonlinear effects in the detector response to hadrons, and residual differences between data and simulation. Furthermore, the jet energy resolution is observed to be worse in data than in simulation, and so additional corrections are applied to the simulation to smear the jets to match the data p_{T} resolution. Typical jet energy resolutions are about 15–20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV at central $|\eta|$.

The PF algorithm first reconstructs muons before identifying the electrons and isolated photons. From the set of remaining PF elements, ECAL and HCAL clusters that are not linked to any tracks are identified as photons and neutral hadrons, respectively. The HCAL clusters that remain and that are linked to a track (and sometimes even an ECAL cluster) are identified as charged hadrons. The momenta of charged hadrons are measured using the track momentum, while the energies of hadrons are measured using the HCAL clusters. Charged hadrons with tracks associated to pileup vertices are removed from the set of PF candidates used for jet clustering.

Jet clustering

From the set of all PF candidates, jets are clustered using a sequential recombination algorithm that is adapted for use in hadron colliders [113]. The algorithm uses an interparticle distance metric d_{ij} and a particle-beam splitting metric d_{iB} to determine how the clustering proceeds. The recombination algorithm works in general as follows:

1. Calculate d_{ij} and d_{iB} for all final state particles (PF candidates), and find the minimum value of each.

2. If the global minimum value is a d_{iB}, then identify particle i as a jet, remove it from the list of all particles, and return to Step 1.

3. If the global minimum value is a d_{ij}, combine particles i and j by their momenta, and return to Step 1.

4. Iterate until all particles have been declared jets.

There are three different daughter algorithms of the above recombination algorithm that are commonly used in particle colliders, each differing in their choice of the two metrics. These are the k_T algorithm, the Cambridge-Aachen (C-A) algorithm, and the anti-k_T algorithm, which are distinguished by:

$$d_{ij} = \begin{cases} \min(p_{T,i}^2, p_{T,j}^2)\Delta R_{ij}^2/R^2, & k_\mathrm{T} \text{ algorithm} \\ \Delta R_{ij}^2/R^2, & \text{C-A algorithm} \\ \min(p_{T,i}^{-2}, p_{T,j}^{-2})\Delta R_{ij}^2/R^2, & \text{anti-}k_\mathrm{T} \text{ algorithm} \end{cases} \tag{3.5}$$

and

$$d_{iB} = \begin{cases} p_{T,i}^2, & k_\mathrm{T} \text{ algorithm} \\ 1, & \text{C-A algorithm} \\ p_{T,i}^{-2}, & \text{anti-}k_\mathrm{T} \text{ algorithm.} \end{cases} \tag{3.6}$$

The quantities p_T and ΔR are used in these metrics because they are longitudinally boost-invariant and thus best suited for use in hadron colliders. The parameter R functions as an angular cut-off and corresponds to the size of the jet cone; two particles separated by a distance $\Delta R_{ij} > R$ will never be combined in a given iteration of the algorithm, regardless of their p_{T}.

All three of these algorithms are infrared and collinear safe: the same set of hard jets will be found if the event is modified by a collinear splitting or the addition of a soft emission. The k_{T} algorithm operates by clustering the soft radiation first, effectively an attempt to recreate the parton shower in the reverse direction. This approach produces irregular jet areas that depend on the detailed distribution of soft radiation in the event and thus are difficult to calibrate. The C-A algorithm clusters particles based only on their angular separation and serves as a compromise between diminishing sensitivity to soft radiation and reflecting the structure of the parton shower.

The anti-k_{T} algorithm, however, abandons the attempt to reflect the parton shower structure and instead begins the clustering with the hardest particles, finding the most energetic cores of jets first. This produces regular jet areas that extend out to a radius R from the jet core and are much more easily calibrated in experiments. All jets used in the analysis of this dissertation are clustered with the anti-k_{T} algorithm [114, 115]. This dissertation work uses jets clustered with distance parameters of $R = 0.4$ (AK4 jets) and $R = 0.8$ (AK8 jets). In general, the AK4 jets are meant to reconstruct the decay of a single hard quark or gluon, while the AK8 jets are used to reconstruct the hadronic decays of boosted Higgs or vector bosons.

Jet flavor identification (b tagging)

Jets arising from the decay of heavy-flavor (HF) quarks can be distinguished from light quark or gluon decays as a result of their relatively longer lifetimes. A b quark, whose identification in an event is crucial in the search for X $\rightarrow$ HH, has a lifetime on the order of 1.5 ps and so typically travels from a few mm to a cm in the detector before its initial decay. This phenomenon gives rise to displaced tracks relative to the PV that originate from some secondary vertex (SV). The displacement of tracks at the SV is characterized by the impact parameter, defined as the distance between the PV and the tracks at their points of closest approach. HF quarks also have larger mass and harder fragmentation than light quarks and massless gluons, and as a result, the decay products of HF hadrons have larger p_{T} relative to the jet axis on average. Furthermore, the decays of HF quarks can sometimes contain a muon or an electron. Figure 3.11 depicts the decay of a b quark into a jet. Algorithms that identify jet flavor take advantage of these phenomena and the fine segmentation of the Si tracker; these are of utmost importance in this analysis to identify $\mathrm{H} \to \mathrm{b\bar{b}}$ decays and top quark decays.

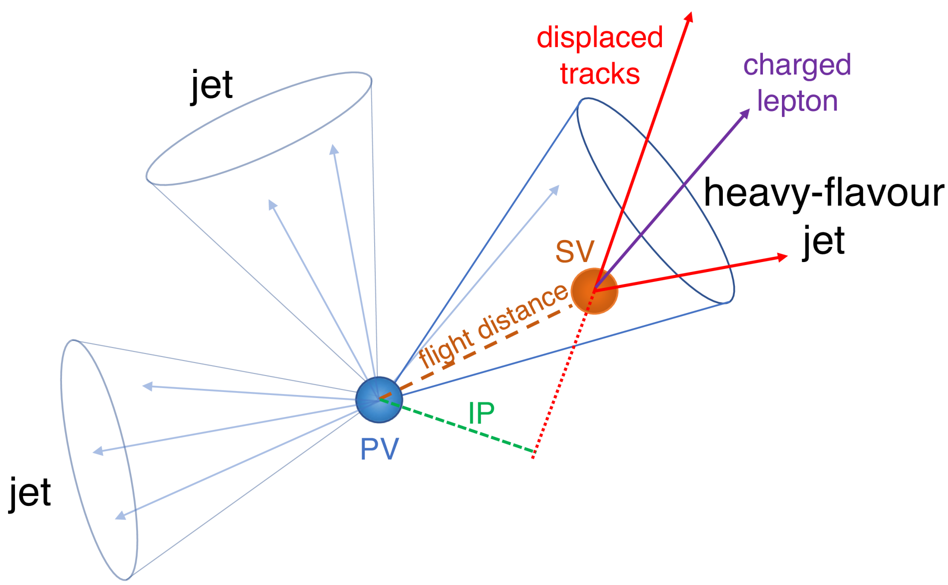

Figure 3.11: Depiction of a jet resulting from a heavy-flavor quark decay, showing the secondary vertex (SV), a charged lepton in the jet from the quark decay, and a displaced track and its impact parameter (IP) relative to the primary vertex (PV) [12].

In the analysis, jets originating from the decay of a single b quark, particularly those involved in top quark decays, are identified using the DeepJet classifier [116], which is compatible with AK4 jets. This is a deep neural network (DNN) based classifier that uses hundreds of variables related to a given jet divided into four categories: global variables, charged PF candidate features, neutral PF candidate features, and SV features. The global variables include jet-level information like the kinematics and the number of tracks and

SVs. Pileup information is included in each category to help the classifier learn how this correlates with HF decay. The PF candidate inputs include kinematic information and track fit information for charged hadrons that includes quality and displacement. Information related to the SVs includes the displacement, fit quality, mass, and energy, and p_{T}. The approach of DeepJet is to use the full information of all particles, tracks, and SVs in the jet. This significantly improves the performance over other methods that instead use a small subset of the charged jet constituents that pass stringent quality criteria in order to provide the cleanest and simplest environment.

Large boosted jets that contain b quarks in the analysis, particularly those arising from the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay, are identified using the mass-decorrelated DeepAK8 $\mathrm{Z/H} \to \mathrm{b\bar{b}}$ tagger [117]. This tagger is designed particularly to discriminate AK8 jets with substructure arising from two b quarks against large light quark or gluon jets. It uses a DNN architecture that takes as input a large variety of information related to up to 100 PF candidates and up to 7 secondary vertices (SVs) in a jet. This information includes the p_{T}, energy deposit, charge, and angular separation from the jet and subjet axes, and tracking information (for charged particles only) of each PF particle. Information on the SVs includes their kinematics, displacement, and reconstruction quality. The DNN is built to prevent itself from extracting features correlated to the jet mass for the optimization, and it thus does not sculpt the background mass distribution. The output of the $\mathrm{b\bar{b}}$ tagger is a discriminating variable $D_{\mathrm{Z/H\to b\bar{b}}}$ that can have a value from 0 to 1, with larger values a stronger indication of $\mathrm{b\bar{b}}$ substructure.

3.3.6 Missing transverse momentum reconstruction

In pp collisions, noninteracting particles such as neutrinos or possibly other exotic particles may be produced and escape CMS undetected. Because of the beam geometry, the presence of such particles can be inferred from the momentum imbalance in the transverse plane due to all visible particles. This missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, whose magnitude is denoted $p_{\mathrm{T}}^{\mathrm{miss}}$, is reconstructed as the negative vector sum of all PF candidate momenta in an event. The $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is impacted by effects in the jet energy scale and resolution, and so the JEC are propagated into the calculation to correct $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$:

$$\vec{p}_{\mathrm{T}}^{\text{miss, corr}} = \vec{p}_{\mathrm{T}}^{\mathrm{miss}} - \sum_{\text{jets}} (\vec{p}_{\text{T, jet}}^{\,\text{corr}} - \vec{p}_{\text{T, jet}}), \tag{3.7}$$

where the sum runs over only jets with corrected $p_{\mathrm{T}} > 15$ GeV. If a muon is found inside the jet cone, its four-momentum is subtracted from the jet momentum when the deriving the corrections and then added back into the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ sum.

CHAPTER 4

Event Selection in the Search for HH Resonances

In this chapter, the full selection of events is described. The phase space considered with these selections is optimized for signal significance in the presence of the SM background. The strategy for event selection begins with choosing CMS data sets that contain events with particles reconstructed as in Section 3.3. These data sets are presented in Section 4.1. The trigger selection is described in Section 4.2; these are the set of software-based triggers described in Section 3.3.1 that we deem to be useful for collecting signal events. With the collection of events that fired our triggers of interest in hand, Section 4.3 describes the process for selecting and reconstructing important particles and intermediate states in the full X $\rightarrow$ HH decay chain. We aim to select high quality leptons and reconstruct the decays of the bosons in each event. Finally, Section 4.4 describes the set of selections used to further discriminate signal from background events. The selections are also used to categorize the events before using the mass distributions as input into the modeling procedure, described in Chapter 5.

4.1 Data sets

This search uses a data set of interesting events from pp collisions recorded by the CMS detector. As described in Section 3.3.1, a system of hardware-based triggers is used to process the event data in real-time, after which more computationally expensive techniques in offline software are used to reconstruct the event fully. The data used in this search was reprocessed with the optimal calibrations available at the time of analysis.

The data used in this analysis are those recommended by the Data Quality Monitoring (DQM) group within the CMS collaboration. The DQM group is responsible for monitoring operational information from all subdetector systems and determining whether any data were recorded under problematic online conditions. As such, the DQM group may deem certain sets of data unsuitable for offline analysis due to issues such as detector misalignment or non-operation. The DQM-certified set of events that is considered in this analysis constitutes an integrated luminosity of 137.6 fb^{-1}, a measurement with an uncertainty of 1.6%. This total integrated luminosity for this period contains 36.3 fb^{-1} from 2016 data-taking, 41.5 fb^{-1} from 2017, and 59.7 fb^{-1} from 2018.

The use of samples of simulated events from pp collisions is crucial to optimizing our understanding of the data and allowing us to make useful predictions. Events are simulated using Monte Carlo techniques to simulate various stages of an event, from the matrix element of the fundamental process, to the parton showering and hadronization processes, and through the detector and electronics response. The simulated events are then processed for offline use with the same calibrations as if they were collected as data by the CMS detector. Simulated samples are organized into the 2016, 2017, and 2018 years of data-taking such that the state of the CMS detector and the trigger menu in a particular year is properly reflected in the distribution of events.

Both signal and SM background processes are simulated, allowing us to study how best to discriminate one from the other in our search for a resonance. The simulated signal and background events are also the springboard from which we build our models of the data (described in Chapter 5). Signal samples are simulated for both spin-0 and spin-2 resonances produced through gluon fusion at leading order (LO) and at a variety of values of mass m_{X} from 800 GeV to 4.5 TeV. For such resonances X, the signal processes simulated are $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}VV^*}$ (where $\mathrm{V} = \mathrm{W/Z}$) and $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}}\tau\tau$, and the widths of the resonances are set to a narrow 1 MeV, which is much smaller than the experimental resolution. The background samples used include simulation of all SM processes determined to have an effect on the analysis; processes that are not included have a negligible effect. The simulated background processes considered are shown in Tab. 4.1, along with the order of the perturbative matrix element calculation and the corresponding SM cross sections of the samples, used to normalize the background to the integrated luminosity of the relevant data-taking year.

Table 4.1: SM background processes for which simulated samples are used. Background processes that are not listed here have negligible contributions in the analysis phase space. The cross section and the corresponding order in its calculation are provided for each process.

Process	Order of cross section	Cross section [pb]
QCD multijet	LO	2108083.65
$\mathrm{W} + \text{jets}$ ($\mathrm{W} \to \ell\nu$)	NNLO	2144.75
$\mathrm{Z}/\gamma^* + \text{jets}$ ($\mathrm{Z}/\gamma^* \to \ell\ell$)	NNLO	505.03
$\mathrm{t\bar{t}}$	NNLO	828.15
diboson (WW, ZZ, WZ)	NLO	80.70
single top	NLO	236.37
associated $\mathrm{t\bar{t}}$ ($\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}H}$)	NLO	1.94

4.2 Triggering events

We begin building the collection of events within which we search for signal by selecting events that fire the high-level software-based CMS triggers deemed capable of collecting our hypothetical signal events. Due to the boost of the Higgs bosons in many signal models, leptons are frequently produced with high p_{T} in the final state, suggesting that single-lepton triggers are ideal for this analysis. However, signal leptons in the 1ℓ channel can often be surrounded by considerable hadronic activity, which renders these single-lepton triggers inefficient in some cases, as these triggers have tighter lepton isolation (for muons) or identification (for electrons) criteria than what is used for offline analysis. To recover this lost efficiency, we also employ multi-object triggers that contain both leptons and jets, with looser lepton reconstruction criteria than the single-lepton triggers and with some baseline amount of hadronic activity in the transverse plane (H_{T}). Unfortunately, these multi-object triggers were not consistently operating throughout periods of 2016 and 2017 data-taking, diminishing their effectiveness at times. Therefore, we use a series of supplemental triggers to restore the remaining efficiency. Purely hadronic triggers are used to recover efficiency for events with high H_{T}, and additional triggers are used to recover efficiency for events with high-p_{T} leptons. For high-p_{T} muons, $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers in which muons are not included in the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation are used. For high-p_{T} electrons, a high-p_{T} single-photon trigger is used. An event is collected with the logical OR of all triggers used, and events are filtered such that no single event that is in multiple data sets is collected more than once.

Trigger efficiencies are measured in data and simulation for a suite containing all supplementary triggers and one flavor of lepton triggers with a tag-and-probe method using $e\mu$ $\mathrm{t\bar{t}}$ events. We use $\mathrm{t\bar{t}}$ events because the lepton and jet multiplicities resemble those in signal. These $\mathrm{t\bar{t}}$ events are required to contain one lepton (probe) that passes the standard offline analysis selection criteria (Section 4.3.1) and another lepton of different flavor (tag) that passes tighter identification, isolation, and p_{T} criteria, helping to optimize the purity of the sample. To measure the electron efficiency, we collect events that fire an isolated single-muon trigger and contain a tag muon and a probe electron. We then obtain the efficiency for the event to pass any of the electron or supplementary triggers. We repeat the same procedure to measure the muon efficiency, but in this case collecting events that fire an isolated single-electron trigger and contain a tag electron. Events in the efficiency measurement are also re-weighted to account for periods of inactive triggers such as the multi-object triggers discussed above. The trigger efficiency for signal events in the 1ℓ channel is over 96% at $m_{\mathrm{X}} = 0.8$ TeV and increases to $> 99\%$ above approximately $m_{\mathrm{X}} = 1.0$ TeV. In the 2ℓ channel, the trigger efficiency is $> 99\%$ for all m_{X}.

We choose offline electron and muon p_{T} thresholds of 30 GeV and 27 GeV, respectively, to be efficient with respect to the primary single-lepton triggers. All events are also required to have offline $H_{\mathrm{T}} > 400$ GeV. This selection is chosen because a significant amount of background can be rejected without spoiling sensitivity to an 0.8 TeV signal and because the H_{T} dependence of the trigger suites is modeled well by the simulation for this region. We calculate H_{T} using all jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 5$. The efficiency curves exhibit differences between the measurements in data and simulation, so we correct the simulated efficiencies to match those measured in data by applying multiplicative scale factors (SFs) that are dependent on H_{T} in the 1ℓ channel and on both H_{T} and p_{T} in the 2ℓ channel to the simulated events.

In the 1ℓ channel, the trigger scale factors are parameterized as a function of H_{T} in order to capture the turn-on transition from the lepton triggers to the hadronic triggers. A smaller effect is also the dependence of the scale factors on lepton p_{T}. Since signal and the $e\mu$ $\mathrm{t\bar{t}}$ events used to measure the efficiencies have different lepton p_{T} spectra, we estimate the dependence of the scale factors on the lepton p_{T} by shifting the probe lepton offline p_{T} threshold up and down by 5 GeV. A flat $\pm$ 2% systematic uncertainty conservatively covers the scale factor variations over the full range of H_{T} and is applied to the normalization of 1ℓ events in signal. An example of the scale factor measurement and the estimation of the uncertainty is shown in Fig. 4.1.

In the 2ℓ channel, the trigger efficiencies are handled in a similar manner but with some added complexity. These events are collected with the same set of triggers as in the 1ℓ channel, and we require that at least one of the leptons pass the same offline p_{T} criteria used for the lepton in the 1ℓ channel ($p_{\mathrm{T}} > 30$ GeV for electrons and 27 GeV for muons). Since the leading lepton in p_{T} will be more likely than the other lepton to pass the trigger, we use the same method as in the 1ℓ channel to calculate trigger efficiencies as a function of H_{T} for this lepton. In this case, we use the 2ℓ identification and isolation criteria to measure the efficiencies for this lepton.

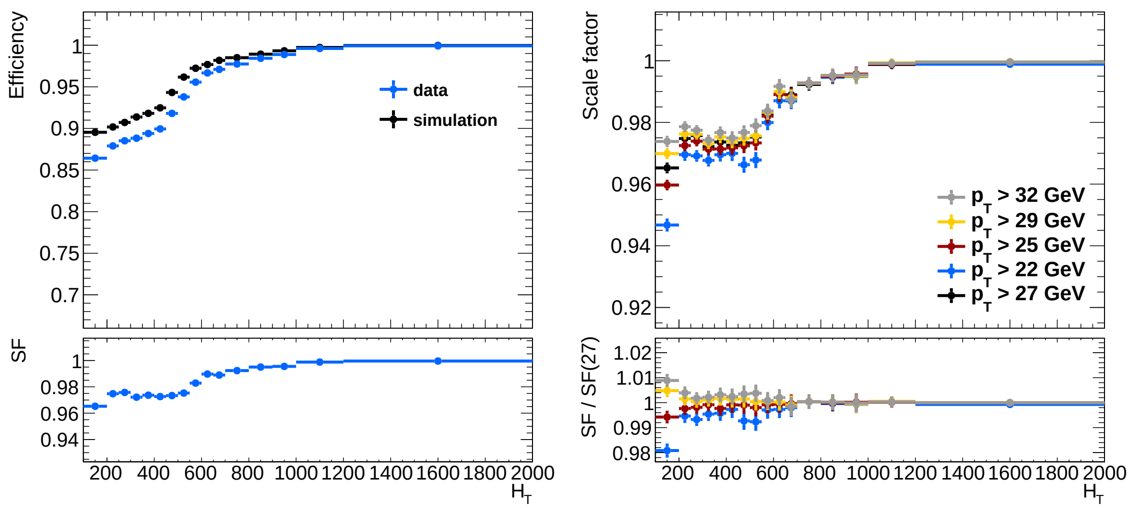

Figure 4.1: Scale factor (SF) and systematic uncertainty measurement for muons in the 1ℓ channel for the 2018 year of data-taking. Left: the efficiencies ϵ for data and simulation are shown on the top pane, and the SF ($\epsilon_{\mathrm{data}}/\epsilon_{\mathrm{sim}}$) is shown in the bottom pane. This SF is applied to simulated events as a function of H_{T} to properly reflect the trigger efficiencies in data. Right: SFs calculated with different p_{T} thresholds for the muons offline in the top pane, and their variations relative to the measurement with the nominal threshold of $p_{\mathrm{T}} > 27$ GeV. Noting that the analysis phase space is $H_{\mathrm{T}} > 400$ GeV, the maximal variation in the analysis scale factors is about 1% relative to the nominal threshold. The maximal variation among all the trigger measurements in the 1ℓ channel is $+2\%$, leading us to place a flat $\pm$ 2% uncertainty on all 1ℓ channel events.

In rare cases, the highest p_{T} lepton will fail to fire the trigger but the sub-leading lepton will fire the trigger successfully. To account for this effect, we also measure the efficiency for the sub-leading lepton to pass any of the single-lepton triggers, excluding the lepton-H_{T} cross triggers, as a function of lepton p_{T}. The combined trigger efficiency for 2ℓ events can then be given as a function of the H_{T} and the sub-leading lepton p_{T} as[1]:

$$\epsilon(H_{\mathrm{T}}, p_{\mathrm{T}}) = \epsilon_1(H_{\mathrm{T}}) + \epsilon_2(p_{\mathrm{T}}) - \epsilon_1(H_{\mathrm{T}}) * \epsilon_2(p_{\mathrm{T}}) \tag{4.1}$$

The subscript of the efficiency in the equation above refers to the lepton used for the measurement ("1" for higher p_{T} lepton and "2" for sub-leading lepton).

The efficiency for the leading lepton to pass the trigger includes the efficiency for the event to fire the H_{T} triggers. An assumption in our 2ℓ method is that the sub-leading lepton would not pass the lepton-H_{T} multi-object triggers (which we make because it is difficult to disentangle the lepton leg from the H_{T} leg when including these triggers in the measurement). We estimate our sensitivity to this assumption by recalculating the efficiencies and scale factors after including these multi-object triggers in the sub-leading lepton efficiency measurement. A systematic uncertainty of $\pm$ 3% conservatively covers any variation in the SFs due to this procedure and is applied to the normalization of the 2ℓ events in signal. Example plots of the scale factor and estimate of the systematic uncertainty for 2ℓ events are shown in Figs. 4.2 and 4.3, respectively.

4.3 Reconstruction of the HH decay chain

In order to reconstruct the full decay chain of a signal event, selections are imposed on the set of basic reconstructed objects in the CMS detector introduced in Section 3.3. The selections described in this section are chosen such that these basic objects can be optimally associated

[1]This expression is derived assuming the two measurements are independent. Then, the probability for either lepton to pass would be one minus the probability for neither to pass: $1-(1-\epsilon_1)(1-\epsilon_2) = \epsilon_1+\epsilon_2-\epsilon_1\epsilon_2$.

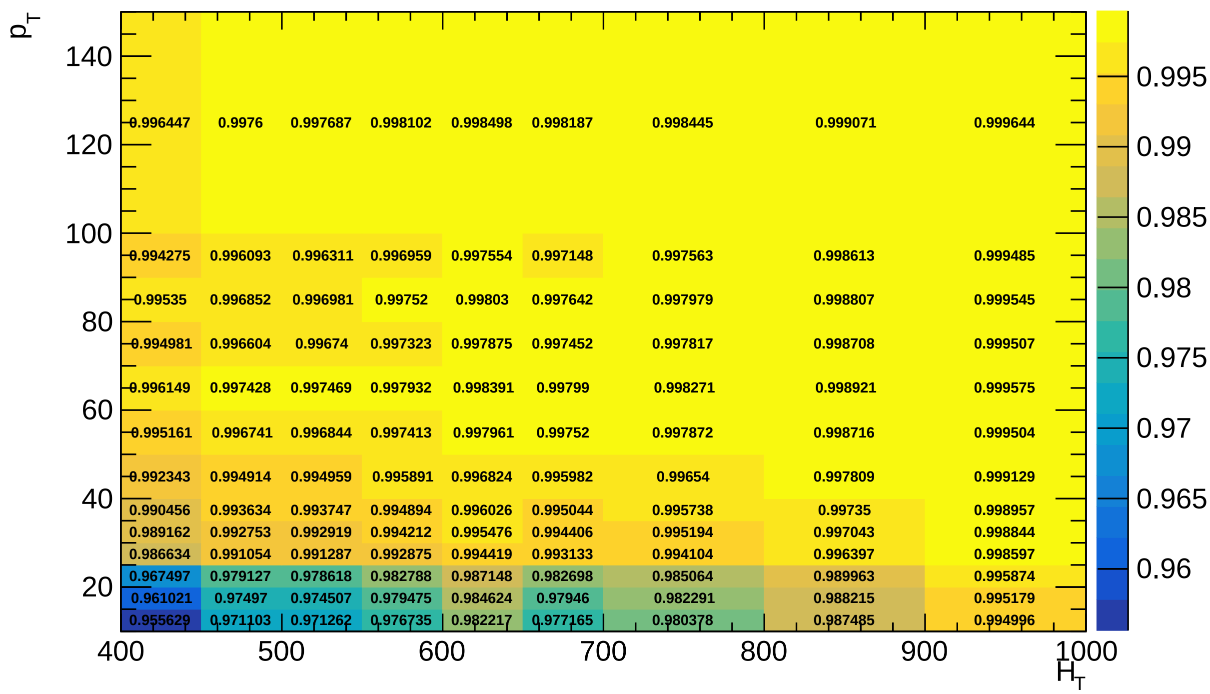

Figure 4.2: The trigger scale factor $\epsilon_{\rm data}/\epsilon_{\rm sim}$ for dilepton $e\mu$ events in 2016, where the electron e is the highest-$p_{\rm T}$ lepton. The plots are zoomed in for ease of viewing: there are lower bounds of 400 GeV in $H_{\rm T}$ and 10 GeV in sub-leading lepton $p_{\rm T}$, which correspond to the analysis phase space. The scale factors above 1000 GeV in $H_{\rm T}$ and 150 GeV in $p_{\rm T}$ are closer to 1.0 than the bins on display.

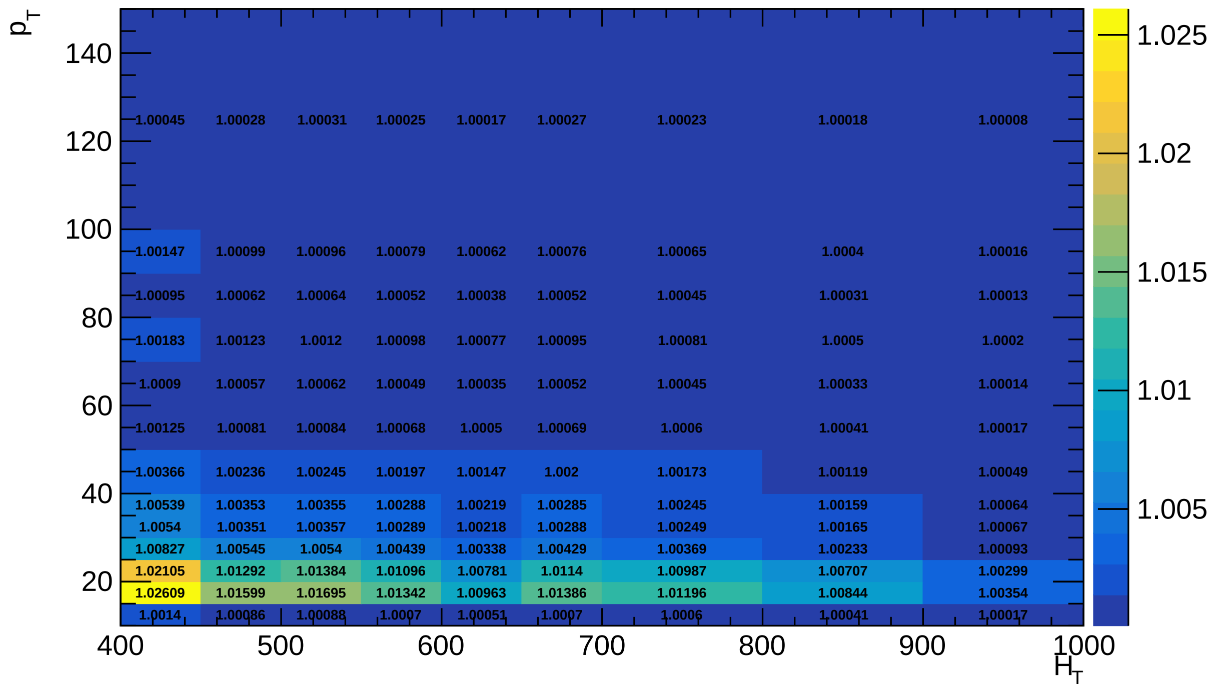

Figure 4.3: The variation in the scale factor for dilepton $e\mu$ events in 2016 due to including the lepton-H_{T} multi-object triggers in the numerator of the efficiency measurement. A maximal variation of 2.6% can be seen in the figure, which is within the $\pm 3\%$ uncertainty applied for this measurement. The plots are zoomed in for ease of viewing: there are lower bounds of 400 GeV in H_{T} and 10 GeV in sub-leading lepton p_{T}, which correspond to the analysis phase space. The systematic variation above 1000 GeV in H_{T} and 150 GeV in p_{T} are closer to 1.0 than the bins on display.

to the intermediate or final state particles in the decay of X $\rightarrow$ HH.

4.3.1 Leptons

Electrons and muons are considered in the search for leptons resulting from a signal decay. To be considered a candidate, the lepton must pass offline acceptance (p_{T} and η), identification, and isolation requirements. Lepton identification seeks to determine whether the detector signature is truly produced by a lepton or whether other particles such as hadrons or photons are misreconstructed as leptons.

As mentioned in Section 4.2, muons are required to have $p_{\mathrm{T}} > 27$ GeV and electrons to have $p_{\mathrm{T}} > 30$ GeV. Each event, regardless of channel, is required to have one such lepton passing these p_{T} criteria. In the 2ℓ channel, the second lepton must only have $p_{\mathrm{T}} > 10$ GeV. Muons in both channels are required to have $|\eta| < 2.4$ in order to be within the acceptance of both the tracker and muon system. Electrons in the 2ℓ channel are likewise required to have $|\eta| < 2.5$, but in the 1ℓ channel electrons are restricted to $|\eta| < 1.479$, which corresponds to the acceptance of the ECAL barrel region. Many misidentified electrons can arise from the QCD multijet background and are located in the CMS endcaps. Restricting lepton $|\eta|$ to be in the ECAL barrel suppresses this effect dramatically and bears a tolerable loss in signal acceptance, as seen in Fig. 4.4.

The identification of signal-like electrons relies on a variety of variables related to the shower shape in the calorimeters and the track kinematics and fit quality [112]. These are necessary to discriminate against background sources such as photon conversions, hadrons misidentified as electrons, and secondary electrons from b or c decays.

Some important variables commonly used in electron identification are described here. The H/E ratio, defined as the ratio of HCAL energy in a cone of radius $\Delta R = 0.15$ around the electron SC, is primarily used to discriminate electrons from hadrons. For a genuine electron, the measured hadronic energy H can be due to HCAL electronics noise, pileup, or leakage

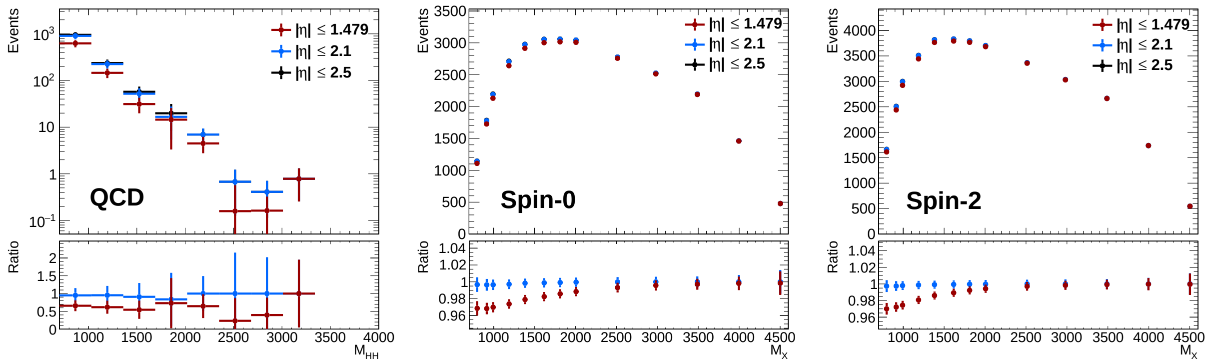

Figure 4.4: Comparisons of the QCD multijet background (left) and signal (middle and right) yields for several choices of the $|\eta|$ requirement on electrons in the1ℓ channel. The background plot is shown as a function of reconstructed resonance mass m_{HH}, and the signal plots are shown as a function of the generated resonance mass m_{X}. The chosen working point is $|\eta| < 1.479$ because we can reject 30% to 40% of the QCD background with losing much sensitivity to any signal (from 3% at low mass to none at high mass). The ratios in the bottom pane of each plot are evaluated with respect to the selection $|\eta| < 2.5$.

of the electron energy through the inter-module gaps (usually due to high p_{T} electrons). An important shower shape variable is $\sigma_{i\eta i\eta}$, the second moment of the log-weighted distribution of crystal energies in η, calculated in the 5×5 crystal matrix around the most energetic crystal in the SC. This distribution for an electron shower is expected to be narrow for an electron shower and broad for showers from two photons from neutral hadron decays within a jet. The shower width in η is more discriminating because the magnetic field already causes the SC profile to be spread out in ϕ, as described in Section 3.3. Other important variables using tracker information include $|1/E - 1/p|$, where E is the SC energy and p is the track momentum at the point of closest approach to the PV. Other variables involve the angular separation between tracks and seed objects or energy clusters, where minimal separation is indicative of a good electron. These include $|\Delta\eta_{\mathrm{seed}}|$, defined as $|\eta_{\mathrm{seed}} - \eta_{\mathrm{track}}|$, and $|\Delta\phi_{\mathrm{in}}| = |\phi_{SC} - \phi_{\mathrm{track}}|$, where the ϕ of the energy-weighted SC position is used instead of the seed cluster ϕ.

Because of the unique background composition in each of the 1ℓ and 2ℓ channels, different lepton selection criteria are used to optimize the signal acceptance against background. In the 2ℓ channel, a menu of several identification selections is applied using some of the variables described above. This set of selections corresponds to the medium working point of the cut-based electron identification criteria in CMS, which is on average 80% efficient for identifying prompt electrons. The full set of selections is shown in Table 4.2. In the 1ℓ channel, a multivariate discriminator that uses these and other related variables in a boosted decision tree (BDT) is optimal for selecting prompt electrons that may be accompanied by nearby hadronic activity (such as the jet from the $\mathrm{W} \to \mathrm{q\bar{q}}$ decay). This BDT is trained on $\mathrm{Z}/\gamma^* + \mathrm{jets}$ events with $\mathrm{Z}/\gamma^* \to \ell\ell$ using prompt electrons as signal and unmatched plus nonprompt electrons as background. The working point we use for identification corresponds to an approximate 90% efficiency.

Table 4.2: Selections in the medium cut-based identification criteria used for electrons in the 2ℓ channel. The electron SC energy is given by E_{SC}, while ρ corresponds to the pileup contamination density in the event.

Variable	Barrel ($\lvert\eta\rvert < 1.479$)	Endcap ($\lvert\eta\rvert > 1.479$)
H/E	$< 0.046 + (1.16 + 0.0324\rho)/E_{SC}$	$< 0.0275 + (2.52 + 0.183\rho)/E_{SC}$
$\sigma_{i\eta i\eta}$	< 0.0106	< 0.0387
$\Delta\eta_{\mathrm{seed}}$	< 0.0032	< 0.00632
$\Delta\phi_{\mathrm{in}}$	< 0.0547 rad	< 0.0394 rad
$\lvert 1/E - 1/p\rvert$	$< 0.184\ \mathrm{GeV}^{-1}$	$< 0.0721\ \mathrm{GeV}^{-1}$
Missing hits	< 1	< 1
Conversion veto	true	true

The identification of muons relies primarily on tracker information from both the Si tracker and muon systems. This information includes the track fit χ^2, the number and location of hits per track, and the degree of matching between tracker tracks and standalone tracks (for the case of global muons). One important variable, the muon segment compatibility, is calculated by propagating a tracker track to the muon system and evaluating the number of matched segments and the precision of matching. A discriminator with values from 0.0 to 1.0 indicates a high degree of compatibility with a value closer to 1.0. A kink finding algorithm is also used to evaluate the quality of a track by breaking it into two separate tracks, which are determined to have some degree of compatibility.

Using these criteria, two different sets of selections (working points) are used to identify signal-like muons in the 1ℓ and 2ℓ channels. In the 2ℓ channel, a muon must be selected by the PF algorithm and either be a tracker or global muon, as described in Section 3.3.3. This corresponds to a loose muon ID criteria that is designed to select prompt muons and muons from light- and heavy-flavor decays. In the 1ℓ channel, where the backgrounds are larger and muon misidentification is more probable, tighter criteria are used. A medium ID is used in this case, optimized for prompt muons and only those from heavy-flavor decay now. A medium muon is a loose muon that additionally has a tracker track that uses hits from more than 80% of the inner tracker layers traversed. Additionally, there are muon segment compatibility and track fit constraints. If a muon is only a tracker muon, then the muon segment compatibility is required to be > 0.451. If the muon is both a tracker and global muon, the segment compatibility is only required to be > 0.303, but also the global fit χ^2/dof is required to be < 3, the position match between tracker and standalone muon tracks must have $\chi^2 < 12$, and the kink-finding algorithm's maximum χ^2 is required to be < 20. More detailed information can be found in Ref. [8].

Constraints are also set on the impact parameter (IP) of charged lepton tracks with respect to the PV, serving to suppress the contribution of background leptons from heavy-flavor quark decays such as those from b and c quarks. These quarks typically travel a few mm (up to 1 cm for b quarks) before decaying, so for a leptonic decay the lepton track will be displaced from the PV. One set of IP selections are applied to the tracks of both electrons and muons, chosen to tolerate a minimal loss in signal efficiency. These include constraining the longitudinal and transverse displacements of the lepton track IP from the PV to be less than 0.1 cm and 0.05 cm, respectively. In the 1ℓ channel, the significance of the three-dimensional measurement b_{IP} of the IP, defined as $|b_{\mathrm{IP}}|/|db_{\mathrm{IP}}|$, is also constrained to be fewer than 4.0 standard deviations. However, this constraint is not applied in the 2ℓ channel in order to remain sensitive to the $\mathrm{HH} \to \mathrm{b\bar{b}}\tau\tau$ final state. The τ leptons have a lifetime of about 1/5 of that of a b quark; lepton tracks from a leptonically decaying τ lepton are still displaced from the PV but not so much that the longitudinal and transverse IP displacements are larger than the cut values. The IP significance distribution is shifted toward large enough values that a cut on this variable is not useful considering the low background levels that already exist in the 2ℓ channel.

Because of their clean detector signatures relative to other particles, leptons are often expected to be isolated. A lepton's relative isolation, computed with the PF candidates, is gauged by considering the summed scalar p_{T} of all particles within a cone of radius R in $\eta - \phi$ space around the lepton: $I_{\mathrm{rel}} = \sum_i p_{\mathrm{T}}^i / p_{\mathrm{T}}^{\mathrm{lep}}$, where i runs over all particles within an angular distance R of the lepton.

Because of the significant boost of the Higgs bosons, final-state particles are typically collimated, which can cause high quality leptons to fail isolation requirements, particularly in the 1ℓ channel where the lepton of interest is close to or even overlapping with the jet from the $\mathrm{W} \to \mathrm{q\bar{q}}$ decay. Mini-isolation can mitigate this effect by employing an isolation cone radius that varies as a function of the lepton p_{T} in GeV [118]:

$$R_{\mathrm{mini-isolation}} = \begin{cases} 0.2, & p_{\mathrm{T}} < 50 \text{ GeV} \\ 10 \text{ GeV}/p_{\mathrm{T}}, & 50 < p_{\mathrm{T}} < 200 \text{ GeV} \\ 0.05, & p_{\mathrm{T}} > 200 \text{ GeV}. \end{cases} \tag{4.2}$$

Larger p_{T} is indicative of larger event boost, a situation in which the $\mathrm{q}\bar{\mathrm{q}}$ jet is closer on average to the lepton, which is not expected to be isolated well. Between 50 and 200 GeV, the mini-isolation cone radius shrinks with a $1/p_{\mathrm{T}}$ dependence considering a less isolated environment. In particular, this dependence is inspired by considering the two-body decay of a particle with mass M and large p_{T} to two daughter particles, where the angular separation of the daughters is roughly $\Delta R \approx 2M/p_{\mathrm{T}}$. The isolation calculation also involves corrections for identifying and excluding particles from pileup interactions. This isolation strategy is necessary and particularly powerful for the 1ℓ channel, where we require $I_{\mathrm{rel}} < 0.2$. In the 2ℓ channel, it also performs well when maximizing signal acceptance relative to background rejection. A tighter selection in this channel of $I_{\mathrm{rel}} < 0.15$ is applied because there is not expected to be a jet near any leptons.

The criteria for lepton selection in the 1ℓ channel is distinct from the 2ℓ channel in order to optimize signal efficiency and background rejection for each decay channel. We enforce that the two sets of events that comprise each channel be disjoint. The rate for true 2ℓ events to pass the full 1ℓ channel selection is larger than the rate for true 1ℓ events to pass the full 2ℓ selection. So, to maximize the purity of each channel, priority is given to the 2ℓ channel; any event that has exactly two opposite-sign leptons that pass the 2ℓ lepton criteria is selected for the 2ℓ channel. If an event has a lepton that passes the 1ℓ lepton criteria and also has fewer than two leptons that pass the 2ℓ lepton criteria (thus failing the selection), then the event is selected for the 1ℓ channel. In this case, the highest-p_{T} lepton that passes the 1ℓ criteria is chosen for HH reconstruction. If neither of these conditions are met by the lepton candidates in an event, then the event is not used for analysis.

The lepton efficiencies due to reconstruction, identification, and isolation are shown in Fig. 4.5 for various signal hypotheses of mass m_{X}. These efficiencies are measured in simulation by matching reconstructed leptons to true particle-level leptons within a tight $\Delta R < 0.1$. The presence of a jet in the 1ℓ channel signal results in universally lower lepton selection efficiencies than in the 2ℓ channel signal. For 1ℓ muons, the inefficiency due to the jet is

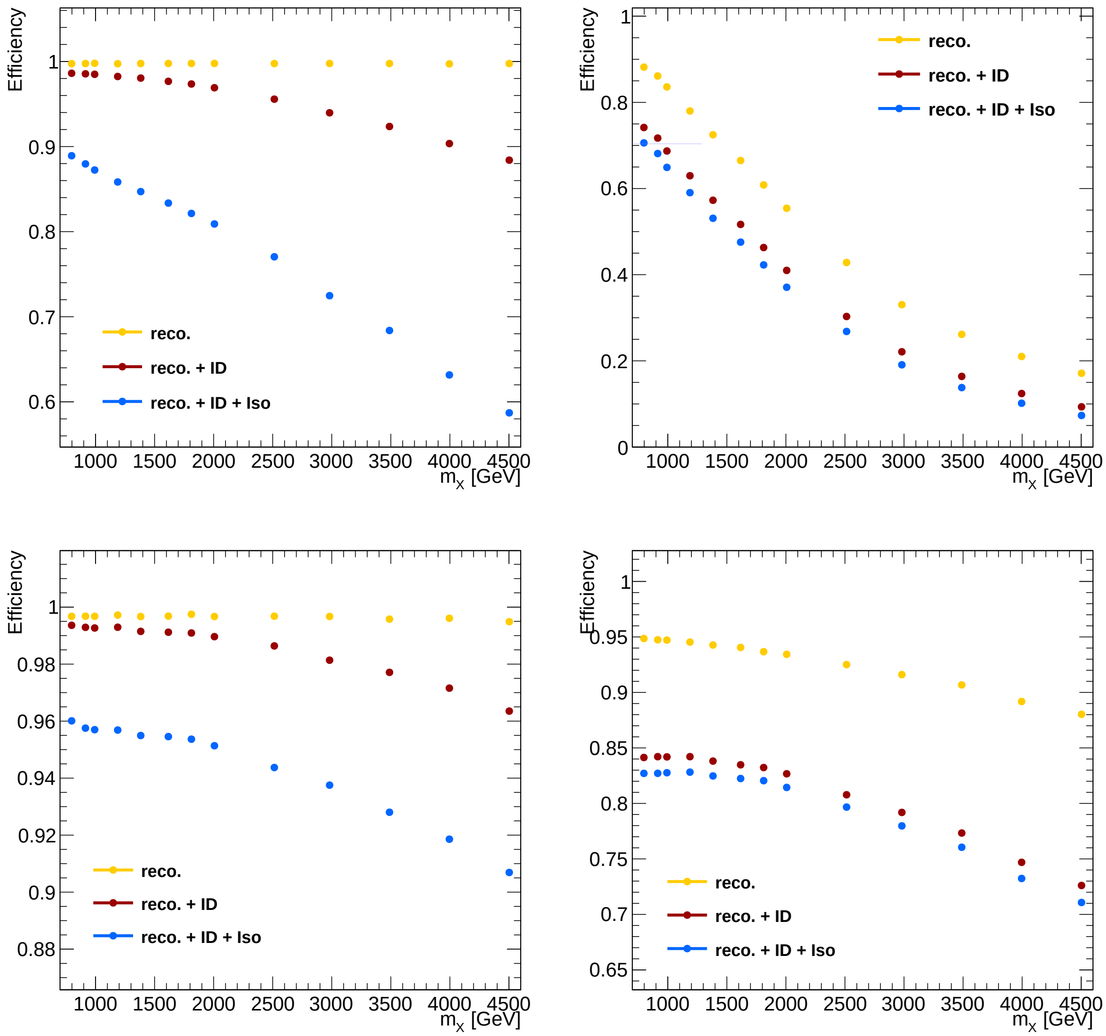

Figure 4.5: Lepton selection efficiencies in signal for reconstruction ("reco."), identification ("ID"), and isolation ("Iso"), applied sequentially in this order. The muon (electron) efficiencies are shown in the left (right) column, and the top (bottom) row shows the 1ℓ (2ℓ) channel.

manifest in the isolation selection, wherein this jet's energy contaminates the lepton isolation cone. The muon reconstruction (Section 3.3.3) and ID relies heavily on information from the muon systems, which is largely devoid of hadronic energy, and so these efficiencies are not as strongly dependent on jet proximity.

