

SEARCH FOR EXOTIC HIGGS DECAYS TO LIGHT
NEUTRAL SCALARS IN FINAL STATES WITH
BOTTOM QUARKS AND TAU LEPTONS

KA YU STEPHANIE KWAN

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

NOT YET RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PHYSICS
ADVISER: ISOBEL OJALVO

MAY 2024

© Copyright by Ka Yu Stephanie Kwan, 2024.

All Rights Reserved

Abstract

Open questions in particle physics may be addressed by the existence of an extended Higgs sector beyond the Higgs boson with mass 125 GeV discovered at the Large Hadron Collider (LHC) by the CMS and ATLAS experiments in 2012, may answer remaining open questions in particle physics. Many properties of a potential extended Higgs sector remain unconstrained by current measurements, making direct searches of exotic Higgs decays are a powerful probe of new physics. We present a search performed at the CMS experiment for exotic decays of a Higgs boson with mass 125 GeV to two light neutral pseudoscalars which respectively decay to two bottom quarks and two tau leptons ($h \rightarrow aa \rightarrow bb\tau\tau$), which is combined with a different analysis to the final state with two bottom quarks and two muons, with results interpreted in Two Higgs Doublet Models extended with a singlet scalar (2HDM+S). We also present preliminary work on a new search of 125 GeV Higgs boson decays to two light pseudoscalars which can have different masses ($h \rightarrow a_1a_2$) with final states containing bottom quarks and tau leptons. This “asymmetric” decay scenario has not been studied to date at the CMS experiment and is predicted by Two Real Singlet Models (TRSM).

Acknowledgements

Placeholder acknowledgements.

Placeholder dedication

Contents

Abstract	iii
Acknowledgements	iv
List of Tables	xi
List of Figures	xii
1 Introduction	1
1.1 History of the Standard Model	1
1.2 The Standard Model as a gauge theory	3
1.2.1 Gauge invariance	3
1.2.2 Local gauge symmetries	4
1.3 The Higgs Mechanism	6
1.4 Two-Higgs Doublet Models	8
1.5 Two Real Singlet Model	11
2 The Large Hadron Collider and the CMS Experiment	14
2.1 The Large Hadron Collider	14
2.2 Luminosity and pileup	16
2.3 The High-Luminosity LHC	17
2.4 The CMS Detector	18
2.5 Sub-detectors of CMS	20
2.5.1 Inner tracking system	20

2.5.2	ECAL	22
2.5.3	HCAL	23
2.5.4	Muon detectors	24
2.5.5	The High-Level Trigger	26
2.5.6	Particle reconstruction	27
2.5.7	Data storage and computational infrastructure	27
3	The CMS Phase-1 Level-1 Trigger	28
3.1	The Phase-1 Level-1 Trigger	28
4	The Phase-2 Upgrade of CMS	32
4.1	High-Luminosity LHC and CMS	32
4.2	The Phase-2 Level-1 Trigger	32
4.3	Standalone Barrel Calorimeter electron/photon reconstruction	35
4.3.1	Phase-2 geometry of the ECAL Barrel trigger	35
4.3.2	Phase-2 electron/photon reconstruction algorithm	36
5	Event reconstruction	45
5.1	Taus	45
5.1.1	Decay modes	45
5.1.2	Hadron plus strips (HPS) reconstruction	46
5.1.3	DeepTau	48
5.2	Muons	48
5.3	Electrons	50
5.4	Jets	51
5.5	B jets	52
6	Datasets and Monte Carlo samples	54
6.1	Datasets used	54

6.2	Monte Carlo samples	54
6.3	Embedded samples	55
7	Event selection	56
7.1	General procedure for all channels	56
7.2	Event selection in the $\mu\tau_h$ channel	57
7.3	Event selection in the $e\tau_h$ channel	58
7.4	Event selection in the $e\mu$ channel	60
7.5	Extra lepton vetoes in all channels	61
8	Scale factors and corrections	64
8.1	Tau energy scale	64
8.2	Muon energy scale	65
8.3	Electron energy scale	65
8.4	τ_h identification efficiency	66
8.5	$e \rightarrow \tau_h$ and $\mu \rightarrow \tau_h$ energy scales and misidentification efficiency . . .	67
8.6	Electron and muon ID, isolation, and tracking	67
8.7	Recoil corrections	68
8.8	Drell-Yan corrections	68
8.9	Pileup reweighing	68
8.10	Trigger efficiencies	69
8.11	Top p_T spectrum reweighing	69
8.12	B-tagging efficiency	69
8.13	Jet energy resolution and jet energy smearing	70
9	Background estimation	71
9.1	Z+jets	71
9.2	W+jets	72
9.3	$t\bar{t}$ + jets	72