For electrons in the 1ℓ channel, the inefficiency due to the jet arises in the reconstruction process detailed in Section 3.3.4. The electron reconstruction procedure involves a flat H/E cut that considers the HCAL energy within an angular distance $\Delta R < 0.15$ from the energy-weighted center of the electron SC relative to the electron SC energy. The cone area created with $\Delta R < 0.15$ contains about 9 HCAL towers. This intrinsic isolation cut in the reconstruction is often tighter than the mini-isolation that is used in the lepton selection; the mini-isolation cone area is smaller than the H/E cone area for lepton $p_{\mathrm{T}} \gtrsim 67$ GeV. In signal, energy from the jet often falls within the cone area, preventing the successful reconstruction of a good electron. The severity of this inefficiency led to additional studies on how to mitigate this effect, as it also posed a general problem for heavy mass searches that involve boosted decay topologies. A simple solution was found wherein the electron reconstruction also considers the energy of the HCAL tower directly behind the electron SC central position relative to the other HCAL towers within the H/E cone area. The modified procedure was found to be capable of increasing the electron reconstruction efficiency of $\mathrm{X} \to \mathrm{HH} \to \mathrm{b\bar{b}WW^*} \to \mathrm{b\bar{b}}e\nu_e\mathrm{q\bar{q}}$ by about 50% at $m_{\mathrm{X}} = 1$ TeV to over 100% at 3 TeV. Unfortunately, the data reprocessing campaign that included this change in the electron reconstruction software began deep into the analysis process covered in this dissertation, and so the data used in these results does not benefit from the recovered efficiency. However, future searches like this one will benefit from the increased electron reconstruction efficiency.

The reconstruction, ID, and isolation efficiencies each have associated systematic uncertainties that account for discrepancies with the efficiencies in data, typically measured with $\mathrm{Z}/\gamma^* \to \ell\ell$ events. Scale factors are applied to the event weights to correct the simulated efficiencies to the data as well, and the uncertainties on the scale factors factor into the total systematic uncertainty. However, the $\mathrm{Z}/\gamma^* \to \ell\ell$ events used to measure reconstruction and ID efficiencies cannot adequately describe the lepton efficiencies in our analysis phase space, since our efficiencies are heavily influenced by the jet proximity in the 1ℓ channel. Additional uncertainties associated with the lepton identification and isolation are applied to the signal yield, discussed more in Section 5.2.

4.3.2 Jets

This analysis considers a number of different jets that are clustered from the PF candidates. As described in Section 3.3.5, all jets are clustered using the anti-k_T algorithm with two different values for the distance parameter: $R = 0.4$ and 0.8. The different jets used in this analysis are:

AK8 jets: Boosted $\mathrm{H} \to \mathrm{b\bar{b}}$ have collimated decays that are typically contained within $R = 0.8$. These large-R jets are used as initial seeds for $\mathrm{H} \to \mathrm{b\bar{b}}$ identification before applying further selections on the jet substructure and b tagging.

Lepton-subtracted (LS) AK8 jets: Used only in the 1ℓ channel analysis, these are large-R jets used specifically to reconstruct the $\mathrm{W} \to \mathrm{q\bar{q}}$ decays that contain within its jet cone the lepton from the $\mathrm{W} \to \ell\nu$ decay. Before clustering the jets, the lepton is removed from the PF collection so that it is not included when calculating jet substructure quantities. Leptons that pass the 1ℓ channel identification and IP criteria described in Section 4.3.1 are removed from the collection of PF particles. This lepton subtraction is not applied for the $\mathrm{H} \to \mathrm{b\bar{b}}$ reconstruction because it may remove leptons from the decays of the b quarks in $\mathrm{H} \to \mathrm{b\bar{b}}$.

AK4 jets These $R = 0.4$ jets are used primarily to identify b jets from top quark decays, important because $\mathrm{t\bar{t}}$ is the largest background. This identification is used to suppress the $\mathrm{t\bar{t}}$ background.

Trigger AK4 jets These $R = 0.4$ jets are used to calculate the event H_T. Looser constraints are applied to these jets than to the AK4 jets listed above.

As described in Section 3.3.5, all jets are corrected for non-uniform reconstruction. Jets are also corrected for the presence of contributions from pileup with different strategies for AK4 and AK8 jets [119]. For AK4 jets, charged hadrons that are associated to a pileup vertex are removed completely from the jet momentum with the Charged-Hadron Subtraction (CHS) technique within the PF algorithm [10]. On the other hand, AK8 jets are calibrated for pileup contributions using the the Pileup Per Particle Identification (PUPPI) algorithm [120]. Rather than remove PF candidates associated with pileup from the jet computations, the PUPPI algorithm first calculates a local shape variable α for each PF candidate in the jet, used to distinguish parton shower-like radiation from pileup radiation. Using this shape α and other variables like p_{T}, a weight with values from 0.0 to 1.0 is assigned to each jet constituent in order to rescale the momentum of each. Ideally, particles from the hard scatter receive a weight of 1.0, while those from pileup receive a weight of 0.0. Furthermore, all jets except the trigger AK4 jets are cleaned with a set of criteria meant to reject fake, poorly reconstructed, and noise jets while maintaining a 98–99% efficiency for real jets [121]. This criteria corresponds to a tight jet ID working point, defined with selections on the set of jet constituents:

- Neutral hadron fraction < 0.90
- Neutral EM fraction < 0.90
- Charged hadron fraction > 0.00
- Charged EM fraction < 0.99
- Number of constituents ≥ 2
- Number of charged constituents ≥ 1

Table 4.3 contains a summary of the corrections and acceptance selections for each of the four classes of jets used.

Table 4.3: Summary of jet collection properties.

Collection	R	Pileup corrections	p_{T} selection	$\|\eta\|$ selection	Jet ID
AK8 jets	0.8	PUPPI	> 50 GeV	< 2.4	tight
LS AK8 jets	0.8	PUPPI	> 50 GeV	< 2.4	tight
AK4 jets	0.4	CHS	> 30 GeV	< 2.4	tight
Trigger AK4 jets	0.4	CHS	> 30 GeV	< 5.0	none

4.3.3 Boson reconstruction

In order to best understand the entire HH decay chain, it is necessary to reconstruct the Higgs and W bosons that are intermediate states of the event. The hadronically decaying Higgs and W bosons are reconstructed as the aforementioned $\mathrm{b\bar{b}}$ jet and $\mathrm{q\bar{q}}$ jet, respectively. Bosons are first reconstructed and selected with a baseline set of criteria. Then, additional criteria are imposed to classify the event into a search region or control region. Boson candidate selection is defined such that there can be only one $\mathrm{b\bar{b}}$ jet in the event and additionally only one $\mathrm{q\bar{q}}$ jet in the 1ℓ channel.

Understanding the jet substructure is crucial for measurement and identification of hadronically decaying bosons. In this dissertation, we are concerned with two such states: $\mathrm{H} \to \mathrm{b\bar{b}}$ in both channels and $\mathrm{W} \to \mathrm{q\bar{q}}$ in the 1ℓ channel. Identifying these states relies on the reconstructions of the local energy deposits of each of the two quarks within the AK8 jet, referred to as subjets. Reconstruction of the subjets is necessary to discriminate signal against background jets and to mitigate contamination from initial state radiation and radiation from the underlying event and pileup. Subjet reconstruction is performed with the soft-drop (SD) algorithm [122], a version of the modified mass-drop tagger [123], which removes soft, wide-angle radiation from a jet with an iterative declustering procedure.

The SD algorithm is described here. For each AK8 jet, the PF constituents of that jet alone are reclustered using the C-A algorithm rather than the anti-k_{T} algorithm to form a pairwise clustering tree with an angular-ordered structure. The algorithm then unwinds iteratively with the following procedure:

1. Divide the jet j into two subjets j_1 and j_2 by undoing the last step of the C-A clustering. Order the jets such that $m_{j_1} > m_{j_2}$.

2. Check if the subjets pass the SD condition, defined as:

$$\text{Soft-drop condition: } \frac{\max\left(p_{\mathrm{T}1}, p_{\mathrm{T}2}\right)}{p_{\mathrm{T}1} + p_{\mathrm{T}2}} > z_{\mathrm{cut}} \left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}, \tag{4.3}$$

   where ΔR_{12} is the angular distance between j_1 and j_2, R_0 is the clustering distance parameter (0.8 in our case), z_{cut} is the SD threshold, and β is an angular exponent. If the subjets pass the SD condition, then deem jet j to be the final SD jet.

3. Otherwise, if the SD condition is not fulfilled, redefine j to be the jet j_1 (with larger p_{T}) and iterate the procedure, returning to Step 1.

By construction, the SD condition fails for wide-angle, soft radiation. The parameters z_{cut} and β are tuned to control the degree of jet grooming; for the hadronically decaying bosons in this dissertation, the values used are $z_{\mathrm{cut}} = 0.1$ and $\beta = 0$. By iteratively declustering the jet into pairs of subjets, the goal is that the proper subjets will be found once a sufficiently large drop in the invariant mass of j_1 and j_2 is observed relative to the mass of j. The use of this procedure not only discriminates two-body decays from a massive object from large decays of gluons or quarks, but it also improves the mass resolution when jet mass is taken to be the invariant mass of the subjets.

We now review the reconstruction and selection of the various intermediate-state bosons.

$\mathrm{H} \to \mathrm{b\bar{b}}$

The $\mathrm{b\bar{b}}$ jet is chosen from at most the two leading AK8 jets in p_{T} that have $p_{\mathrm{T}} > 200$ GeV in the event. Of these two jets, the $\mathrm{H} \to \mathrm{b\bar{b}}$ candidate is chosen as the highest p_{T} jet that is sufficiently far away from certain objects in the event; in the 1ℓ channel, these objects are the lepton and the $\mathrm{q\bar{q}}$ jet, while in the 2ℓ channel, these are the two leptons:

- 1ℓ channel: $\Delta R(\mathrm{b\bar{b}}\text{ jet}, \mathrm{q\bar{q}}\text{ jet}) > 1.6$ AND $|\Delta\phi(\mathrm{b\bar{b}}\text{ jet}, \ell)| > 2.0$
- 2ℓ channel: $|\Delta\phi(\mathrm{b\bar{b}}\text{ jet}, \vec{p}_{\ell\ell})| > 2.0$ AND $|\Delta\phi(\mathrm{b\bar{b}}\text{ jet}, \ell)| > 0.8$ for each ℓ

Furthermore, this jet candidate is required to contain two SD subjets each satisfying $p_{\mathrm{T}} > 20$ GeV and $|\eta| > 2.4$.

The SD mass $m_{\mathrm{b\bar{b}}}$ of this jet is one of the two search variables of this analysis. The raw $m_{\mathrm{b\bar{b}}}$ without any corrections is on average less than $m_{\mathrm{H}} = 125$ GeV due to the energy loss from b quark decay, and this energy loss is also p_{T}-dependent. We therefore correct $m_{\mathrm{b\bar{b}}}$ to remove this offset but also primarily to remove the p_{T} dependence to ensure a uniform response for hypothetical signals with different m_{X}. The correction is obtained from signal simulation as a function of p_{T} and η and is applied as follows to obtain the corrected mass $m_{\mathrm{b\bar{b}}}^{\mathrm{corr}}$:

$$m_{\mathrm{b\bar{b}}}^{\mathrm{corr}} = m_{\mathrm{b\bar{b}}} * \mathrm{C}_{\mathrm{SD}}(p_{\mathrm{T}} * \mathrm{C}_{p_{\mathrm{T}}}(p_{\mathrm{T}}, \eta), \eta), \tag{4.4}$$

The correction $\mathrm{C}_{p_{\mathrm{T}}}$ accounts for the jet p_{T} response, calibrating the jet p_{T} so that it matches the particle-level p_{T} on average. Then, in bins of p_{T} and η, the $m_{\mathrm{b\bar{b}}}$ spectrum is fit with a resonant template (the same used in the background modeling, described in detail in Section 5.1) to obtain a measurement of the $m_{\mathrm{b\bar{b}}}$ resonance. The correction C_{SD} is the known Higgs boson mass of 125 GeV divided by this measured value. The effect on the measured $m_{\mathrm{b\bar{b}}}$ due to various corrections is demonstrated as a function of m_{X} in Fig. 4.6. The correction reduces the $m_{\mathrm{b\bar{b}}}$ p_{T} dependence while producing the average mass accurately. There is a single correction, given the p_{T} and $|\eta|$, applied to the jet regardless of the year of the run. While this correction was derived using only 2016 simulation, the corrected behavior is consistent for 2016–2018, demonstrated in Fig. 4.6.

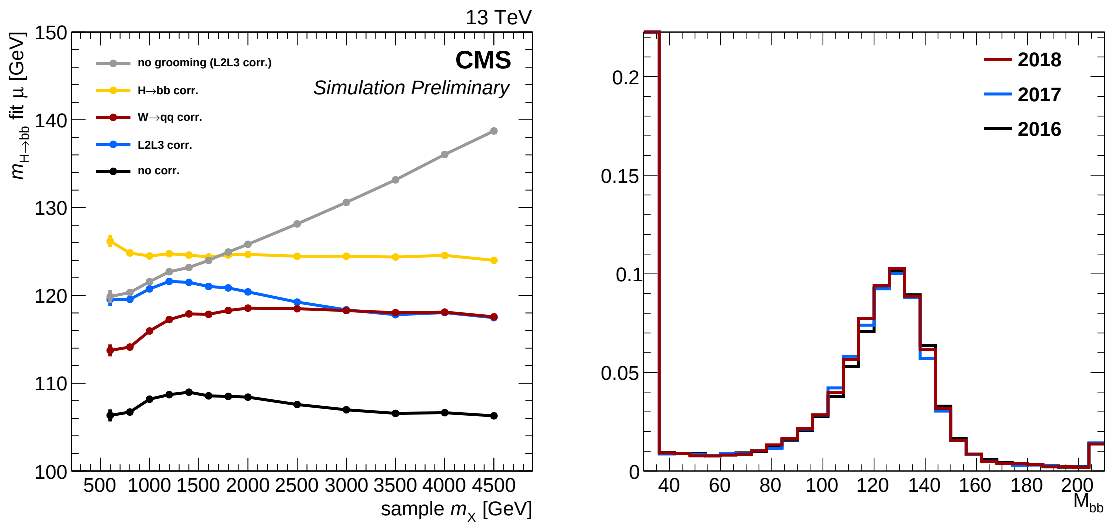

Figure 4.6: Left: the fitted $m_{\mathrm{b\bar{b}}}$ as a function of hypothetical m_{X} for various schemes of jet correction. The AK8 jet SD mass before applying jet energy corrections is given by "no corr." The gray curve is obtained by applying standard CMS jet energy corrections to the raw jet mass. The blue curve corresponds to the SD mass of the AK8 jet in the gray curve. The red curve shows the SD mass when applying corrections to the blue curve that were measured particularly for SD jets from $\mathrm{W} \rightarrow \mathrm{q\bar{q}}$ decays. The relative mass reduction here is unsuitable for $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$ decays. The proper correction described in this work and applied to the blue curve produces the yellow curve, which puts $m_{\mathrm{b\bar{b}}}$ at 125 GeV on average across the full range of m_{X}. Right: the corrected MC distribution for a spin-0 signal of $m_{\mathrm{X}} = 1.0$ TeV compared for each year of data-taking. The spike in events below 20 GeV correspond to jets for which the SD algorithm overgrooms and removes too much energy.

Finally, as first mentioned in Section 3.3.5, the quark flavor content of the $\mathrm{b\bar{b}}$ jet is evaluated with the mass-decorrelated DeepAK8 $\mathrm{Z/H} \rightarrow \mathrm{b\bar{b}}$ tagger. The use of this tagger to construct search and control regions will be discussed in Section 4.4.

$\mathbf{W \to q\bar{q}}$

An AK8 jet encapsulating the decay of a hadronic W boson is only reconstructed in the 1ℓ channel. Only the LS AK8 jet closest in ΔR to the selected lepton is considered as the candidate $\mathrm{q\bar{q}}$ jet, and this jet must also satisfy the following criteria:

- $\Delta R(\mathrm{q\bar{q}\ jet,\ lepton}) < 1.2$
- $p_{\mathrm{T}} > 50$ GeV
- Two subjets each with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$

The ΔR requirement reduces backgrounds such as W+jets while maintaining high signal efficiency. Two subjets with p_{T} high enough to be consistent with at least a moderately boosted $\mathrm{W \to q\bar{q}}$ decay are required. An identification scheme to distinguish between two-prong and single-prong substructure in this jet is also used and will be discussed in Section 4.4.

Single-lepton $\mathbf{H \to WW^*}$

To reconstruct the Higgs boson four-momentum in the $\mathrm{H \to WW^* \to \ell\nu q\bar{q}}$ decay chain from the visible and invisible decay products, a likelihood-based technique is employed. For each event, values for the following five parameters are extracted by maximizing a likelihood function:

$\vec{p}_{\nu}$: the three components of the neutrino momentum. The $p_{\mathrm{T}}^{\mathrm{miss}}$ can help us infer the transverse components of the neutrino momentum, but it offers no insight into the longitudinal component. The other physics objects in the event help to provide more information on these momentum components.

$V_{\mathrm{q\bar{q}}}$: a boolean indicator of whether the $\mathrm{q\bar{q}}$ jet favors a larger or smaller mass than the $\mathrm{W \to \ell\nu}$ boson decay. Because $m_{\mathrm{H}} < 2m_{\mathrm{W}}$, typically one of the W bosons will be on-shell with a mass near the true mass of 80 GeV, and the other will be off-shell with a much lower mass ($\approx$ 30 GeV on average). This parameter encodes this on/off-shell information by gauging which of the bosons has lower mass.

$\boldsymbol{R_{q\bar{q}}}$**:** a jet response correction, a multiplicative scale factor on the p_{T} of the $q\bar{q}$ jet. Because this jet has a large energy resolution, the jet p_{T} is allowed to vary. Neither the mass nor the angle of the jet is altered in the fit.

With these parameters determined, the $H \rightarrow WW^*$ four-momentum can be fully determined as the summed four-momentum of the neutrino (p_ν), the $q\bar{q}$ jet (modified by $R_{q\bar{q}}$), and the lepton. This method assumes that the lepton is reconstructed perfectly, and so the lepton momentum is not varied in the fit.

The likelihood function is constructed with six probability density functions (pdfs) $P(x|\vec{y})$ estimated from signal simulation, where x is the corresponding observable for the pdf, and $\vec{y}$ represents the set of free parameters associated to that pdf. For example, m_{jet} is one such observable, and $V_{q\bar{q}}$ is one such free parameter. These pdfs are represented as one-dimensional histograms P and act as constraints on the free parameters mentioned above. The full likelihood function is:

$$\mathcal{L} = P(m_{\mathrm{jet}}|V_{q\bar{q}})P(p_{\mathrm{T,jet}}|R_{q\bar{q}}, V_{q\bar{q}})P(m_{\ell\nu q\bar{q}}|\vec{p}_\nu, R_{q\bar{q}}, V_{q\bar{q}})P(m_{\ell\nu}|\vec{p}_\nu, V_{q\bar{q}})P(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}|\vec{p}_\nu, V_{q\bar{q}}). \quad (4.5)$$

Below, we provide specific details for each of the pdf factors in the likelihood function. Any variable that is denoted as "true" is a quantity taken from particle-level information in simulation. Separate sets of pdfs are derived for each year of data-taking to account for differences in reconstruction among the years.

$\boldsymbol{P(m_{\mathrm{jet}}|V_{q\bar{q}})}$: the soft-drop mass of the $q\bar{q}$ jet. This pdf is implemented very coarsely with only two bins so that the likelihood $\mathcal{L}$ does not depend strongly on the precise modeling of the soft-drop mass shape. Figure 4.8 shows this distribution binned more finely than implemented for signal events, categorized by which W boson has the lower true mass. The division at 60 GeV yields good separation between the two categories.

$\boldsymbol{P(p_{\mathrm{T,jet}}|R_{\mathrm{q\bar{q}}}, V_{\mathrm{q\bar{q}}})}$: this pdf is constructed as the jet p_{T} response, parameterized as the ratio of the reconstructed $\mathrm{q\bar{q}}$ jet p_{T} to the true jet p_{T}. With this parameterization, the dependence of the pdf for the true $p_{\mathrm{T,jet}}$ on m_{X} of the signal is eliminated. By fitting for this jet response correction, we can better determine the true momentum of the $\mathrm{q\bar{q}}$ jet, whose energy measurement has large uncertainties. This distribution is shown in Fig. 4.8.

$\boldsymbol{P(m_{\ell\nu}|\vec{p}_{\nu}, V_{\mathrm{q\bar{q}}})}$: the mass distribution of the $\mathrm{W} \to \ell\nu$ decay, whose probability distribution is built using the true neutrino three-momentum and the reconstructed lepton momentum. Examples of this distribution for both the on- and off-shell cases of the $\mathrm{q\bar{q}}$ jet can be seen in Fig. 4.7.

$\boldsymbol{P(m_{\ell\nu\mathrm{q\bar{q}}}|\vec{p}_{\nu}, R_{\mathrm{q\bar{q}}}, V_{\mathrm{q\bar{q}}})}$: the mass distribution of the reconstructed $\mathrm{H} \to \mathrm{WW}^* \to \ell\nu\mathrm{q\bar{q}}$ decay, whose probability distribution is built using the true neutrino three-momentum, the reconstructed lepton momentum, and a hybrid $\mathrm{q\bar{q}}$ jet momentum constructed with the particle-level p_{T}, the reconstructed η and ϕ, and a fixed mass of either 30 or 81 GeV depending on whether the $\mathrm{W} \to \mathrm{q\bar{q}}$ decay is truly off- or on-shell, respectively. Examples of this distribution for both the on-shell and off-shell cases of the $\mathrm{q\bar{q}}$ jet can be seen in Fig. 4.7.

$\boldsymbol{P(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}|\vec{p}_{\nu}, V_{\mathrm{q\bar{q}}})}$: the product of two pdfs, each corresponding to one component of the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$:

$$P(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}|\vec{p}_{\nu}, V_{\mathrm{q\bar{q}}}) = P(p_{\mathrm{T},\parallel}^{\mathrm{miss}}|\vec{p}_{\nu}, V_{\mathrm{q\bar{q}}})P(p_{\mathrm{T},\perp}^{\mathrm{miss}}|\vec{p}_{\nu}, V_{\mathrm{q\bar{q}}}). \tag{4.6}$$

This missing momentum in signal events can be modeled without a strong dependence on m_{X} by considering the reference frame of the momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{reco}}$, given in

Eq. (4.7):

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{reco}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} + (\vec{p}_{\ell} + \vec{p}_{\mathrm{q\bar{q}\ jet}})_{\mathrm{T}}. \tag{4.7}$$

Rotating the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ into the frame of $\vec{p}_{\mathrm{T}}^{\,\mathrm{reco}}$, we can nicely model each of the parallel and orthogonal components of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ relative to $\vec{p}_{\mathrm{T}}^{\,\mathrm{reco}}$. The parallel component arises from jet mismeasurements, particularly in the energy of the $\mathrm{b\bar{b}}$ jet that tends to be approximately opposite $\vec{p}_{\mathrm{T}}^{\,\mathrm{reco}}$ in the transverse plane. The orthogonal component is smaller and arises from underlying event interactions and pileup. As implemented in Eq. (4.5), the pdfs are constructed as the "extra" missing momentum relative to the neutrino momentum, $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} - \vec{p}_{\mathrm{T},\nu}$, and are parameterized not to have strong dependence on m_{X} of signal. The parallel component is dependent on m_{X} because the $\mathrm{b\bar{b}}$ jet mismeasurement is affected by the event boost, so this extra momentum is modeled as $(p_{\mathrm{T},\parallel}^{\mathrm{miss}} - p_{\parallel}^{\nu})/m_{\mathrm{X}}$ to remove such dependence. The orthogonal component does not depend on the event boost, so this momentum is modeled as $(p_{\mathrm{T},\perp}^{\mathrm{miss}} - p_{\perp}^{\nu})$. These distributions are derived in signal simulation using the true neutrino momentum and the reconstructed $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, and examples are given in Fig. 4.9.

The pdfs that comprise Eq. (4.5) are built such that the dependence on m_{X} is suppressed, but some residual dependence remains. To further ensure minimal dependence of these pdfs on the event boost, we also parameterize each P in Eq. (4.5) as a function of $p_{\mathrm{T}}^{\mathrm{reco}}$ in Eq. (4.7). Two sets of pdfs are produced, one at low $p_{\mathrm{T}}^{\mathrm{reco}}$ ($\lesssim 600$ GeV) and one at high $p_{\mathrm{T}}^{\mathrm{reco}}$ ($\gtrsim 1400$ GeV), and then event-by-event, the histogram P of the pdf is obtained by interpolating between the histograms at the two regimes of $p_{\mathrm{T}}^{\mathrm{reco}}$. This interpolation is performed linearly as a function of the event $p_{\mathrm{T}}^{\mathrm{reco}}$. The forms of the P are all dependent on whether the hadronically decaying W boson is heavier than the leptonically decaying W, so each term is dependent on the free parameter $V_{\mathrm{q\bar{q}}}$. Correlations among the observables in the likelihood were studied and found not to affect the result significantly.

In the first iteration of this analysis [72], this reconstruction was performed by constrain-

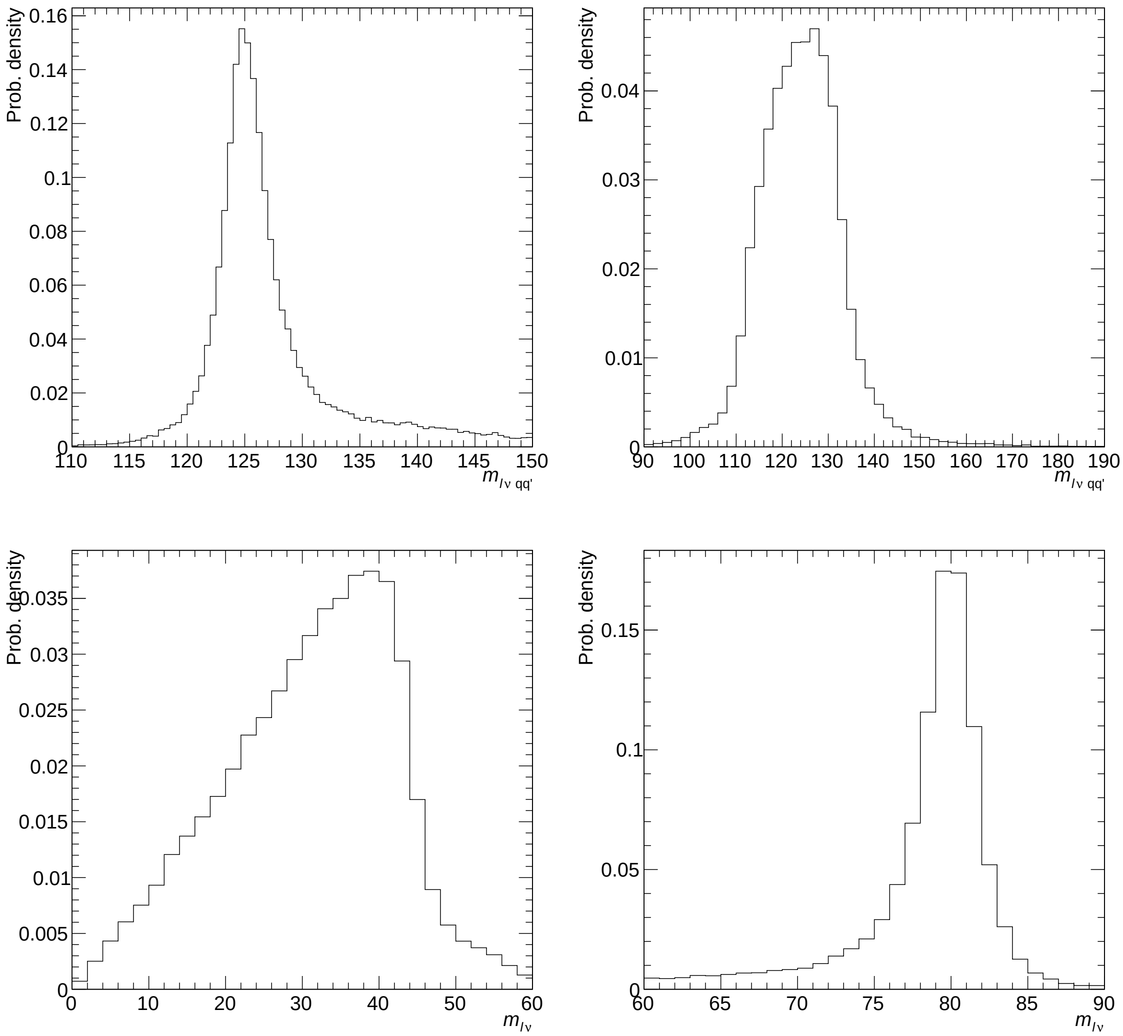

Figure 4.7: Examples histograms of the pdfs in the likelihood of Eq. (4.5) for the observables $m_{\ell\nu\mathrm{q\bar{q}}}$ and $m_{\ell\nu}$ and the on-shell (left) and virtual (right) $\mathrm{W} \to \mathrm{q\bar{q}}$ cases. The top row shows $P(m_{\ell\nu\mathrm{q\bar{q}}}|\vec{p}_\nu, R_{\mathrm{q\bar{q}}}, V_{\mathrm{q\bar{q}}})$ and the bottom shows $P(m_{\ell\nu}|\vec{p}_\nu, V_{\mathrm{q\bar{q}}})$. These histograms correspond to the low $p_{\mathrm{T}}^{\mathrm{reco}}$ set of pdfs.

ing the $\mathrm{H} \to \mathrm{WW}^*$ mass at 125 GeV and solving for $\vec{p}_\nu$ with a simple second-order equation, using the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ directly for the transverse components of $\vec{p}_\nu$. The signal m_{HH} resolution is similar between the previous and the current method, but the background mass spectrum is softer with the current likelihood method, as can be seen in Fig. 4.10. We take advantage

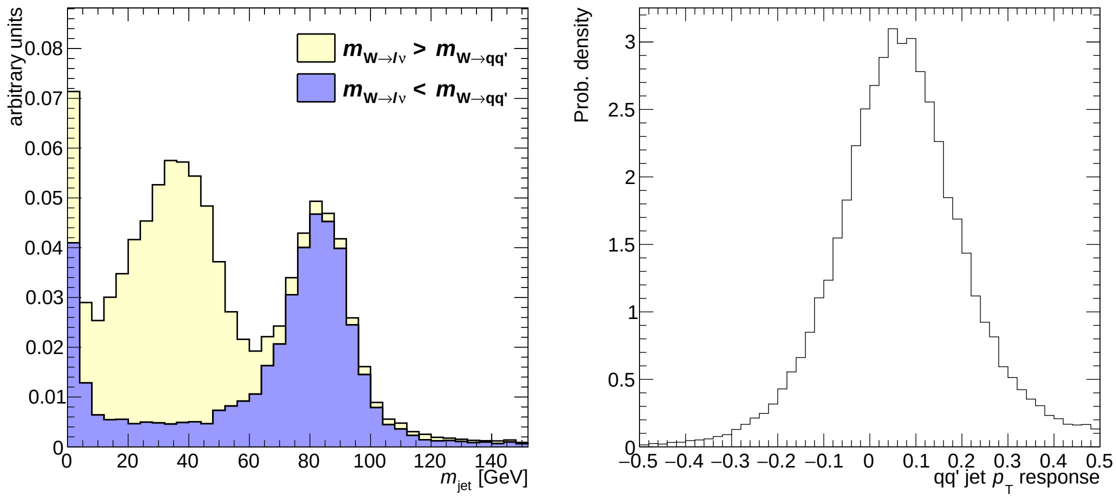

Figure 4.8: Example histograms of the pdfs in the likelihood of Eq. (4.5) for the observables m_{jet} and R_{jet}. Left: comparison of the m_{jet} distribution for the on-shell and virtual $\mathrm{W} \to \mathrm{q\bar{q}}$ cases. The pdfs $P(m_{\mathrm{jet}}|V_{\mathrm{q\bar{q}}})$ are created from this distribution rebinned to only two bins divided at 60 GeV. The histograms are binned so coarsely as to remain insensitive to the precise soft-drop mass modeling. Right: the pdf of the jet p_{T} $P(p_{\mathrm{T,jet}}|R_{\mathrm{q\bar{q}}}, V_{\mathrm{q\bar{q}}})$, constructed as the jet response function, for the virtual $\mathrm{W} \to \mathrm{q\bar{q}}$ and low $p_{\mathrm{T}}^{\mathrm{reco}}$ cases.

of this fact using an alternative likelihood $\mathcal{L}_{\mathrm{alt}}$, which is less constrained by the intermediate masses. Instead of fitting for the neutrino p_z, we include the masses of the $\mathrm{H} \to \mathrm{WW}^*$ and $\mathrm{W} \to \ell\nu$ decays as free parameters. Thus, difference between the change from the nominal to the alternative likelihood is given by:

$$\mathcal{L}_{\mathrm{alt}} = P(m_{\mathrm{jet}}|V_{\mathrm{q\bar{q}}})P(p_{\mathrm{T,jet}}|R_{\mathrm{q\bar{q}}}, V_{\mathrm{q\bar{q}}})P(m_{\ell\nu\mathrm{q\bar{q}}}|m_{\mathrm{WW}^*}, V_{\mathrm{q\bar{q}}})P(m_{\ell\nu}|m_{\mathrm{W}\to\ell\nu}, V_{\mathrm{q\bar{q}}})P(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}|\vec{p}_{\mathrm{T}}^{\,\nu}, V_{\mathrm{q\bar{q}}}), \tag{4.8}$$

where m_{WW^*} and $m_{\mathrm{W}\to\ell\nu}$ are the free parameters for the masses of the $\mathrm{H} \to \mathrm{WW}^*$ and $\mathrm{W} \to \ell\nu$ decays, respectively. We then use both likelihoods to construct a discriminating

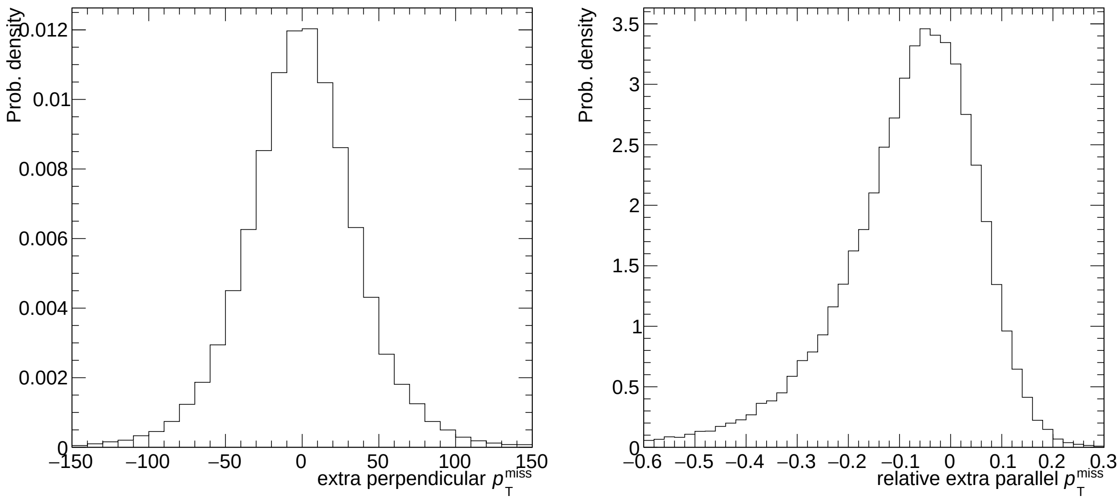

Figure 4.9: Example histograms of the pdfs in the likelihood of Eq. (4.5) for the $p_{\mathrm{T}}^{\mathrm{miss}}$ observables. Left: $P\ (p_{\mathrm{T},\parallel}^{\mathrm{miss}}\ |\vec{p}_\nu, V_{\mathrm{q\bar{q}}})$, the extra $p_{\mathrm{T}}^{\mathrm{miss}}$ relative to the neutrino momentum parallel to $\vec{p}_{\mathrm{T}}^{\mathrm{\,reco}}$. This distribution is modeled as $(p_{\mathrm{T},\parallel}^{\mathrm{miss}} - p_\parallel^\nu)/m_{\mathrm{X}}$ to minimize the dependence on signal m_{X}. Right: $P\ (p_{\mathrm{T},\perp}^{\mathrm{miss}}\ |\vec{p}_\nu, V_{\mathrm{q\bar{q}}})$, the extra $p_{\mathrm{T}}^{\mathrm{miss}}$ relative to the neutrino momentum orthogonal to $\vec{p}_{\mathrm{T}}^{\mathrm{\,reco}}$. This distribution is modeled as $(p_{\mathrm{T},\perp}^{\mathrm{miss}} - p_\perp^\nu)$ and does not have any inherent dependence on m_{X}, as the extra $p_{\mathrm{T}}^{\mathrm{miss}}$ in this direction arises from underlying event interactions such as pileup. The pdfs are shown for the virtual $\mathrm{W} \to \mathrm{q\bar{q}}$ and low $p_{\mathrm{T}}^{\mathrm{reco}}$ cases.

variable between signal and background:

$$D_{\ell\nu\mathrm{q\bar{q}}} \equiv -2\log \mathcal{L}/\mathcal{L}_{\mathrm{alt}}, \tag{4.9}$$

where $\mathcal{L}$ is the likelihood described in Eq. (4.5) and $\mathcal{L}_{\mathrm{alt}}$ is the alternative likelihood. The use of $D_{\ell\nu\mathrm{q\bar{q}}}$ to discriminate between signal and background is discussed in Section 4.4.

Dilepton $\mathbf{H \to WW^*}$ and $\mathbf{H \to \tau\tau}$

In the 2ℓ channel, the Higgs boson that does not decay to $\mathrm{b\bar{b}}$ is reconstructed by summing the dilepton four-momentum $p_{\ell\ell}$ and the invisible four-momentum p_{inv} due to neutrinos, which

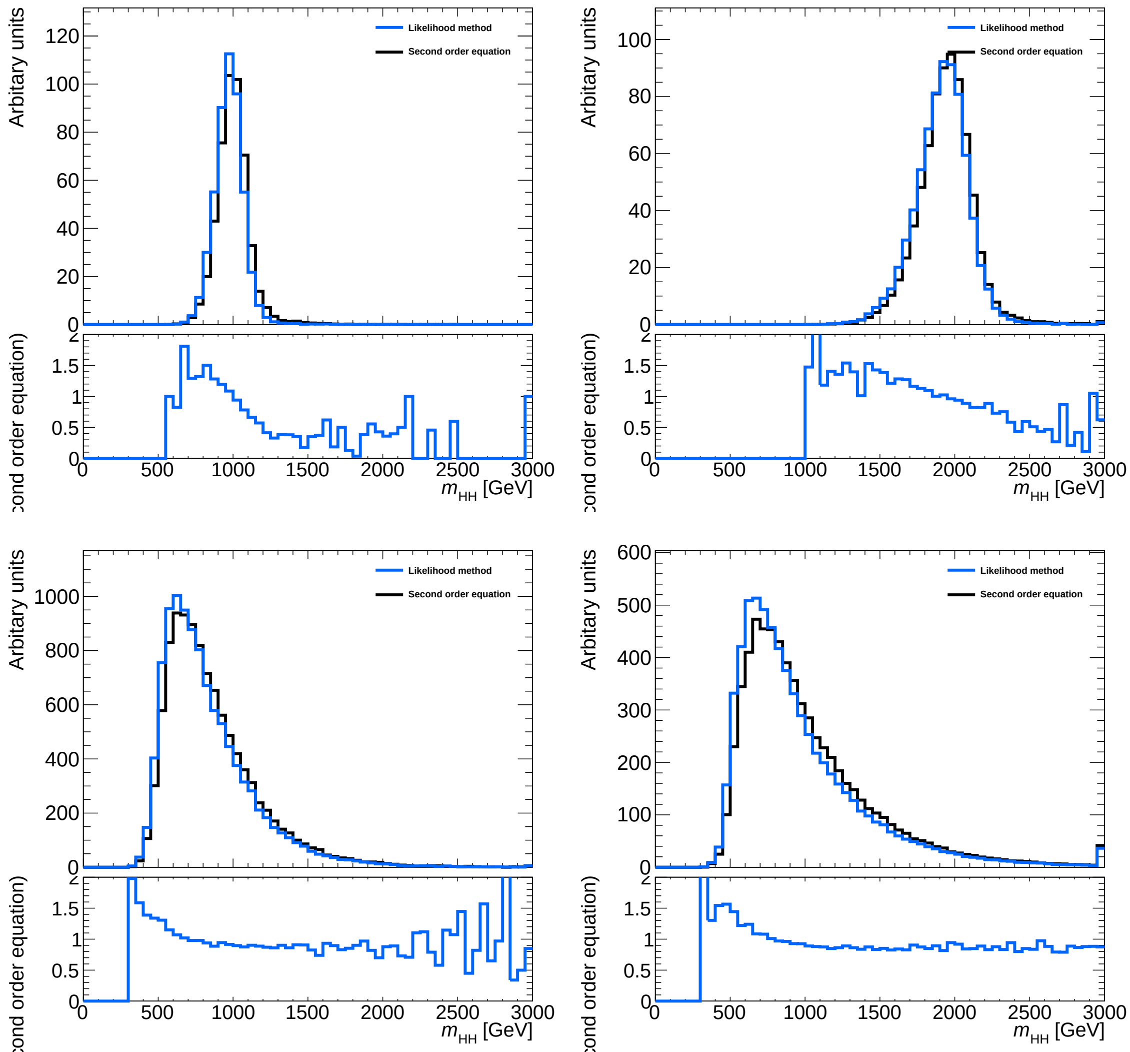

Figure 4.10: The reconstructed $m_{\ell\nu\mathrm{q}\bar{\mathrm{q}}}$ obtained by either minimizing $-2\log\mathcal{L}$ or solving a second-order equation. Top: 1 TeV (left) and 2 TeV (right) signals. Bottom: $\mathrm{t}\bar{\mathrm{t}}$ (left) and $\mathrm{W}+\mathrm{jets}$ (right) background processes.

must be determined. The absence of a $\mathrm{q}\bar{\mathrm{q}}$ jet and the presence of larger $p_{\mathrm{T}}^{\mathrm{miss}}$ in this channel limits the effectiveness of a likelihood-based technique for Higgs boson reconstruction. Instead, we make simple assumptions of the decay kinematics, motivated by the simulation of signal, to determine p_{inv}. First, the transverse components of p_{inv} are taken directly from $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. Second, because of the boost of the Higgs decay, we assume the polar angle θ of

p_{inv} to be equal to that of $p_{\ell\ell}$: $\theta_{\text{inv}} = \theta_{\ell\ell}$. The distribution of $\theta_{\text{inv}} - \theta_{\ell\ell}$ is narrow due to the large boost of the H boson, even in the low-mass regime, as can be seen in Fig. 4.11. With this constraint, the z-component of p_{inv} is obtained. Finally, the invariant mass of all neutrinos in the event is estimated to be 55 GeV on average from the true distribution from simulation, seen in Fig. 4.11. The analysis is not sensitive to the choice of reconstruction solution because $p_{\mathrm{T}}^{\text{miss}}$ is on the order of the $\mathrm{b\bar{b}}$ jet resolution, and thus there is limited useful information in the $\vec{p}_{\mathrm{T}}^{\text{miss}}$ to reconstruct the Higgs decay more precisely.

4.4 Event selection

After reconstructing all intermediate and final states in the HH decay chain, the set of events that remain are then filtered with selections on discriminating variables that target signal-like decay topologies and reject events with background-like signatures. Some of these selections are common between the 1ℓ and 2ℓ channels, and some are unique to each one. Some variables are also used to categorize events before modeling, discussed in Sec. 5.1, and noted when applicable in this section.