9.4	Single top	73
9.5	Diboson	73
9.6	Standard Model Higgs	73
9.7	Jet faking τ_h for the $\mu\tau_h$ and $e\tau_h$ channels	74
9.8	QCD multijet background for the $e\mu$ channel	75
10	Event categorization	77
10.1	B-tag jet multiplicity	77
10.2	DNN-based event categorization	77
11	Systematic uncertainties	81
11.1	Uncertainties associated with physics objects	82
11.1.1	Uncertainties in the lepton energy scales	82
11.1.2	Uncertainties from other lepton corrections	83
11.1.3	Uncertainties from jet energy scale and resolution	83
11.1.4	Uncertainties from b-tagging scale factors	84
11.1.5	Uncertainties from MET	85
11.2	Uncertainties associated with samples used	85
11.3	Other uncertainties	86
11.4	Pulls and impacts	86
12	Signal extraction from the di-tau mass	88
12.1	Estimation of the $\tau\tau$ mass	88
12.1.1	Original SVFit “standalone”: maximum likelihood	89
12.1.2	“Classic SVFit” with matrix element	90
12.1.3	FastMTT: optimized SVFit	90
12.2	Signal extraction	91

13 Results	92
13.1 Results from $bb\tau\tau$	92
13.2 Combination procedure with $bb\mu\mu$ final state	93
13.3 Results from combination	93

List of Tables

7.1	High-Level Trigger (HLT) paths used to select data and simulation events in 2016 for the three $\tau\tau$ channels.	61
7.2	High-Level Trigger (HLT) paths used to select data and simulation events in 2017 for the three $\tau\tau$ channels.	62
7.3	High-Level Trigger (HLT) paths used to select data and simulation events in 2018 for the three $\tau\tau$ channels. In 2018 a HLT trigger path using the hadron plus strips (HPS) tau reconstruction algorithm became available.	62
8.1	Energy scales applied to genuine hadronic tau decays τ_h by data-taking year/era and decay mode, along with systematic errors.	65
8.2	Energy scales and systematic errors applied to genuine muons.	65
8.3	Energy scales and systematic errors applied to electrons in embedded samples by data-taking year/era.	66
8.4	Tau ID efficiency for the DeepTau vs. jet medium working point, with central, up, and down values for 2018, binned in the tau p_T	67
10.1	Event categorization based on DNN scores for events with exactly 1 b-tag jet (1bNN), for the three $\tau\tau$ channels and three eras.	80
10.2	Event categorization based on DNN scores for events with 2 b-tag jets (2bNN), for the three $\tau\tau$ channels and three eras.	80

List of Figures

1.1	Table of Standard Model particles showing the grouping of the fermions into three generations of matter and the bosons, responsible for carrying the three fundamental forces in the Standard Model. The masses, charges, and spins of the particles are shown. The antimatter counterparts of the fermions are not shown. The possible interactions between the fermions and gauge bosons are highlighted.	2
1.2	An illustration of the Higgs potential.	7
1.3	Branching ratios of a singlet-like pseudoscalar in Type II 2HDM+S for $\tan \beta = 0.5$ (left) and $\tan \beta = 5$ (right).	10
1.4	Benchmark plane BP1 for benchmark scenario 1, for the decay signature $h_{125} \rightarrow h_1 h_2$ with $h_{125} \equiv h_3$, defined in the (M_1, M_2) plane. . . .	13
2.1	Aerial view of the Large Hadron Collider (LHC).	15
2.2	Distribution of the mean number of inelastic collisions per bunch crossing (pileup) in data, for proton-proton collisions in 2016-2018	18
2.3	Sketch of particle trajectories of muons, electrons, charged and neutral hadrons, and photons in a transverse cross-section of the CMS detector.	19
2.4	Cross section of the current Phase-1 CMS tracker.	21
2.5	Longitudinal view of the CMS detector showing the hadron calorimeter barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeters.	23

2.6	Layout of the CMS barrel muon drift tube (DT) chambers in one of the five wheels.	25
3.1	Dataflow for the Phase-1 Level-1 Trigger [CITE], which is implemented in custom hardware and is responsible for reducing the event rate from the LHC bunch crossing frequency of 400 MHz (bunch crossings every 25 ns) to a maximum rate of 100 kHz. In Phase-1, the Level-1 Trigger has access to information from the calorimeter and muon detectors.	29
4.1	Functional diagram of the CMS L1 Phase-2 upgraded trigger design.	33
4.2	Summary of the links between the trigger primitives, the trigger objects, the Level-1 algorithms, and the physics channels in the Phase-2 menu.	36
4.3	Schematic of the geometry of the Phase-2 ECAL barrel in the Regional Calorimeter Trigger (RCT), showing the division of the barrel region into 36 Regional Calorimeter Trigger (RCT) cards (<i>red</i>). Each card spans 17×4 towers in $\eta \times \phi$ (<i>green</i>), and each tower is 5×5 in single crystals in $\eta \times \phi$. Towers in the overlap region (<i>shaded yellow</i>) are read out to both the barrel and endcap.	39
4.4	Schematic of two example RCT cards in the negative eta (<i>top</i>) and positive eta (<i>center</i>) regions of the ECAL barrel. Each RCT card is divided into five regions: four regions are of size 3×4 towers in $\eta \times \phi$ (<i>bottom left</i>), and a fifth smaller overlap region of size 2×4 towers (<i>bottom right</i>). Each tower is 5×5 ($\eta \times \phi$) in crystals.	40

4.5	Schematic of the Phase-2 ECAL barrel in the Global Calorimeter Trigger (GCT), which will process the outputs of the Regional Calorimeter Trigger (RCT) in three cards (<i>magenta highlights</i>). Each card in the GCT processes the equivalent of sixteen RCT cards, with the center twelve being unique to that GCT card (<i>shaded pink</i>), and the remaining four processed in overlap with the other GCT cards.	41
4.6	Illustration of an example electron/photon (e/γ) cluster in the Phase-2 Level-1 Trigger standalone barrel e/γ reconstruction, in a region of 15×20 crystals (3×4 towers). Each small pink square is one crystal, the highest-granularity ECAL trigger primitives available to the L1 Trigger in Phase-2. The core cluster consists of the energy sum in a 3×5 window of crystals, (<i>shaded light blue</i>) centered around the seed crystal (<i>red</i>). Bremsstrahlung corrections are checked in the adjacent 3×5 windows in the ϕ direction (<i>shaded light yellow</i>). The relative energies in windows of size 2×5 and 5×5 in crystals (<i>dashed dark blue and dark red</i>) are used to compute shower shape variables to identify true e/γ objects. Lastly, an isolation sum is computed in a window of size 7×7 in towers (not shown in figure).	42
4.7	Efficiency of the standalone barrel e/γ reconstruction, as a function of the true electron's transverse momentum p_T	43
4.8	Rates of the standalone barrel e/γ reconstruction measured as a function of the minimum energy (E_T) required of the reconstructed e/γ object in each event.	44
8.1	Electron/photon energy scale factors and uncertainties for 2018. . . .	66
9.1	Leading-order Feynman diagrams of Higgs production.	74
11.1	Top sixty impacts for the combination of all channels and years. . . .	87

12.1 Distributions of $m_{\tau\tau}$ reconstructed by the classic SVFit algorithm, and masses of visible tau decay products (before SVFit)	91
13.1 Postfit final $m_{\tau\tau}$ distributions in the $\mu\tau_h$ channel	95
13.2 Postfit final $m_{\tau\tau}$ distributions in the $e\tau_h$ channel	96
13.3 Postfit final $m_{\tau\tau}$ distributions in the $e\mu$ channel	97
13.4 95% CL exclusion limits on $B(h \rightarrow aa \rightarrow bb\tau\tau)$ in %	98
13.5 Observed 95% CL upper limits on $B(h \rightarrow aa)$ in %, for the $bb\mu\mu$ channel using the full Run 2 integrated luminosity of 138 fb^{-1} for different 2HDM+S models	99
13.6 Observed 95% CL upper limits on $B(h \rightarrow aa)$ in %, for the $bb\tau\tau$ channel using the full Run 2 integrated luminosity of 138 fb^{-1} for different 2HDM+S models	99
13.7 Observed 95% CL upper limits on $\mathcal{B}(h \rightarrow aa)$ in %, for the combination of $bb\mu\mu$ and $bb\tau\tau$ channels using the full Run 2 integrated luminosity of 138 fb^{-1} for different 2HDM+S models	100
13.8 Observed 95% CL upper limits on $\mathcal{B}(h \rightarrow aa)$ in %, for the combination of $bb\mu\mu$ and $bb\tau\tau$ channels using the full Run 2 integrated luminosity of 138 fb^{-1} for Type II (<i>left</i>), Type III (<i>middle</i>), and Type IV (<i>right</i>) 2HDM+S in the $\tan\beta$ vs. m_a phase space	100
13.9 95% limits on $\frac{\sigma(h)}{\sigma_{\text{SM}}} \times \mathcal{B}(h \rightarrow aa)$ in the 2HDM+S type-I scenario for exotic Higgs decay searches, performed with data collected at 13 TeV	101
13.10 95% limits on $\frac{\sigma(h)}{\sigma_{\text{SM}}} \times \mathcal{B}(h \rightarrow aa)$ in the 2HDM+S type-II scenario with $\tan\beta = 2.0$ for exotic Higgs decay searches, performed with data collected at 13 TeV	101
13.11 95% limits on $\frac{\sigma(h)}{\sigma_{\text{SM}}} \times \mathcal{B}(h \rightarrow aa)$ in the 2HDM+S type-III scenario with $\tan\beta = 2.0$ for exotic Higgs decay searches, performed with data collected at 13 TeV	102