In order to determine whether the selected $\mathrm{b\bar{b}}$ jet truly arises from a two-prong decay to b quarks, we employ a tagging tool called the DeepAK8 mass-decorrelated (MD) $\mathrm{Z/H} \to \mathrm{b\bar{b}}$ tagger, used primarily to discriminate $\mathrm{b\bar{b}}$ jets against jets from QCD multijet events. This was first introduced and discussed in detail in Section 3.3.5. All events in this analysis are required to have $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} > 0.8$, a selection with an efficiency for $\mathrm{X} \to \mathrm{HH}$ signal of $> 80\%$ and a total background rejection rate of $\approx 95\%$ on the $\mathrm{b\bar{b}}$ jet. The discriminator $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}}$ is also used to categorize events into either a loose (bL) or a tight (bT) category, defined by $0.8 \leq D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} < 0.97$ and $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} \geq 0.97$, respectively. Most of the signal is in the tight categories, while most of the background falls into the loose categories, making the tight categories the most sensitive to signal. The distribution $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}}$ is shown in Fig. 4.12 for all events entering the 1ℓ and 2ℓ channels. Discrepancies between the tagging efficiencies mea-

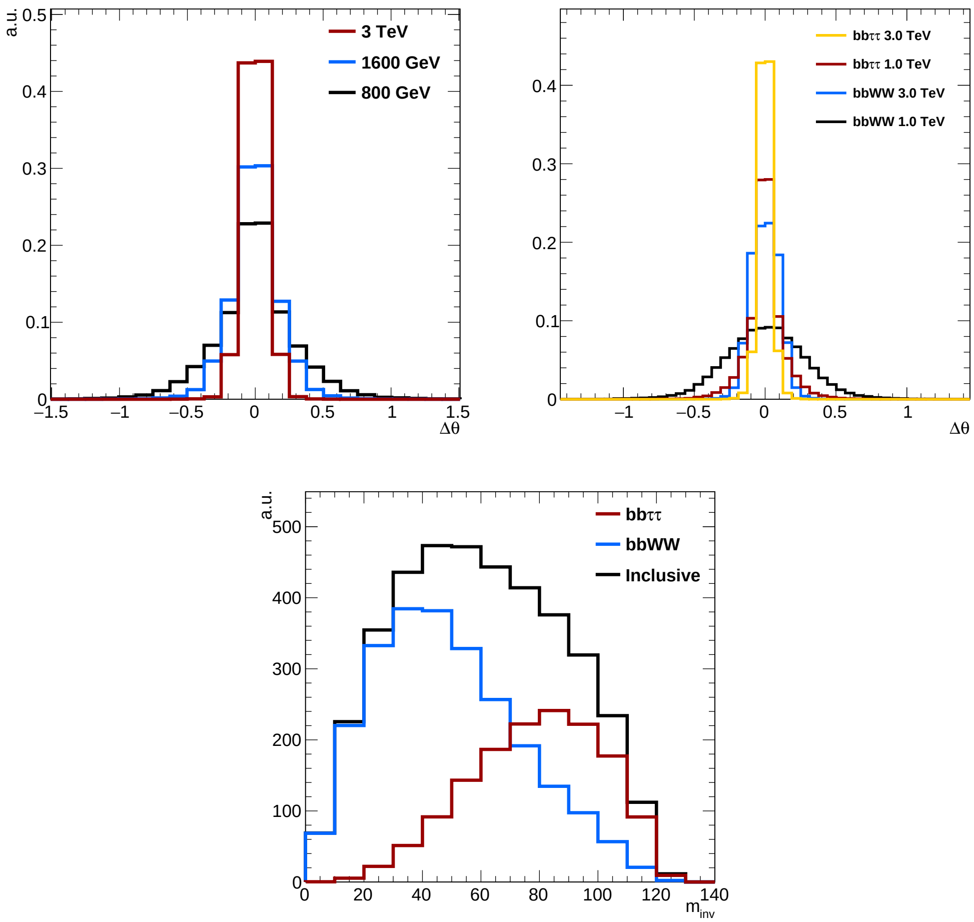

Figure 4.11: Top row: the true $\Delta\theta = \theta_{\ell\ell} - \theta_{\text{inv}}$ for 2ℓ signals. The left plot compares different m_{X} for spin-0 events containing both $\text{b}\bar{\text{b}}\text{WW}^*$ and $\text{b}\bar{\text{b}}\tau\tau$. The right plot compares $\text{b}\bar{\text{b}}\text{WW}^*$ against $\text{b}\bar{\text{b}}\tau\tau$ for a low and high m_{X} value. This distribution is narrower in $\text{b}\bar{\text{b}}\tau\tau$ than in $\text{b}\bar{\text{b}}\text{WW}^*$. Bottom: the invisible invariant mass m_{inv} for each 2ℓ decay mode and the combined 2ℓ signal for spin-0. This distribution has negligible dependence on both the spin and resonance mass of the signal.

sured in simulation and data are accounted for with multiplicative scale factors to the event weights (such that the simulated efficiency matches the data) and systematic uncertainties on the signal yield (discussed further in Section 5.2.2).

In background events from $\text{t}\bar{\text{t}}$ production, the largest background contributor, we often

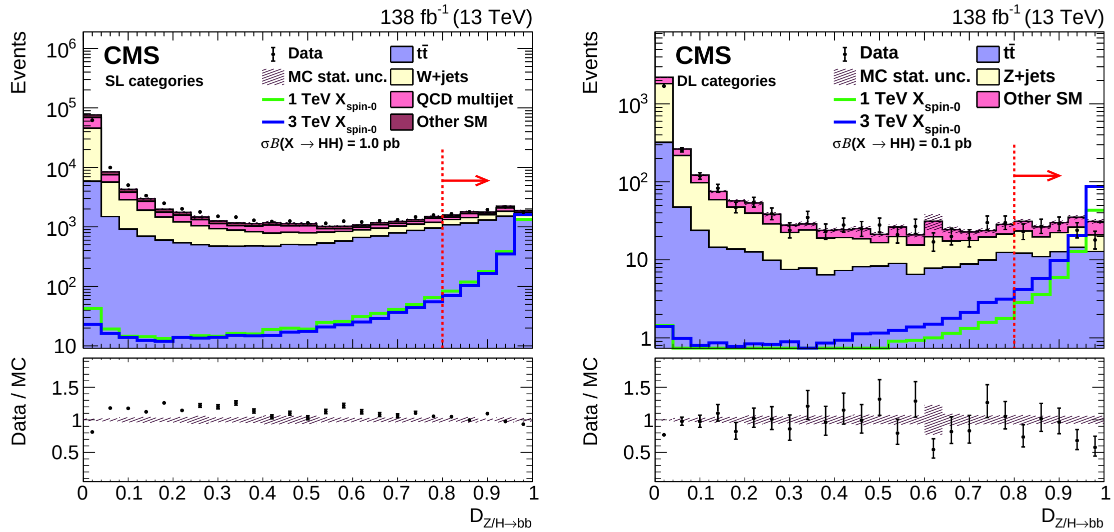

Figure 4.12: The discriminator distribution $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}}$ of the DeepAK8-MD Z/H $\to$ $\mathrm{b\bar{b}}$ tagger for the 1ℓ (left) and 2ℓ (right) channels. The data are shown as the black points with error bars, and two signal hypotheses, for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV, are shown as solid lines. The colored, stacked histograms show the important SM background processes along with the aggregate distribution of the subdominant SM backgrounds. The cross section of the signal process is arbitrarily set for ease of viewing. The red dashed line and arrow indicate that the selected region of the distribution is $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}} \geq 0.8$.

misidentify localized hadronic activity around one of the b quarks as the $\mathrm{b\bar{b}}$ jet. An AK4 jet found separated from the $\mathrm{b\bar{b}}$ jet cone and consistent with the decay of one b quark is a strong indicator that an event arises from $\mathrm{t\bar{t}}$ and not a signal. AK4 jets from b quarks are identified with the DeepJet b tagger, detailed in Section 3.3.5. Any event that includes a b-tagged jet and is $\Delta R > 1.2$ away from the $\mathrm{b\bar{b}}$ jet is excluded from the search. The operating point of the tagger used in this analysis corresponds to an efficiency of ≈80% for tagging b jets and a misidentification rate of ≈1% for light-flavor jets. Scale factors are applied to match the simulation efficiencies to data, and systematic uncertainties are applied to the signal yield (Section 5.2.2).

With the b jet veto in this search, $\mathrm{t\bar{t}}$ is rejected at a rate of 70–75%, and the signal efficiency is $> 80\%$. In the 1ℓ channel, most of the signal inefficiency is due to the misidentification of a c jet from a hadronic W decay as a b jet. However, if the veto required that the b jet be far from the $\mathrm{q\bar{q}}$ jet as well, the ability to reject $\mathrm{t\bar{t}}$ would be spoiled because the search targets boosted topologies, and the b quark is often close to a hadronic W decay. In this case, the veto would not be nearly as effective at suppressing $\mathrm{t\bar{t}}$.

In the 1ℓ channel, there are several other discriminating selections that related to various elements of the $\mathrm{H} \to \mathrm{WW}^*$ decay. Because of the boost in the HH system, the angular distribution of the Higgs bosons tends to be central and low in $|\eta|$. Background jets are also typically produced at high $|\eta|$ than the H bosons. We target such central topologies by requiring the ratio of the $\mathrm{H} \to \mathrm{WW}^*$ candidate p_{T} and m_{HH}, referred to simply as p_{T}/m, to be > 0.3. We also filter for high-quality $\mathrm{H} \to \mathrm{WW}^*$ candidates, discussed in Sec. 4.3, by requiring the likelihood ratio discriminator $D_{\ell\nu\mathrm{q\bar{q}}}$ to be ≤ 11.0. To determine how consistent the $\mathrm{q\bar{q}}$ jet reconstruction is with the two-prong jet substructure expected from $\mathrm{W} \to \mathrm{q\bar{q}}$, the discriminating metric $N-$subjettiness [124] is used. Given N jet axes $\hat{n}_k$, $N-$subjettiness is defined as:

$$\tau_N = \frac{\sum_{i \in J} p_{\mathrm{T}}^i \min\left(\Delta R_{ik}\right)}{\sum_{i \in J} p_{\mathrm{T}}^i R_0}, \tag{4.10}$$

where i runs over all particles (PF candidates) in the jet J, k runs over the N jet axes, ΔR_{ik} is the angular distance between particle i and jet axis k, and R_0 is the jet cone radius (0.8 for the $\mathrm{q\bar{q}}$ jet in our case). Lower values of $N-$subjettiness more strongly support the hypothesis of an N-prong jet substructure. We use the ratio of the 2-subjettiness to the 1-subjettiness, denoted τ_2/τ_1, to discriminate the $\mathrm{q\bar{q}}$ jet from single-prong decays by requiring that $\tau_2/\tau_1 \leq 0.75$ for all 1ℓ events in the search region. Scale factors for signal event weights and uncertainties have been measured for the τ_2/τ_1 working points with $\mathrm{W} \to \mathrm{q\bar{q}}$ decays from a $\mathrm{t\bar{t}}$-enriched set of events [125]. The systematic uncertainties are applied to the signal normalization (Section 5.2.2).

Both the discriminators τ_2/τ_1 and $D_{\ell\nu\mathrm{q\bar{q}}}$, shown in Fig. 4.14 are used together to categorize 1ℓ events based on the reconstructed $\mathrm{H} \to \mathrm{WW}^*$ decay kinematics. There are two operating

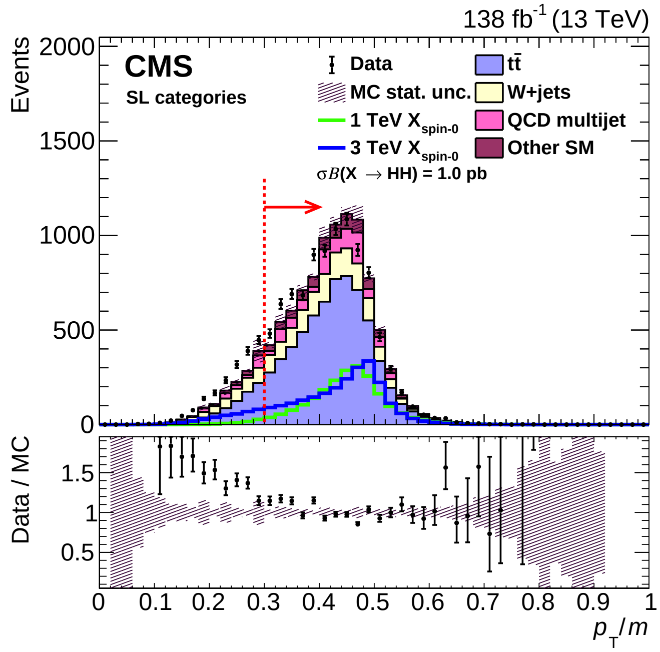

Figure 4.13: The discriminating variable p_{T}/m, defined as the ratio of the $\mathrm{H} \to \mathrm{WW}^*$ candidate p_{T} and m_{HH}. The data are shown as the black points with error bars, and two signal hypotheses, for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV, are shown as solid lines. The colored, stacked histograms show the important SM background processes along with the aggregate distribution of the subdominant SM backgrounds. The cross section of the signal process is arbitrarily set for ease of viewing. The red dashed line and arrow indicate that the selected region of the distribution is $p_{\mathrm{T}}/m \geq 0.3$.

points for each of these two variables used in the categorization. In addition to $D_{\ell\nu\mathrm{q}\bar{\mathrm{q}}} \leq 11.0$, a tighter value of 2.5, chosen with the goal of optimizing $S/\sqrt{B}$, is used to divide events into sets with different purities. Likewise, in addition to $\tau_2/\tau_1 \leq 0.75$, there is a tighter working point of 0.45, except for events from 2016, for which the working point is 0.55 because the reconstruction conditions of the CMS detector differed between these two periods of data-taking. We construct a high purity (HP) category that contains events for which either $D_{\ell\nu\mathrm{q}\bar{\mathrm{q}}} \leq 2.5$ or $\tau_2/\tau_1 \leq 0.45$ (0.55 in 2016), and we construct a low purity (LP) category with all other 1ℓ events that enter the search region.

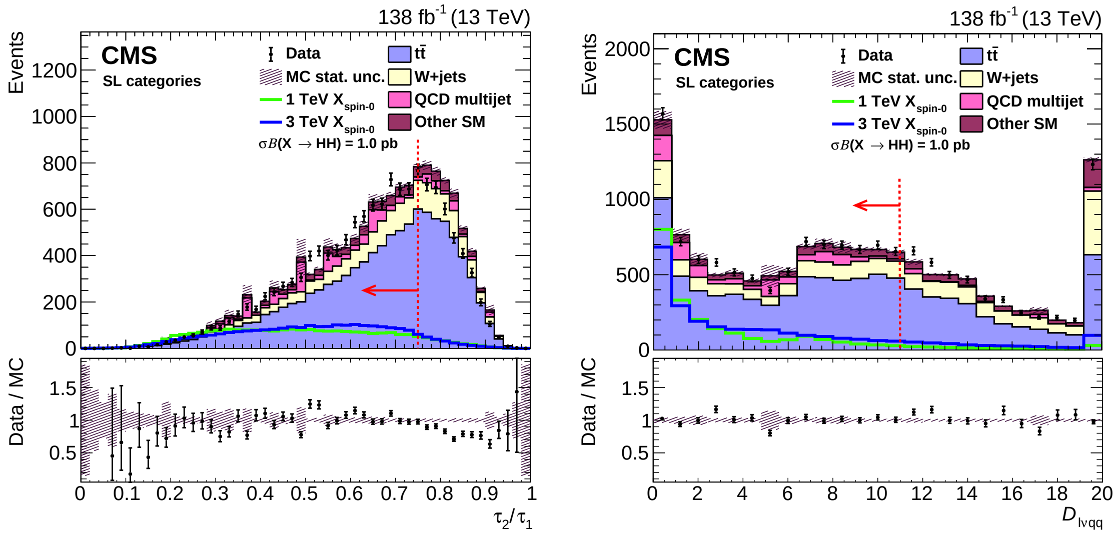

Figure 4.14: The distributions of τ_2/τ_1 (left) and $D_{\ell\nu\mathrm{q\bar{q}}}$ (right) for 1ℓ channel events. The data are shown as the black points with error bars, and two signal hypotheses, for $m_\mathrm{X} = 1.0$ and 3.0 TeV, are shown as solid lines. The colored, stacked histograms show the important SM background processes along with the aggregate distribution of the subdominant SM backgrounds. The cross section of the signal process is arbitrarily set for ease of viewing. The red dashed lines and arrows indicate the selected regions in the analysis, corresponding to $\tau_2/\tau_1 \leq 0.75$ and $D_{\ell\nu\mathrm{q\bar{q}}} \leq 11.0$. The HP category for $\mathrm{H} \to \mathrm{WW}^*$ purity contains events that have either $D_{\ell\nu\mathrm{q\bar{q}}} \leq 2.5$ or $\tau_2/\tau_1 \leq 0.45$ (0.55 in 2016). The LP category contains all other 1ℓ events in the search region.

In the 2ℓ channel, signal-like topologies are targeted by requiring the leptons and $p_\mathrm{T}^\mathrm{miss}$ to be collimated and localized in $\eta-\phi$ space. This is accomplished by requiring $p_\mathrm{T}^\mathrm{miss} \geq 85$ GeV, $\Delta\phi(\vec{p}_\mathrm{T}^\mathrm{\,miss}, \vec{p}_{\ell\ell}) \leq \pi/2$, and $\Delta R_{\ell\ell} \leq 1.0$. The first two are primarily used to suppress $\mathrm{Z}/\gamma^* + \mathrm{jets}$, wherein the $p_\mathrm{T}^\mathrm{miss}$ does not arise from neutrinos, while the latter primarily suppresses contributions from $\mathrm{t\bar{t}}$, in which the leptons are typically produced opposite each other in the transverse plane. The spectrum of the dilepton invariant mass $m_{\ell\ell}$ in signal rises and falls within the low-mass continuum and Z boson resonance regions of the $\mathrm{Z}/\gamma^* + \mathrm{jets}$ spectrum, and so this background is suppressed with $6 \leq m_{\ell\ell} \leq 75$ GeV. Each of the four aforementioned discriminating selections in the 2ℓ channel is shown in Fig. 4.15. The selections imposed in the 1ℓ and 2ℓ channels that have just been discussed are summarized in Tab. 4.4 and 4.5, respectively.

Table 4.4: 1ℓ channel search region cuts to discriminate signal against background

Description	Selection
$\mathrm{H} \to \mathrm{b\bar{b}}$ purity	$D_{\mathrm{Z/H \to b\bar{b}}} \geq 0.8$
No extra b-jets	No AK4 b-tagged jets resolved from $\mathrm{b\bar{b}}$ jet
Event centrality	$p_{\mathrm{T}}/m \geq 0.3$
$\mathrm{H} \to \mathrm{WW}^*$ purity	$D_{\ell\nu\mathrm{q\bar{q}}} \leq 11$
$\mathrm{W} \to \mathrm{q\bar{q}}$ two-prong substructure	$\tau_2/\tau_1 \leq 0.75$

Table 4.5: 2ℓ channel search region cuts to discriminate signal against background

Description	Selection
$\mathrm{H} \to \mathrm{b\bar{b}}$ purity	$D_{\mathrm{Z/H \to b\bar{b}}} \geq 0.8$
No extra b-jets	No AK4 b-tagged jets resolved from $\mathrm{b\bar{b}}$ jet
Exclude Z resonance	$6 \leq m_{\ell\ell} \leq 75$ GeV
Leptons in close proximity	$\Delta R_{\ell\ell} \leq 1.0$
Leptons near $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$	$\lvert\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\ell\ell})\rvert \leq \pi/2$
Significant missing energy	$p_{\mathrm{T}}^{\mathrm{miss}} \geq 85$ GeV

Extra selections are needed to mitigate the effects of HCAL modules that suffered a power outage during the 2018 data-taking run. Two sectors of the $\eta < 0$ endcap HCAL system, known as HEM15 and HEM16 were inactive from run 319077 through the end of 2018. The affected region of the detector spanned $\eta < -1.479$ and $-1.55 < \phi < -0.9$. Two consequences of this issue are considered for calibrating the set of events in the analysis.

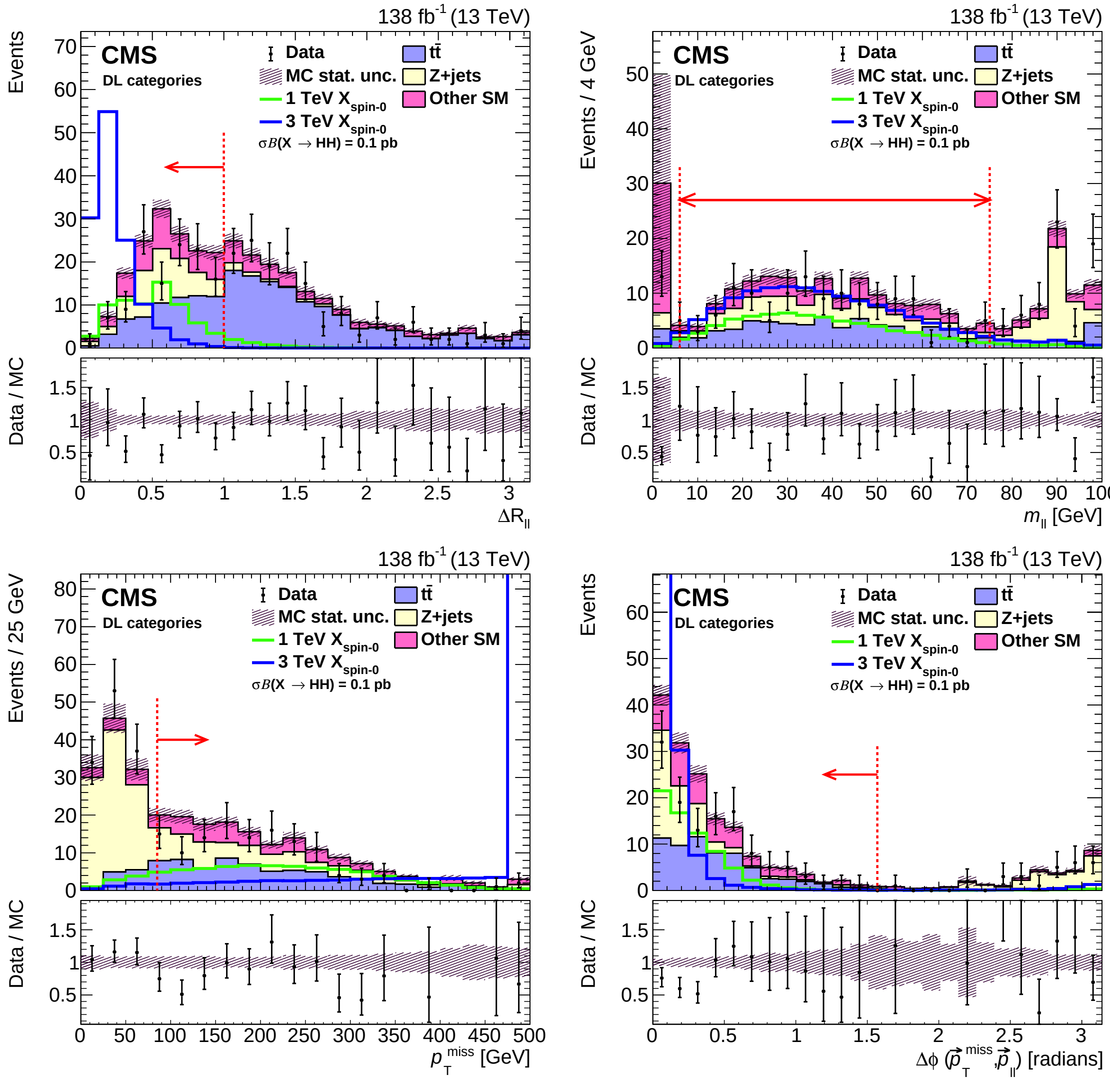

Figure 4.15: Distributions of discriminating variables in the 2ℓ channel: $\Delta R_{\ell\ell}$ (top left), $m_{\ell\ell}$ (top right), $p_{\mathrm{T}}^{\mathrm{miss}}$ (bottom left), and $\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\ell\ell})$ (bottom right). The data are shown as the black points with error bars, and two signal hypotheses, for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV, are shown as solid lines. The colored, stacked histograms show the important SM background processes along with the aggregate distribution of the subdominant SM backgrounds. The cross section of the signal process is arbitrarily set for ease of viewing. The red dashed lines and arrows indicate the selected regions of the analysis, corresponding to $\Delta R_{\ell\ell} \leq 1.0$, $6 \leq m_{\ell\ell} \leq 75$ GeV, $p_{\mathrm{T}}^{\mathrm{miss}} \geq 85$ GeV, and $\Delta\phi(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\ell\ell}) \leq \pi/2$.

First, jet energies are undermeasured in this region of the detector. Second, an excess of hadrons are misreconstructed as electrons, since the H/E (discussed in Section 3.3.4) for electron candidates in this region is 0.0, and thus this particular selection is fulfilled for every candidate. Since the affected region of the detector is beyond the acceptance for electrons in the 1ℓ channel, the effect is only important in the 2ℓ channel. In data, this effect is mitigated by excluding any 2ℓ event in the affected run range that contains an electron in the critical region of the detector. In simulation, for each electron in the critical region in a given event, a multiplicative event weight is applied that is equal to the ratio of the 2018 luminosity before run 319077 and the full 2018 luminosity. This produces a negligible change in both the signal and expected background normalizations. The effect of excluding these electrons from the 2ℓ channel data is demonstrated in Fig. 4.16. The undermeasurement of jet energies in the HEM region is accounted for by including an extra systematic uncertainty, discussed in Section 5.2.2.

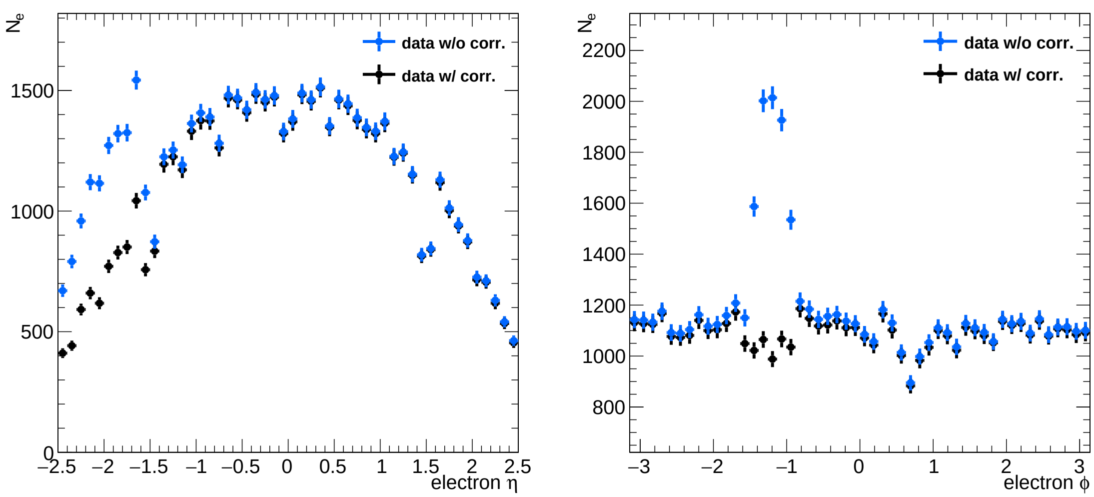

Figure 4.16: The η (left) and ϕ (right) distributions of electrons in the 2ℓ channel data in 2018. In each plot, we show the data before removing (blue) and after removing (black) the affected HEM region for runs ≥ 319077.

With the full selection described in this section applied to the analysis, the expected SM background composition is described in Table 4.6, and the expected yields for hypothetical signals assuming $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH}) = 1.0$ pb are provided in Table 4.7. The event categorization mentioned earlier constructs 12 search region (SR) categories using the lepton flavor content, $D_{\mathrm{Z/H\to b\bar{b}}}$, and additionally in the 1ℓ channel, τ_2/τ_1 and $D_{\ell\nu\mathrm{q\bar{q}}}$. The categorization is summarized in Tables 4.8 and 4.9.

Table 4.6: SM background composition in the SR. The total event yields, corresponding to an integrated luminosity of 137.6 fb^{-1}, are 9267.1 (154.2) events in the 1ℓ (2ℓ) channel. The table shows the fraction of these events that is attributed to each process.

Background process	1ℓ channel	2ℓ channel
$\mathrm{t\bar{t}}$ 0ℓ	1.4%	0.0%
$\mathrm{t\bar{t}}$ 1ℓ	63.1%	25.8%
$\mathrm{t\bar{t}}$ 2ℓ	4.4%	15.2%
single top	3.1%	13.4%
W + jets	15.1%	2.1%
Z/γ^* + jets	2.2%	32.1%
QCD multijet	9.1%	0.0%
diboson	0.9%	7.8%
$\mathrm{t\bar{t}V}$	0.6%	2.0%
$\mathrm{t\bar{t}H}$	0.2%	1.6%

For spin-0 scenarios, the selection efficiency for 1ℓ channel events to pass the criteria of any of the 12 SR categories is 9% at $m_{\mathrm{X}} = 0.8$ TeV. The efficiency increases with m_{X} up to 23% at $m_{\mathrm{X}} = 1.5$ TeV because the Higgs boson decays become more collimated. Above 1.5 TeV, the selection efficiency decreases to a minimum of 14% at $m_{\mathrm{X}} = 4.5$ TeV for two main reasons: the b tagging efficiency degrades for high-p_{T} jets as does the lepton isolation for extremely collimated Higgs boson decays. For 2ℓ channel events, the combined selection

Table 4.7: Expected signal yields assuming $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH}) = 1.0$ pb and an integrated luminosity of 137.6 fb^{-1}. Yields are provided for all mass points m_{X} for which simulated samples are used and for all events that enter any one of the search categories in the 1ℓ and 2ℓ channels.

m_{X}	1ℓ channel events		2ℓ channel events	
	spin-0	spin-2	spin-0	spin-2
800 GeV	1087.52	1581.64	323.376	472.929
900 GeV	1705.7	2417.75	536.054	739.579
1.0 TeV	2101.72	2890.48	686.561	932.013
1.2 TeV	2546.2	3310.15	877.561	1134.56
1.4 TeV	2730.41	3508.64	1014.11	1259.71
1.6 TeV	2776.75	3481.56	1099.44	1333.04
1.8 TeV	2764.54	3435.66	1141.48	1372.07
2.0 TeV	2743.79	3352.06	1182.67	1389.14
2.5 TeV	2506.06	3058.3	1257.88	1436.87
3.0 TeV	2288.65	2759.15	1282.15	1468.92
3.5 TeV	2005.43	2443.82	1289.92	1456.19
4.0 TeV	1810.1	2176.67	1287.41	1427.12
4.5 TeV	1613.96	1928.06	1257.25	1401.98

efficiency to pass the criteria of any category is 9% at $m_{\mathrm{X}} = 0.8$ TeV, increases sharply with m_{X} to 30% at $m_{\mathrm{X}} = 1.5$ TeV, and then increases more slowly to 36% at $m_{\mathrm{X}} = 4.5$ TeV. The efficiency grows over the full range of m_{X} because the leptons, with no nearby jet, become easier to select at high p_{T}.

Table 4.8: The 1ℓ channel event categorization and corresponding category labels. All combinations of the two lepton flavors, two $\mathrm{b\bar{b}}$ jet tagging, and two $\mathrm{H} \to \mathrm{WW}^*$ decay purity selections are used to form eight independent event categories. The lower τ_2/τ_1 working point is 0.55 (0.45) in 2016 (2017–2018).

Categorization type	Selection	Category label
Lepton flavor	Electron	e
	Muon	μ
$\mathrm{b\bar{b}}$ jet tagging	$0.8 < D_{\mathrm{Z/H \to b\bar{b}}} < 0.97$	bL
	$D_{\mathrm{Z/H \to b\bar{b}}} > 0.97$	bT
$\mathrm{H} \to \mathrm{WW}^*$ purity	$0.45(0.55) < \tau_2/\tau_1 < 0.75$ or $2.5 < D_{\ell\nu\mathrm{q\bar{q}}} < 11.0$	LP
	$\tau_2/\tau_1 < 0.45(0.55)$ and $D_{\ell\nu\mathrm{q\bar{q}}} < 2.5$	HP

Table 4.9: The 2ℓ channel event categorization and corresponding category labels. All combinations of the two lepton flavors and two $\mathrm{b\bar{b}}$ jet tagging selections are used to form four independent event categories.

Categorization type	Selection	Category label
Lepton flavor	Two electrons or two muons	SF
	One electron and one muon	OF
$\mathrm{b\bar{b}}$ jet tagging	$0.8 < D_{\mathrm{Z/H \to b\bar{b}}} < 0.97$	bL
	$D_{\mathrm{Z/H \to b\bar{b}}} > 0.97$	bT

CHAPTER 5

Modeling X $\rightarrow$ HH Signal and Background

5.1 Modeling

The search evaluates the existence of signal by performing a two-dimensional (2D) simultaneous maximum likelihood fit of the data in the 12 event categories listed in Tables 4.8 and 4.9. The data are binned in two dimensions, $m_{\mathrm{b\bar{b}}}$ and m_{HH}, within the ranges $30 < m_{\mathrm{b\bar{b}}} < 210$ GeV and $700 < m_{\mathrm{HH}} < 5050$ GeV. The expected signal is uniquely characterized by mass resonances in both dimensions: a resonance in the $m_{\mathrm{b\bar{b}}}$ dimension near the Higgs boson mass $m_{\mathrm{H}} = 125$ GeV and another resonance at an unknown mass *a priori* in the m_{HH} dimension. The lower bound of $m_{\mathrm{b\bar{b}}}$ is chosen because the SD modeling is not well understood below this value and has larger uncertainty. However, we include such low $m_{\mathrm{b\bar{b}}}$ in order to use this sideband to constrain the background in the modeling. We extend $m_{\mathrm{b\bar{b}}}$ up to 210 GeV so that the parameter space contains the peak of the top quark mass. The lower bound of m_{HH} is chosen such that the total background spectrum in this dimension is monotonically falling, and the upper bound is chosen to be a few hundred GeV above the largest mass data point observed in any event that fulfills the physics object selection in Section 4.3.

In the $m_{\mathrm{b\bar{b}}}$ dimensions, the bin width used in the fit is a constant 6 GeV, but in the m_{HH} dimensions the bin widths for fitting are variable: 25 GeV width within $700 < m_{\mathrm{HH}} < 2000$ GeV, 50 GeV within $2000 < m_{\mathrm{HH}} < 3100$ GeV, and 75 GeV in the region $m_{\mathrm{HH}} > 3100$ GeV. In both dimensions, the bin widths are chosen to be smaller than the mass resolutions of the signal; a typical signal contains about 8 bins under the peak, particularly 4 along the $m_{\mathrm{b\bar{b}}}$ dimension and 2 along m_{HH}. Furthermore, the variable bin width significantly improved computing time without sacrificing sensitivity to any signal.

The signal and background mass distributions are modeled with independent 2D templates for each event category and are constructed to be smoother than the distributions directly from MC simulation only. Methods are used to increase the statistical power of the templates relative to that of the simulation in the category of interest. Shape and normalization uncertainties are included in the fit in order to account for differences in the spectra of data and simulation.

5.1.1 Background component classification

To ultimately perform the fit to data, the background model is divided into components, and a 2D template is generated in the ($m_{\mathrm{b\bar{b}}}$, m_{HH}) mass plane for each independently. The shape and normalization of each component is then varied in the fit in each event category.

In both the 1ℓ and 2ℓ channels, the background is dominated by contributions from $\mathrm{t\bar{t}}$ production. Subdominant contributions come from $\mathrm{W}+\text{jets}$ with $\mathrm{W}\to\ell\nu$ and QCD multijet processes in the 1ℓ channel, while contributions from $\mathrm{Z}/\gamma^*+\text{jets}$ with $\mathrm{Z}/\gamma^*\to\ell\ell$ are almost as significant as $\mathrm{t\bar{t}}$ in the 2ℓ channel. Instead of splitting by SM process, the background is classified by the set of generator-level quarks from the hadronic decay of a top quark or W boson that fall within $\Delta R < 0.8$ of the $\mathrm{b\bar{b}}$ jet axis. There are several advantages of this approach to the classification. First, we are able to carefully model background components with $m_{\mathrm{b\bar{b}}}$ resonances apart from those without. Second, we are able to more reliably produce high quality templates with increased statistical power (discussed in Section 5.1.2). This classification is only important for dividing the simulation samples before building the background templates. The background components are described below and summarized in Table 5.1:

1. $\boldsymbol{m_{\mathrm{t}}}$ **background:** All events in which the b quark and both W quarks from the

top quark decay fall within the $b\bar{b}$ jet cone. The $m_{b\bar{b}}$ distribution of these events is characterized by a peak around the top quark mass. This background is primarily composed of $t\bar{t}$.

2. **m_W background:** All events not within the m_t category but that have both quarks from a hadronic W decay inside the $b\bar{b}$ jet cone. The $m_{b\bar{b}}$ distribution of these events is characterized by a peak around the W boson mass. This background is also primarily composed of $t\bar{t}$.

3. **lost t/W background:** All events not within the preceding two categories but that have at least one quark from a hadronically decaying top quark or W boson inside the $b\bar{b}$ jet cone. The $m_{b\bar{b}}$ spectrum of these events does not contain any resonant peaks, as the full quark content of the top quark or W boson is not encapsulated by the $b\bar{b}$ jet. The dominant SM process in this background is $t\bar{t}$.

4. **q/g background:** All remaining events that do not fall within the previous three categories and so have no generator-level quarks from a hadronically decaying top quark or W boson within the $b\bar{b}$ jet. In the 1ℓ channel, W+jets ($W \to \ell\nu$) and QCD multi-jet processes are the main sources of this background, while in the 2ℓ channel, Z/γ^*+jets ($Z/\gamma^* \to \ell\ell$) is the primary contributor with some contribution from dilepton $t\bar{t}$.

The background composition for each of the two channels is described in Table 5.2 and is shown in simulation for the two search variables in Fig. 5.1. The lost t/W background is the largest contributor in the 1ℓ channel, while the q/g background accounts for the majority of the total 2ℓ channel background. The m_W background is the smallest component in both channels.

Table 5.1: The four background components with their kinematical properties and defining number of generator-level quarks within $\Delta R < 0.8$ of the $\mathrm{b\bar{b}}$ jet axis.

Bkg. category	Dominant SM processes	Resonant in $m_{\mathrm{b\bar{b}}}$	Number of quarks
m_{t}	$\mathrm{t\bar{t}}$	top quark mass	3 from top quark
m_{W}	$\mathrm{t\bar{t}}$	W boson mass	2 from W boson
lost t/W	$\mathrm{t\bar{t}}$	No	1 or 2
q/g	V + jets and QCD multijet	No	0

Table 5.2: The composition of the SM background in the SR broken down into the background components described in Table 5.1. The table shows the fraction of all 1ℓ and 2ℓ events (separately) that is attributed to each process.

Background process	1ℓ channel	2ℓ channel
q/g	28.4%	51.9%
lost t/W	43.8%	30.7%
m_{W}	4.3%	5.0%
m_{t}	23.5%	12.4%

5.1.2 Template construction strategy

For each of the four background components, a unique template in the ($m_{\mathrm{b\bar{b}}}$, m_{HH}) mass plane is produced for each of the 12 SR categories (yielding 48 independent background templates). For each background component, a small set of inclusive templates that have more statistical power than the set of events in the individual SR categories is produced. These inclusive templates are made by combining events in multiple SR categories and by relaxing certain selections. In this procedure, care is taken to ensure that the inclusive shape remains consistent with the shape for the full selection in the categories of interest.

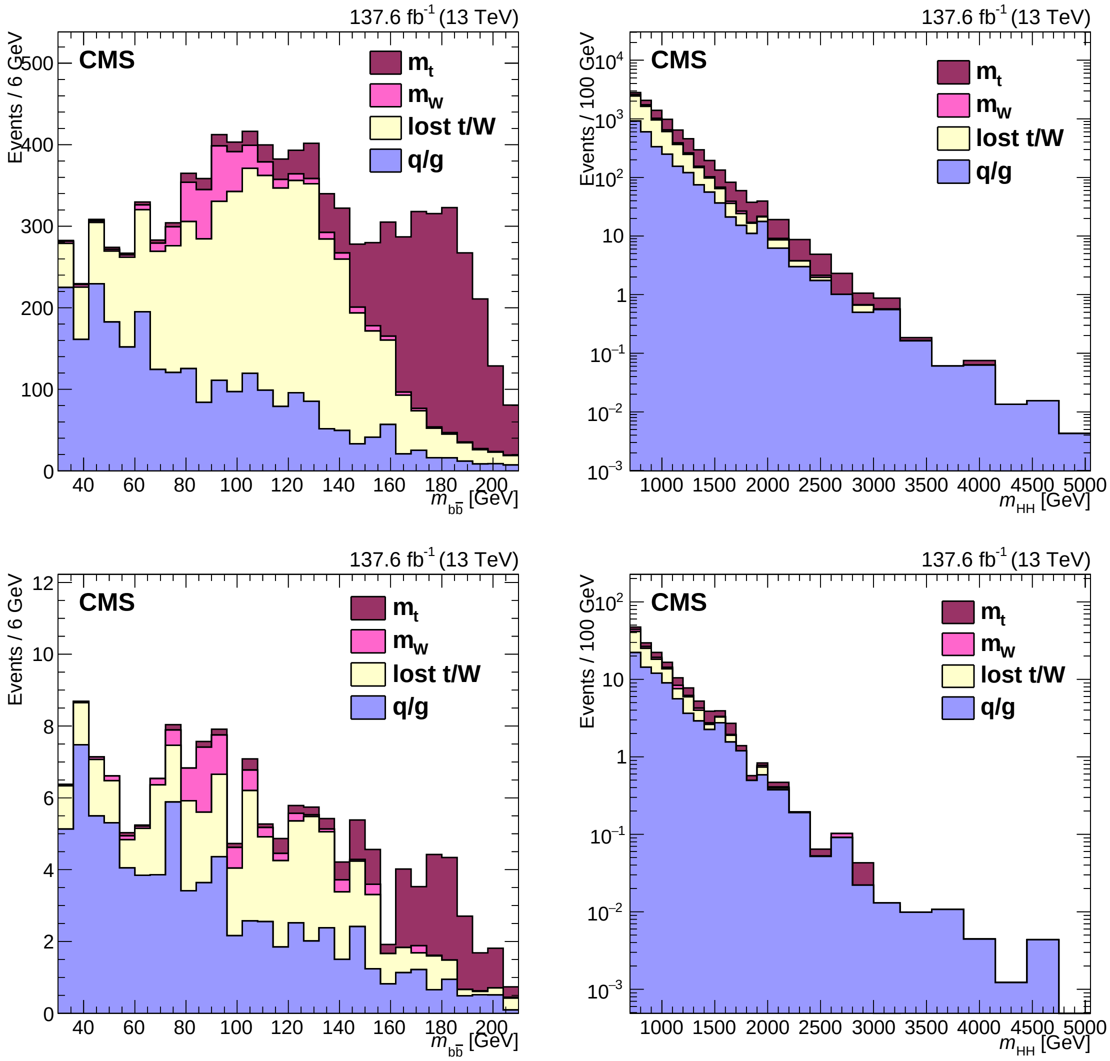

Figure 5.1: Simulated SR mass variables for the 1ℓ (top) and 2ℓ (bottom) channels showing the background category composition for each. The simulation is normalized to the integrated luminosity of the data.

Then, for each of the 12 SR categories, the inclusive templates are fit to the simulated mass distributions to produce a unique template. This fit is performed in a similar manner and with a similar parameterization of the template shape as is done for the fit to data, described in more detail later in Section 5.1.4. The background templates and associated systematic uncertainties are ultimately validated by fitting to the data in dedicated control regions (CRs), a procedure that will be described in Section 6.1.

In two cases we derive corrections to certain processes in simulation that ultimately allow for easier modeling. While these corrections should in principle have no impact on the analysis result because mismodeling of the background normalization and p_{T} spectrum is accounted for in the fit, it is helpful to more accurately predict the pre-fit background spectrum in order to optimize the selection strategy. These two calibrations are now reviewed below.

5.1.2.1 $\mathrm{t\bar{t}}$ p_{T} mismodeling

It is known within CMS that simulation overpredicts the data in high-p_{T} $\mathrm{t\bar{t}}$ event [126, 127]. The effects of this mismodeling are studied in a signal-depleted, $\mathrm{t\bar{t}}$-enriched region that is similar to the SR of the analysis. An inclusive correction is measured for all $\mathrm{t\bar{t}}$ simulation as a function of resonance mass m_{HH}, as this quantity is correlated to the $\mathrm{t\bar{t}}$ p_{T}.

To derive this correction, we consider both 1ℓ and 2ℓ events that satisfy the trigger selection and have all objects from the HH decay chain reconstructed, as described in Section 4.3. Selections involving the b tagging are imposed to obtain a $\mathrm{t\bar{t}}$-enriched region: $D_{\mathrm{Z/H\to b\bar{b}}} > 0.8$ (as in the SR) and the requirement of an AK4 b jet that is $\Delta R > 1.2$ from the $\mathrm{b\bar{b}}$ jet (disjoint from the SR). The $\mathrm{t\bar{t}}$ purity is further enhanced by requiring $m_{\mathrm{b\bar{b}}} > 80$ GeV, reducing the contributions of non-$\mathrm{t\bar{t}}$ backgrounds.

The background and data composition in the measurement region, as well as the ratio distribution from which the scale factor is extracted, are shown in Fig. 5.2. The correction is derived by fitting the difference between the number of data events and non-$\mathrm{t\bar{t}}$ MC events, divided by the number of $\mathrm{t\bar{t}}$ MC events: $(N_{\mathrm{data}} - N_{\mathrm{non\text{-}t\bar{t}}})/N_{\mathrm{t\bar{t}}}$. We fit the combined 2016–2018 data and simulation with a polynomial to extract a smooth correction as a function of m_{HH}.

In Ref. [72], a flat correction was used to scale down the $\mathrm{t\bar{t}}$ normalization. In this

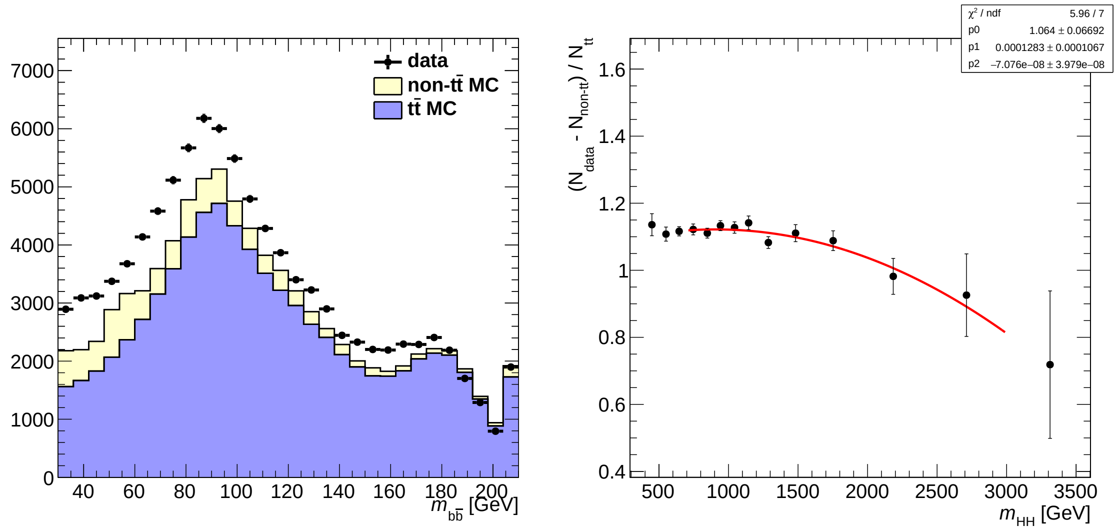

Figure 5.2: Left: Simulation and data distributions of $m_{b\bar{b}}$ in the control region used to calculate the inclusive $t\bar{t}$ scale factor. We perform the measurement for $m_{b\bar{b}} \geq 80$ GeV to enhance the purity of the sample. Right: the scale factor measurement as a function of m_{HH}. The red line is the extracted correction using a second-order polynomial.

dissertation, scale factors that are used to correct the simulated $b\bar{b}$ jet tagging efficiencies from DeepAK8 (Section 3.3.5) also account for the $t\bar{t}$ p_T mismodeling and help to match the simulation and data. Still, measuring our own corrections is useful to more precisely match the simulation to the data.