Chapter 1

Introduction

The Standard Model is the current prevailing theoretical framework that encompasses all known elementary particles to date and describes their interactions, yet falls short of describing open problems in physics. Here, we introduce the Standard Model (Section 1.1) and provide a mathematical motivation of the SM a gauge theory (Section 1.2). We introduce the Higgs mechanism (Section 1.3), and outline two groups of theoretical extensions to the Standard Model that feature extended Higgs sectors (Sections 1.4 and 1.5).

1.1 History of the Standard Model

The building blocks of our modern-day understanding of particle physics were established over the course of decades by experimental discoveries and theoretical advances, culminating in the development of a theoretical framework known as the Standard Model (SM). In the 1880s, the electron was the first subatomic particle to be identified, through measurements of particles produced by ionizing gas. By the 1930s, atoms were known to consist mostly of empty space, with protons and neutrons concentrated at the center and orbited by electrons. Spurred by advances in particle accelerator technology, the experimental discoveries of the positron, the muon, and

the pion, painted an increasingly complicated picture of particle physics that could not be described solely with atomic physics [1].

In the absence of a theoretical framework describing these particles, in the 1960s and 1970s physicists and mathematicians developed the Standard Model to describe and encompass these fundamental particles and the forces that govern their interactions. The particle content of the Standard Model is shown in Fig. 1.1: they are grouped into fermions, which comprise all known matter, and bosons, which mediate the interactions between particles.

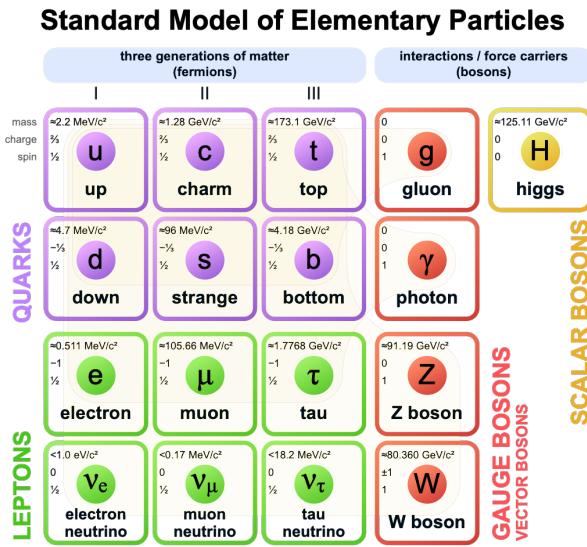


Figure 1.1: Table of Standard Model particles showing the grouping of the fermions into three generations of matter and the bosons, responsible for carrying the three fundamental forces in the Standard Model. The masses, charges, and spins of the particles are shown. The antimatter counterparts of the fermions are not shown. The possible interactions between the fermions and gauge bosons are highlighted.

Fermions consist of quarks and leptons, and are grouped into three generations. For example, the electron belongs to the first generation of leptons. The second and third generation counterparts of the electron are the muon and the tau lepton, and are over 200 and 30,000 times heavier than the electron respectively. Bosons are force carriers; the interaction of fermions with bosons corresponds to fundamental forces. The Standard Model describes the electromagnetic force, the strong nuclear force,

and the weak nuclear force.