5.1.2.2 Smoothing the QCD multijet simulation

In the 1ℓ channel, extra effort is undertaken to smooth the QCD multijet simulation before including it in the template-building procedure. This simulated process has large statistical fluctuations because large enough samples cannot be produced with the needed effective integrated luminosity (note the interaction cross section for this process in Table 4.1). However, the $b\bar{b}$ jet reconstruction in the QCD multijet simulation is similar to that in W + jets, and the W + jets simulation has much more statistical power. Both processes contribute significantly to the q/g background, with AK8 jets that are misidentified as b jets. As a result, both processes exhibit very similar falling shapes in the $m_{b\bar{b}}$ spectrum and similar $b\bar{b}$ jet tagging distributions of $D_{Z/H\to b\bar{b}}$. The similarity in the shapes of these distributions is seen in Fig. 5.3.

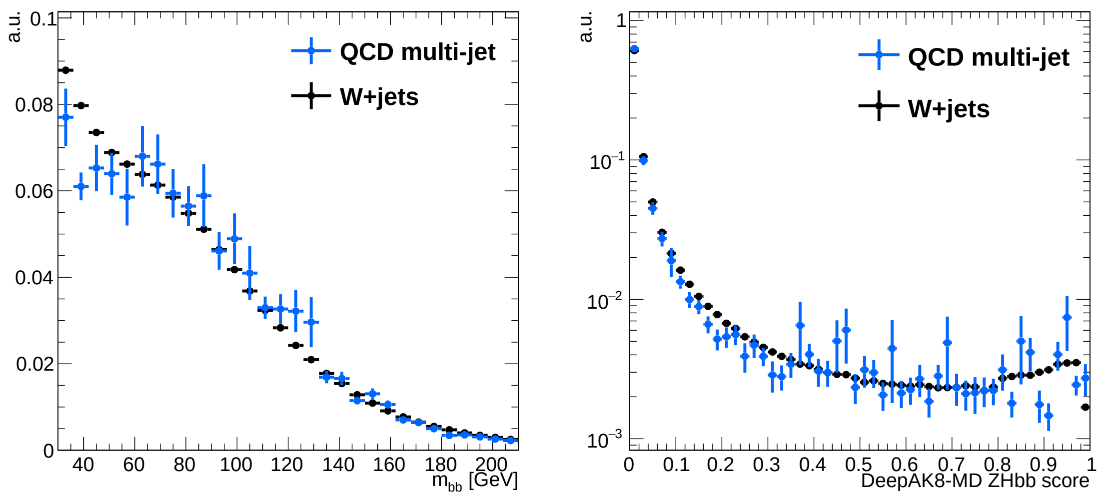

Figure 5.3: Distributions of $m_{b\bar{b}}$ (left) and $D_{Z/H\to b\bar{b}}$ (right) for QCD multijet and $W+\text{jets}$ simulation. Considering the statistical uncertainty of the QCD multijet simulation, the shapes between both processes are consistent in both plots.

Instead of using the QCD simulation directly in the q/g background modeling, a combined distribution is created by measuring the ratio of QCD to $W+\text{jets}$ event yields as a function of m_{HH} and then using these corrections to scale up the $W+\text{jets}$ simulation. Corrections and distributions are obtained for each lepton flavor and $H\to WW^*$ purity category, since the $b\bar{b}$ jet tagging between $W+\text{jets}$ and QCD is equivalent. The distributions are fit with a second-order polynomial, allowing us to obtain smooth corrections that are parameterized in the same way as the m_{HH} shape systematics are (discussed later in Section 5.1.4). This distribution is used as an input to the q/g background modeling to account for both processes. Thus, the corrections account for residual differences in the relative multijet fraction. Example measurements from this procedure are shown in Fig. 5.4. The correction is only needed for the 1ℓ channel simulation because the QCD multijet contribution to the 2ℓ channel background is negligible.

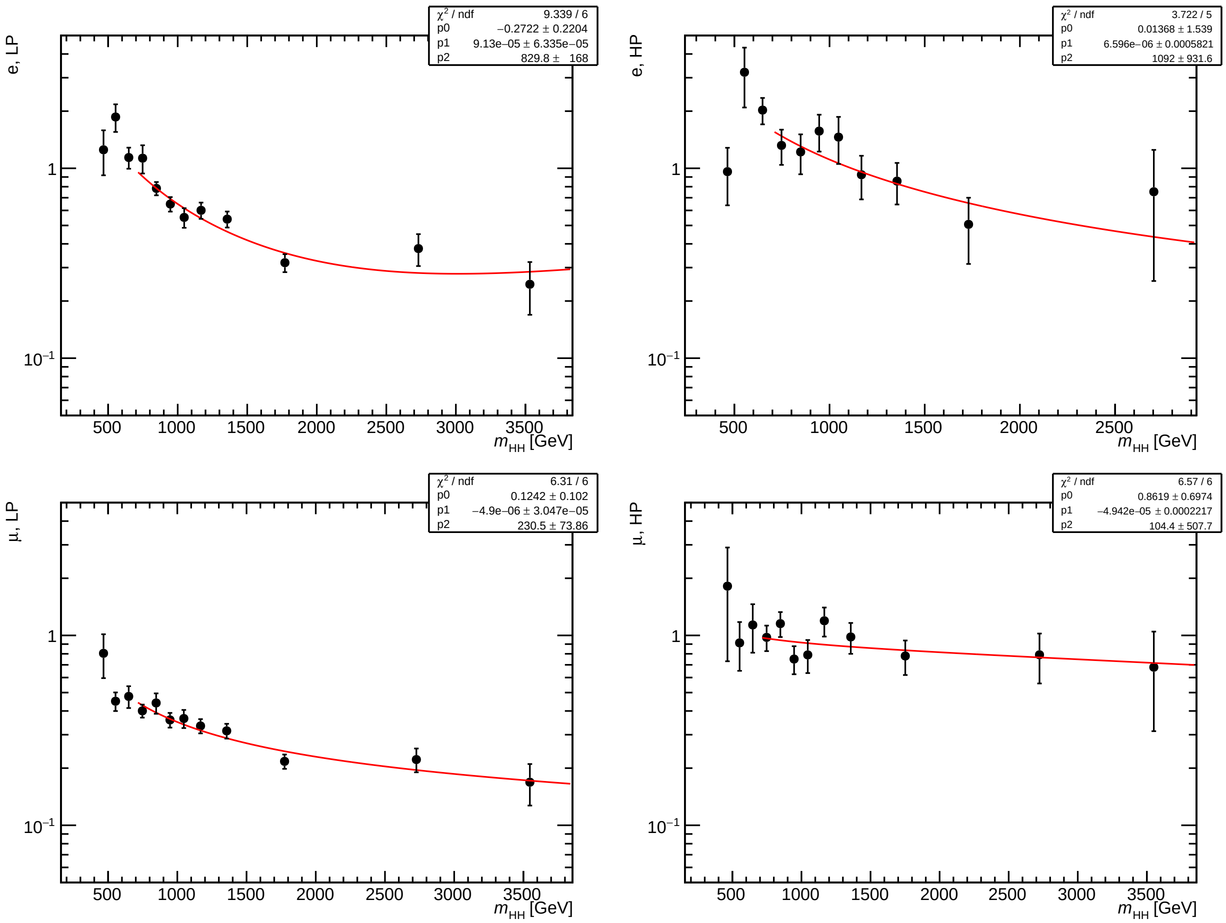

Figure 5.4: Fitted scale factors used to perform the QCD simulation smoothing. Four measurements are performed, with no selection on $D_{\mathrm{Z/H}\to\mathrm{b\bar{b}}}$, for the selections e LP, μ LP, e HP, and μ HP. The y-axis of each plot is the ratio of QCD to W+jets simulation, binned in m_{HH} along the x-axis.

5.1.3 Models

Due to the background component classification in Section 5.1.1 and the nature of the signal, both resonant and nonresonant mass shapes need to be modeled. This is accomplished with two different methods.

Nonresonant mass shapes are modeled using kernel density estimation (KDE) [128–130]. KDE is a non-parametric density estimation technique that provides smoothing to an input data set and is useful for modeling distributions that are difficult to be described by analytical functions, as is the case for the nonresonant shapes seen in the background m_{HH} and $m_{\mathrm{b\bar{b}}}$ spectra (lost t/W and q/g). For weighted data (the case for our simulated events), the univariate estimate $\hat{f}(x)$ of a distribution $f(x)$ is:

$$\hat{f}(x) = \frac{1}{\tilde{n}h} \sum_{i=1}^{n} w_i K\left(\frac{x - x_i}{h}\right). \tag{5.1}$$

This sum runs over n events of a weighted data set with values x_i of the variable x that have weights w_i; the sum of all weighted events is $\tilde{n} = \sum_{i=1}^{n} w_i$. The kernel $K(t)$ is a function used to perform the smearing of the input data, and the width of this function is controlled by the parameter h_i, known as the bandwidth. A multivariate estimate $\hat{f}(\vec{x})$ can be constructed in a d-dimensional space as:

$$\hat{f}(\vec{x}) = \frac{1}{\tilde{n}} \sum_{i=1}^{n} w_i \prod_{j=1}^{d} \frac{1}{h_j} K\left(\frac{x_j - x_{ij}}{h_j}\right) \tag{5.2}$$

with different bandwidths (h_j) in each dimension j.

In our background modeling, a Gaussian kernel is used because of its smooth and symmetric properties. The choice of kernel is not nearly as important as the choice of bandwidth, however, as this dictates the degree of smoothing that will be applied, with larger bandwidths yielding more smoothing. An optimal bandwidth h_{opt} can be obtained in general by minimizing the mean integrated square error of $\hat{f}(x)$, which has a dependence in d dimensions of

$$h_{\text{opt}} \sim n_{\text{eff}}^{-1/(d+4)}, \tag{5.3}$$

where n_{eff} is given by:

$$n_{\text{eff}} = \frac{(\sum_{i=1}^{n} w_i)^2}{\sum_{i=1}^{n} w_i^2}. \tag{5.4}$$

In the case of 1D data that is not normally distributed, Silverman [129] recommends that the optimal bandwidth be:

$$h_{\text{opt}} = 0.9 \min{(\sigma, \text{IQR}/1.34)} n_{\text{eff}}^{-1/5}, \tag{5.5}$$

where σ is the standard deviation (we calculate it with the sample standard deviation), and IQR (interquartile range) is the distance between the 75% and 25% quartiles of the input distribution. In practice, we also multiply this bandwidth by a constant scale factor to allow us to tune and optimize the density estimate for each input distribution. For multidimensional distributions, the factor in front of $n_{\text{eff}}^{-1/(d+4)}$ is difficult to obtain, so a reasonable initial value, motivated by the 1D situation, is used for each dimension j:

$$h_{\text{opt,j}} = \sigma_j n_{\text{eff}}^{-1/(d+4)}. \tag{5.6}$$

As in the 1D case, in practice an independent constant scale factor is applied to each $h_{\text{opt,j}}$.

The discussion so far has featured a global bandwidth h that is applied everywhere along one dimension in the estimate. This is not ideal to use when the input distribution spans many orders of magnitude in event density. A better strategy is to provide less smoothing (smaller h) in regions of high event density, where we care about the fine structure of the local distribution, and more smoothing in regions of low event density, where the statistical uncertainty is large. This is accomplished by using an adaptive bandwidth that is a function of the local event density of the input distribution. In the 1D case, the density estimate is now:

$$\hat{f}(x) = \frac{1}{\tilde{n}} \sum_{i=1}^{n} \frac{w_i}{h_i} K\left(\frac{x - x_i}{h_i}\right), \tag{5.7}$$

where h_i is given by:

$$h_i = h \left(\frac{g}{\tilde{f}(x_i)}\right)^{1/2}. \tag{5.8}$$

Resonant mass shapes are seen in both dimensions for signal and in the $m_{\mathrm{b\bar{b}}}$ dimension for the m_{W} and m_{t} backgrounds. In the resonant case, the distributions can be described analytically using a common parametric function in HEP — the double-sided Crystal Ball (DCSB) function [131]. This parametric function, which is continuous and differentiable, contains a Gaussian core that describes a mass peak with power law tails on both sides of the core that account for the effects of energy loss. In general, the DSCB has 6 parameters and is defined as:

$$\mathrm{DSCB}(m; \mu, \sigma, \alpha_1, \alpha_2, n_1, n_2) = \begin{cases} Ne^{-\frac{\beta(m)^2}{2}}, & -\alpha_1 \leq \beta(m) \leq -\alpha_2 \\ Ne^{-\frac{\alpha_1^2}{2}} \left(\frac{\alpha_1}{n_1}\left(\frac{n_1}{\alpha_1} - \alpha_1 - \beta(m)\right)\right)^{-n_1}, & \beta(m) < -\alpha_1 \\ Ne^{-\frac{\alpha_2^2}{2}} \left(\frac{\alpha_2}{n_2}\left(\frac{n_2}{\alpha_2} - \alpha_2 - \beta(m)\right)\right)^{-n_2}, & \beta(m) > \alpha_2, \end{cases} \tag{5.9}$$

where N is a normalization constant, $\beta(m) = (m - \mu)/\sigma$ is the Gaussian argument term that contains the mean μ and the width σ of the Gaussian core, and the α_1 (α_2) and n_1 (n_2) parameters control the low (high) mass tails of the function. Mass distributions corresponding to the top quark, W boson, and HH resonance are fit with the DSCB, and the fitted μ and σ parameters are interpreted as the measured mass scale and resolution, respectively.

5.1.4 Background modeling

The background templates are modeled using conditional probabilities of $m_{\mathrm{b\bar{b}}}$ as a function of m_{HH} so that the templates include the correlation of these two variables. The full 2D template is modeled as:

$$P_{\mathrm{bkg}}(m_{\mathrm{b\bar{b}}}, m_{\mathrm{HH}}) = P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{HH}}, \theta_1)P_{\mathrm{HH}}(m_{\mathrm{HH}}|\theta_2), \tag{5.10}$$

where $P_{\mathrm{b\bar{b}}}$ is a 2D conditional probability distribution, P_{HH} is a 1D probability distribution, and θ_1 and θ_2 are sets of nuisance parameters used to account for background shape uncertainties. The sets θ_1 and θ_2 do not share any common nuisance parameters.

For each background component, the P_{HH} templates are produced by applying KDE to 1D m_{HH} histograms of the falling spectra and fitting the high mass tails with exponentials. The $P_{\mathrm{b\bar{b}}}$ conditional templates are implemented as 2D histograms and are constructed differently depending on whether or not the background component is resonant in the $m_{\mathrm{b\bar{b}}}$ dimension. For the nonresonant backgrounds (q/g and lost t/W), the $P_{\mathrm{b\bar{b}}}$ are produced by applying KDE to the 2D histograms of the ($m_{\mathrm{b\bar{b}}}$, m_{HH}) mass distributions, with the exponential component in the high mass m_{HH} tails. For the resonant backgrounds (m_{W} and m_{t}), the $m_{\mathrm{b\bar{b}}}$ spectra are first fit with a DSCB function in bins of m_{HH}. Then, the parameters of the DSCB are each interpolated as a function of m_{HH} in order to obtain a smooth model for $m_{\mathrm{b\bar{b}}}$ dependent on m_{HH}.

In the case of both $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{HH}})$ and $P_{\mathrm{HH}}(m_{\mathrm{HH}})$, the templates are initially constructed from a superset of the events in any given SR category. The supersets are chosen by verifying that the shape of their mass distributions is consistent with the shape of the subset SR category distributions. This method allows us to obtain initial templates (referred to as inclusive templates) with higher statistical power than what is available in the SR categories, provided that the shapes remain reasonably consistent. For example, the $P_{\mathrm{HH}}(m_{\mathrm{HH}})$ exhibits no dependence on the lepton categorization in either the 1ℓ or 2ℓ channels, and so the inclusive templates are built from an event superset that contains both lepton flavor categories. Examples of the inclusive templates are shown in Figs. 5.5 and 5.6 for 1D $P_{\mathrm{HH}}(m_{\mathrm{HH}})$ templates. Projections of the nonresonant 2D conditional templates into $m_{\mathrm{b\bar{b}}}$ and m_{HH} are shown in Fig. 5.7. Examples of an $m_{\mathrm{b\bar{b}}}$ fit used to make the resonant $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|mhh)$ are shown in Fig. 5.8. An example of the interpolation of the DSCB parameters to obtain a smooth $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|mhh)$ is shown in Fig. 5.9. Once in hand, the inclusive templates are then fit to the simulated mass distribution in each SR category to obtain final distinct templates in each category. Examples of projections of the final 2D templates into the $m_{\mathrm{b\bar{b}}}$ dimension are shown in Figs. 5.10, 5.11, 5.12, and 5.13.

This methodology is also particularly powerful for modeling the 2ℓ search categories, which have extremely low statistics available, particularly for the resonant background components. The same four background components are used to describe both the 1ℓ and 2ℓ mass distributions, and the 1ℓ distributions are verified to be consistent with the 2ℓ distributions within the large statistical uncertainties. The inclusive $P_{\mathrm{b\bar{b}}}$ conditional templates are constructed from a set of events that includes both 1ℓ and 2ℓ events, and so common inclusive $P_{\mathrm{b\bar{b}}}$ templates are used for both the 1ℓ and 2ℓ categories. The inclusive P_{HH} templates are sufficiently constructed for the 1ℓ and 2ℓ channels independently.

As described in Section 5.1.3, the KDE modeling is tuned using two parameters: h_{SF} is the bandwidth scale factor that modulates the width of the Gaussian kernel, and h_{co} corresponds to the upper bandwidth cutoff and is a multiplicative factor to the median bandwidth. The h_{SF} and h_{co} are tuned independently for each of the $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{HH}})$ and $P_{\mathrm{HH}}(m_{\mathrm{HH}})$ templates of each background component.

5.1.5 Signal modeling

The signal templates are also modeled using 2D conditional probabilities and 1D probabilities, except with the ordering of $m_{\mathrm{b\bar{b}}}$ and m_{HH} flipped:

$$P_{\mathrm{signal}}(m_{\mathrm{b\bar{b}}}, m_{\mathrm{HH}}|m_{\mathrm{X}}) = P_{\mathrm{HH}}(m_{\mathrm{HH}}|m_{\mathrm{b\bar{b}}}, m_{\mathrm{X}}, \theta_1')P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{X}}, \theta_2'). \tag{5.11}$$

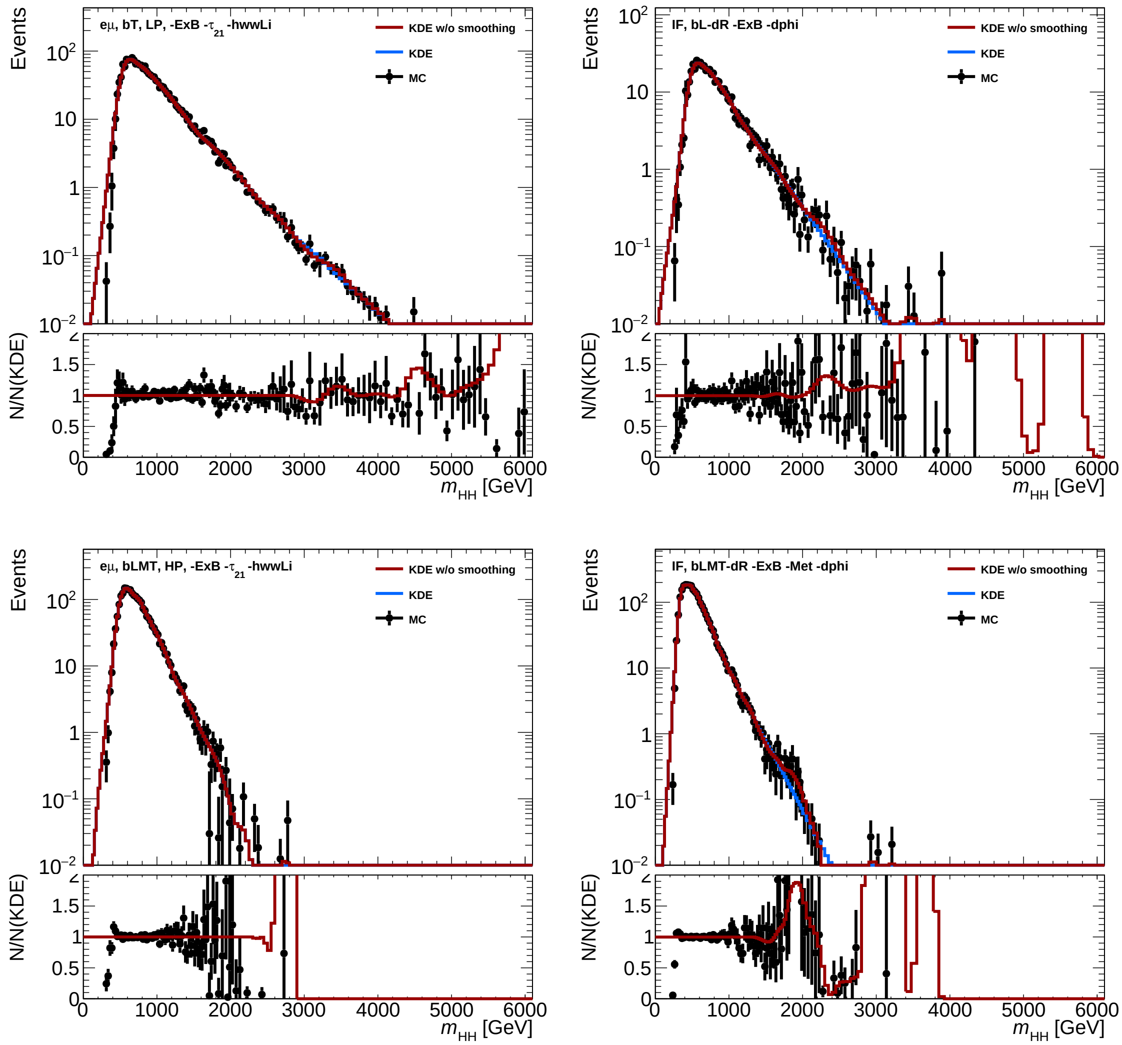

Figure 5.5: Some of the initial inclusive 1D templates $P_{\mathrm{HH}}(m_{\mathrm{HH}})$ shown with and without exponential tail smoothing compared to the input MC simulation from which they are constructed. Examples from the 1ℓ (2ℓ) channel are shown in the left (right) column. The q/g (lost t/W) backgrounds are shown on the top (bottom) row. These inclusive templates are then used to fit the simulation of each SR category to obtain the final $P(m_{\mathrm{HH}})$ that are inputs to building the full 2D model.

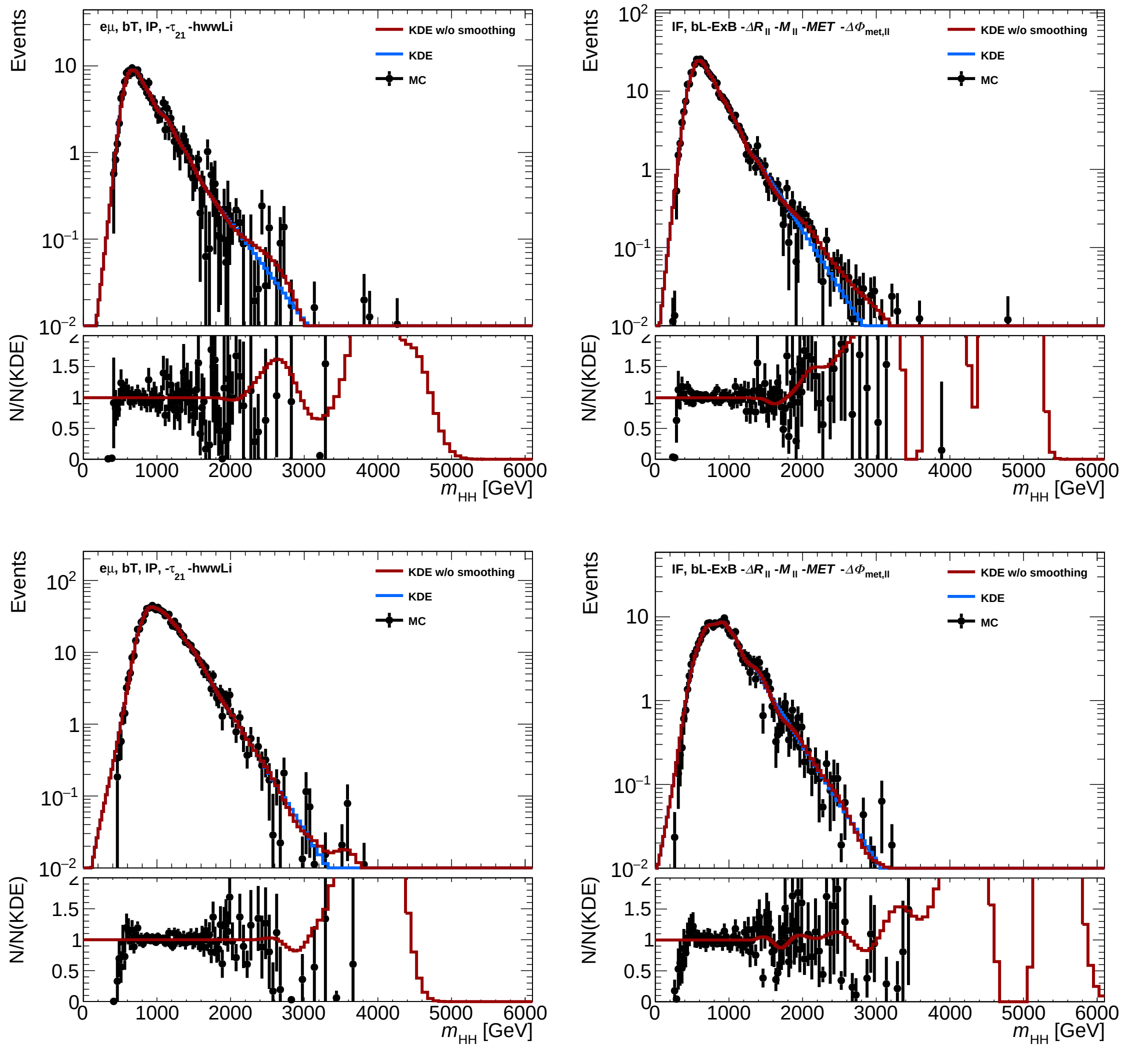

Figure 5.6: Some of the initial inclusive 1D templates $P_{\mathrm{HH}}(m_{\mathrm{HH}})$ shown with and without exponential tail smoothing compared to the input MC simulation from which they are constructed. Examples from the 1ℓ (2ℓ) channel are shown in the left (right) column. The m_{W} (m_{t}) backgrounds are shown on the top (bottom) row. These inclusive templates are then used to fit the simulation of each SR category to obtain the final $P(m_{\mathrm{HH}})$ that are inputs to building the full 2D model.

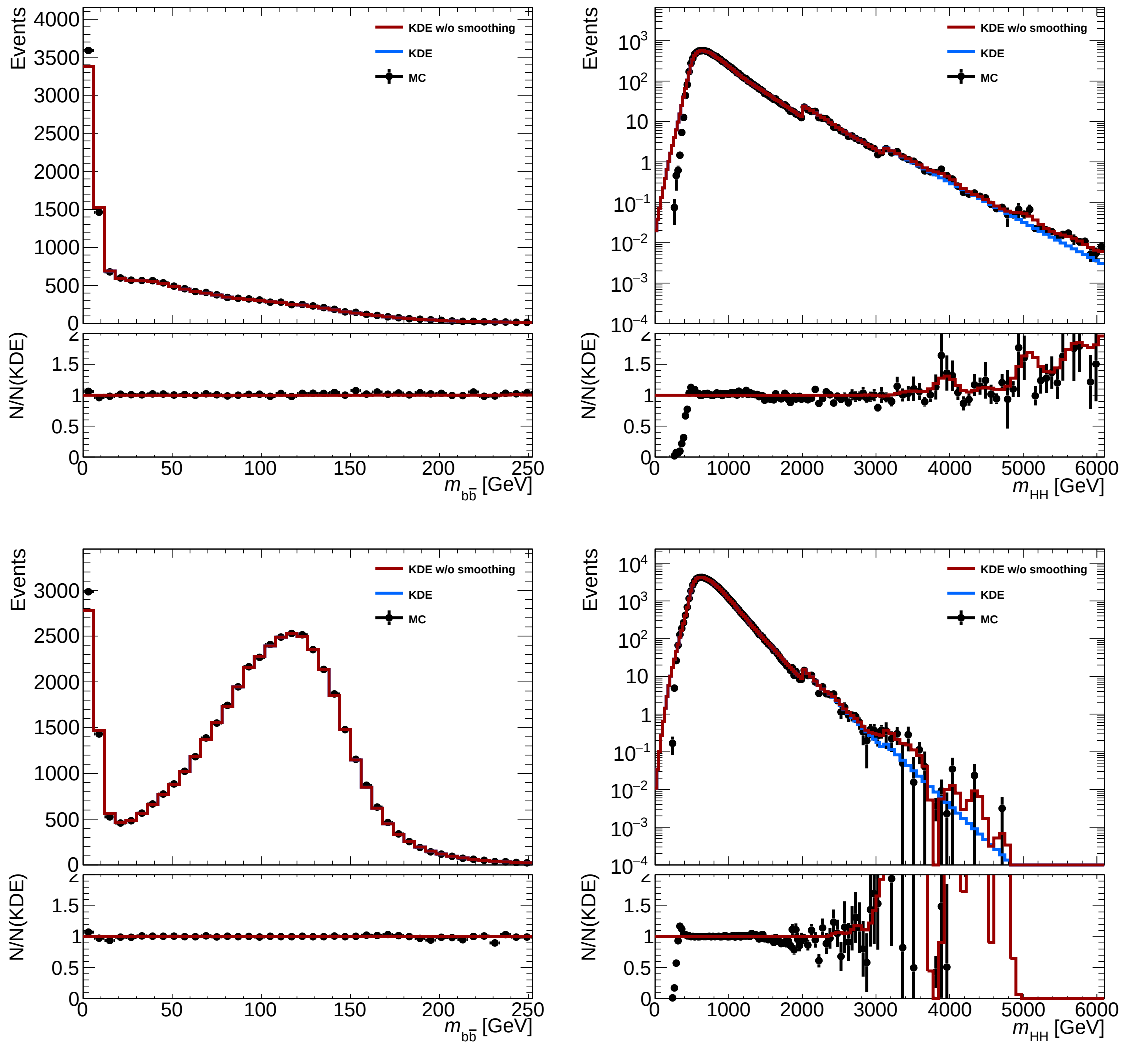

Figure 5.7: Projections of the inclusive 2D conditional template $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{HH}})$ onto $m_{\mathrm{b\bar{b}}}$ (left column) and m_{HH} (right column) for the q/g (top row) and lost t/W (bottom row) backgrounds. These inclusive 2D templates are combined with the corresponding 1D inclusive P_{HH} template and then fit to the simulation in each SR category to obtain the full 2D model for each.

The sets θ_1' and θ_2' do not share any common nuisance parameters. However, θ_2' from Eq. (5.11) and θ_1 from Eq. (5.10) do share two nuisance parameters corresponding to the scale

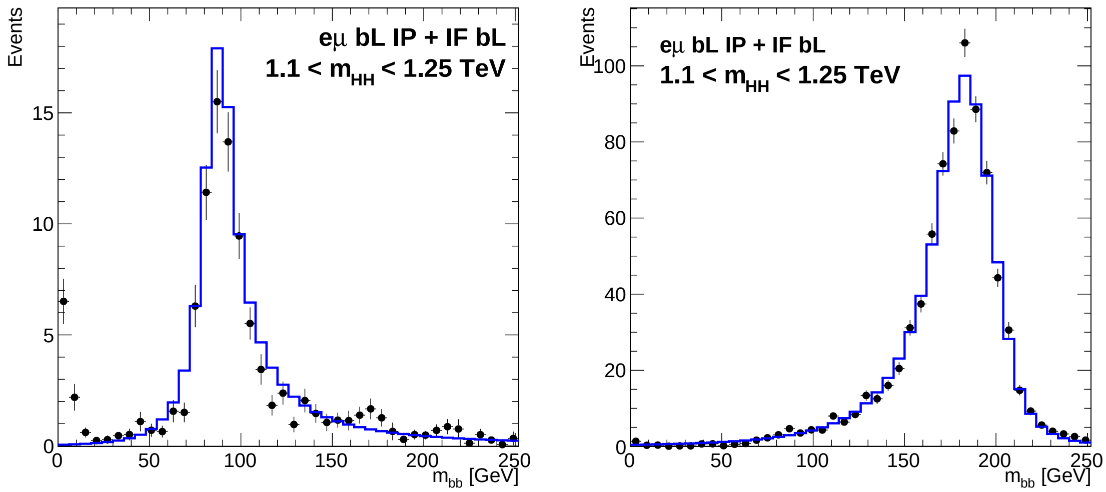

Figure 5.8: DSCB fits of the $m_{\mathrm{b\bar{b}}}$ spectrum in the bin $1.1 < m_{\mathrm{HH}} < 1.25$ TeV for the m_{W} (left) and m_{t} (right) backgrounds.

and resolution uncertainties of SD jet resonances in the $m_{\mathrm{b\bar{b}}}$ dimension. This is discussed in more detail in Section 5.2.

The P_{signal} distributions are first obtained for discrete m_{X} values by fitting histograms of the signal mass distributions. The mass shapes for spin-0 and spin-2 signals are nearly identical, and so the modeling is performed on the combined set of events and applied to both spin hypotheses. Inclusive templates are constructed independently for the 1ℓ and 2ℓ channels, by appropriately combining categories and relaxing selections (as in Section 5.1.4), and are subsequently fit to the simulation distributions in each SR category to obtain the final templates.

Models continuous in m_{X} are then produced by interpolating the fit parameters. The 1D $P_{\mathrm{b\bar{b}}}$ templates are created by fitting the $m_{\mathrm{b\bar{b}}}$ spectra with a DSCB function; the mass resolution in $m_{\mathrm{b\bar{b}}}$ is slightly larger than 10%, largest at low mass. The modeling of events in the bL category also contains an additive exponential component to model the small fraction of signal events with no resonant peak in the distribution.

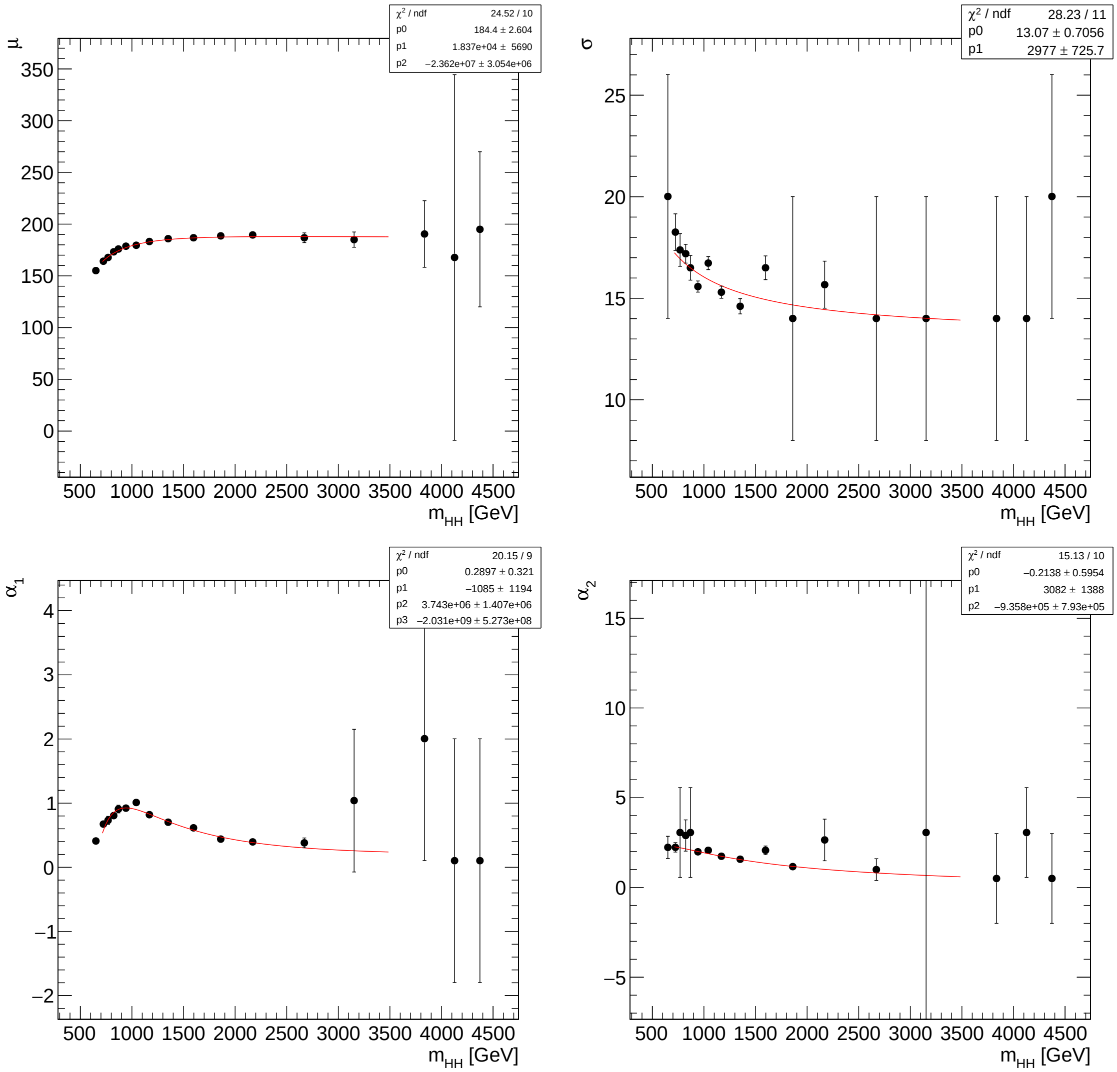

Figure 5.9: Interpolation of the DSCB parameters from the fits of the $m_{\mathrm{b\bar{b}}}$ spectra in bins of m_{HH} for the m_{t} background. The y-axis of each plot contains the name of the parameter. A smooth inclusive parameterization as a function of m_{HH} is obtained that is the inclusive $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{HH}})$ template. The inclusive selection here combines all events in the 1ℓ and 2ℓ channels in the bL $\mathrm{b\bar{b}}$ jet tagging category. This inclusive template is then refit to the $m_{\mathrm{b\bar{b}}}$ spectrum in the relevant bin of m_{HH} and relevant SR categories to obtain the final $P_{\mathrm{b\bar{b}}}(m_{\mathrm{b\bar{b}}}|m_{\mathrm{HH}})$ template.

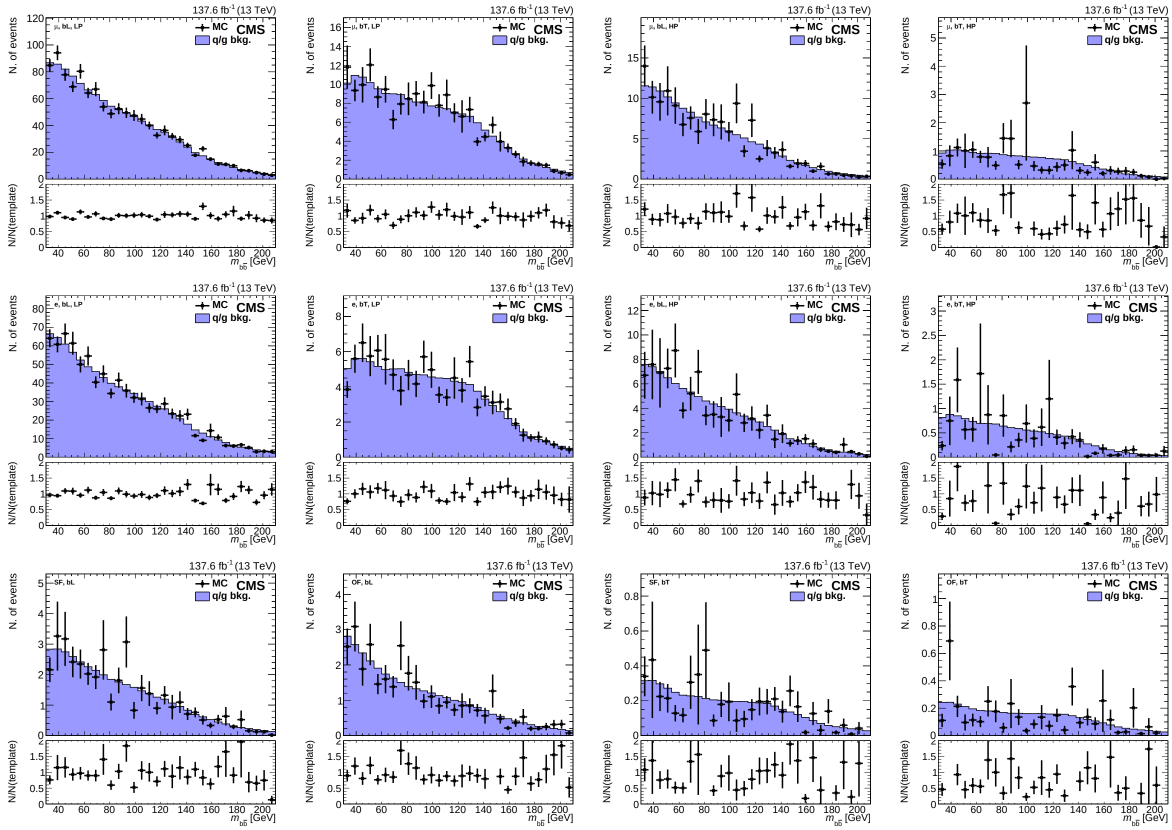

Figure 5.10: Projections of the full fitted q/g background model into $m_{\mathrm{b\bar{b}}}$ for each SR category. From top to bottom, the rows show the 1ℓ muon, 1ℓ electron, and then 2ℓ categories. From left to right, the 1ℓ rows show the SR categories bL LP, bT LP, bL HP, bT HP, and the 2ℓ rows show SF bL, OF bL, SF bT, OF bT.

Examples of fitted distributions are shown in Figs. 5.14 and 5.15.

The 2D P_{HH} templates are built to account for correlations between m_{HH} and $m_{\mathrm{b\bar{b}}}$. These m_{HH} distributions are also modeled with a DSCB function, but with an additional linear dependence on $m_{\mathrm{b\bar{b}}}$, parameterized by $\Delta_{\mathrm{b\bar{b}}} = (m_{\mathrm{b\bar{b}}} - \mu_{\mathrm{b\bar{b}}})/\sigma_{\mathrm{b\bar{b}}}$. Here, $\mu_{\mathrm{b\bar{b}}}$ and $\sigma_{\mathrm{b\bar{b}}}$ are the mean and width parameters, respectively, in the fit to the $m_{\mathrm{b\bar{b}}}$ spectra. To accomplish this, the mean parameter μ_{HH} along m_{HH} in the DSCB function fit is then

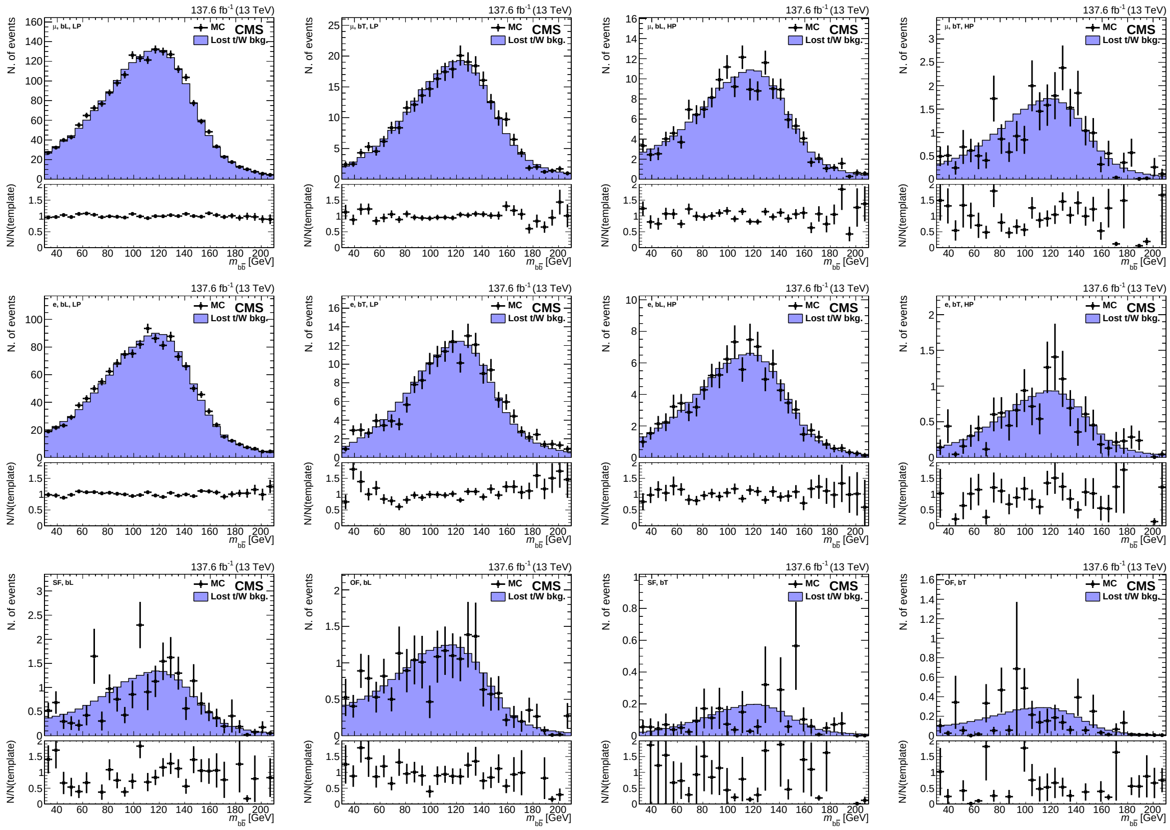

Figure 5.11: Projections of the full fitted lost t/W background model into $m_{b\bar{b}}$ for each SR category. From top to bottom, the rows show the 1ℓ muon, 1ℓ electron, and then 2ℓ categories. From left to right, the 1ℓ rows show the SR categories bL LP, bT LP, bL HP, bT HP, and the 2ℓ rows show SF bL, OF bL, SF bT, OF bT.

$$\mu_{\mathrm{HH}} = \mu_0(1 + \mu_1 \Delta_{b\bar{b}}), \tag{5.12}$$

where μ_0 and μ_1 are fit parameters. Doing this, we can account for mismeasurements of the $b\bar{b}$ jet that result in mismeasurements of m_{HH}. The resolution of the m_{HH} resonance, denoted as σ_{HH}, is also dependent on $m_{b\bar{b}}$ such that

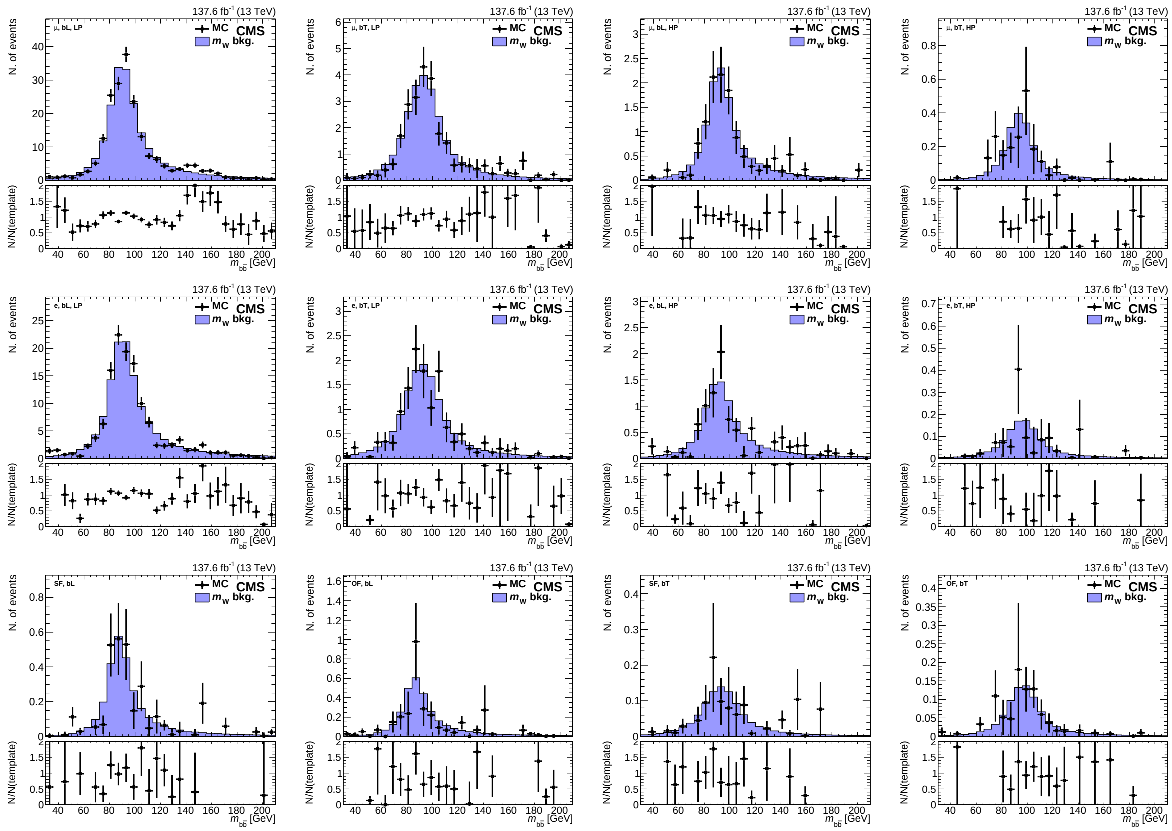

Figure 5.12: Projections of the full fitted m_{W} background model into $m_{\mathrm{b\bar{b}}}$ for each SR category. From top to bottom, the rows show the 1ℓ muon, 1ℓ electron, and then 2ℓ categories. From left to right, the 1ℓ rows show the SR categories bL LP, bT LP, bL HP, bT, HP, and the 2ℓ rows show SF bL, OF bL, SF bT, OF bT.

$$\sigma_{\mathrm{HH}} = \begin{cases} \sigma_0(1 + \sigma_1 \Delta_{\mathrm{b\bar{b}}}), & \Delta_{\mathrm{b\bar{b}}} < 0 \\ \sigma_0, & \Delta_{\mathrm{b\bar{b}}} > 0 \end{cases} \tag{5.13}$$

where σ_0 and σ_1 are fit parameters. In the case that the SD algorithm produces an undermeasurement of $m_{\mathrm{b\bar{b}}}$ by removing too much energy from the Higgs boson decay, the correlation increases, and the m_{HH} resolution grows wider. For $|\Delta_{\mathrm{b\bar{b}}}| > 2.5$, we use the value

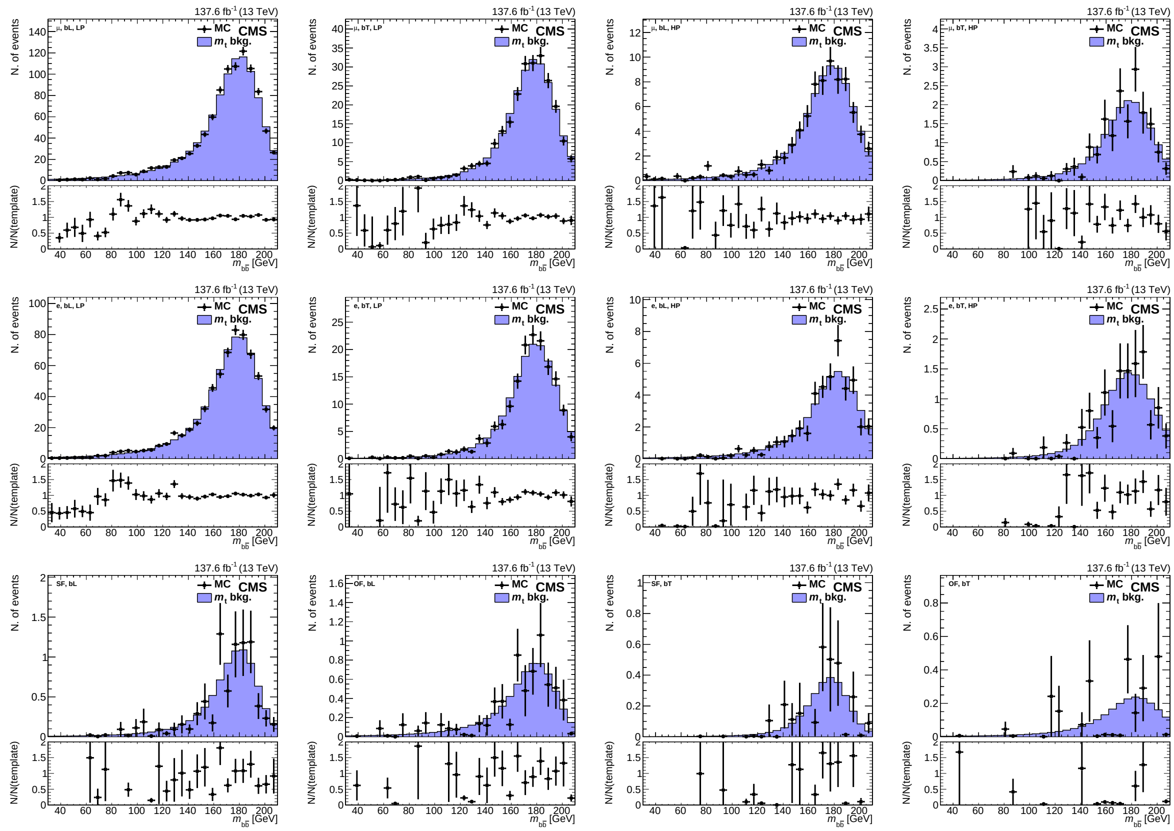

Figure 5.13: Projections of the full fitted m_{t} background model into $m_{\mathrm{b\bar{b}}}$ for each SR category. From top to bottom, the rows show the 1ℓ muon, 1ℓ electron, and then 2ℓ categories. From left to right, the 1ℓ rows show the SR categories bL LP, bT LP, bL HP, bT, HP, and the 2ℓ rows show SF bL, OF bL, SF bT, OF bT.

at the boundary since the correlation does not hold for severe mismeasurements. The m_{HH} resolution is $\approx 5\%$.

The product of the acceptance and efficiency for $\mathrm{X} \to \mathrm{HH}$ events to fall into any of the individual search categories is taken from simulation. As done for the signal shape parameters, the event yields are interpolated along m_{X} to obtain the efficiencies. Uncertainties in the relative acceptances and in the integrated luminosity of the sample are included in the 2D maximum likelihood fit that is used to obtain confidence intervals on the $\mathrm{X} \to \mathrm{HH}$

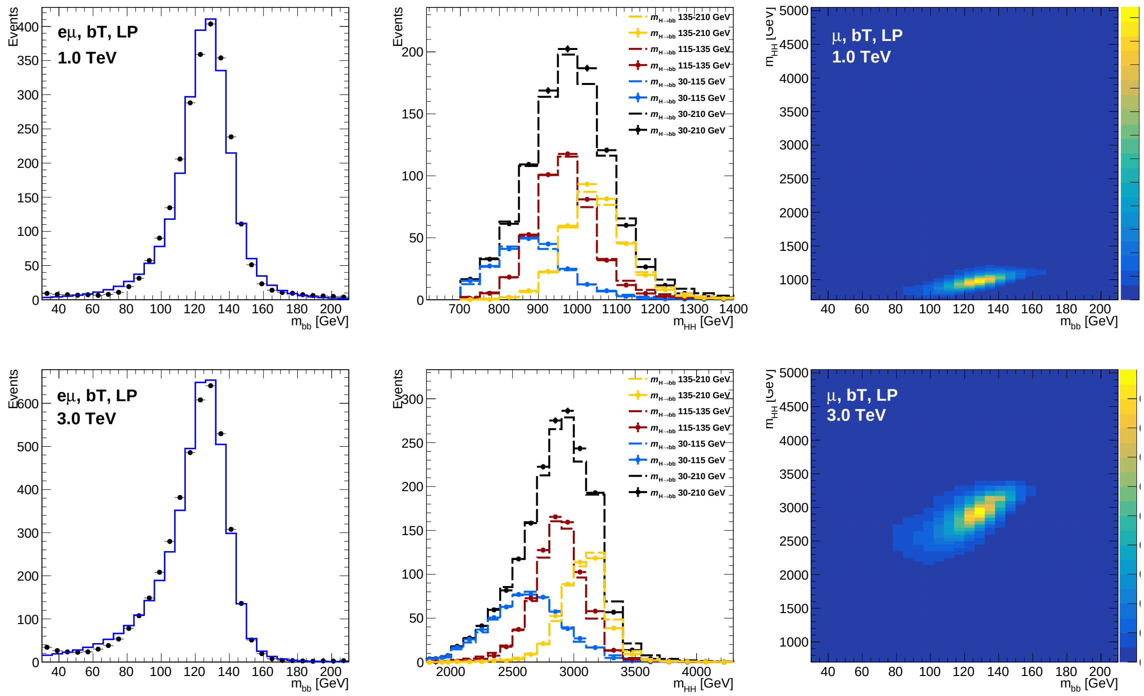

Figure 5.14: Left column: example 1D inclusive $P(m_{\mathrm{b\bar{b}}})$ templates for the 1ℓ channel μ bT LP category, showing the fitted template and the simulation in the fitting region. Middle column: projections of the 2D fitted templates and simulation into m_{HH}, shown for various ranges of $m_{\mathrm{b\bar{b}}}$ to demonstrate the correlations. Right column: the full 2D model for one such SR category and resonance mass m_{X}. The top (bottom) row shows a signal at 1.0 (3.0) TeV. The selection and generated mass are shown in the upper left corner for each plot. The $P(m_{\mathrm{b\bar{b}}})$ shown in the left column are those used to construct the full model for the distinct SR category in the bins to their right.

process cross section. The event yields, assuming a $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH}) = 1.0$ pb are shown for spin-0 (spin-2) in Figs. 5.16 and 5.17 (5.18 and 5.19).

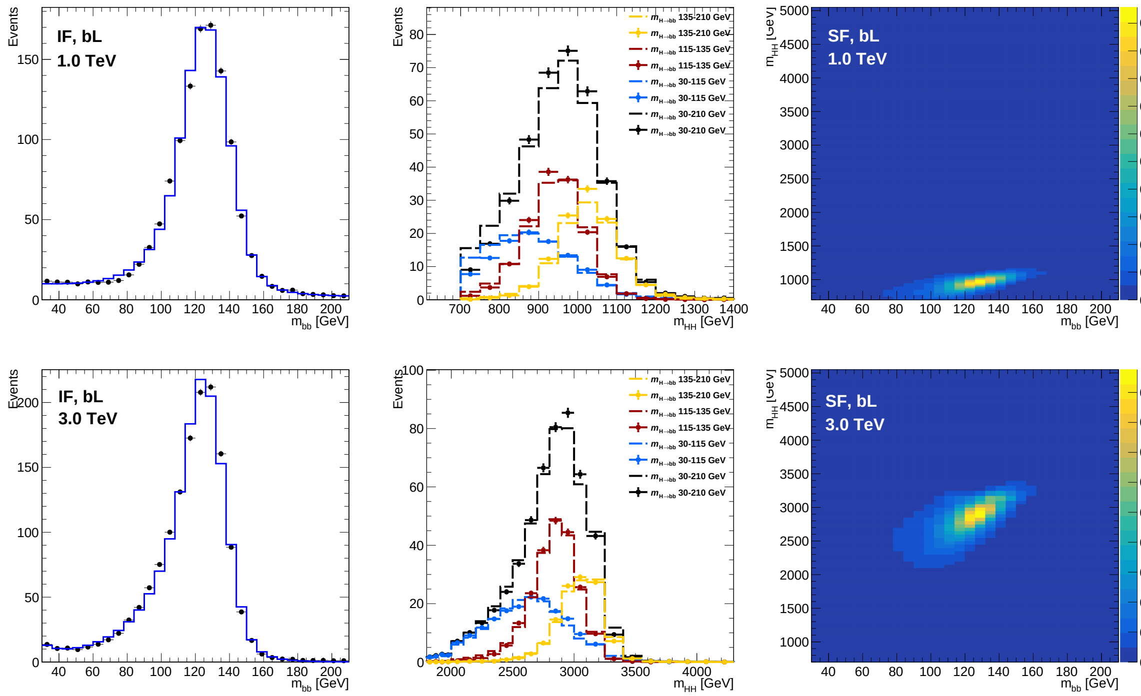

Figure 5.15: Left column: example 1D inclusive $P(m_{\mathrm{b\bar{b}}})$ templates for the 2ℓ channel SF bL category, showing the fitted template and the simulation in the fitting region. Middle column: projections of the 2D fitted templates and simulation into m_{HH}, shown for various ranges of $m_{\mathrm{b\bar{b}}}$ to demonstrate the correlations. Right column: the full 2D model for one such SR category and resonance mass m_{X}. The top (bottom) row shows a signal at 1.0 (3.0) TeV. The selection and generated mass are shown in the upper left corner for each plot. The $P(m_{\mathrm{b\bar{b}}})$ shown in the left column are those used to construct the full model for the distinct SR category in the bins to their right.

5.2 Systematic uncertainties and fitting method

Systematic uncertainties that affect the normalization and shape of the signal and background are modeled with nuisance parameters in the 2D maximum likelihood fit to data. These uncertainties are used to account for distortions in the shape and normalization, which correspond to possible mismodeling of the MC simulation, particularly in the scales and res-

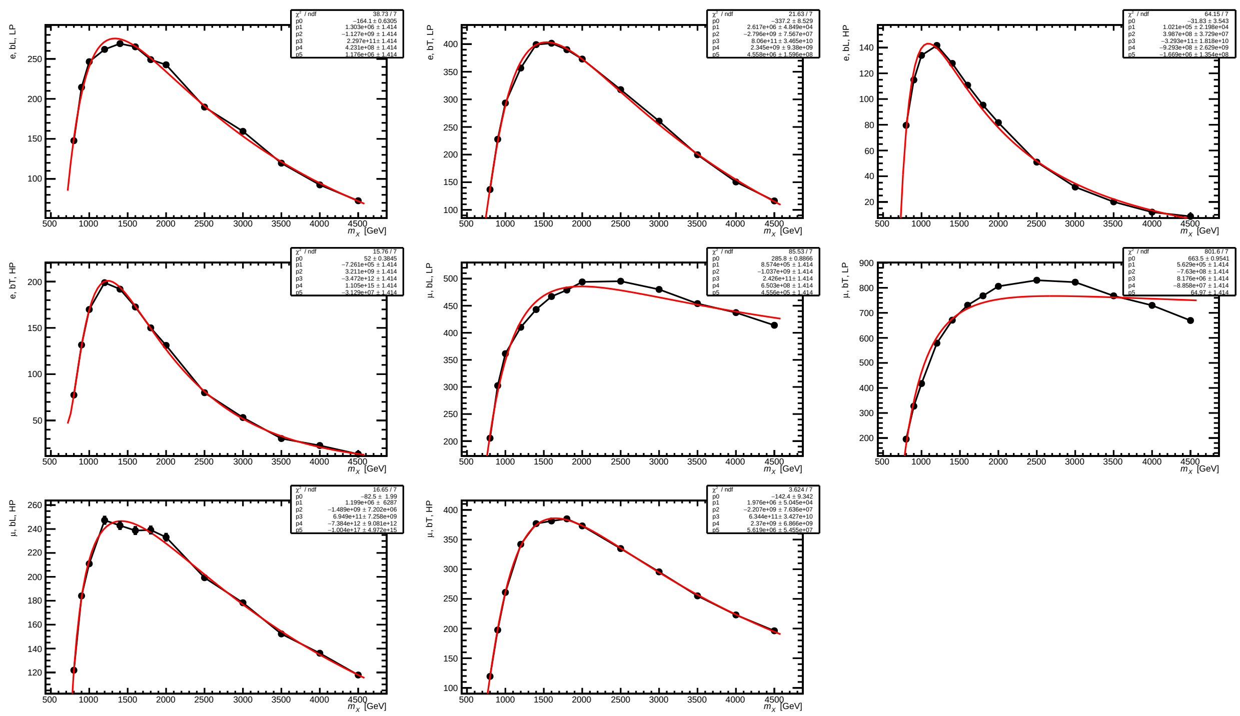

Figure 5.16: Expected event yields for a spin-0 resonance interpolated between various m_{X} for each category in the 1ℓ channel.

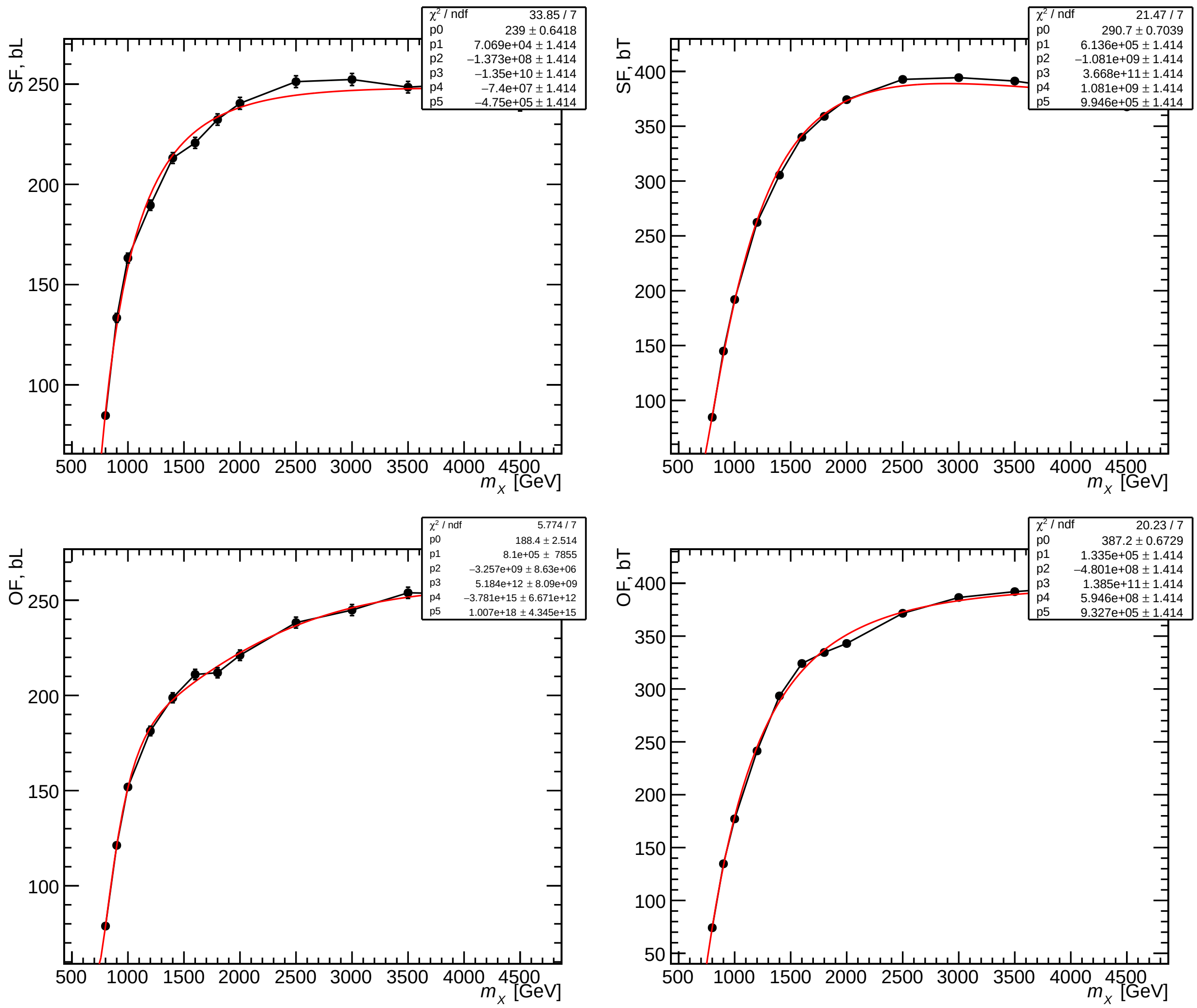

Figure 5.17: Expected event yields for a spin-0 resonance interpolated between various m_{X} for each category in the 2ℓ channel.

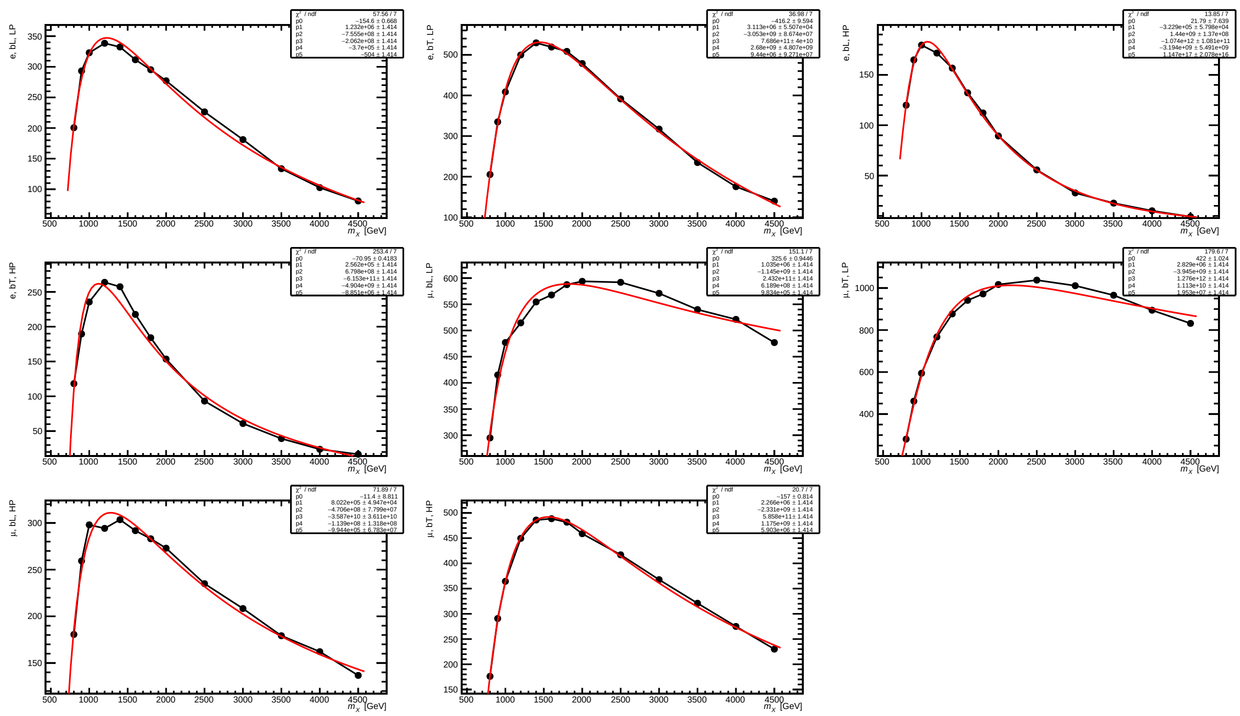

Figure 5.18: Expected event yields for a spin-2 resonance interpolated between various m_{X} for each category in the 1ℓ channel.

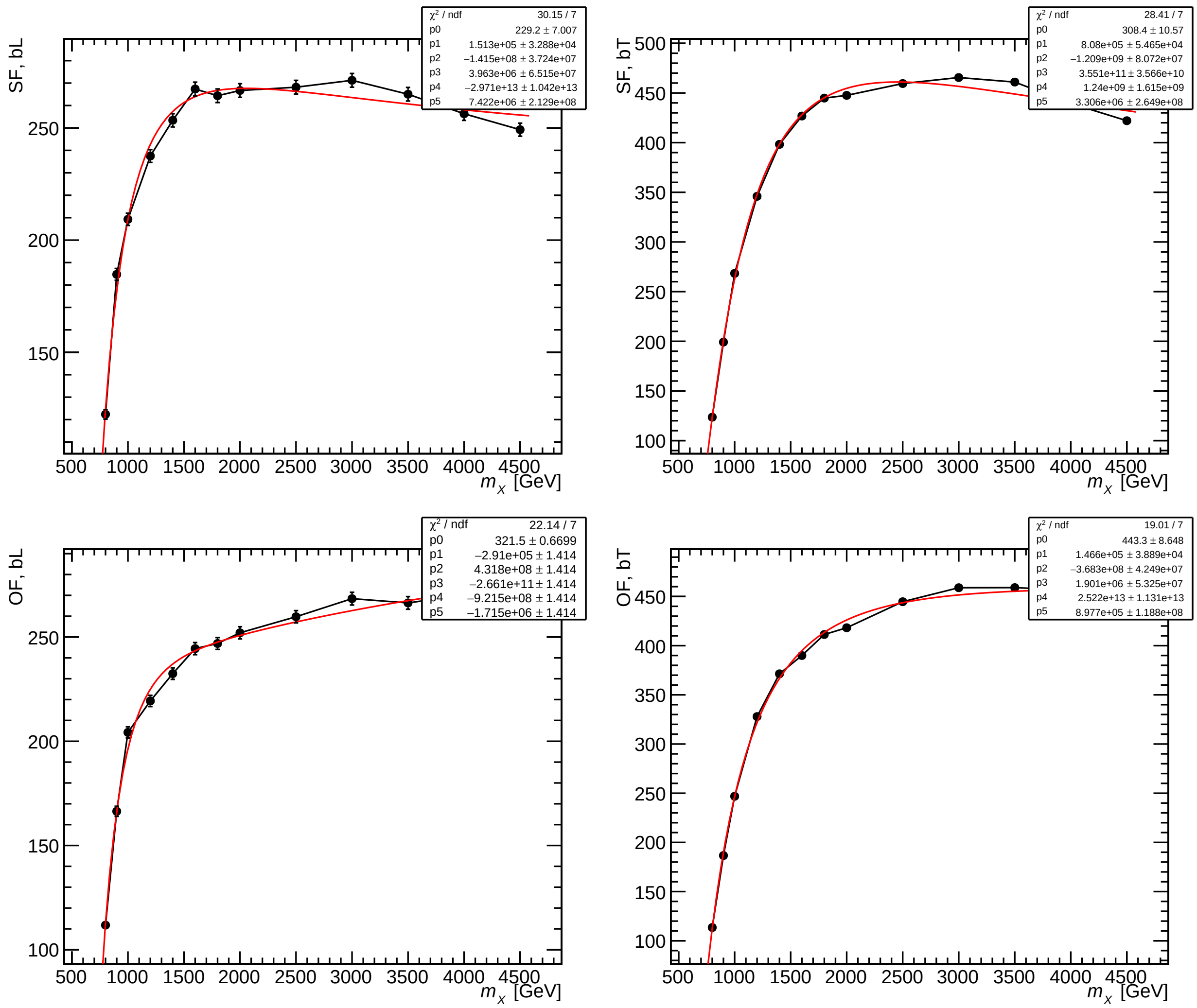

Figure 5.19: Expected event yields for a spin-2 resonance interpolated between various m_{X} for each category in the 2ℓ channel.

olutions of the particle masses. Loose constraints are placed on the nuisance parameters, corresponding to conservative initial estimates of the uncertainties, before performing the fit to the data. These pre-fit estimates are motivated by discrepancies between data and simulation in the CRs, and the analysis results are not sensitive to their initial sizes, provided that they are large enough to fit the data well. Ultimately, the nuisance parameters are constrained in the fit to data, and the final uncertainties come from the fit.

Nuisance parameters for the shape uncertainties have Gaussian constraints, while those for normalization uncertainties have log-normal constraints (a typical choice of pdf for multiplicative corrections such as efficiencies). In certain cases, a single nuisance parameter may affect both the normalization and shape of a mass resonance, in which case the nuisance parameter constraint is Gaussian. There are a total of 129 nuisance parameters in the full signal + background (S+B) model: 102 of these parameters affect the background only, 25 affect the signal only, and 2 affect both the signal and background. All uncertainties are summarized in Tables 5.3 and 5.4 for the background and signal, respectively. Detailed descriptions of the uncertainties are given in the rest of this section.

5.2.1 Background uncertainties

The systematic uncertainties for the background are chosen by considering possible discrepancies between data and simulation due to features such as the relative background compositions or the jet energy scale. The studies in the two CRs described in Section 6.1.4 are used to verify that the chosen uncertainties are able to cover such differences. The fitted values of the nuisance parameters do not strongly depend on the size of the pre-fit uncertainty estimates, verified with tests that inflate the chosen initial uncertainties and then perform a fit to the data.

Shape uncertainties for the nonresonant mass distributions, which only feature in the background, are implemented by including alternative templates in the fit. The alternative templates are constructed by "morphing" a nominal template along one of the mass

Table 5.3: Background systematic uncertainties included in the maximum likelihood fit. The uncertainty types with "normalization" correspond to uncertainties on the background yield, while all others are uncertainties on the background shape. The N_{p} column indicates the number of nuisance parameters used to model the uncertainty. In the last two columns, σ_{I} refers to the initial estimate of the uncertainty, and σ_{C} refers to the constrained uncertainty obtained post-fit. For the q/g, $\mathrm{t\bar{t}}$, and lost t/W shape uncertainties, "scale" uncertainties are those implemented with alternative templates with $f(m) = m$, and "inverse scale" from $f(m) = 1/m$.

Uncertainty type	Processes	N_{p}	σ_{I}	$\sigma_{\mathrm{C}}/\sigma_{\mathrm{I}}$
SD jet $m_{\mathrm{b\bar{b}}}$ scale	m_{W}, m_{t}, signal	2	0.54%, 2.0% (m_{t})	98%, 19% (m_{t})
SD jet $m_{\mathrm{b\bar{b}}}$ resolution	m_{W}, m_{t}, signal	2	8.6%, 17.2% (m_{t})	95%, 25% (m_{t})
q/g normalization	q/g	12	50% (1ℓ), 100% (2ℓ)	37–78%
q/g m_{HH} scale	q/g	10	$\pm 0.5\ m_{\mathrm{HH}}/\mathrm{TeV}$	78–99%
q/g m_{HH} inverse scale	q/g	10	$\pm 1.4\ \mathrm{TeV}/m_{\mathrm{HH}}$	64–99%
q/g $m_{\mathrm{b\bar{b}}}$ scale	q/g	4	$\pm 0.00375\ m_{\mathrm{b\bar{b}}}/\mathrm{GeV}$	81–99%
q/g $m_{\mathrm{b\bar{b}}}$ inverse scale	q/g	4	$\pm 15\ \mathrm{GeV}/m_{\mathrm{b\bar{b}}}$	77–99%
Lost t/W $m_{\mathrm{b\bar{b}}}$ scale	lost t/W	4	$\pm 0.003\ m_{\mathrm{b\bar{b}}}/\mathrm{GeV}$	71–99%
Lost t/W $m_{\mathrm{b\bar{b}}}$ inverse scale	lost t/W	4	$\pm 18\ \mathrm{GeV}/m_{\mathrm{b\bar{b}}}$	88–99%
top normalization	lost t/W, m_{W}, m_{t}	12	35% (1ℓ), 70% (2ℓ)	19–68%
top relative normalization	lost t/W, m_{W}, m_{t}	8	35% (1ℓ), 70% (2ℓ)	9–96%
top m_{HH} scale	lost t/W, m_{W}, m_{t}	12	$\pm 0.25\ m_{\mathrm{HH}}/\mathrm{TeV}$	84–99%
top m_{HH} relative scale	lost t/W, m_{W}, m_{t}	8	$\pm 0.25 * m_{\mathrm{HH}}/\mathrm{TeV}$	74–99%
top m_{HH} inverse scale	lost t/W, m_{W}, m_{t}	12	$\pm 0.7 * \mathrm{TeV}/m_{\mathrm{HH}}$	61–99%

dimensions: modified event weights that include multiplicative parameters proportional to a function of the mass dimension are applied to the nominal distributions. For each morphing variable $f(m)$, two alternative templates are created corresponding to one upward and one downward shift in the shape of the nominal template. The change in the bin contents n_i of the template histograms due to the morphing for alternative template production takes the form:

$$\text{up:} \qquad n_i \to n_i(1 + \alpha_f f(m)) \tag{5.14}$$

$$\text{down:} \qquad n_i \to \frac{n_i}{(1 + \alpha_f f(m))}. \tag{5.15}$$

The coefficient α_f is tuned independently for each $f(m)$ used in the modeling. The $f(m)$ used to produce the alternative templates for fitting in this analysis are $f(m) = m$ and $1/m$ for each of $m_{\mathrm{b\bar{b}}}$ and m_{HH}. Thus, all the morphing parameterizations used for alternative templates are $\propto m_{\mathrm{b\bar{b}}}$, $1/m_{\mathrm{b\bar{b}}}$, m_{HH}, and $1/m_{\mathrm{HH}}$.

Relative to the nominal template, each set of up and down alternative templates corresponds to a $\pm 1\sigma$ deviation of the shape in the metric $f(m)$. Fitting the nominal template to some data sample requires the input of the nominal template and all the desired sets of up and down templates. With the input histograms, the fitted shapes (event fractions in each histogram bin) are obtained by interpolating with a spline between the $\pm 1\sigma$ alternative templates and extrapolating linearly beyond them [132]. In the fit to the data, the nuisance parameters for the shape uncertainties are Gaussian multiplicative factors that encode the fitted shape relative to the nominal and alternative templates: all shape nuisances have an initial value of 0 (corresponding to the nominal template) and a Gaussian width of 1 (corresponding to the up and down histograms).

Systematic uncertainties for the shape of resonant mass distributions ($m_{\mathrm{b\bar{b}}}$ for the m_{W} and m_{t} backgrounds) account for deviations in the scale and resolution of the mass resonances from the simulation to data. These systematics are implemented with nuisance parameters that function as multiplicative scale factors to the μ and σ parameters of the DSCB function that is used to fit the mass peak. Overall, for both the resonant and nonresonant cases, none of the background shape uncertainties are correlated between $m_{\mathrm{b\bar{b}}}$ and m_{HH} in order to simplify the modeling. Since the analysis searches for a resonant signal peak on top of a falling background distribution, the most precise modeling of the falling shape is not necessary. Correlating the $m_{\mathrm{b\bar{b}}}$ and m_{HH} systematics would not significantly impact the sensitivity relative to the adopted strategy.

Many of the background systematic uncertainties are correlated among the three background components whose compositions are dominated by $\mathrm{t\bar{t}}$ production: the lost t/W, m_{W}, and m_{t} backgrounds. These systematics are referred to as the top background systematics. Discrepancies in the global $\mathrm{t\bar{t}}$ normalization will manifest as discrepancies in these three backgrounds in the same manner. However, differences in the $\mathrm{b\bar{b}}$ jet tagging and $\mathrm{t\bar{t}}$ p_{T} spectrum (which is effectively m_{HH}) can introduce differences in the shape and relative composition of the three top backgrounds. For this reason, there are additional nuisance parameters that allow these three backgrounds to vary in independent or anti-correlated ways relative to each other. Nuisance parameters for the q/g background uncertainties are independent of the uncertainties for the other three backgrounds, as its composition is dominated by different SM processes.

Each systematic uncertainty will now be detailed for the top backgrounds and then the q/g background. Each description contains the name of the uncertainty in bold, the number of nuisance parameters used to model the uncertainty, the initial pre-fit size of the uncertainty, and any other necessary information. This is all summarized in Table 5.3.

q/g normalization (12) A 50% uncertainty is applied. There is one independent nuisance parameter for each SR category.

top normalization (12) A 35% uncertainty is applied. There is one independent nuisance parameter for each SR category, applied to each of the lost t/W, m_{W}, and m_{t} backgrounds with 100% correlation to account for the total normalization of backgrounds derived from $\mathrm{t\bar{t}}$ and other processes with top quarks.

top relative normalization (8) A 30% uncertainty is applied to pairs of the three top backgrounds with 100% anti-correlation to allow their normalizations to vary against each other. This kind of uncertainty accounts for differences in the relative composition of these backgrounds in data and simulation. There is one systematic uncertainty that anti-correlates the lost t/W and m_{t} background normalizations and another that anti-correlates the m_{W} and m_{t} backgrounds. Studies in simulation show that the fractional composition of these backgrounds as part of the total top background in each channel is only sensitive to the $\mathrm{b\bar{b}}$ jet tagging, and so for each of these two systematics, one nuisance parameter is included per $\mathrm{b\bar{b}}$ jet tagging category independently in each of the 1ℓ and 2ℓ channels.

q/g $\boldsymbol{m_{\mathrm{b\bar{b}}}}$ shape (8) The $m_{\mathrm{b\bar{b}}}$ shape is allowed to vary between alternative templates constructed with $f(m) = m_{\mathrm{b\bar{b}}}$ and $1/m_{\mathrm{b\bar{b}}}$. The $m_{\mathrm{b\bar{b}}}$ shape exhibits some dependence on the $\mathrm{b\bar{b}}$ jet reconstruction but is independent of the other event categorization criteria. Thus, there is one nuisance parameter for each $f(m)$ and $\mathrm{b\bar{b}}$ jet tagging category in each of the 1ℓ and 2ℓ channels. Comparisons of the nominal q/g background template $m_{\mathrm{b\bar{b}}}$ shape and the variations due to these $m_{\mathrm{b\bar{b}}}$ systematic uncertainties are shown in the left plots of Figs. 5.20 and 5.21.

lost t/W $\boldsymbol{m_{\mathrm{b\bar{b}}}}$ shape (8) Same as for the q/g $m_{\mathrm{b\bar{b}}}$ shape systematics above. Comparisons of the nominal lost t/W background $m_{\mathrm{b\bar{b}}}$ shape and the variations due to these $m_{\mathrm{b\bar{b}}}$ systematic uncertainties are shown in Figs. 5.22 and 5.23.

resonant $\boldsymbol{m_{\mathrm{b\bar{b}}}}$ shape (4) For the m_{W} and m_{t} backgrounds, the uncertainties on the scale (μ) and resolution (σ) of the $m_{\mathrm{b\bar{b}}}$ DSCB shape are dominated by the effects of SD algorithm mismodeling. Uncertainties in the SD mass have been measured for each year of data-taking within CMS using a sample of two-prong hadronic W/Z decays [125]. These uncertainties are used for the m_{W} background, but they cannot describe the uncertainties for three-prong jets from top quark decays. Thus, independent nuisance parameters are used for the m_{W} and m_{t} backgrounds. The uncertainties in the

$m_{\rm W}$ background are correlated with the signal since both $\rm H \to b\bar{b}$ and $\rm W \to q\bar{q}$ are two-prong decays, and the SD algorithm behaves similarly. The pre-fit estimates of the scale and resolution uncertainties respectively are 0.54% and 8.6% for the $m_{\rm W}$ background and 2% and 17.6% for the $m_{\rm t}$ background. The sizes of the $m_{\rm t}$ uncertainties are inflated to account for the inability of the CMS measurements to describe three-prong top jets. Comparisons of the nominal top background $m_{\rm b\bar{b}}$ shapes and the variations due to these SD systematic uncertainties are shown in Figs. 5.22 and 5.23.

q/g $\boldsymbol{m_{\rm HH}}$ shape (20) The $m_{\rm HH}$ shape is allowed to vary between alternative templates constructed with $f(m) = m_{\rm HH}$ and $1/m_{\rm HH}$. The $m_{\rm HH}$ shape exhibits some dependence on all SR category criteria except the lepton flavor in the 2ℓ channel. Thus, for each $f(m)$, there is one independent nuisance parameter for each 1ℓ category and each $\rm b\bar{b}$ jet tagging category in the 2ℓ channel. Comparisons of the nominal q/g background template $m_{\rm HH}$ shape and the variations due to these $m_{\rm HH}$ systematic uncertainties are shown in the right plots of Figs. 5.20 and 5.21.

top $\boldsymbol{m_{\rm HH}}$ shape (24) The $m_{\rm HH}$ shape is allowed to vary between alternative templates constructed with $f(m) = m_{\rm HH}$ and $1/m_{\rm HH}$ among the lost t/W, $m_{\rm W}$, and $m_{\rm t}$ backgrounds with 100% correlation. There is one independent nuisance parameter for each SR category and each $f(m)$ for the set of the these three backgrounds. Comparisons of the nominal top background $m_{\rm HH}$ shapes and the variations due to these $m_{\rm HH}$ systematic uncertainties are shown in the lower plots of Figs. 5.22 and 5.23.

top $\boldsymbol{m_{\rm HH}}$ relative scale (8) The $m_{\rm HH}$ shape is allowed to vary between alternative templates from $f(m) = m_{\rm HH}$ only for pairs of the three top backgrounds with 100% anti-correlation. This is motivated by the fact that the $\rm t\bar{t}$ boost is somewhat correlated with the background type: the $m_{\rm t}$ background exhibits a harder $m_{\rm HH}$ spectrum than both the lost t/W and $m_{\rm W}$ backgrounds (which are similar). Nuisance parameters are included that allow the scale of the $m_{\rm t}$ background $m_{\rm HH}$ spectrum to float independently of the lost t/W and m_W backgrounds and vice-versa for the lost t/W and m_W components together. Accounting for differences in the relative top background composition due to the $\mathrm{b\bar{b}}$ jet tagging, there is one nuisance parameter for each of the aforementioned pairs, each $\mathrm{b\bar{b}}$ jet tagging category, and each of the 1ℓ and 2ℓ channels.

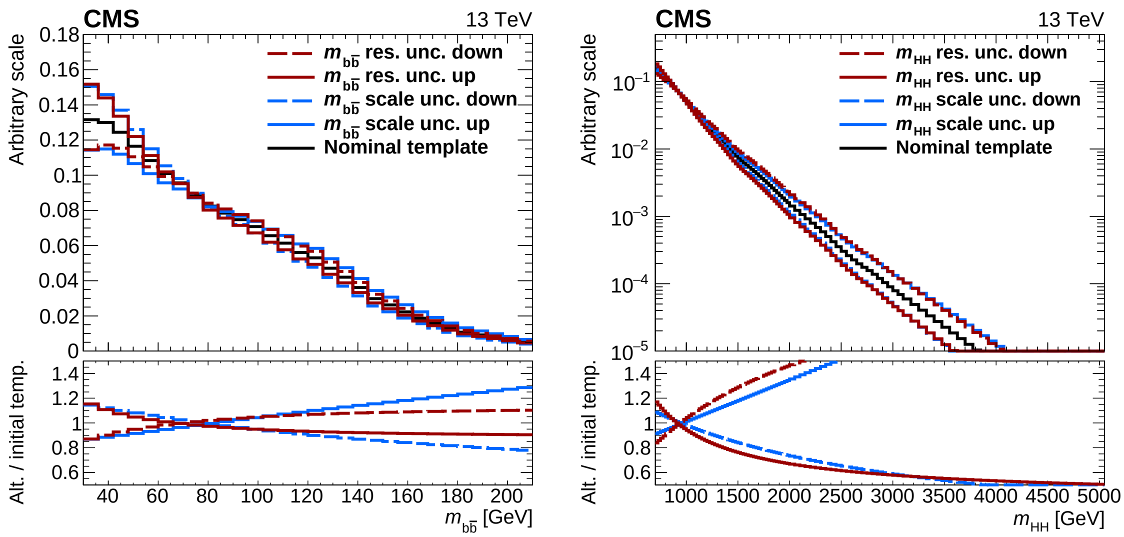

Figure 5.20: Projections of the nominal q/g background template and templates morphed by 1σ of the pre-fit shape uncertainties for the category μ bL LP. Left: the $m_{\mathrm{b\bar{b}}}$ projection with shifts of the $m_{\mathrm{b\bar{b}}}$ systematic uncertainties. Right: the m_{HH} projection with shifts of the m_{HH} systematic uncertainties. Each curve corresponds to shifting the systematic of interest by 1σ while keeping the others fixed.

5.2.2 Signal uncertainties

Systematic uncertainties are applied to the normalization of the signal and to the shape of its mass peak in both the $m_{\mathrm{b\bar{b}}}$ and m_{HH} dimensions. As done for the background, descriptions are now provided that contain the name of the uncertainty in bold, the number of nuisance parameters used to model the uncertainty, the estimate of the uncertainty, and any other necessary information. Unless otherwise noted, the signal uncertainty is 100% correlated among all the SR categories. The signal uncertainty information is summarized in Table 5.4.

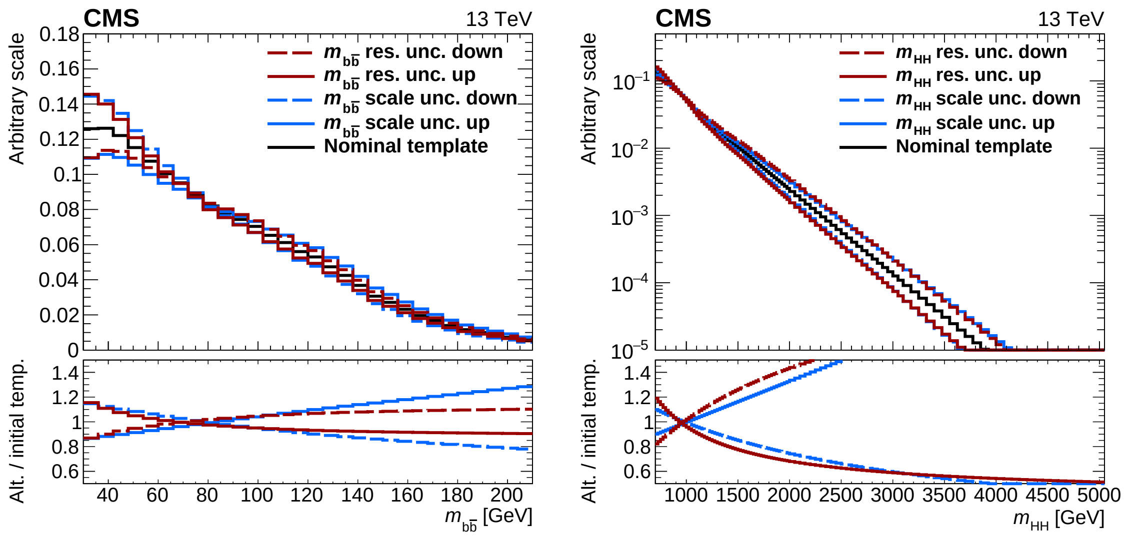

Figure 5.21: Projections of the nominal q/g background template and templates morphed by 1σ of the pre-fit shape uncertainties for the category SF bL. Left: the $m_{b\bar{b}}$ projection with shifts of the $m_{b\bar{b}}$ systematic uncertainties. Right: the m_{HH} projection with shifts of the m_{HH} systematic uncertainties. Each curve corresponds to shifting the systematic of interest by 1σ while keeping the others fixed.

Luminosity (1) A flat 1.8% uncertainty on the 2016–2018 measured integrated luminosity of 137.6 fb^{-1} is applied to the signal normalization [133–135].

PDF and $\boldsymbol{\alpha_S}$ (1) A 2.0% (2.5%) uncertainty is applied to the spin-0 (spin-2) signal to account for uncertainties in the choice of parton distribution functions (PDFs) and QCD scale α_S in the simulation. The uncertainty is obtained by considering variations of the scale and PDF sets. The envelope of the scale variations is taken as the scale uncertainty, and the standard deviation of the distribution of PDF set weights is taken as the PDF uncertainty. The total uncertainty is obtained from adding the PDF and scale uncertainties in quadrature.

Pileup (1) A 1.0% (0.6%) uncertainty is applied to all 1ℓ (2ℓ) categories to account for uncertainties in the data pileup profile that is used to reweight simulated events. The

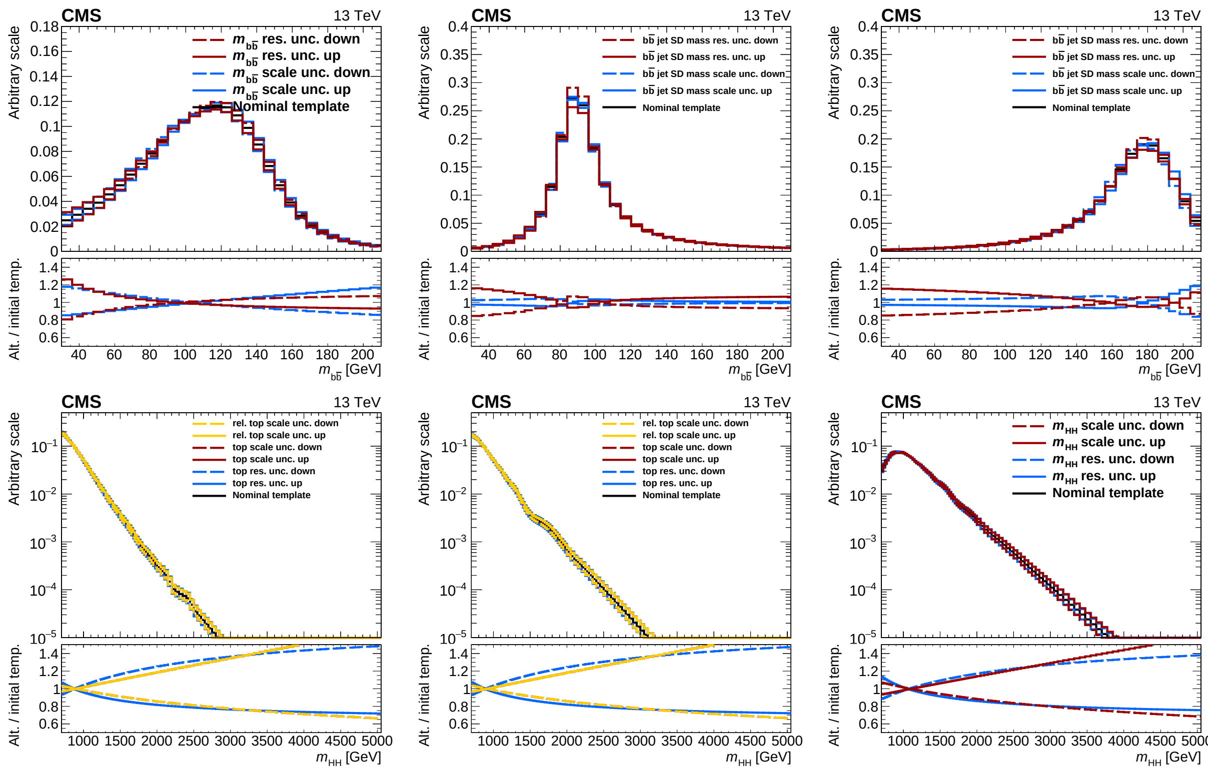

Figure 5.22: Projections of the nominal and $\pm 1\sigma$ alternative background templates for the category μ bL LP. The top (bottom) row contains the $m_{b\bar{b}}$ (m_{HH}) projections. From left to right, the lost t/W, m_{W}, and m_{t} backgrounds are shown.

data pileup profiles used, seen earlier in Fig. 3.10, were measured using a minimum bias cross section $\sigma_{\text{min-bias}}$ of 69.2 mb. Alternate pileup profiles corresponding to shifts in $\sigma_{\text{min-bias}}$ up and down by 4.6% are used to reweight the signal events, and the relative differences in the signal yield relative to the nominal $\sigma_{\text{min-bias}}$ are considered as the uncertainty.

L1 ECAL prefiring (1) A 0.5% uncertainty is applied to the signal normalization to account for the L1 ECAL prefiring rate observed in 2016–2017 data [136]. A gradual timing shift was observed in the ECAL that caused high $|\eta|$ trigger primitives to be associated to the previous bunch crossing (BX), which would veto selection of the correct BX since two consecutive BXes cannot fire the trigger. Simulated events in 2016 and

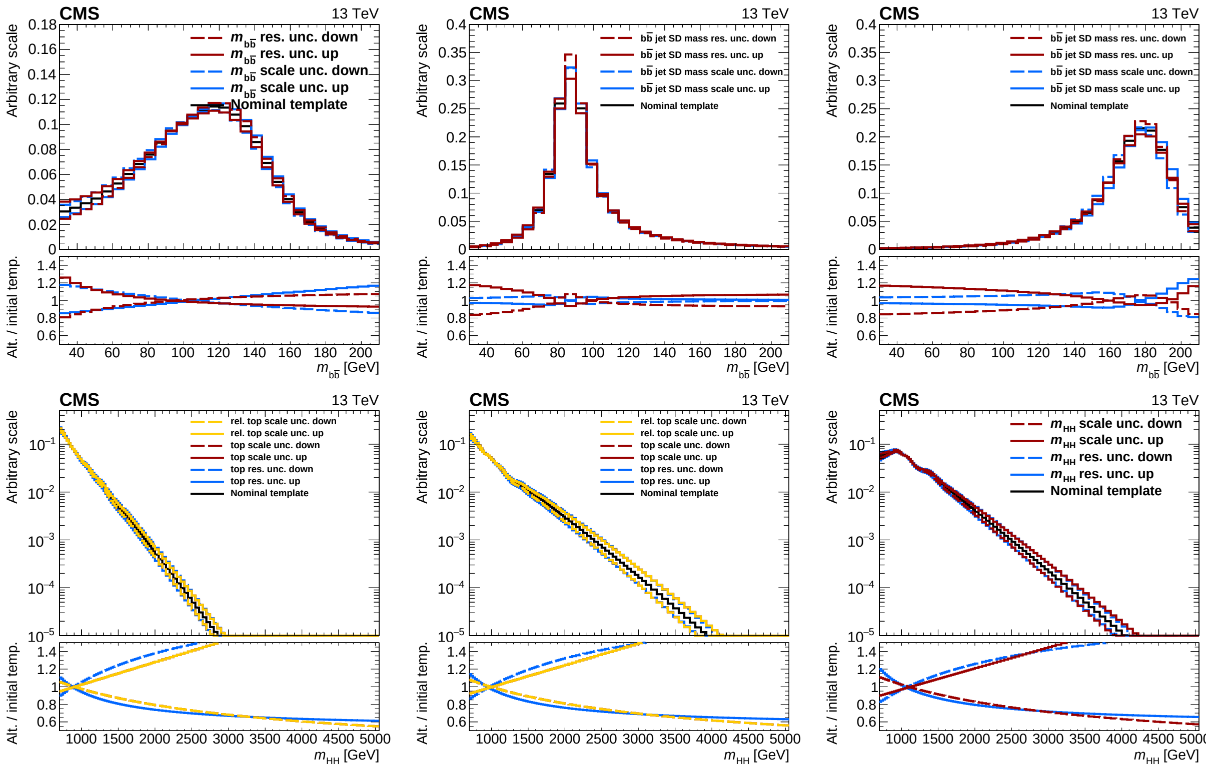

Figure 5.23: Projections of the nominal and $\pm 1\sigma$ alternative background templates for the category SF bL. The top (bottom) row contains the $m_{b\bar{b}}$ (m_{HH}) projections. From left to right, the lost t/W, m_W, and m_t backgrounds are shown.

2017 receive an event weight that simulates this effect. The uncertainty is evaluated by applying alternative event weights to the signal that correspond to shifts in the prefiring probability.

Trigger efficiency (6) One nuisance parameter per trigger efficiency measurement described in Section 4.2 (e and μ, ee, $\mu\mu$, $e\mu$ and μe). A 2% uncertainty is applied for 1ℓ categories, and a 3% uncertainty for 2ℓ categories.

Electron reconstruction efficiency (1) Uncertainties were evaluated in a $Z/\gamma^* \rightarrow \ell\ell$ sample and then propagated through the full analysis selection. The pre-fit uncertainty is gauged by analyzing the relative change in the signal yield. The applied

Table 5.4: Signal systematic uncertainties included in the maximum likelihood fit. The N_{p} column indicates the number of nuisance parameters used to model the uncertainty. In the "Uncertainty values" column, some uncertainties are noted to affect both the yield (Y) and m_{HH} shape (S for scale, R for resolution) of the signal. All other uncertainties, except the SD jet mass uncertainties, are uncertainties in the signal yield.

Uncertainty type	N_{p}	Uncertainty values
SD jet $m_{\mathrm{b\bar{b}}}$ scale	1	0.54%
SD jet $m_{\mathrm{b\bar{b}}}$ resolution	1	8.6%
Integrated uminosity	1	1.6%
PDFs+scales	1	spin-0: 2.0%, spin-2: 2.5%
Trigger	6	1ℓ: 2.0%, 2ℓ: 3.0%
Pileup	1	1ℓ: 1.0%, 2ℓ: 0.6%
Electron reconstruction	1	1ℓ: 0.5%, 2ℓ: < 0.8%
Electron identification	2	1ℓ: 4.2%, 2ℓ: < 2.6%
Muon identification	2	1ℓ: 2.3%, 2ℓ: < 2.3%
Electron isolation	1	1ℓ: 6.0%, 2ℓ: 3.0% for each electron
Muon isolation	1	1ℓ: 6.0%, 2ℓ: 2.0% for each muon
Jet energy scale	1	Y: 2.0%, $S(m_{\mathrm{HH}})$: 0.8%, $R(m_{\mathrm{HH}})$: 3.0%
Jet energy resolution	1	Y: 0.5%. $S(m_{\mathrm{HH}})$: 0.3%, $R(m_{\mathrm{HH}})$: 4.0%
Unclustered energy	1	Y: 0.5%, $R(m_{\mathrm{HH}})$: 1.5%
Other detector effects	2	Y: 0.6%, $R(m_{\mathrm{HH}})$: 1.0%
AK4 jet b tag efficiency	1	< 4.0%
AK4 jet b tag misidentification rate	1	< 2.5%
$\mathrm{b\bar{b}}$ jet tagging	1	bL: 8.5%, bT: 11.5%
$\mathrm{q\bar{q}}$ jet τ_2/τ_1 efficiency	1	LP: 26%, HP: 6.7%

uncertainty is 0.5% for 1ℓ electron categories 0.8% for 2ℓ OF categories. In 2ℓ SF events, the fraction of ee ($\mu\mu$) events is a flat 46% (54%) across all signal m_X. An uncertainty of 0.37% is applied to 2ℓ SF events considering the ee fraction.

Lepton identification efficiency (4) One nuisance parameter is included for each of the four unique identification criteria used for electrons and muons in the 1ℓ and 2ℓ channels. The uncertainties were evaluated in a $Z/\gamma^* \to \ell\ell$ sample and then propagated through the full analysis selection. However, these measurements do not consider the impact parameter constraints, and so an extra 2% uncertainty is added in quadrature with the uncertainties from the $Z/\gamma^* \to \ell\ell$ measurements. Uncertainties of 4.2% for 1ℓ electron categories and 2.3% for 1ℓ muon categories are applied. In the 2ℓ channel, the electron uncertainty is 2.6%, and the muon uncertainty is 2.3%. The ee and $\mu\mu$ fractions are considered when applying the uncertainties to the SF categories.

Lepton isolation efficiency (2) A 6% uncertainty is applied to 1ℓ categories, while for 2ℓ categories a 2% uncertainty is applied for each electron and 3% for each muon. These large uncertainties compensate for the lack of a proper dedicated measurement for all of 2016–2018 and are also motivated by uncertainties measured for SUSY searches for our mini-isolation working points but with respect to different lepton ID criteria.

AK4 jet b tag efficiency (1) Systematic uncertainties are provided by DeepJet and propagated through the analysis phase space, where their effect on the signal normalization is analyzed. In the 2ℓ channel, the applied uncertainty is a flat 0.5%. In the 1ℓ channel, the uncertainties are larger. A flat 1% uncertainty is applied to the electron categories. For muon categories, an m_X-dependent monotonic uncertainty is applied, which is approximately 1% at low m_X and increases to 3% (4%) at high m_X for LP (HP) categories.

AK4 jet b tag misidentification (1) Systematic uncertainties are similarly provided by the DeepJet group, as in the efficiency description above. A flat uncertainty of 2%

(2.5%) is applied to the 1ℓ (2ℓ) categories.

$\mathrm{b\bar{b}}$ jet tagging efficiency (1) Systematic uncertainties are provided by the DEEPAK8 group and propagated through the analysis phase space, where their effect on the signal normalization is analyzed. One nuisance parameter is used to model the uncertainty, which has a different size depending on the $\mathrm{b\bar{b}}$ jet tagging category: 8.5% for bL and 11.5% for bT. The $\mathrm{b\bar{b}}$ jet tagging efficiency is the largest source of uncertainty on the signal normalization in the analysis.

τ_2/τ_1 selection efficiency (1) Applied only in the 1ℓ channel, scale factors and their associated uncertainties for the τ_2/τ_1 categorization were measured using the $\mathrm{W} \to \mathrm{q\bar{q}}$ peak in a $\mathrm{t\bar{t}}$-enriched sample for each year of data-taking [125]. The scale factors and uncertainties for each year are combined into measurements for the full 2016–2018 data. The low purity τ_2/τ_1 scale factor and uncertainty is 1.116 ± 0.294, while the high purity measurement is 0.99 ± 0.067. Furthermore, the uncertainties between the two τ_2/τ_1 working points are anti-correlated. By construction of the $\mathrm{H} \to \mathrm{WW}^*$ purity categories, the HP categories only contain high purity τ_2/τ_1 events, but the LP categories contain a mix of both low and high purity τ_2/τ_1 events. The fraction of LP events that fulfill only the low-purity τ_2/τ_1 selection is flat across the full range of signal m_{X} at 74.5%. The uncertainty applied to the normalization of LP categories takes this fraction and the anti-correlation into account.

Jet energy scale (1) The energies of all jets are shifted up and down by one standard deviation of their CMS-measured uncertainties, and the shifts in jet energy are propagated to the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. Relative changes in the normalization and signal shape are analyzed. We implement a normalization uncertainty of 2%, an m_{HH} scale uncertainty of 0.8%, and an m_{HH} resolution uncertainty of 3%. The effects of the jet energy scale on the $m_{\mathrm{b\bar{b}}}$ shape pale in comparison to the uncertainties from the SD mass modeling, so we assume no effect on the $m_{\mathrm{b\bar{b}}}$ shape.

Jet energy resolution (1) The energy resolution of all jets is smeared up and down by one standard deviation of their CMS-measured uncertainties, and the changes are propagated to the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. Relative changes in the normalization and signal shape are analyzed. We implement a normalization uncertainty of 0.5%, an m_{HH} scale uncertainty of 0.3%, and an m_{HH} resolution uncertainty of 4%. The effects of the jet energy scale on the $m_{\mathrm{b\bar{b}}}$ shape pale in comparison to the uncertainties from the SD mass modeling, so we assume no effect on the $m_{\mathrm{b\bar{b}}}$ shape.

Unclustered energy resolution (1) The missing energy is shifted by one standard deviation of its CMS-measured uncertainty, and the effects on the signal normalization and shape are evaluated. We implement a normalization uncertainty of 0.5% and an m_{HH} resolution uncertainty of 1.5%. The effects of the jet energy scale on the $m_{\mathrm{b\bar{b}}}$ shape pale in comparison to the uncertainties from the SD mass modeling, so we assume no effect on the $m_{\mathrm{b\bar{b}}}$ shape.

HEM 15/16 Effect (1) The effects of the HEM 15/16 detector modules that suffered a power outage in the 2018 run is described in Section 4.4. The under-measurement of jet energies in the HEM region are evaluated with the following test: for any jet with the tight ID, $p_{\mathrm{T}} > 15\mathrm{GeV}$, and $-1.57 < \phi_{\mathrm{jet}} < -0.87$, the four-momentum of the jet is scaled down either by 20% if $-2.5 < \eta_{\mathrm{jet}} < -1.3$ or by 35% if $-3.0 < \eta_{\mathrm{jet}} < -2.5$. The changes are then propagated to the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, and the effects are analyzed in the full analysis phase space. We observe and place a 0.3% uncertainty on the signal normalization and a 1.0% uncertainty on the m_{HH} resolution.

Resonant $m_{\mathrm{b\bar{b}}}$ shape (2) Because of the two-prong decay structure of both $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$ and $\mathrm{W} \rightarrow \mathrm{q\bar{q}}$, the same $m_{\mathrm{b\bar{b}}}$ uncertainties are applied to the signal as are applied to the m_{W} background with 100% correlation. In this way, the signal shape is constrained by the background. The scale and resolution uncertainties are 0.54% and 8.6%, respectively. The SD mass of the $\mathrm{b\bar{b}}$ jet does not directly enter the calculation of m_{HH}, which instead takes the ungroomed jet mass as part of the $b\bar{b}$ jet four-momentum. As a result, these uncertainties have a negligible effect on the m_{HH} shape of signal compared to the jet energy scale and resolution.

CHAPTER 6

Statistical Analysis and Results

6.1 Statistical analysis and fit validation

In this section, we detail a host of statistical tests used both to validate the models and to evaluate the existence of signal. Before evaluating the model performance in the SR, tests of the modeling are executed in the CRs, described in Section 6.1.4. With confidence that the modeling behavior is understood and under control in the CRs, the tests are then executed in the SR. The modeling is checked by analyzing the fitted values and constrained uncertainties relative to the pre-fit estimates — the so-called "pulls" — of the nuisance parameters from a fit to data. More importantly for validating the modeling, goodness-of-fit (GOF) tests, described in Section 6.1.2, are performed. The signal modeling is also tested for bias by artificially injecting signal into a toy data set and then fitting the model to extract the signal strength.

6.1.1 Likelihood function

The statistical methods employed in the analysis all rely on maximum likelihood estimation. Given the CMS data, a likelihood function is constructed with the pdfs of the models as

$$\mathcal{L}(\mathbf{n}|r\mathbf{s}+\mathbf{b},\boldsymbol{\theta}) = P(\mathbf{n}|r\mathbf{s}+\mathbf{b}) \cdot p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}), \tag{6.1}$$

where the bold terms are understood as vectors representing bins in a histogram or the set of all nuisance parameters. The data are represented by $\mathbf{n}$, the signal shape by $\mathbf{s}$ (normalized to 1.0 pb), and the shapes and normalizations of each of the four backgrounds by $\mathbf{b}$. The nuisance parameters by $\boldsymbol{\theta}$, with $\tilde{\boldsymbol{\theta}}$ corresponding to the initial values. The parameter of interest is r, which corresponds to the global signal strength. This likelihood $\mathcal{L}$ corresponds to the full signal + background (S+B) hypothesis. The background-only (B-only) hypothesis corresponds to setting $r = 0$.

The nuisance parameters are assumed to be independent from each other, so their joint pdf is the product of the individual pdfs for each:

$$p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}) = \prod_i p_i(\tilde{\theta}_i|\theta_i). \tag{6.2}$$

As discussed earlier, the nuisance parameters for shape (normalization) uncertainties have Gaussian (log-normal) pdfs. For binned distributions, P in Eq. (6.1) is given by the product of the Poisson probabilities for the contents in every bin of the histograms:

$$P(\mathbf{n}|r\mathbf{s} + \mathbf{b}) = \prod_i \frac{(rs_i + b_i)^{n_i}}{n_i!} e^{-(rs_i+b_i)}, \tag{6.3}$$

where i runs over all histogram bins. The maximum likelihood fits are also performed simultaneously in all categories of the analysis, and so the full likelihood for the analysis is the product of the likelihoods in each of the event categories (12 in the SR).

6.1.2 Goodness-of-fit tests

GOF tests are used to evaluate the compatibility of the post-fit models with the data. These tests are employed in both the CRs and the SRs, and they are performed by constructing a test statistic that is a function of all the information of the data and the models.

The test statistic used in the GOF is a likelihood ratio, which has several appealing properties. They are independent of the the metric in which both the parameters and the data are described. They are also useful for comparing two hypotheses for reasons such as the Neyman-Pearson Lemma [137]. The likelihood ratio is constructed using the maximum $\mathcal{L}$ from the fit and the saturated likelihood $\mathcal{L}_{\text{sat}}$ as the alternative likelihood. The saturated model has a total number of parameters equal to the total number of histogram bins and is thus able to fit the data exactly. This useful for providing a denominator in the construction of the test statistic:

$$t_{\text{GOF}} = -2\ln\left(\mathcal{L}_{\text{max}}/\mathcal{L}_{\text{sat}}\right) \equiv -2\ln\lambda. \tag{6.4}$$

In the limit of Gaussian data (large data in the histogram bins), t_{GOF} takes the form of the χ^2 GOF test statistic.

The goal of the GOF test is to estimate the distribution of the test statistic t_{GOF} and then analyze it in the context of the observed test statistic t_0 in data. The distribution of t_{GOF} is estimated by first performing a B-only fit to the data to obtain post-fit values for all the nuisance parameters. Toy data sets are then constructed by sampling the pdfs of the nuisance parameters around their post-fit values. For each toy experiment, the nominal model is fit to the toy data, and the t_{GOF} is calculated. The distribution $f(t)$ of these toy t_{GOF} can be used to compute a p-value for the observed as:

$$p = \int_{t_0}^{\infty} f(t)dt. \tag{6.5}$$

Two GOF tests with slight differences are used to analyze the model performance. The global GOF using $\mathcal{L}$ from Eq. (6.1) is computed using software provided by the CMS Higgs combination group. This software has two undesirable features. First, it is difficult to analyze the t_{GOF} distribution in certain subsets of the analysis phase space (as a function of mass or category). Second, the toy data sets are generated by sampling the nuisance parameters from pdfs constructed around the post-fit values but with the pre-fit widths. It would be more appropriate to sample from pdfs using the post-fit constrained uncertainties on the nuisance parameters as the widths. This is accomplished with software developed specifically for this analysis that calculates an alternative GOF. The alternative GOF uses post-fit widths in the nuisance parameter pdfs, but the test statistic is calculated from a likelihood that only includes the Poisson terms for the histogram bins (no nuisance terms). The tests that rely on this alternative GOF also are only used with the B-only model, so the test statistic takes the form:

$$t_{\text{alt}} = -2\ln\lambda = 2\sum_i (b_i - n_i + n_i \ln(n_i/b_i)). \tag{6.6}$$

In the rest of this section, this GOF test will be referred to as the reduced GOF, while the nominal one will be referred to as the global GOF.

6.1.3 Hypothesis testing

The likelihood ratio as a test statistic, introduced in Section 6.1.2, has further usefulness in hypothesis testing procedures that may be used to quantify the presence or absence of a signal in the observed data set. Within the context of this search, we are interested in two hypotheses that correspond to whether or not signal is present in the data, respectively referred to as H_s and H_b.

Measuring upper limits, which constitute a one-sided confidence interval on the cross section of a signal, involves finding a value for the signal strength r at which the H_s hypothesis can be excluded in favor of H_b. The computation of this r value relies on a profile likelihood ratio test statistic:

$$q_r \equiv -2\ln\frac{\mathcal{L}(n|r,\hat{\boldsymbol{\theta}}_r)}{\mathcal{L}(n|\hat{r},\hat{\boldsymbol{\theta}})} = -2\ln\frac{\mathcal{L}(n|r,\hat{\boldsymbol{\theta}}_r)}{\mathcal{L}_{\max}}, \qquad \text{with } 0 \leq \hat{r} < r. \tag{6.7}$$

In the likelihood ratio, the numerator corresponds to the profile likelihood, where for some pre-specified value of r, $\hat{\boldsymbol{\theta}}_r$ contains the values of the nuisance parameters that maximize the likelihood. The denominator corresponds to the global maximum likelihood, where $\hat{r}$ and $\hat{\boldsymbol{\theta}}$ are the typical maximum likelihood values for the parameter of interest and nuisance parameters. Larger values of q_r indicate a greater degree of incompatibility between the data and the signal hypothesis with strength r. Restricting $\hat{r} > 0$ forces this confidence interval to be one-sided.

Upper exclusion limits are calculated using the modified frequentist approach CL_s [138, 139] in conjunction with the test statistic q_r. The CL_s method makes use of two p-values associated with the B-only and S+B hypotheses:

$$p_r = P(q_r > q_r^{\mathrm{obs}} | H_s) = \int_{q_r^{\mathrm{obs}}}^{\infty} f(q_r|r)\ \mathrm{d}q_r \tag{6.8}$$

$$1 - p_b = P(q_0 > q_0^{\mathrm{obs}} | H_b) = \int_{q_0^{\mathrm{obs}}}^{\infty} f(q_0|0)\ \mathrm{d}q_0. \tag{6.9}$$

These p-values correspond to the probability to observe test-statistic values q_r that are less compatible than the observed value q_r^{obs} under the B-only (p_b) and S+B (p_r) hypotheses. The test statistic q_r is particularly useful because the asymptotic form of its pdf $f(q_r|r)$ can be described analytically [140]. This is used to drastically reduce the computation time relative to estimating $f(q_r|r)$ with toy MC experiments. With these two p-values, the quantity CL_s is then defined as:

$$\mathrm{CL}_s = \frac{p_r}{1 - p_b}. \tag{6.10}$$

The criterion for exclusion of a signal is interpreted from CL_s: a signal of strength r can be excluded at a confidence level (CL) of α for $\mathrm{CL}_s(r) < 1 - \alpha$. The upper limits that are reported as results in Section 6.2 are calculated at a 95% CL. The limit is obtained by performing the global fit and then scanning over values of r, calculating $f(q_r|r)$ with asymptotic formulae and then $\mathrm{CL}_s(r)$ until the value of r is found for which $\mathrm{CL}_s(r) = 0.05$.

An excess of signal could be quantified using the test statistic q_0, defined as:

$$q_0 = -2 \ln \frac{\mathcal{L}(n|0, \hat{\boldsymbol{\theta}}_0)}{\mathcal{L}(n|\hat{r}, \hat{\boldsymbol{\theta}})}, \qquad \text{with } \hat{r} > 0, \tag{6.11}$$

corresponding to the test statistic used in Eq. (6.7) with $r = 0$. The excess can then be quantified with the p-value measuring the probability to observe a q_0 as or more extreme than the observed value q_0^{obs} from data under the B-only hypothesis H_b:

$$p_0 = P(q_0 > q_0^{\text{obs}}|H_b). \tag{6.12}$$

This p_0 effectively measures the probability for the background to fluctuate upward to produce an excess at least as large as the one measured in the observed data. The p-value can be translated into a significance Z by using a one-sided Gaussian integral:

$$p_0 = \frac{1}{\sqrt{2\pi}} \int_Z^\infty e^{-x^2/2} \mathrm{d}x. \tag{6.13}$$

Typically, in high energy physics it is convention to require $Z \geq 5$ to claim a discovery.

6.1.4 Validation with control regions

The background modeling is first validated with the use of two CRs consisting of sets of events that are disjoint from the SR events but that are constructed with a similar set of selections. The initial sizes of the systematic uncertainties, detailed in Section 5.2, are estimated from the discrepancies between the data and the simulated backgrounds in the CRs. The validation procedure involves building the smooth background models in the CRs in the same manner as is done for the SR. The full background model is then fit to the 2D data in the CRs. By performing the full fit with all systematics in the CRs, we are able to evaluate the modeling methods while remaining blind to the critical region of data in the SR.

The CRs are constructed such that the amount of expected signal in the data compared to the SR is heavily reduced and such that the background contributions are larger than the levels in the SR. Each of the two CRs is enriched in a subset of the four background components, and so the tests executed in each CR target the validation of the modeling of these components. The first CR targets the background components that are largely born from $\mathrm{t\bar{t}}$ production — the lost t/W, m_{W}, and m_{t} backgrounds — and is named the Top CR. The second CR targets the q/g background and is named the Non-top CR.

The background-enriched, signal-depleted CRs are constructed primarily by considering the b tagging selections in the SR. The Top CR is constructed from the SR selection by inverting the b jet veto — requiring that there be a b-tagged AK4 jet that is $\Delta R > 1.2$ away from the $\mathrm{b\bar{b}}$ jet for a $\mathrm{t\bar{t}}$-enriched sample. Furthermore, relaxing the p_{T}/m selection in the 1ℓ channel and requiring $\Delta R_{\ell\ell} > 0.4$ instead of $\Delta R_{\ell\ell} < 1.0$ increases both the purity and statistical power of the Top CR samples. The Non-top CR is constructed by selecting a region of the $D_{\mathrm{Z/H\to b\bar{b}}}$ distribution that is enriched in light quark or gluon jets like those typically seen in events from the q/g background. Rather than $D_{\mathrm{Z/H\to b\bar{b}}} > 0.8$ as in the SR, we require $0.01 < D_{\mathrm{Z/H\to b\bar{b}}} < 0.04$. The region $D_{\mathrm{Z/H\to b\bar{b}}} < 0.01$ is not used because of heavy mismodeling of the $\mathrm{b\bar{b}}$ tagging in this region. As is the case for the q/g background in the SR, this CR is enriched in events from $\mathrm{V+jets}$ and QCD multijet processes.

The background composition of the top and non-top CRs is detailed in Table 6.1. The composition of the non-top CR is at least 89% composed of q/g background events, while the top CR is composed of backgrounds with top quarks by over 85%. The simulated distributions of the top (non-top) CR before the modeling process are shown in Fig. 6.1 (6.2). One can see from these figures that the top CR is more similar in composition to the SR.

The fitted distributions can be visualized relative to the CR data and the pre-fit model in Figs. 6.3, 6.4, 6.5, and 6.6, which show projections into each mass dimension separately for the 1ℓ and 2ℓ channels. In the non-top CR fit to data, the m_{t} background is not included in the fit, since its contributions are negligible and technical difficulties were encountered trying to include it. The non-top CR is only used to validate the q/g background modeling anyway. Furthermore, since the non-top CR is selected from low $D_{\mathrm{Z/H\to b\bar{b}}}$ values, it does not have any $\mathrm{b\bar{b}}$ jet tagging categories and thus has only 6 (instead of 12) categories.

The fits behave well by eye. The behavior of the nuisance parameters in a fit of the B-only model to data is demonstrated with their pulls, shown in Figs. 6.7 and 6.8 for the top CR and in Figs. 6.9 and 6.10 for the non-top CR. The systematic uncertainties are universally constrained by the fit, and the fitted values are almost all within the initial uncertainty estimates. The pulls indicate that there is not much tension in the model. Results from the GOF tests are shown in Fig. 6.11. In both CRs, the observed value of t_{GOF} is within the central 68% interval of the expected distribution, indicating that the models can sufficiently describe the data.

Overall, the statistical tests of the CR fits indicate that the parameterization of the background systematic uncertainties and their initial sizes can sufficiently model the data. This lends confidence in applying these methods to the background modeling in the SR.

6.1.5 Fit results in the search region

The results of fits of the B-only and S+B models to the data in the SR categories are presented in this section. Projections into $m_{\mathrm{b\bar{b}}}$ and m_{HH} of the full 2D fit that compare the post-fit models, the data, and some simulated signal hypotheses are shown in Figs. 6.12 and 6.13. The total event yield in each SR category is described in Table 6.2, with numbers provided for the observed data, the expected SM background contribution pre- and post-fit, and the associated post-fit uncertainty on the yield.

The pulls of the nuisance parameters from the 2D fit are shown in Fig. 6.14, and the nuisances appear to be constrained and behave well within the initial uncertainty estimates. There is not much tension in the SR fit, and the similar behavior of the B-only and S+B fits is an indicator of the absence of signal in the data. The impact of each nuisance parameter on the fitted signal strength is also studied. The test is performed by fixing each parameter, one at a time, to its post-fit $\pm 1\sigma$ values and then profiling all others. The change in the fitted signal strength $\hat{r}$ from the full fit to the profiled fit is analyzed. Tests were performed using the observed data and using pseudodata with various fixed amounts of signal. An

Table 6.1: Background compositions of the top and non-top CRs. For each CR, the fractional contribution of each background process is given in each of the 1ℓ and 2ℓ channels. These numbers are given both for all considered SM processes and for the four background components in the modeling. The number in parentheses below each channel and CR in the top row of table cells is the expected total event yield corresponding to an integrated luminosity of 137.6 fb^{-1}.

	Top CR		Non-top CR	
Background process	1ℓ channel	2ℓ channel	1ℓ channel	2ℓ channel
Events =	23546.3	367.7	24493.0	713.5
$\mathrm{t\bar{t}}$ 0ℓ	0.8%	0.0%	0.2%	0.0%
$\mathrm{t\bar{t}}$ 1ℓ	83.1%	40.4%	11.0%	2.9%
$\mathrm{t\bar{t}}$ 2ℓ	5.6%	38.5%	0.4%	13.1%
single top	2.7%	8.8%	1.3%	2.1%
$\mathrm{W}+\mathrm{jets}$	0.9%	0.2%	50.8%	3.9%
$\mathrm{Z}/\gamma^*+\mathrm{jets}$	0.2%	4.4%	7.0%	66.2%
QCD multijet	5.5%	1.7%	27.4%	0.0%
diboson	0.1%	0.6%	1.6%	11.4%
$\mathrm{t\bar{t}V}$	0.7%	2.9%	0.1%	0.2%
$\mathrm{t\bar{t}H}$	0.4%	2.3%	0.0%	0.1%
q/g	10.3%	37.5%	89.0%	93.7%
lost t/W	54.6%	44.4%	3.1%	2.8%
m_{W}	5.8%	4.7%	7.2%	3.1%
m_{t}	29.2%	13.4%	0.7%	0.4%

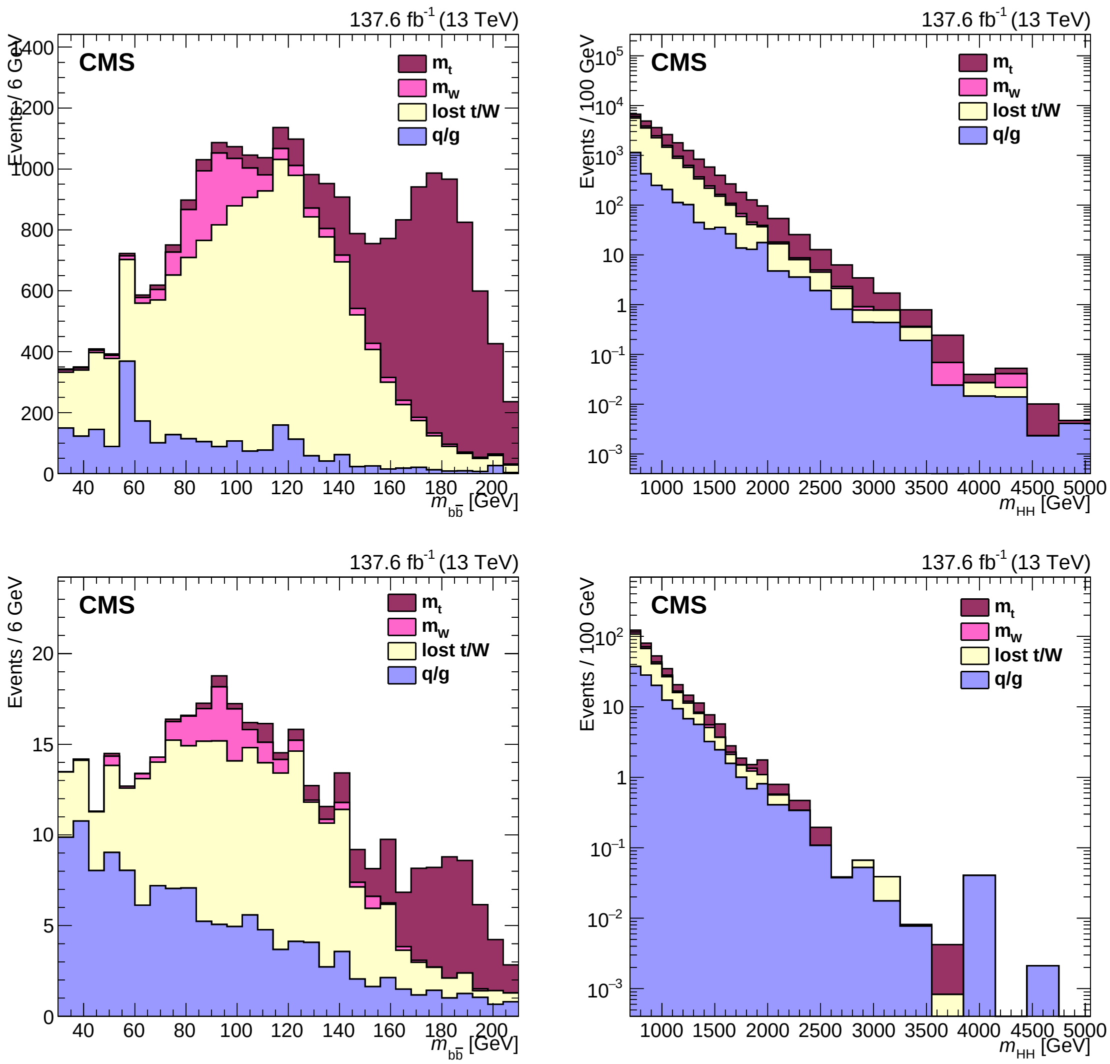

Figure 6.1: Top CR simulated distributions of $m_{\mathrm{b\bar{b}}}$ (left) and m_{HH} (right) in the 1ℓ (top) and 2ℓ (bottom) channel.

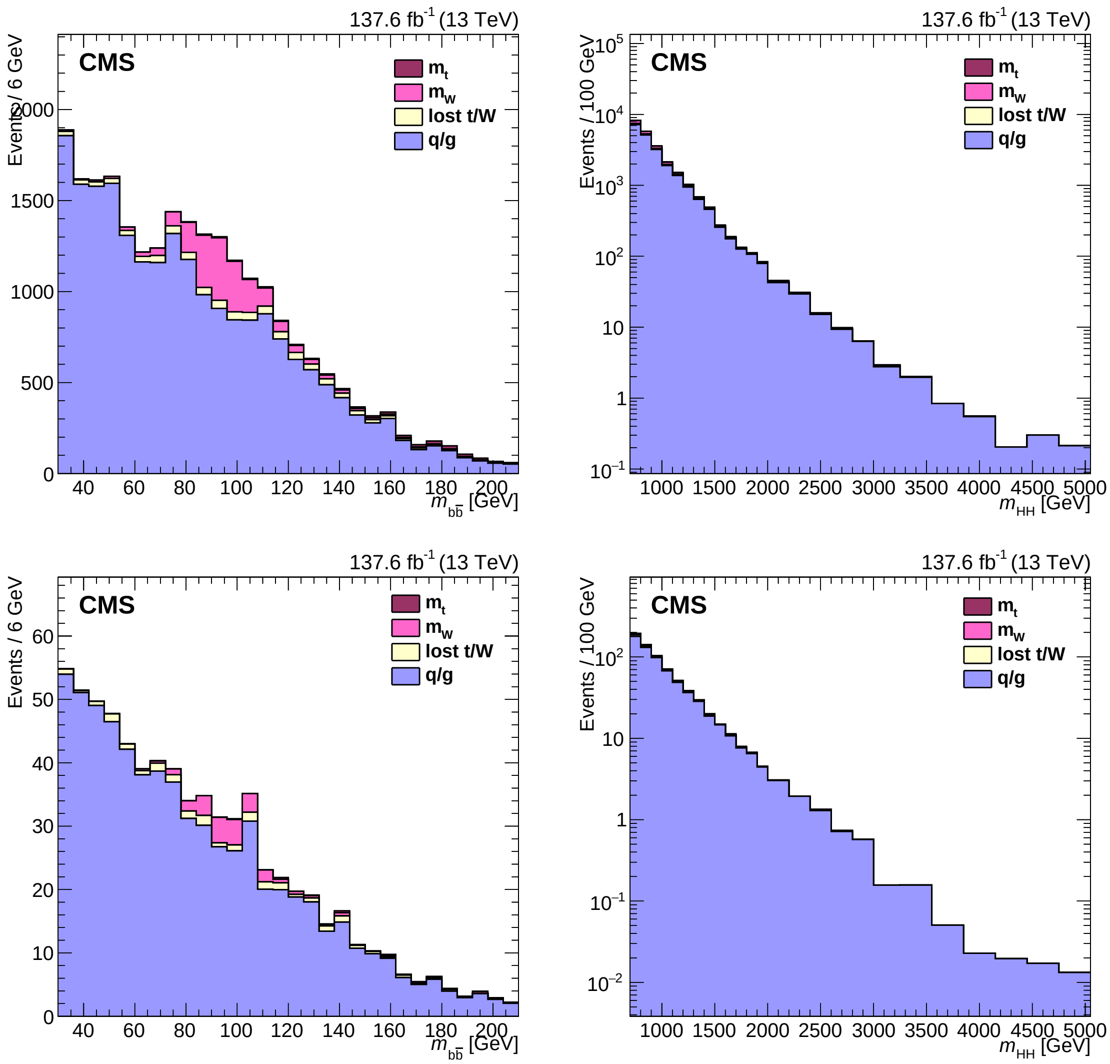

Figure 6.2: Non-top CR simulated distributions of $m_{\mathrm{b\bar{b}}}$ (left) and m_{HH} (right) in the 1ℓ (top) and 2ℓ (bottom) channel.

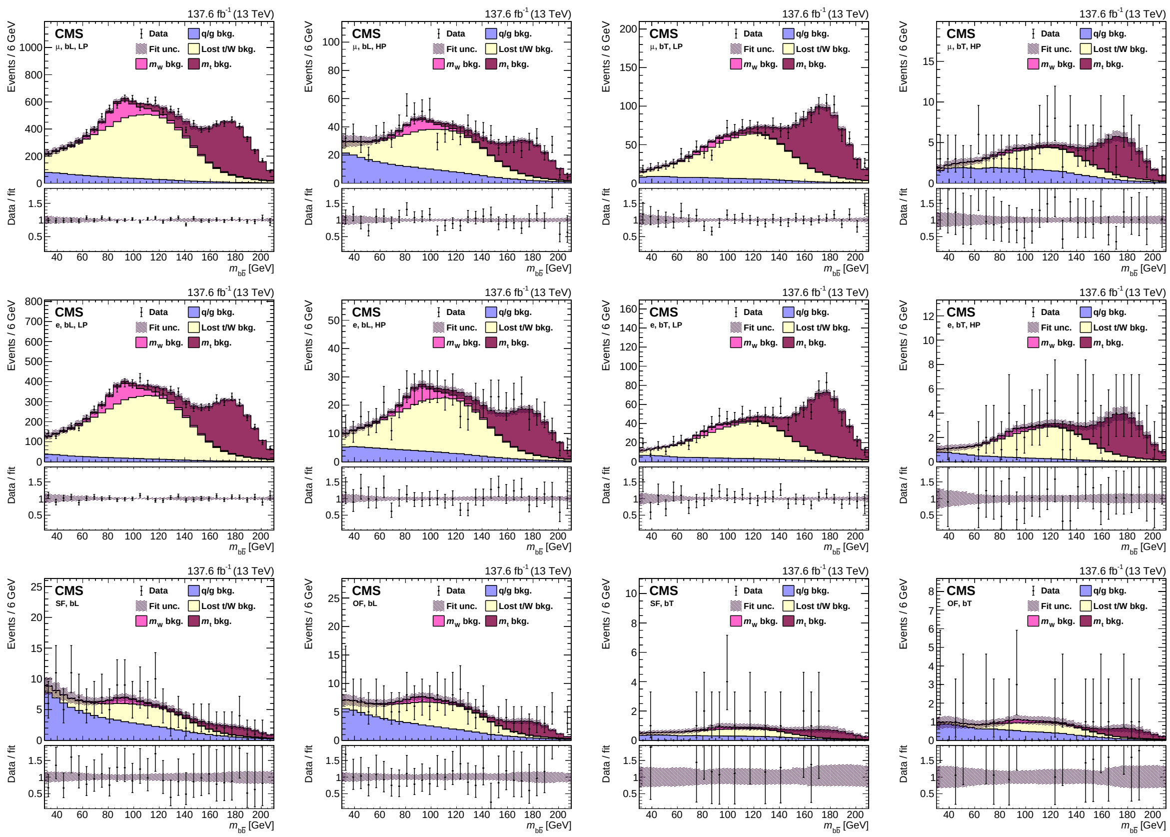

Figure 6.3: Post-fit $m_{b\bar{b}}$ distributions of the top CR, with the pre-fit model and data overlayed. The category names for each plot are indicated in the upper-left corner. The filled histograms are the post-fit model, and the hashed band depicts the fit uncertainty.

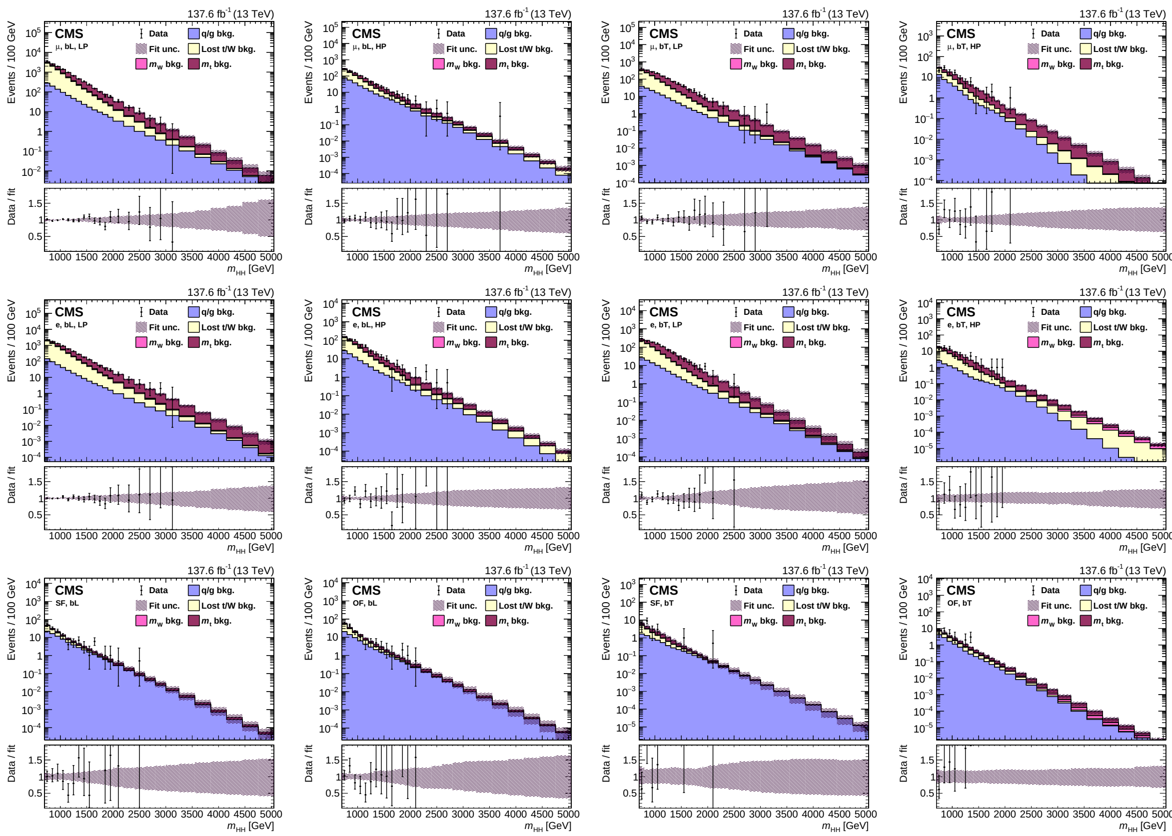

Figure 6.4: Post-fit m_{HH} distributions of the top CR, with the pre-fit model and data overlayed. The category names for each plot are indicated in the upper-left corner. The filled histograms are the post-fit model, and the hashed band depicts the fit uncertainty.

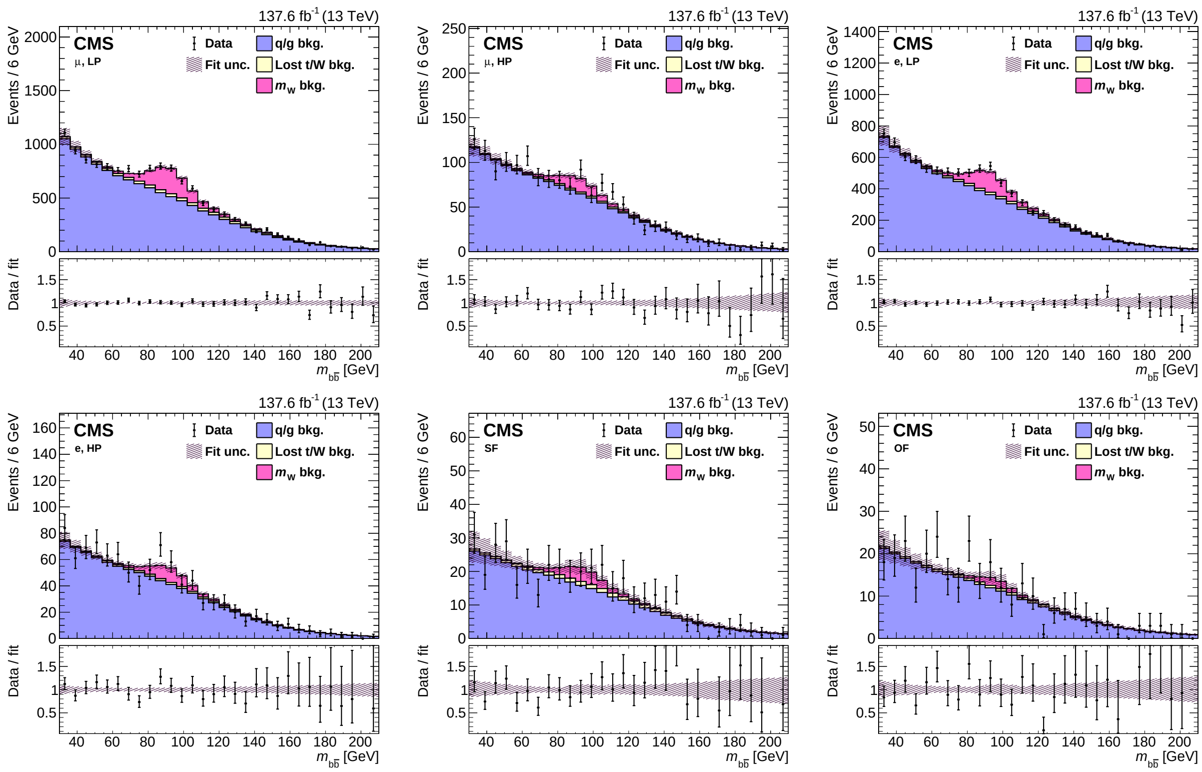

Figure 6.5: Post-fit $m_{b\bar{b}}$ distributions of the non-top CR, with the pre-fit model and data overlayed. The category names for each plot are indicated in the upper-left corner. The filled histograms are the post-fit model, and the hashed band depicts the fit uncertainty.

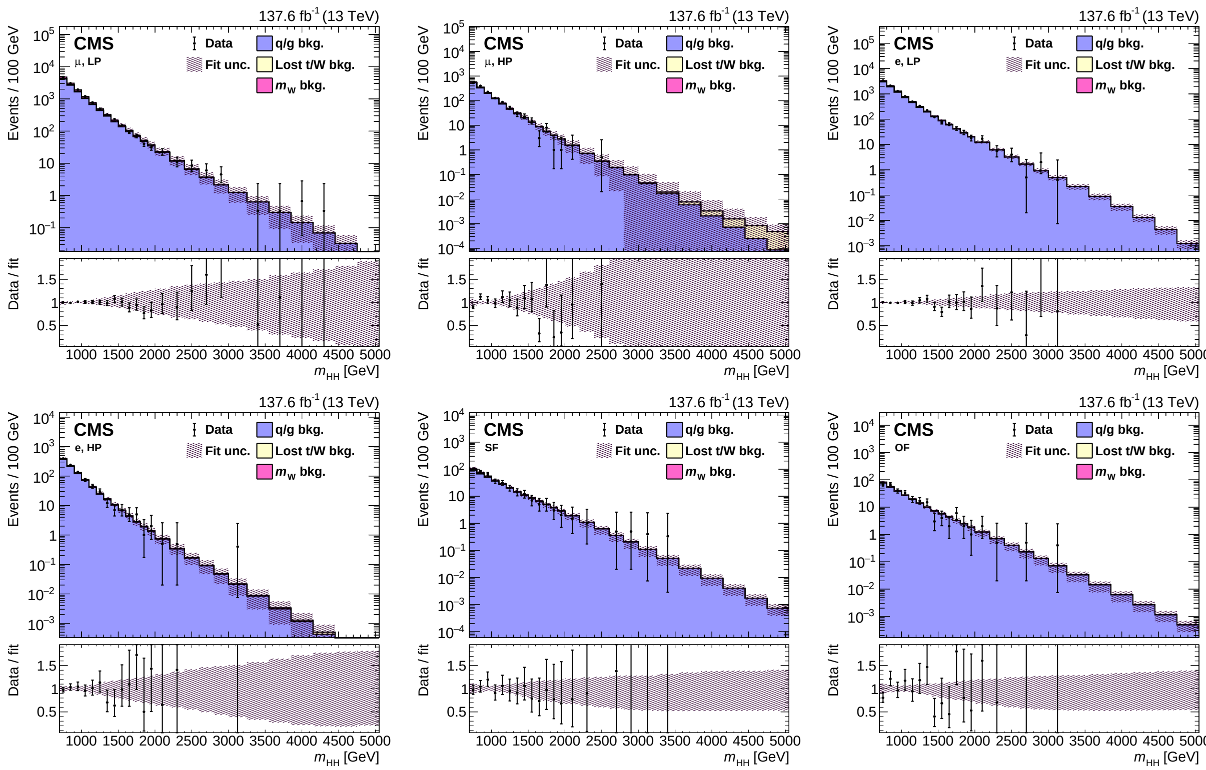

Figure 6.6: Post-fit m_{HH} distributions of the non-top CR, with the pre-fit model and data overlayed. The category names for each plot are indicated in the upper-left corner. The filled histograms are the post-fit model, and the hashed band depicts the fit uncertainty.

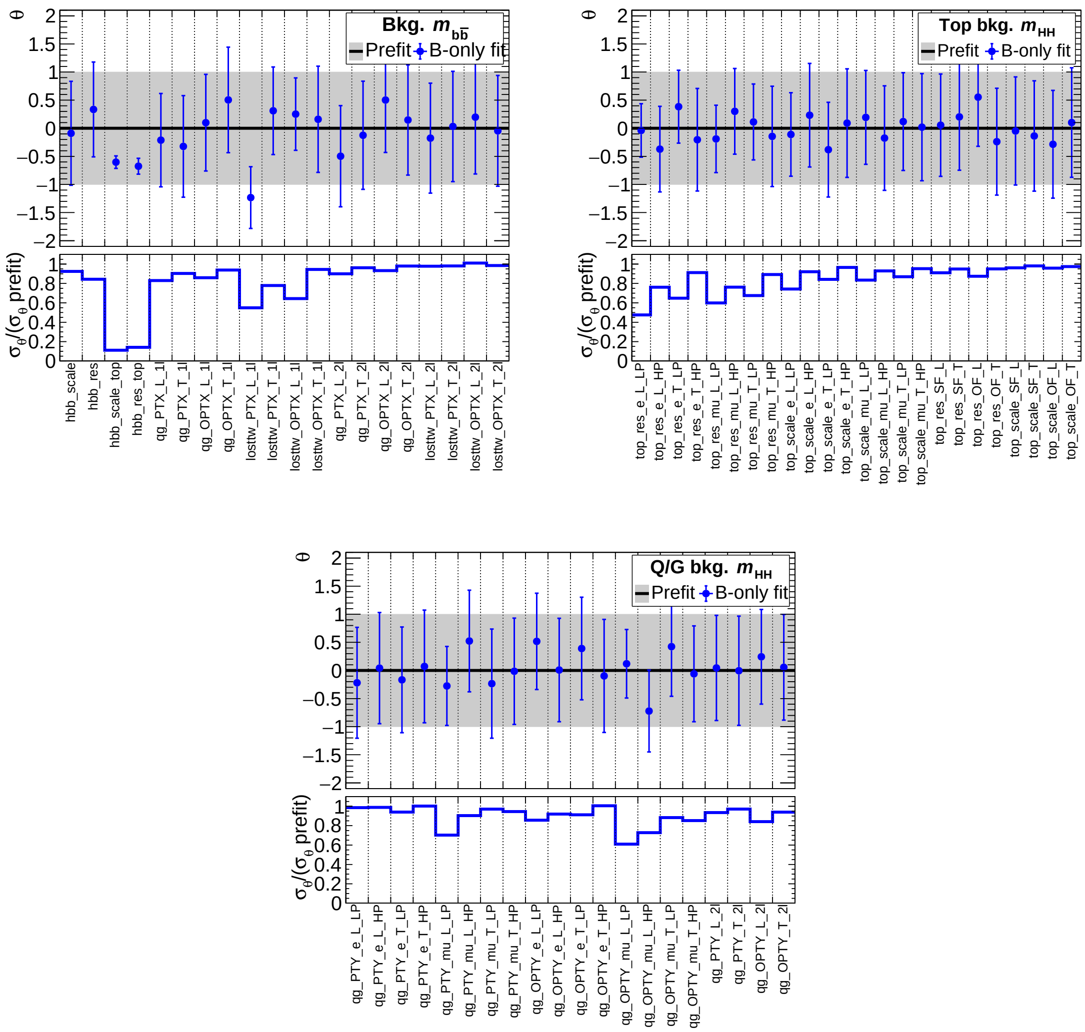

Figure 6.7: The pulls of the nuisance parameters for the background shape from the fit in the top CR, and the ratio of post-fit and pre-fit uncertainties in the bottom panels. The y-axis is normalized such that all pre-fit uncertainties have a size of 1.0. The systematics labelled "PTX" are used to morph the background templates $\propto m_{\mathrm{b\bar{b}}}$, while "OPTX" are those for $1/m_{\mathrm{b\bar{b}}}$ morphing. The "PTY" and "OPTY" systematics do the same but for the m_{HH} dimension. The resonant SD jet mass systematics are those labeled "hbb_scale" and "hbb_res."

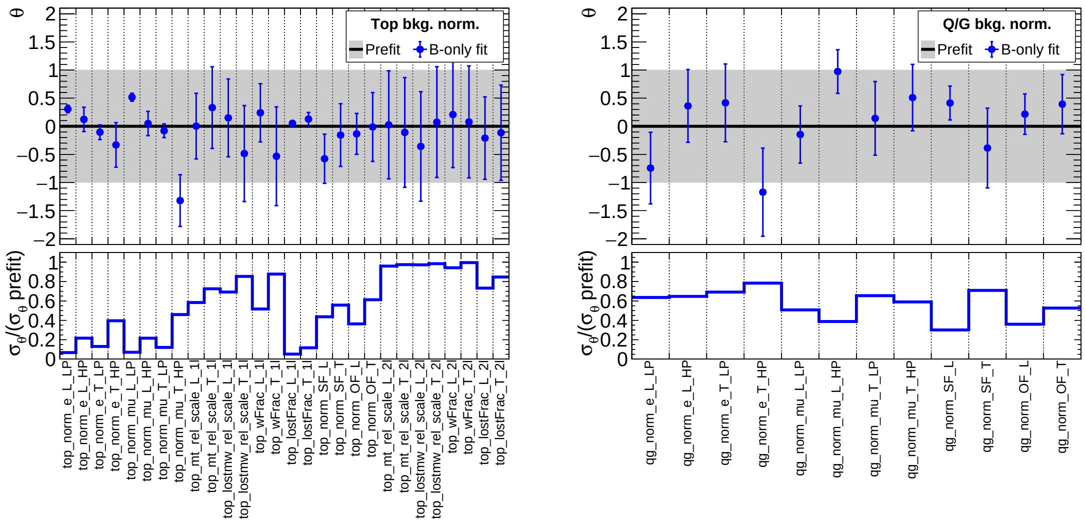

Figure 6.8: The pulls of the nuisance parameters for the background normalization from the fit in the top CR, and the ratio of post-fit and pre-fit uncertainties in the bottom panels. The y-axis is normalized such that all pre-fit uncertainties have a size of 1.0.

example of the 10 most impactful systematics in a pseudodata set with an injected spin-0 1.0 TeV signal is shown in Fig. 6.15. In general, in a data set that contains signal, the $b\bar{b}$ jet tagging uncertainties contribute most to the total uncertainty on the signal strength. At low m_{X}, where the background levels are the largest, the top background normalization uncertainties in the most sensitive categories are the next most impactful, as changes in these low background levels induce larger changes in the signal sensitivity. At high m_{X} where there is much less background, the τ_2/τ_1 and then other signal uncertainties dominate. However, these systematic effects do not contribute as much to the total uncertainty on $\hat{r}$ as the statistical uncertainty does.

The agreement between the data and the post-fit model is also quantified with the GOF tests introduced in Section 6.1.2. These are shown in Fig. 6.16 for the reduced GOF using the B-only model and the global GOF using the S+B model. In all cases, the observed value of t_{GOF} in the data is consistent with the expected distribution from toy experiments.

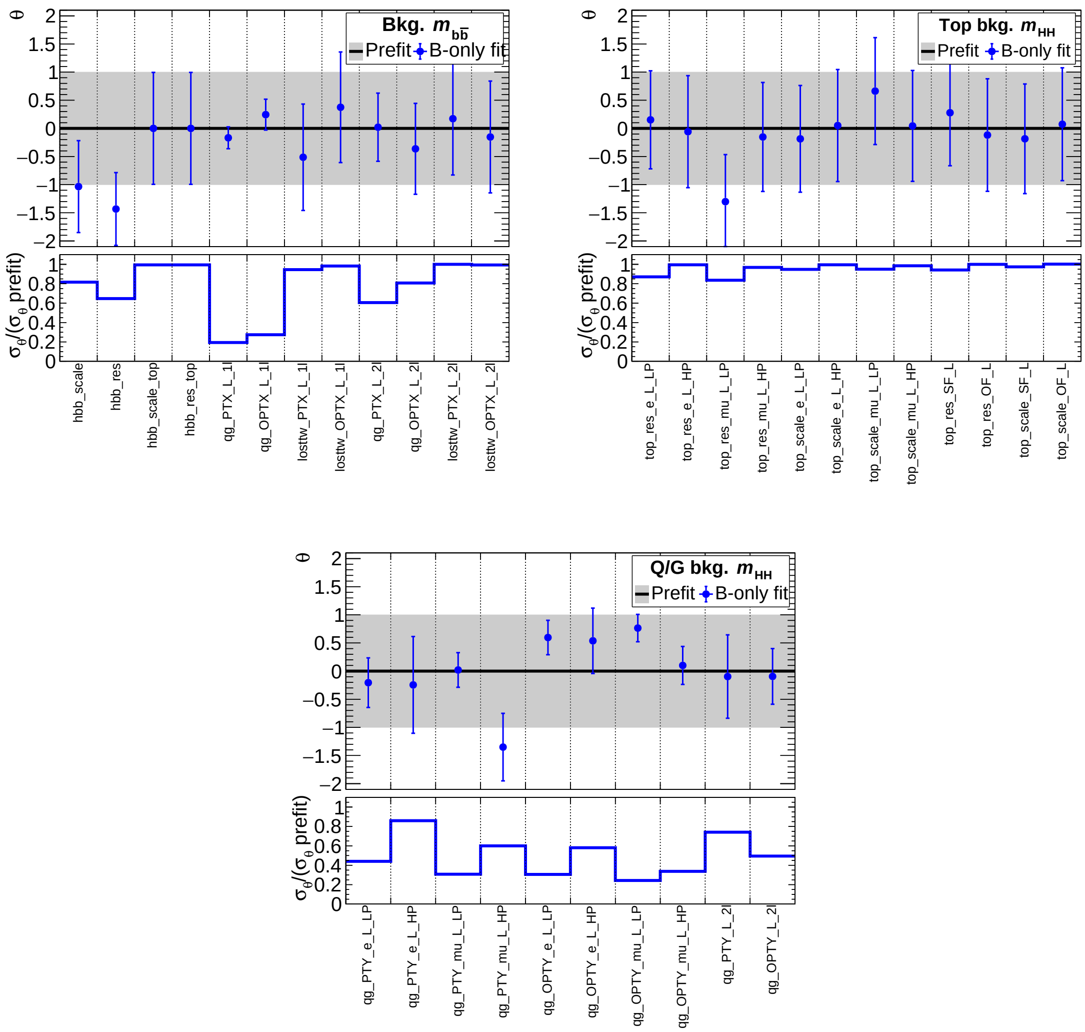

Figure 6.9: The pulls of the nuisance parameters for the background shape from the fit in the non-top CR, and the ratio of post-fit and pre-fit uncertainties in the bottom panels. The y-axis is normalized such that all pre-fit uncertainties have a size of 1.0. The systematics labelled "PTX" are used to morph the background templates as $\propto m_{b\bar{b}}$, while "OPTX" are those for $1/m_{b\bar{b}}$ morphing. The "PTY" and "OPTY" systematics do the same but for the m_{HH} dimension. The resonant SD jet mass systematics are those labeled "hbb_scale" and "hbb_res."

Table 6.2: Event yields broken down by SR category. For each category, shown are the event yields observed in data, expected before and after a fit of the B-only model, and the corresponding relative uncertainty.

Search category	Observed	Expected (pre-fit)	Expected (post-fit)	Post-fit uncertainty
μ bL LP	4542	4362.2	4540.9	1.5%
μ bL HP	417	402.4	416.1	4.8%
μ bT LP	657	731.8	658.5	4.2%
μ bT HP	56	67.0	57.3	10.0%
e bL LP	2945	2973.7	2945.4	1.9%
e bL HP	248	246.1	247.7	5.7%
e bT LP	423	443.0	423.9	4.2%
e bT HP	37	41.0	37.7	14.6%
SF bL	59	70.2	59.6	14.2%
OF bL	50	61.1	50.8	13.5%
SF bT	6	11.3	7.9	31.6%
OF bT	6	11.6	8.1	25.8%

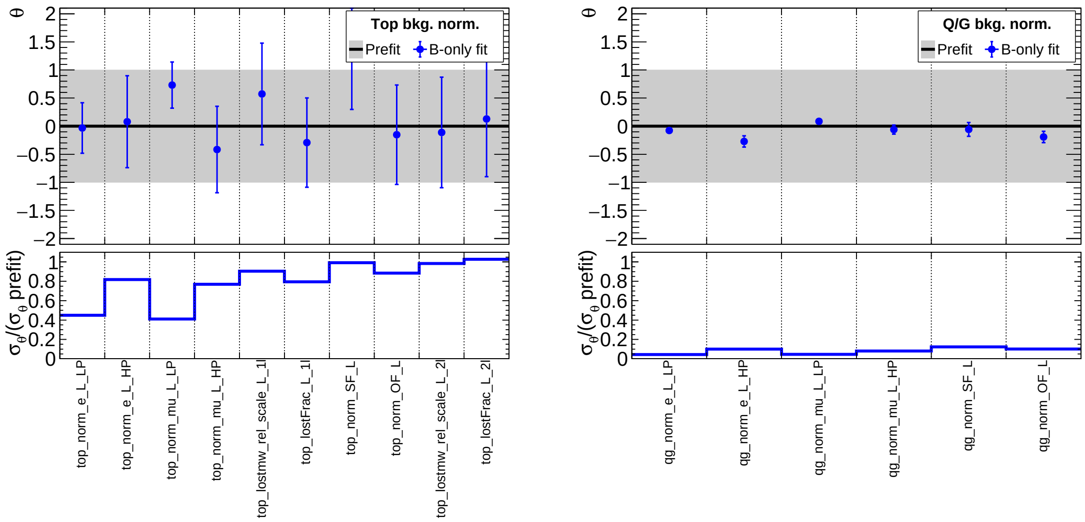

Figure 6.10: The pulls of the nuisance parameters for the background normalization from the fit in the non-top CR, and the ratio of post-fit and pre-fit uncertainties in the bottom panels. The y-axis is normalized such that all pre-fit uncertainties have a size of 1.0.

6.1.6 Signal bias testing

In the event of a possible discovery, it is important to check for and quantify the amount of potential bias in the signal modeling. This is accomplished by performing a series of tests in which signal is injected into and extracted from a data set. To perform the test, the B-only model is first fit to the data. The background nuisance parameters are then set to their post-fit values, and the signal strength is set to some fixed value in order to generate an ensemble of toy data sets from the full S+B model. Then for each toy experiment, a maximum likelihood fit of the pre-fit S+B model to the toy data set is performed, with the signal strength floating. For each toy experiment, the bias in the signal is defined to be:

$$r_{\text{bias}} = \frac{r_{\text{fit}} - r_{\text{inj}}}{\sigma_r} \tag{6.14}$$

The extracted signal strength from the fit and its associated uncertainty are r_{fit} and σ_r, respectively. The fixed value of the injected signal strength is r_{inj}. For the discrete

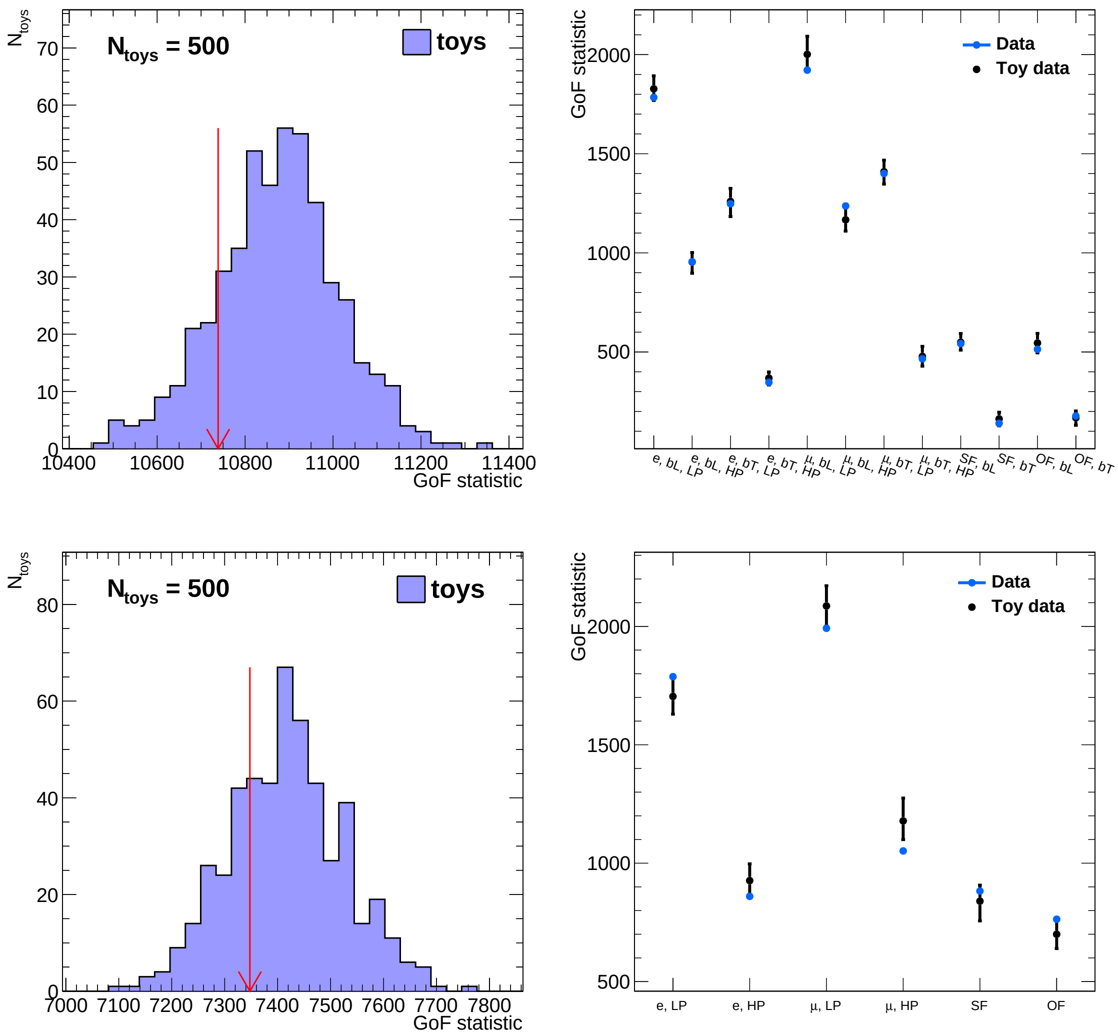

Figure 6.11: Saturated likelihood GOF tests for the top CR (top row) and non-top CR (bottom row). The left column shows the global GOF test, with the histogram showing the distribution from toy experiments and the red arrow indicating the observed t_{GOF}. The right column shows the reduced GOF test projected into each CR category. The blue points indicate the data t_{GOF}, while the black points and bands demonstrate the median and central 68% interval of the toy distribution.

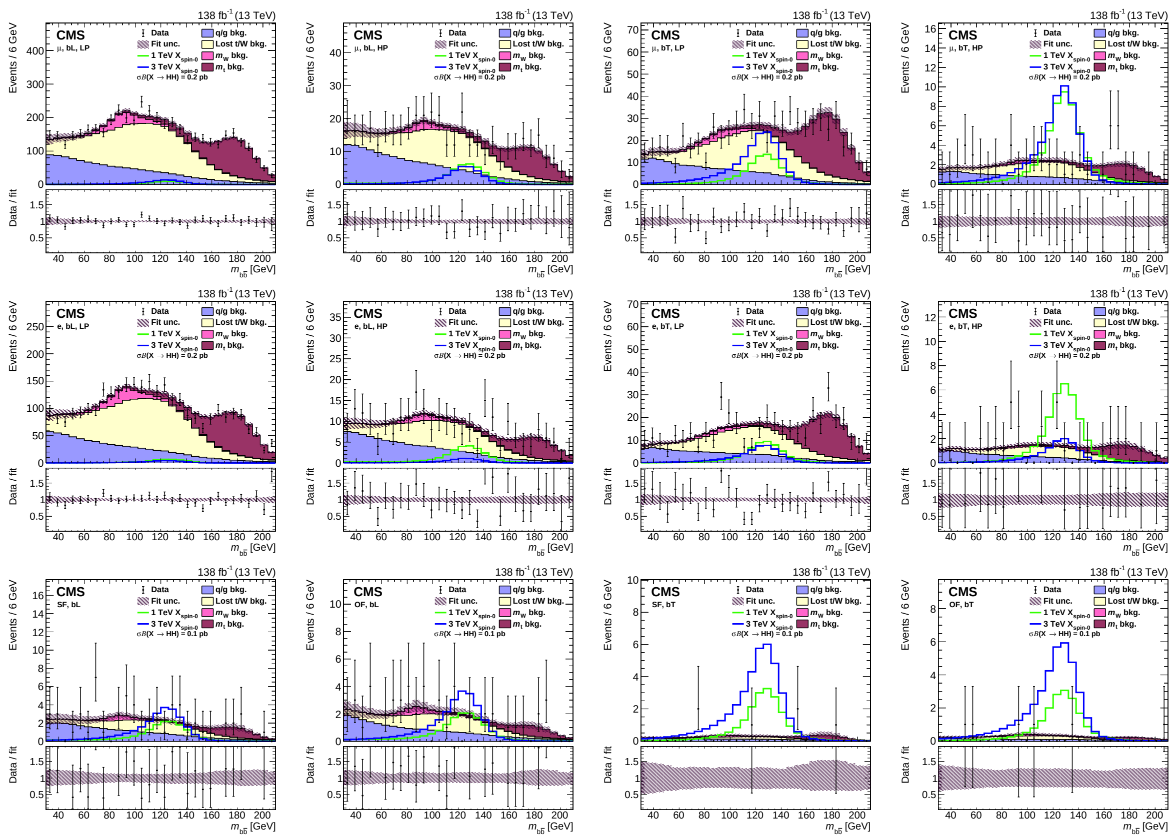

Figure 6.12: The B-only 2D fit result compared to data projected into $m_{\mathrm{b\bar{b}}}$ for both the 1ℓ and 2ℓ channels. The label for each search category is in the upper-left of each plot. The fit result is the filled histogram, with the different colors indicating different background categories. The background shape uncertainty from the fit is shown as the hatched band. Example spin-0 signal distributions for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV are shown as solid lines, with $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH})$ set to 0.2 and 0.1 pb for the 1ℓ and 2ℓ channels, respectively. The bottom panels show the ratio of the data to the fit result.

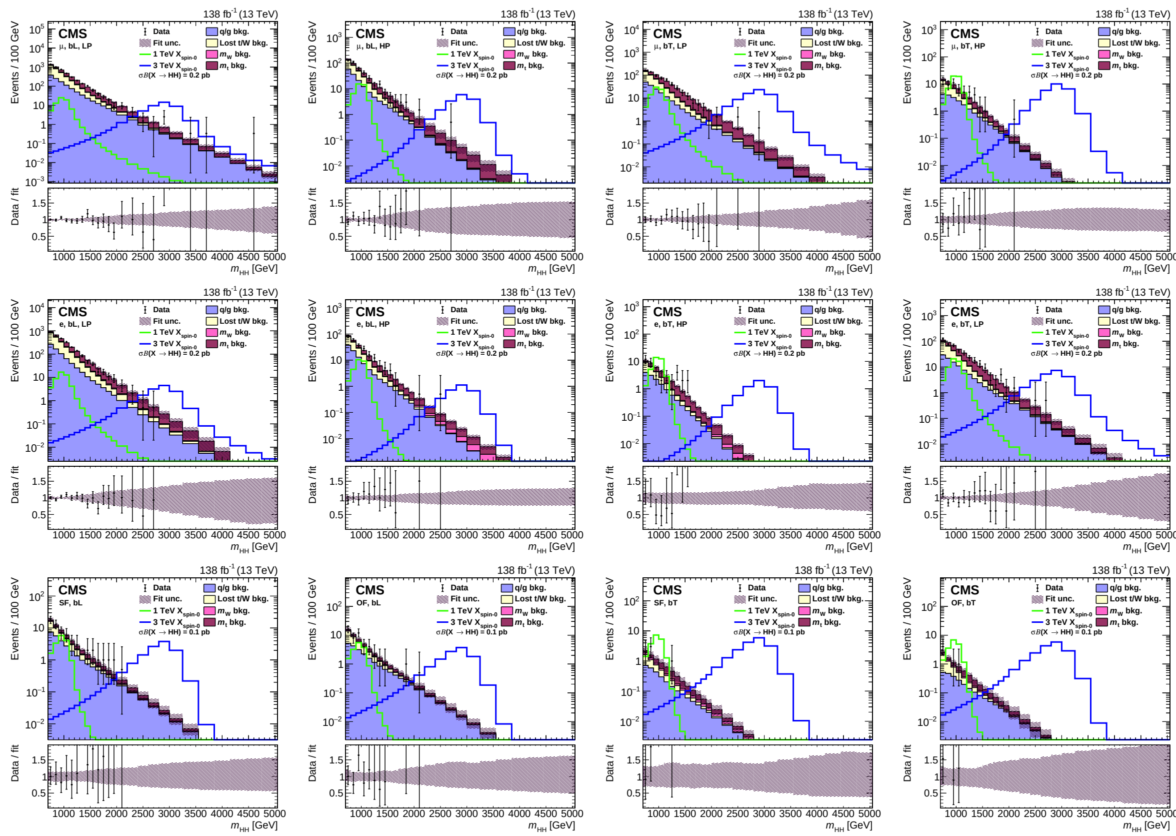

Figure 6.13: The B-only 2D fit result compared to data projected into m_{HH} for both the 1ℓ and 2ℓ channels. The label for each search category is in the upper-left of each plot. The fit result is the filled histogram, with the different colors indicating different background categories. The background shape uncertainty from the fit is shown as the hatched band. Example spin-0 signal distributions for $m_{\mathrm{X}} = 1.0$ and 3.0 TeV are shown as solid lines, with $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH})$ set to 0.2 and 0.1 pb for the 1ℓ and 2ℓ channels, respectively. The bottom panels of each plot show the ratio of the data to the fit result.

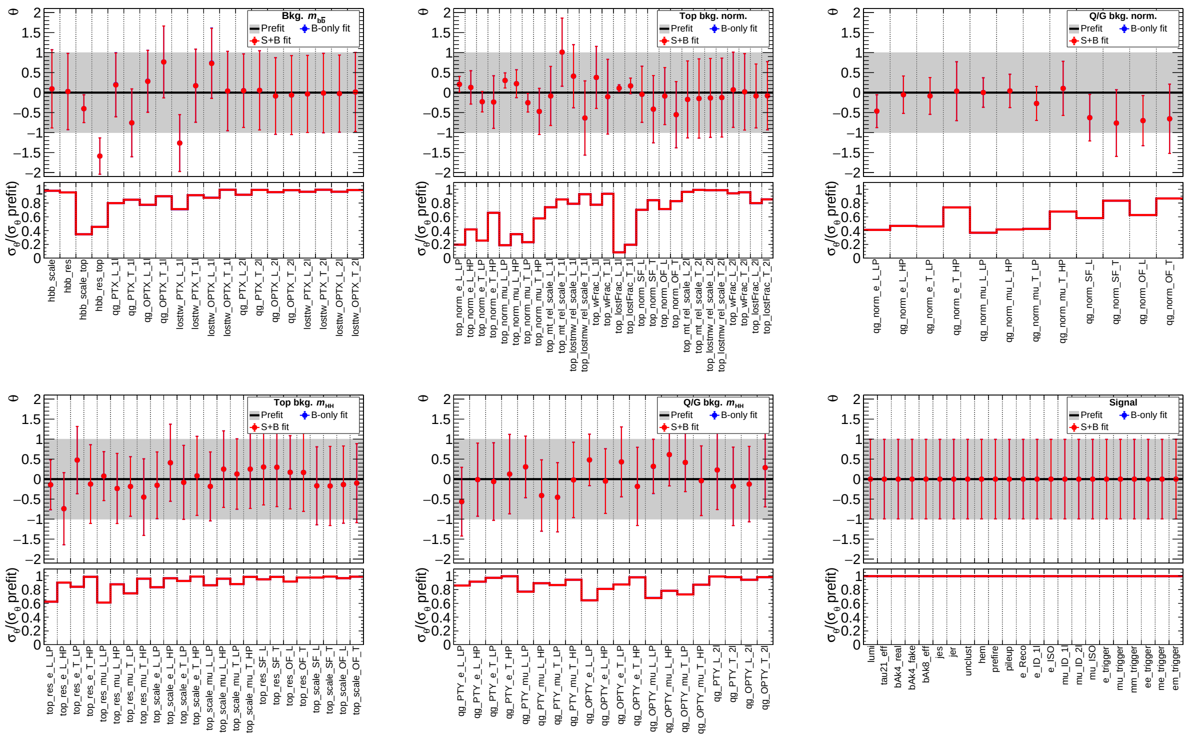

Figure 6.14: The pulls of the nuisance parameters from the B-only and S+B fits in the SR. The bottom panels show the ratio of post-fit and pre-fit uncertainties. The y-axis is normalized such that all pre-fit uncertainties have a size of 1.0. The systematics labelled "PTX" are used to morph the background templates as $\propto m_{\mathrm{b\bar{b}}}$, while "OPTX" are those for $1/m_{\mathrm{b\bar{b}}}$ morphing. The "PTY" and "OPTY" systematics do the same but for the m_{HH} dimension. The resonant SD jet mass systematics are those labeled "hbb_scale" and "hbb_-res."

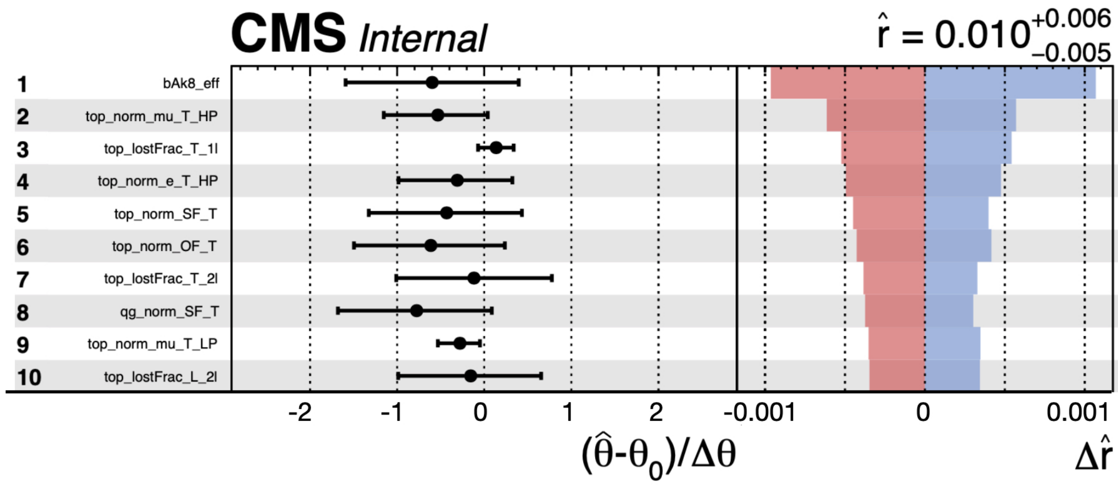

Figure 6.15: Impacts of the systematic uncertainties on the signal strength. The 10 uncertainties that induce the largest change $\Delta\hat{r}$ are shown for a test in which a 1.0 TeV signal is injected with $r_{\mathrm{inj}} = 0.0094$ pb into a toy data. In this particular test, the most impactful uncertainty is from the $\mathrm{b\bar{b}}$ jet tagging efficiency, followed by top background normalization uncertainties.

m_{X} in Table 4.7 and both spin-0 and spin-2 signals, bias tests are performed with injected signal strengths equal to the expected limits r_{excl} and for some integer multiples of these strengths, $2r_{\mathrm{excl}}$ and $5r_{\mathrm{excl}}$. We evaluate the existence of bias by fitting the r_{bias} distributions, which in general are approximately Gaussian, and by analyzing the properties of the raw distributions themselves. An example of one such r_{bias} distribution and its Gaussian fit is shown in Fig. 6.17. Plots that demonstrate the results of the full suite of tests for all m_{X} are shown in Fig. 6.18 and 6.19. Overall, the extracted signal strength and its uncertainty is consistent with the injected signal strength, no evidence of bias is observed.

6.2 Results

When the full combined 2D fit is performed in the SR data, no significant excess is observed. Thus, model-independent upper limits on $\sigma\mathcal{B}(\mathrm{X} \to \mathrm{HH})$ are computed at 95% CL and

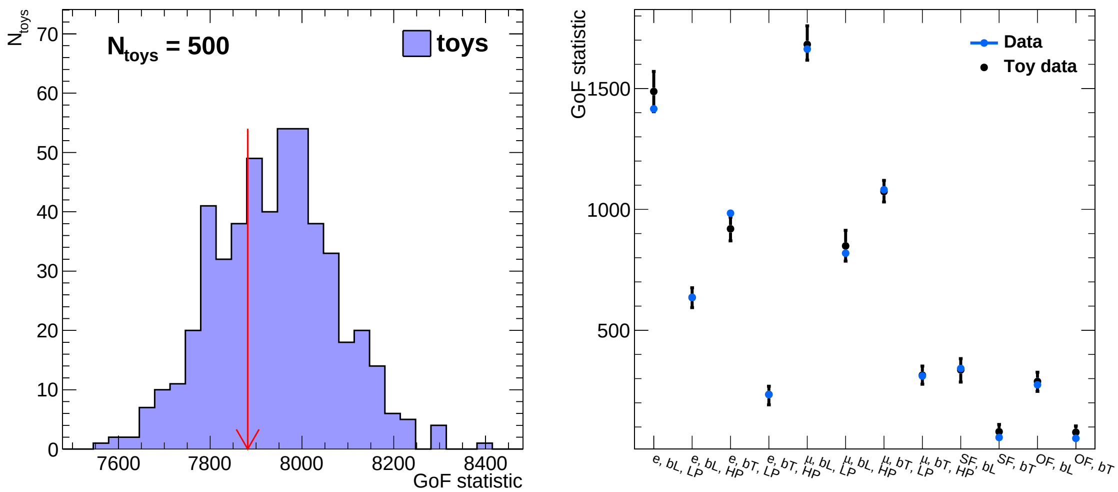

Figure 6.16: Saturated likelihood GOF tests for the SR. The left figure shows the global GOF test, with the histogram showing the distribution from toy experiments and the red arrow indicating the observed t_{GOF}. The right figure shows the reduced GOF test projected into each SR category. The blue points indicate the data t_{GOF}, while the black points and bands demonstrate the median and central 68% interval of the toy distribution.

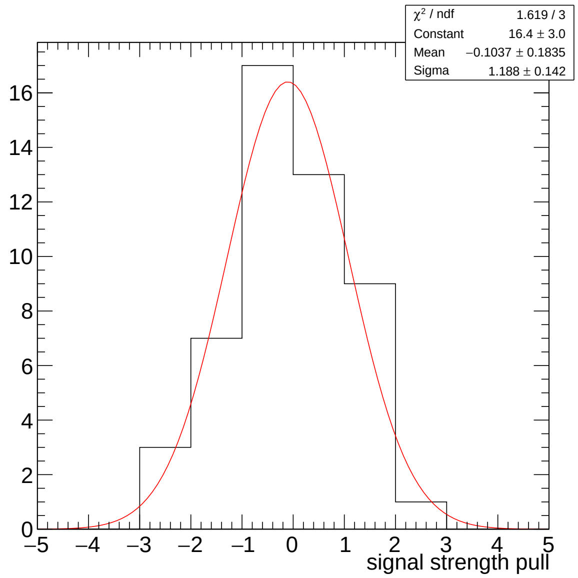

Figure 6.17: The r_{bias} distribution and its Gaussian fit for a $m_{\text{X}} = 1.0$ TeV spin-0 signal with $r_{\text{inj}} = 5r_{\text{excl}} = 46.5$ fb.

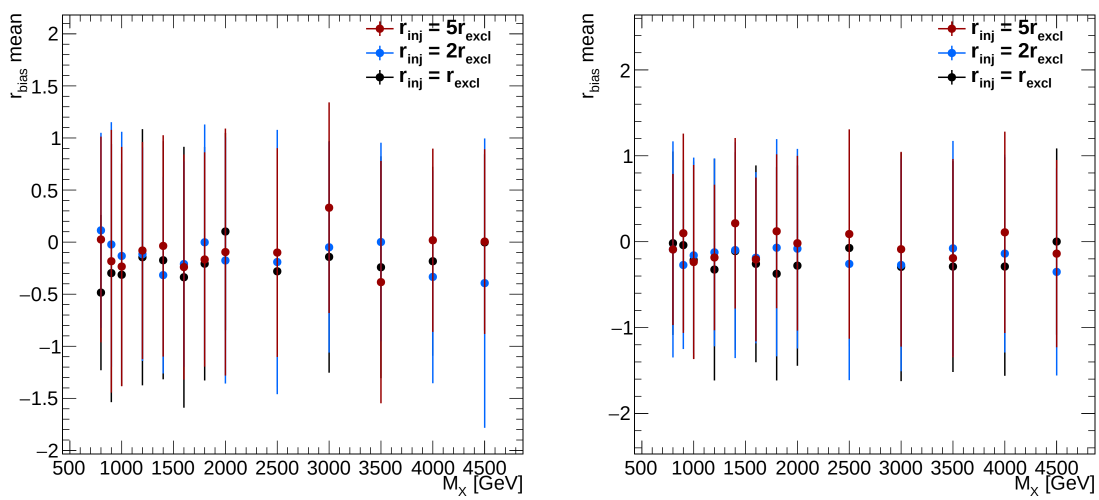

Figure 6.18: Bias test results for the full range of m_{X}. Left: spin-0. Right: spin-2. The plots show the mean (points) and standard deviation (error band) of the raw r_{bias} distributions for each m_{X} and r_{inj}.

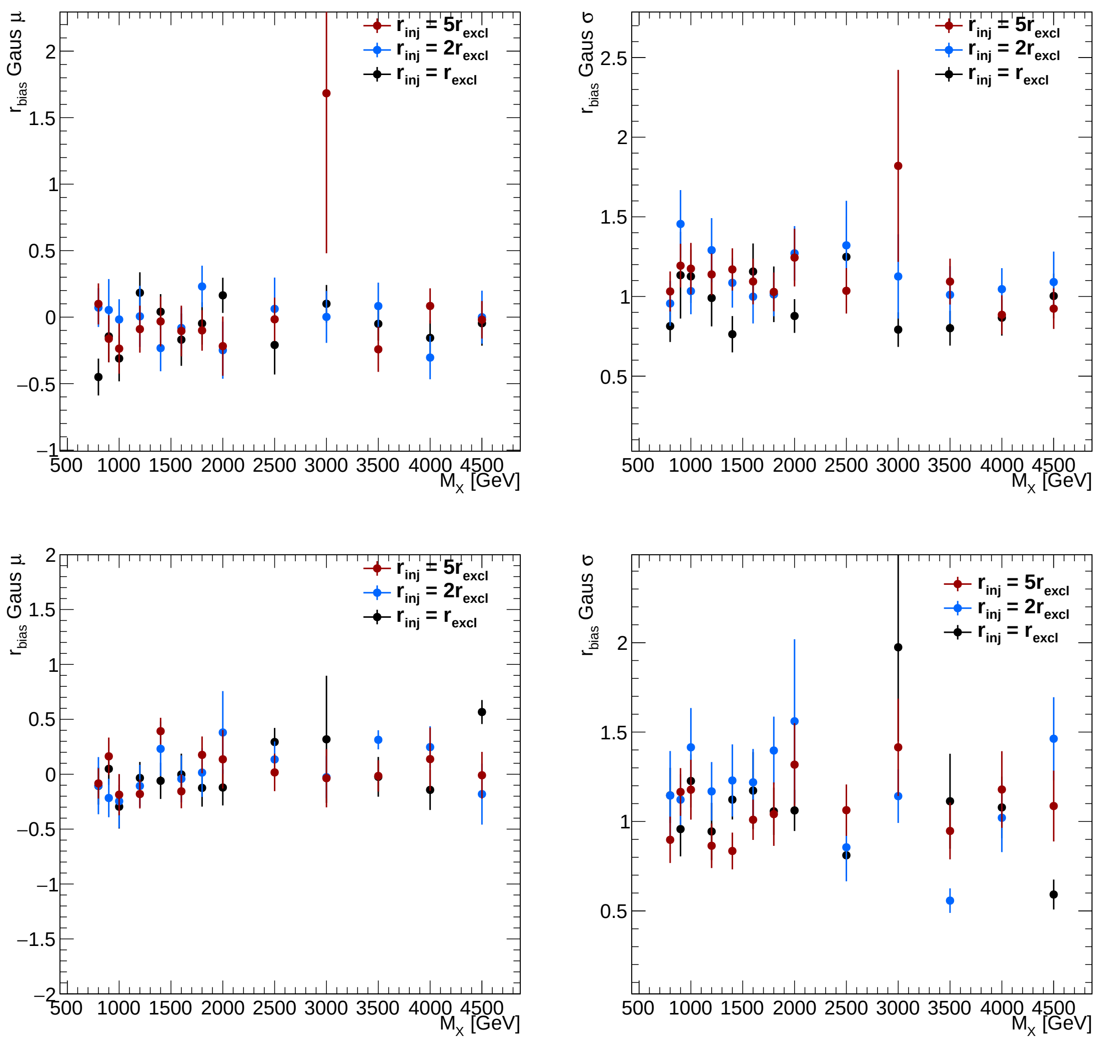

Figure 6.19: Bias test results for the full range of m_{X}. Left: the fitted Gaussian mean μ of the r_{bias} distribution, with the associated error bars indicating the uncertainty on the fitted μ. Right: the fitted Gaussian width σ of the r_{bias} distribution, with the fit uncertainty on the width indicated with the error bars. The spin-0 (spin-2) results are shown in the top (bottom) row.

reported as results. The upper limits are shown in Fig. 6.20 for both signal spin hypotheses as a function of the signal mass m_{X}. For each, the observed limits in data are shown along with the median expected limit and the 1σ and 2σ uncertainty bands. The observed limits are consistent with the expectations. The most tension occurs in $1.0 \lesssim m_{\mathrm{X}} \lesssim 1.2$ TeV where there is a deviation of about 2σ from the expectation. This is downward statistical fluctuation that arises from a deficit of data in some of the most sensitive search categories.

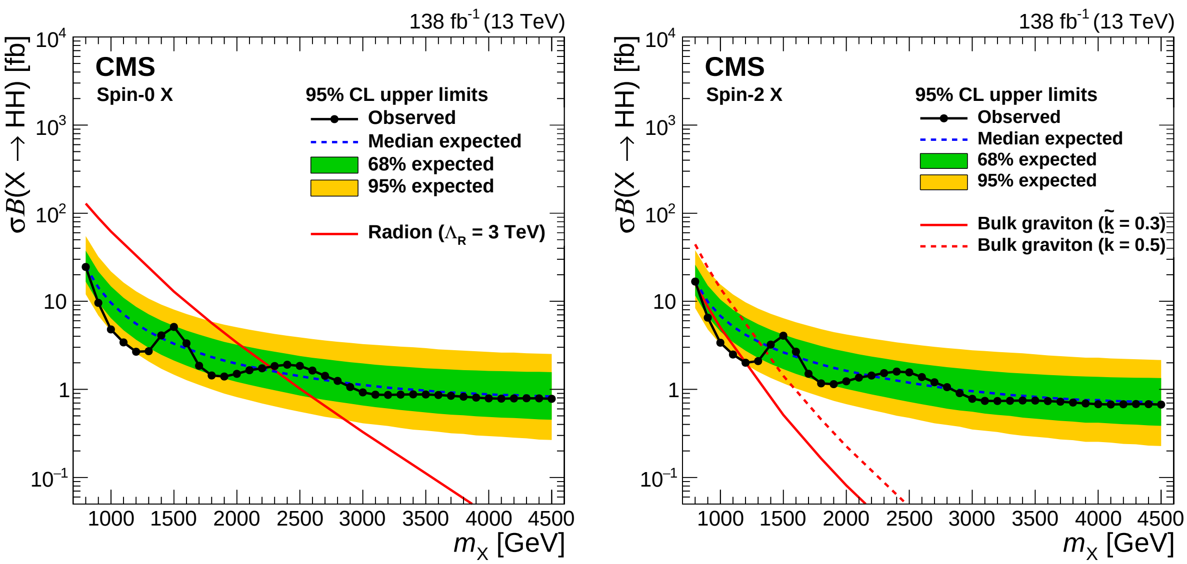

Figure 6.20: Observed and expected 95% CL upper limits on the product of the cross section and branching fraction to HH for a generic spin-0 (left) and spin-2 (right) boson X, as functions of mass. Example radion and bulk graviton predictions are also shown. The HH branching fraction is assumed to be 25% (10%) for the radion (bulk graviton).

The observed limits are also interpreted in the context of the bulk RS models that predict resonances with significant couplings to HH. Typical model parameters are chosen as proposed in Ref. [141]. The red curves assume a 25% branching fraction for a radion decaying to HH and a 10% branching fraction for bulk gravitons to HH. The observed limits exclude the existence of a radion corresponding to a model with UV cutoff $\Lambda_{\mathrm{R}} = 3.0$ TeV with mass $m \lesssim 2.3$ TeV. Bulk graviton production is excluded for a graviton mass of $m \lesssim 1350$

(1200) GeV with $\tilde{k} = 0.3$ (0.5). A general spin-0 signal at $m_X = 0.8$ TeV is excluded for $\sigma\mathcal{B} > 24.5$ fb, with the limit strengthening over the full mass range to $\sigma\mathcal{B} > 0.78$ fb at $m_X = 4.5$ TeV. Spin-2 signals have larger acceptance, and so the exclusion limit on these signals is stronger: at $m_X = 0.8$ TeV, $X \to HH$ production is excluded for $\sigma\mathcal{B} > 16.7$ fb, and at $m_X = 4.5$ TeV for $\sigma\mathcal{B} > 0.67$ fb.

The median expected exclusion limit at 95% CL for each SR category individually is shown in Fig. 6.21. Figure 6.22 compares the limits computed for the 1ℓ channel only, the 2ℓ channel only, and the total analysis.

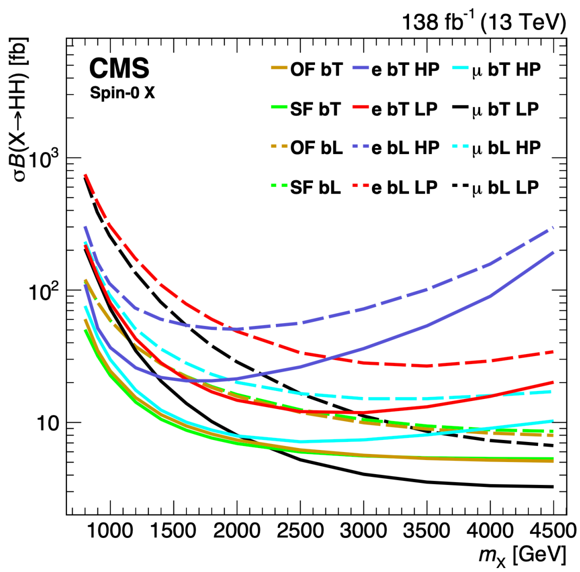

Figure 6.21: Median expected upper limits at 95% CL for each of the 12 search categories individually.

In general, the tight (bT) $b\bar{b}$ jet tagging categories are the most sensitive over the full range of m_X, since these contain the most signal and the least amount of background. The 2ℓ categories are generally more sensitive than most 1ℓ categories since the background yields are much lower in the 2ℓ channel. A notable exception to this trend is the μ bT LP category, which is the most sensitive above approximately 2.5 TeV. At high m_X, the 1ℓ electron categories are the least sensitive because the electron reconstruction efficiency is

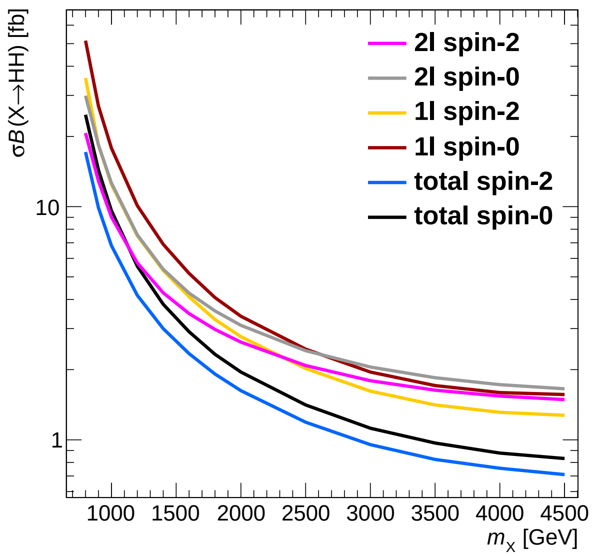

Figure 6.22: Median expected upper limits at 95% CL for the 1ℓ channel only, the 2ℓ channel only, and the total for each of the spin-0 and spin-2 signal hypotheses.

severely degraded, as described in Section 4.3.1. Overall, the 2ℓ channel sensitivity is stronger than the 1ℓ channel for m_{X} below about 2.5 TeV, after which the 1ℓ channel sensitivity becomes slightly stronger, due to the similar reasons as noted before for the individual SR categories.

This work yields the most sensitive upper limits in ATLAS and CMS for $\mathrm{X} \rightarrow \mathrm{HH}$ production with leptons in the final state. The only $\mathrm{X} \rightarrow \mathrm{HH}$ searches that are more sensitive in any subspace of the considered m_{X} and for any spin hypothesis are those in the $\mathrm{b\bar{b}b\bar{b}}$ final state from ATLAS and CMS, each of which extend only up to 3.0 TeV. From 0.8 TeV to 3.0 TeV, the sensitivity is comparable to and in some places stronger than these searches. This work has been the most successful at exploring the largest m_{X} for all $\mathrm{X} \rightarrow \mathrm{HH}$ searches.

6.3 Conclusion

In the previous chapters, a search was presented for heavy BSM resonances decaying to a pair of Higgs bosons. This analysis was undertaken at the beginning of the 2016 run of the LHC, during a wave of excitement for increasing the CMS data set by almost an order of magnitude and exploring resonance masses deeper into the TeV scale. Many searches for spin-0 and spin-2 diboson resonances began around this time, among them searches for HH resonances.

In the context of diboson searches, the HH search is a challenging one, especially in the boosted regime. In this analysis, this is largely due to the $\mathrm{H} \to \mathrm{WW}^* \to \ell\nu\mathrm{q}\bar{\mathrm{q}}$ decay. High-p_T leptons are useful in resonant searches for their extremely high efficiencies for triggering, reconstruction, and identification. Furthermore, such leptons are typically expected to be isolated from deposits of hadronic energy. In this analysis, the leptons are lower in p_T on average than those in other diboson searches because of the extra states in the decay chain (consider $\mathrm{W} \to \ell\nu$ versus $\mathrm{H} \to \mathrm{WW}^* \to \ell\nu\mathrm{q}\bar{\mathrm{q}}$, for example). Furthermore, because of the boost, the jet from $\mathrm{W} \to \mathrm{q}\bar{\mathrm{q}}$ is close to the lepton and can sometimes contain it within the jet area. As a result, the leptons in this analysis are not expected to be well-isolated and their triggering and reconstruction are expected to be less efficient.

An initial search for $\mathrm{X} \to \mathrm{HH} \to \mathrm{b}\bar{\mathrm{b}}\mathrm{WW}^* \to \mathrm{b}\bar{\mathrm{b}}\ell\nu\mathrm{q}\bar{\mathrm{q}}$ was performed using 35.9 fb^{-1} of data from the 2016 run for pp collisions at $\sqrt{s} = 13$ TeV [72] that established $\mathrm{b}\bar{\mathrm{b}}\mathrm{WW}^*$ as a competitive channel. For the full 2016–2018 analysis, the dilepton $\mathrm{b}\bar{\mathrm{b}}\mathrm{WW}^*$ channel was added along with the dilepton decay of $\mathrm{b}\bar{\mathrm{b}}\tau\tau$. These two channels are harmoniously added under a single set of selections that dramatically reduces the amount of expected background.

The entire analysis relies heavily on efficient b tagging with the CMS tracker and efficient selection of electrons and muons. In the 1ℓ channel, this requires special clustering of jets to select the $\mathrm{W} \to \mathrm{q}\bar{\mathrm{q}}$ decay and lepton selection strategies tailored for boosted environments. Further selections are applied in both channels to select signal-like topologies and optimize the signal significance relative to the background.

The analysis evaluates the existence of signal with a 2D maximum likelihood fit of signal and background models to the observed data, a procedure developed in early Run-II diboson searches [59]. The search region is broken down into 12 independent categories, each containing a set of events with distinct signal to background ratios, and the fit is performed simultaneously in these 12 categories. A unique classification of the background modeling was also developed for this analysis that allowed for the creation of templates with more statistical power than available in the search region. The templates and the fit were validated extensively with several different methods.

No excess of a signal X is observed within $0.8 < m_{\mathrm{X}} < 4.5$ TeV, and upper limits on $\sigma\mathcal{B}(\mathrm{X} \rightarrow \mathrm{HH})$ are computed at 95% CL. Over the signal phase space considered, the results in this dissertation are either the most sensitive or competitive with the most sensitive results (typically $\mathrm{b\bar{b}b\bar{b}}$ searches) on $\sigma\mathcal{B}(\mathrm{X} \rightarrow \mathrm{HH})$ for any spin-0 or spin-2 boson X. The results are also interpreted in the context of models of warped extra dimensions, and certain phase space of the model parameters are excluded.

The $\mathrm{HH} \rightarrow \mathrm{b\bar{b}WW^*}$ search has been established as a critical channel in the boosted regime of HH and diboson searches in general, arguably as important as the $\mathrm{b\bar{b}b\bar{b}}$ channel now. The work described in this dissertation has laid the foundation for several future developments in HH searches. The combined $\mathrm{X} \rightarrow \mathrm{HH}$ analysis with all channels in CMS is beginning at the time of this dissertation writing, and the competitiveness of this analysis with $\mathrm{b\bar{b}b\bar{b}}$ is promising for a strong result. While this analysis has only considered narrow width signals with intrinsic resolution much smaller than the detector resolution, the analysis is well constructed to evaluate the existence of large width signals on the order of the detector resolution. In future iterations of this search, the electron efficiency of the 1ℓ channel will also significantly increase due to improvements described in Section 4.3.1 that were motivated by this analysis. This also has implications for many other boosted searches with electrons and jets within CMS, especially as CMS pushes into the energy frontier. Ultimately, the rest of the LHC program calls for about a factor of 20 times more data than has been recorded already. This search can be expected to yield upper limits that are at least 5–20 times more sensitive than they currently are, allowing us to explore $\sigma\mathcal{B}(\mathrm{X} \rightarrow \mathrm{HH})$ below 1 fb over most m_{X} and down below 100 ab at high masses.

REFERENCES

[1] Avelino Vicente Montesinos. *Phenomenology of supersymmetric neutrino mass models*. Phd thesis, 4 2011. arXiv:1104.0831.

[2] Alexandra Oliveira. Gravity particles from warped extra dimensions, predictions for LHC. 2014. arXiv:1404.0102.

[3] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. General Photo. URL: https://cds.cern.ch/record/1621583.

[4] https://cms.cern/detector.

[5] W. Adam et al. The CMS Phase-1 Pixel Detector Upgrade. *JINST*, 16(02):P02027, 2021. arXiv:2012.14304, doi:10.1088/1748-0221/16/02/P02027.

[6] A Benaglia. The CMS ECAL performance with examples. 9(02):C02008–C02008, feb 2014. doi:10.1088/1748-0221/9/02/c02008.

[7] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi:10.1088/1748-0221/3/08/S08004.

[8] A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(06):P06015, 2018. arXiv:1804.04528, doi:10.1088/1748-0221/13/06/P06015.

[9] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined.

[10] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017. arXiv:1706.04965, doi:10.1088/1748-0221/12/10/P10003.

[11] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[12] A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018. arXiv:1712.07158, doi:10.1088/1748-0221/13/05/P05011.

[13] Search for heavy resonances decaying to a pair of boosted Higgs bosons in final states with leptons and a bottom quark-antiquark pair at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2021. URL: https://cds.cern.ch/record/2777173.

[14] Armen Tumasyan et al. Search for heavy resonances decaying to a pair of Lorentz-boosted Higgs bosons in final states with leptons and a bottom quark pair at $\sqrt{s} = 13$ TeV. Submitted to *JHEP*, 2021. arXiv:2112.03161.

[15] David Griffiths. *Introduction to elementary particles*. 2008.

[16] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in β Decay. *Phys. Rev.*, 105:1413–1414, 1957. doi:10.1103/PhysRev.105.1413.

[17] Georges Aad et al. Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:01, 2012. arXiv:1207.7214, doi:10.1016/j.physletb.2012.08.020.

[18] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 gev with the CMS experiment at the LHC. *Phys. Lett. B*, 716:30, 2012. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.

[19] Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV. *JHEP*, 06:081, 2013. arXiv:1303.4571, doi:10.1007/JHEP06(2013)081.

[20] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562, 1998. arXiv:hep-ex/9807003, doi:10.1103/PhysRevLett.81.1562.

[21] Q. R. Ahmad et al. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8b solar neutrinos at the sudbury neutrino observatory. *Phys. Rev. Lett.*, 87:071301, 2001. doi:10.1103/PhysRevLett.87.071301.

[22] Q. R. Ahmad and others. Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory. *Phys. Rev. Lett.*, 89:011301, 2002. doi:10.1103/PhysRevLett.89.011301.

[23] Georges Aad et al. Searches for Higgs boson pair production in the $hh \rightarrow bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector. *Phys. Rev. D*, 92:092004, 2015. arXiv:1509.04670, doi:10.1103/PhysRevD.92.092004.

[24] Georges Aad et al. Search for a new resonance decaying to a W or Z boson and a Higgs boson in the $\ell\ell/\ell\nu/\nu\nu + b\bar{b}$ final states with the ATLAS detector. *Eur. Phys. J. C*, 75:263, 2015. arXiv:1503.08089, doi:10.1140/epjc/s10052-015-3474-x.

[25] Georges Aad et al. Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 75:412, 2015. arXiv:1506.00285, doi:10.1140/epjc/s10052-015-3628-x.

[26] Georges Aad et al. Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state using pp collision data at $\sqrt{s} = 8$ TeV from the ATLAS detector. *Phys. Rev. Lett.*, 114:081802, 2015. arXiv:1406.5053, doi:10.1103/PhysRevLett.114.081802.

[27] Georges Aad et al. Search for WZ resonances in the fully leptonic channel using pp collisions at $\sqrt{s}$ = 8TeV with the ATLAS detector. *Phys. Lett. B*, 737:223, 2014. arXiv:1406.4456, doi:10.1016/j.physletb.2014.08.039.

[28] Morad Aaboud et al. Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the $q\bar{q}^{(\prime)}b\bar{b}$ final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *Phys. Lett. B*, 774:494, 2017. arXiv:1707.06958, doi:10.1016/j.physletb.2017.09.066.

[29] Morad Aaboud et al. Search for WW/WZ resonance production in $\ell\nu qq$ final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 03:042, 2018. arXiv:1710.07235, doi:10.1007/JHEP03(2018)042.

[30] Morad Aaboud et al. Search for resonant WZ production in the fully leptonic final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *Phys. Lett. B*, 787:68, 2018. arXiv:1806.01532, doi:10.1016/j.physletb.2018.10.021.

[31] Morad Aaboud et al. Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *Eur. Phys. J. C*, 78:24, 2018. arXiv:1710.01123, doi:10.1140/epjc/s10052-017-5491-4.

[32] Morad Aaboud et al. Searches for heavy ZZ and ZW resonances in the $\ell\ell qq$ and $\nu\nu qq$ final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 03:009, 2018. arXiv:1708.09638, doi:10.1007/JHEP03(2018)009.

[33] M. Aaboud et al. Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *Eur. Phys. J. C*, 78:293, 2018. arXiv:1712.06386, doi:10.1140/epjc/s10052-018-5686-3.

[34] Morad Aaboud et al. Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector. *JHEP*, 03:174, 2018. arXiv:1712.06518, doi:10.1007/JHEP03(2018)174.

[35] Morad Aaboud et al. Search for diboson resonances with boson-tagged jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *Phys. Lett. B*, 777:91, 2018. arXiv:1708.04445, doi:10.1016/j.physletb.2017.12.011.

[36] Morad Aaboud et al. Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 04:092, 2019. arXiv:1811.04671, doi:10.1007/JHEP04(2019)092.

[37] Morad Aaboud et al. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 01:030, 2019. arXiv:1804.06174, doi:10.1007/JHEP01(2019)030.

[38] Morad Aaboud et al. Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. Lett.*, 121:191801, 2018. arXiv:1808.00336, doi:10.1103/PhysRevLett.121.191801.

[39] Combination of searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ decay channels using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2021-052, 2021. URL: https://cds.cern.ch/record/2786865.

[40] Georges Aad et al. Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Lett. B*, 800:135103, 2020. arXiv:1906.02025, doi:10.1016/j.physletb.2019.135103.

[41] Georges Aad et al. Search for diboson resonances in hadronic final states in 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 09:091, 2019. [Erratum: JHEP 06 (2020) 042]. arXiv:1906.08589, doi:10.1007/JHEP09(2019)091.

[42] Georges Aad et al. Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 07:108, 2020. [Erratum: JHEP 01 (2021) 145, Erratum: JHEP 05 (2021) 207]. arXiv:2001.05178, doi:10.1007/JHEP07(2020)108.

[43] Georges Aad et al. Search for heavy diboson resonances in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 80:1165, 2020. arXiv:2004.14636, doi:10.1140/epjc/s10052-020-08554-y.

[44] Georges Aad et al. Search for resonances decaying into a weak vector boson and a Higgs boson in the fully hadronic final state produced in proton−proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 102:112008, 2020. arXiv:2007.05293, doi:10.1103/PhysRevD.102.112008.

[45] Georges Aad et al. Reconstruction and identification of boosted di-τ systems in a search for Higgs boson pairs using 13 TeV proton-proton collision data in ATLAS. *JHEP*, 11:163, 2020. arXiv:2007.14811, doi:10.1007/JHEP11(2020)163.

[46] Search for resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2021-035, 2021. URL: https://cds.cern.ch/record/2777861.

[47] Albert M Sirunyan et al. Combination of searches for heavy resonances decaying to WW, WZ, ZZ, WH, and ZH boson pairs in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV. *Phys. Lett. B*, 774:533, 2017. arXiv:1705.09171, doi:10.1016/j.physletb.2017.09.083.

[48] Albert M Sirunyan et al. Search for a heavy resonance decaying into a Z boson and a vector boson in the $\nu\overline{\nu}q\overline{q}$ final state. *JHEP*, 07:075, 2018. arXiv:1803.03838, doi:10.1007/JHEP07(2018)075.

[49] Albert M Sirunyan et al. Search for heavy resonances decaying into two Higgs bosons or into a Higgs boson and a W or Z boson in proton-proton collisions at 13 TeV. *JHEP*, 01:051, 2019. arXiv:1808.01365, doi:10.1007/JHEP01(2019)051.

[50] Albert M Sirunyan et al. Search for a heavy resonance decaying into a Z boson and a Z or W boson in $2\ell 2q$ final states at $\sqrt{s} = 13$ TeV. *JHEP*, 09:101, 2018. arXiv:1803.10093, doi:10.1007/JHEP09(2018)101.

[51] Albert M Sirunyan et al. Search for production of Higgs boson pairs in the four b quark final state using large-area jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 01:040, 2019. arXiv:1808.01473, doi:10.1007/JHEP01(2019)040.

[52] Vardan Khachatryan et al. Searches for a heavy scalar boson H decaying to a pair of 125 GeV Higgs bosons hh or for a heavy pseudoscalar boson A decaying to Zh, in the final states with h $\to \tau\tau$. *Phys. Lett. B*, 755:217, 2016. arXiv:1510.01181, doi:10.1016/j.physletb.2016.01.056.

[53] Albert M Sirunyan et al. Search for massive resonances decaying into WW, WZ, ZZ, qW, and qZ with dijet final states at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 97:072006, 2018. arXiv:1708.05379, doi:10.1103/PhysRevD.97.072006.

[54] Albert M Sirunyan et al. Search for heavy resonances that decay into a vector boson and a Higgs boson in hadronic final states at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 77:636, 2017. arXiv:1707.01303, doi:10.1140/epjc/s10052-017-5192-z.

[55] Albert M Sirunyan et al. Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos and b quarks at $\sqrt{s} = 13$ TeV. *JHEP*, 11:172, 2018. arXiv:1807.02826, doi:10.1007/JHEP11(2018)172.

[56] Vardan Khachatryan et al. Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 8 TeV. *Phys. Lett. B*, 749:560, 2015. arXiv:1503.04114, doi:10.1016/j.physletb.2015.08.047.

[57] V. Khachatryan et al. Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos, and b quarks. *Phys. Lett. B*, 768:137, 2017. arXiv:1610.08066, doi:10.1016/j.physletb.2017.02.040.

[58] A. M. Sirunyan et al. Search for a massive resonance decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 781:244, 2018. arXiv:1710.04960, doi:10.1016/j.physletb.2018.03.084.

[59] Albert M Sirunyan et al. Search for a heavy resonance decaying to a pair of vector bosons in the lepton plus merged jet final state at $\sqrt{s} = 13$ TeV. *JHEP*, 05:088, 2018. arXiv:1802.09407, doi:10.1007/JHEP05(2018)088.

[60] Albert M Sirunyan et al. Search for massive resonances decaying into WW, WZ or ZZ bosons in proton-proton collisions at $\sqrt{s} = 13$TeV. *JHEP*, 03:162, 2017. arXiv:1612.09159, doi:10.1007/JHEP03(2017)162.

[61] Albert M Sirunyan et al. Search for ZZ resonances in the $2\ell2\nu$ final state in proton-proton collisions at 13 TeV. *JHEP*, 03:003, 2018. arXiv:1711.04370, doi:10.1007/JHEP03(2018)003.

[62] V. Khachatryan et al. Search for new resonances decaying via WZ to leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 740:83, 2015. arXiv:1407.3476, doi:10.1016/j.physletb.2014.11.026.

[63] V. Khachatryan et al. Search for massive WH resonances decaying into the $\ell\nu\mathrm{b}\overline{\mathrm{b}}$ final state at $\sqrt{s} = 8$TeV. *Eur. Phys. J. C*, 76:237, 2016. arXiv:1601.06431, doi:10.1140/epjc/s10052-016-4067-z.

[64] V. Khachatryan et al. Search for a massive resonance decaying into a Higgs boson and a W or Z boson in hadronic final states in proton-proton collisions at $\sqrt{s} = 8$TeV. *JHEP*, 02:145, 2016. arXiv:1506.01443, doi:10.1007/JHEP02(2016)145.

[65] V. Khachatryan et al. Search for narrow high-mass resonances in proton-proton collisions at $\sqrt{s} = 8$TeV decaying to a Z and a Higgs boson. *Phys. Lett. B*, 748:255, 2015. arXiv:1502.04994, doi:10.1016/j.physletb.2015.07.011.

[66] Albert M Sirunyan et al. Combination of CMS searches for heavy resonances decaying to pairs of bosons or leptons. *Phys. Lett. B*, 798:134952, 2019. arXiv:1906.00057, doi:10.1016/j.physletb.2019.134952.

[67] Albert M Sirunyan et al. Search for a heavy vector resonance decaying to a Z boson and a Higgs boson in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 81:688, 2021. arXiv:2102.08198, doi:10.1140/epjc/s10052-021-09348-6.

[68] Armen Tumasyan et al. Search for heavy resonances decaying to WW, WZ, or WH boson pairs in the lepton plus merged jet final state in proton-proton collisions at $\sqrt{s}$ $= 13$ TeV. Submitted to *Phys. Rev. D.*, 2021. arXiv:2109.06055.

[69] CMS Collaboration. Search for resonant Higgs boson pair production in four b quark final state using large-area jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-B2G-20-004, 2021. URL: https://cds.cern.ch/record/2777083.

[70] Armen Tumasyan et al. Search for heavy resonances decaying to Z($\nu\bar{\nu}$)V(q$\bar{\text{q}}$') in proton-proton collisions at $\sqrt{s} = 13$ TeV. Submitted to *Phys. Rev. D.*, 2021. arXiv:2109.08268.

[71] Armen Tumasyan et al. Search for heavy resonances decaying to ZZ or ZW and axion-like particles mediating nonresonant ZZ or ZH production at $\sqrt{s} = 13$ TeV. Submitted to *JHEP*, 2021. arXiv:2111.13669.

[72] A. M. Sirunyan et al. Search for resonances decaying to a pair of Higgs bosons in the $\mathrm{b\bar{b}q\bar{q}}\ell\nu$ final state in proton-proton collisions at $\sqrt{s} = 13\ TeV$. *JHEP*, 10:125, 2019. arXiv:1904.04193, doi:10.1007/JHEP10(2019)125.

[73] F. Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. 1984.

[74] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508, 1964. doi:10.1103/PhysRevLett.13.508.

[75] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321, 1964. doi:10.1103/PhysRevLett.13.321.

[76] Svend E. Rugh and Henrik Zinkernagel. The quantum vacuum and the cosmological constant problem, 2001. URL: http://philsci-archive.pitt.edu/398/.

[77] Stephen P. Martin. A supersymmetry primer. *Advanced Series on Directions in High Energy Physics*, page 1–98, Jul 1998. URL: http://dx.doi.org/10.1142/9789812839657_0001, doi:10.1142/9789812839657_0001.

[78] Gian F. Giudice. Naturalness after LHC8. Naturalness after LHC-7 and LHC-8. *PoS*, EPS-HEP2013:163. 13 p, Jul 2013. Comments: 13 pages, 1 figure; 2013 EPS Conference Proceedings. URL: https://cds.cern.ch/record/1565743, arXiv:1307.7879.

[79] Aneesh V. Manohar. Introduction to Effective Field Theories. Apr 2018. arXiv:1804.05863, doi:10.1093/oso/9780198855743.003.0002.

[80] Pierre Ramond. Dual theory for free fermions. *Phys. Rev. D*, 3:2415, 1971. doi:10.1103/PhysRevD.3.2415.

[81] Yu. A. Golfand and E. P. Likhtman. Extension of the algebra of Poincaré group generators and violation of P invariance. *JETP Lett.*, 13:323, 1971.

[82] A. Neveu and J. H. Schwarz. Factorizable dual model of pions. *Nucl. Phys. B*, 31:86, 1971. doi:10.1016/0550-3213(71)90448-2.

[83] D. V. Volkov and V. P. Akulov. Possible universal neutrino interaction. *JETP Lett.*, 16:438, 1972.

[84] J. Wess and B. Zumino. A Lagrangian model invariant under supergauge transformations. *Phys. Lett. B*, 49:52, 1974. doi:10.1016/0370-2693(74)90578-4.

[85] J. Wess and B. Zumino. Supergauge transformations in four dimensions. *Nucl. Phys. B*, 70:39, 1974. doi:10.1016/0550-3213(74)90355-1.

[86] Pierre Fayet. Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino. *Nucl. Phys. B*, 90:104, 1975. doi:10.1016/0550-3213(75)90636-7.

[87] Hans Peter Nilles. Supersymmetry, supergravity and particle physics. *Phys. Rep.*, 110:1, 1984. doi:10.1016/0370-1573(84)90008-5.

[88] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, and João P. Silva. Theory and phenomenology of two-higgs-doublet models. *Physics Reports*, 516(1-2):1–102, Jul 2012. URL: http://dx.doi.org/10.1016/j.physrep.2012.02.002, doi:10.1016/j.physrep.2012.02.002.

[89] T. Binoth and J. J. van der Bij. Influence of strongly coupled, hidden scalars on Higgs signals. *Z. Phys. C*, 75:17–25, 1997. arXiv:hep-ph/9608245, doi:10.1007/s002880050442.

[90] Theodor Kaluza. Zum Unitätsproblem der Physik. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, pages 966–972, Jan 1921.

[91] Oskar Klein. The Atomicity of Electricity as a Quantum Theory Law. *Nature*, 118(2971):516, Oct 1926. doi:10.1038/118516a0.

[92] Lisa Randall and Raman Sundrum. A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370, 1999. arXiv:hep-ph/9905221, doi:10.1103/PhysRevLett.83.3370.

[93] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo. Phenomenology of the Randall-Sundrum gauge hierarchy model. *Phys. Rev. Lett.*, 84:2080, 2000. arXiv:hep-ph/9909255, doi:10.1103/PhysRevLett.84.2080.

[94] Kaustubh Agashe, Hooman Davoudiasl, Gilad Perez, and Amarjit Soni. Warped gravitons at the LHC and beyond. *Phys. Rev. D*, 76:036006, 2007. arXiv:hep-ph/0701186, doi:10.1103/PhysRevD.76.036006.

[95] A. Liam Fitzpatrick, Jared Kaplan, Lisa Randall, and Lian-Tao Wang. Searching for the Kaluza-Klein graviton in bulk RS models. *JHEP*, 09:013, 2007. arXiv:hep-ph/0701150, doi:10.1088/1126-6708/2007/09/013.

[96] Walter D. Goldberger and Mark B. Wise. Modulus stabilization with bulk fields. *Phys. Rev. Lett.*, 83:4922, 1999. arXiv:hep-ph/9907447, doi:10.1103/PhysRevLett.83.4922.

[97] O. DeWolfe, D. Z. Freedman, S. S. Gubser, and A. Karch. Modeling the fifth dimension with scalars and gravity. *Phys. Rev. D*, 62:046008, 2000. arXiv:hep-th/9909134, doi:10.1103/PhysRevD.62.046008.

[98] Csaba Csaki, Michael Graesser, Lisa Randall, and John Terning. Cosmology of brane models with radion stabilization. *Phys. Rev. D*, 62:045015, 2000. arXiv:hep-ph/9911406, doi:10.1103/PhysRevD.62.045015.

[99] Csaba Csaki, Michael L. Graesser, and Graham D. Kribs. Radion dynamics and electroweak physics. *Phys. Rev. D*, 63:065002, 2001. arXiv:hep-th/0008151, doi:10.1103/PhysRevD.63.065002.

[100] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3(08):S08001, 2008. doi:10.1088/1748-0221/3/08/s08001.

[101] R Bruce, N Fuster-Martínez, A Mereghetti, D Mirarchi, and S Redaelli. Review of LHC Run 2 Machine Configurations. pages 187–197. 11 p, 2019. URL: https://cds.cern.ch/record/2750415.

[102] V. Karimäki. The CMS tracker system project: Technical Design Report. 1997.

[103] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997. URL: https://cds.cern.ch/record/349375.

[104] *The CMS hadron calorimeter project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997. URL: https://cds.cern.ch/record/357153.

[105] J. G. Layter. *The CMS muon project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997. URL: https://cds.cern.ch/record/343814.

[106] P. Giacomelli. The cms muon detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 478(1):147, 2002. Proceedings of the ninth Int.Conf. on Instrumentation. doi:https://doi.org/10.1016/S0168-9002(01)01739-9.

[107] Roel Aaij et al. Measurement of the inelastic pp cross-section at a centre-of-mass energy of 13 TeV. *JHEP*, 06:100, 2018. arXiv:1803.10974, doi:10.1007/JHEP06(2018)100.

[108] G L Bayatyan, N Grigorian, et al. *CMS The TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical design report. CMS. URL: https://cds.cern.ch/record/706847.

[109] Sergio Cittolin, Attila Rácz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project.* Technical design report. CMS. CERN, Geneva, 2002. URL: https://cds.cern.ch/record/578006.

[110] Vardan Khachatryan et al. The CMS trigger system. *JINST*, 12(01):P01020, 2017. arXiv:1609.02366, doi:10.1088/1748-0221/12/01/P01020.

[111] Rick Field. Min-Bias and the Underlying Event at the LHC. *Acta Phys. Polon. B*, 42:2631, 2011. arXiv:1110.5530, doi:10.5506/APhysPolB.42.2631.

[112] Albert M Sirunyan et al. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *JINST*, 16(05):P05014, 2021. arXiv:2012.06888, doi:10.1088/1748-0221/16/05/P05014.

[113] Jessie Shelton. Jet Substructure. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales*, page 303, 2013. arXiv:1302.0260, doi:10.1142/9789814525220_0007.

[114] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. arXiv:0802.1189, doi:10.1088/1126-6708/2008/04/063.

[115] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012. arXiv:1111.6097, doi:10.1140/epjc/s10052-012-1896-2.

[116] Emil Bols, Jan Kieseler, Mauro Verzetti, Markus Stoye, and Anna Stakia. Jet flavour classification using DeepJet. *JINST*, 15:P12012, 2020. arXiv:2008.10519, doi:10.1088/1748-0221/15/12/P12012.

[117] Tumasyan Sirunyan et al. Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques. *JINST*, 15:P06005, 2020. doi:10.1088/1748-0221/15/06/P06005.

[118] Keith Rehermann and Brock Tweedie. Efficient Identification of Boosted Semileptonic Top Quarks at the LHC. *JHEP*, 03:059, 2011. arXiv:1007.2221, doi:10.1007/JHEP03(2011)059.

[119] Albert M Sirunyan et al. Pileup mitigation at CMS in 13 TeV data. *JINST*, 15:P09018, 2020. arXiv:2003.00503, doi:10.1088/1748-0221/15/09/p09018.

[120] Daniele Bertolini, Philip Harris, Matthew Low, and Nhan Tran. Pileup Per Particle Identification. *JHEP*, 10:059, 2014. arXiv:1407.6013, doi:10.1007/JHEP10(2014)059.

[121] https://twiki.cern.ch/twiki/bin/view/CMS/JetID.

[122] Andrew J. Larkoski, Simone Marzani, Gregory Soyez, and Jesse Thaler. Soft drop. *JHEP*, 05:146, 2014. arXiv:1402.2657, doi:10.1007/JHEP05(2014)146.

[123] Mrinal Dasgupta, Alessandro Fregoso, Simone Marzani, and Gavin P. Salam. Towards an understanding of jet substructure. *JHEP*, 09:029, 2013. `arXiv:1307.0007`, `doi:10.1007/JHEP09(2013)029`.

[124] Jesse Thaler and Ken Van Tilburg. Identifying boosted objects with N-subjettiness. *JHEP*, 03:015, 2011. `arXiv:1011.2268`, `doi:10.1007/JHEP03(2011)015`.

[125] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetWtagging`.

[126] A. M. Sirunyan et al. Measurement of normalized differential $t\bar{t}$ cross sections in the dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 04:060, 2018. `arXiv:1708.07638`, `doi:10.1007/JHEP04(2018)060`.

[127] Vardan Khachatryan et al. Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV. *Phys. Rev. D*, 95:092001, 2017. `arXiv:1610.04191`, `doi:10.1103/PhysRevD.95.092001`.

[128] Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, 27:832, 1956. `doi:10.1214/aoms/1177728190`.

[129] Bernard W. Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, 1986.

[130] David W. Scott. *Multivariate density estimation: theory, practice, and visualization*. John Wiley and Sons, 1992.

[131] M. Oreglia. *A Study of the Reactions $\psi' \to \gamma\gamma\psi$*. Phd thesis, 12 1980.

[132] `https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/`.

[133] Albert M Sirunyan et al. Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. 4 2021. `arXiv:2104.01927`.

[134] CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2018. URL: `https://cds.cern.ch/record/2621960`.

[135] CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2019. URL: `https://cds.cern.ch/record/2676164`.

[136] `https://twiki.cern.ch/twiki/bin/view/CMS/L1ECALPrefiringWeightRecipe`.

[137] `http://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf`.

[138] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999. `arXiv:hep-ex/9902006`, `doi:10.1016/S0168-9002(99)00498-2`.

[139] Alexander L. Read. Presentation of search results: The CL_s technique. *J. Phys. G*, 28:2693, 2002. doi:10.1088/0954-3899/28/10/313.

[140] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. [Erratum: Eur.Phys.J.C 73, 2501 (2013)]. arXiv:1007.1727, doi:10.1140/epjc/s10052-011-1554-0.

[141] Maxime Gouzevitch, Alexandra Oliveira, Juan Rojo, Rogerio Rosenfeld, Gavin P. Salam, and Veronica Sanz. Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production. *JHEP*, 07:148, 2013. arXiv:1303.6636, doi:10.1007/JHEP07(2013)148.