1.2 The Standard Model as a gauge theory

1.2.1 Gauge invariance

Gauge theories of elementary particle interactions originate from a freedom of choice in the mathematical description of particle fields which has no effect on the particles' physical states [2]. The existence and form of the particles' interactions, can be deduced from the existence of physically indeterminate, gaugable quantities.

An example of this gauge invariance is classical physics is the electromagnetic interaction, where the fundamental field is the four-vector potential A^μ [2]. The physical electromagnetic fields and Maxwell's equations arise from the elements of the tensor $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$. Any two choices of A^μ that are related by a transformation of the form

$$A_\mu \rightarrow A_\mu + \partial_\mu \alpha \quad (1.1)$$

for any real, differentiable function $\alpha(x)$, describe the same physical configuration, and has no effect on Maxwell's equations. This "redundancy" in the choice of gauge in Eqn. 1.1 is called a gauge symmetry.

One important consequence of gauge symmetry comes from the application of Noether's theorem, which states that for every global transformation under which the Lagrangian density is invariant, there exists a conserved quantity. If $\mathcal{L}(\Psi(x), \partial_\mu \Psi(x))$ is invariant under the transformation of the wave function $\Psi(x) \rightarrow \Psi'(x)$, where $\Psi'(x) = \Psi(x) + \delta\Psi(x)$, then there exists a conserved current

$$\partial_\mu \left(\frac{\partial \mathcal{L}(x)}{\partial(\partial_\mu \Psi(x))} \delta\Psi(x) \right) = 0 \quad (1.2)$$

In classical mechanics, the conservation of linear momentum, angular momentum, and energy follows from translational invariance, rotational variance, and invariance under translations in time [2]. Likewise, charge conservation can be shown to arise from the invariance of the Dirac Lagrangian density $\mathcal{L}_{\text{Dirac}} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi$ under the particle wavefunction's phase transformation, $\Psi'(x) = \exp(ie\chi)\Psi(x)$. Thus Noether's theorem establishes a correspondence between a gauge symmetry and a conserved internal property (e.g. charge or momentum).

1.2.2 Local gauge symmetries

Interactions between particles arise if we modify the wave function with a phase transformation $\Psi'(x) = \exp(ie\chi)\Psi(x)$, and allow the phase χ to be a function of spacetime [2]. A wave function of the form

$$\Psi'(x) = \exp(ie\chi(x))\Psi(x) \quad (1.3)$$

can be verified to *not* be a solution to the Dirac equation for free particles: $(i\gamma^\mu \partial_\mu - m)\Psi(x) = 0$. This necessitates a modified Dirac equation, where the derivative takes into account that the vector field $V(x)$ needs to be compared at two displaced space-time points in a curvilinear coordinate system:

$$\mathcal{D}_\mu \equiv \lim_{\Delta x^\mu \rightarrow 0} \frac{V_{||}(x + \Delta x) - V(x)}{\Delta x^\mu} \quad (1.4)$$

We define a covariant derivative,

$$D_\mu = \partial_\mu + ieA_\mu \quad (1.5)$$

where $A_\mu(x)$ is a 4-vector potential. Thus the modified Dirac equation reads:

$$(i\gamma^\mu \mathcal{D}_\mu - m) \Psi(x) = 0 \quad (1.6)$$

The simultaneous gauge transformation $A'_\mu(x) = A_\mu(x) - \partial_\mu \chi(x)$ and wavefunction transformation $\Psi'(x) = \exp(ie\chi(x))\Psi(x)$ leaves the covariant-derivative form of the Dirac equation (Eqn 1.1) invariant.

The generalization of this result is as follows: if a theory is invariant for unitary transformations U of the particle states according to

$$\Psi' = U\Psi \quad (1.7)$$

One must define a derivative of the form

$$D^\mu = \partial^\mu + igB^\mu \quad (1.8)$$

to keep the theory invariant under Eqn. 1.7. The four-potential B^μ represents the interacting four-potential which must be added to keep the theory invariant.

In the case of the Standard Model, the theory is built around the gauge transformations $G = SU(3) \times SU(2) \times U(1)$. $SU(3)$ is associated to the strong force (subscripted C); $SU(2)$ is associated to the weak force (subscripted L); and $U(1)$ is hypercharge (subscripted Y). The gauge-covariant derivative is

$$\mathcal{D}_\mu = \partial_\mu - ig'B_\mu \frac{Y}{2} - igW_\mu^\alpha \frac{\tau_a}{2} - ig_s G_\mu^k \frac{\lambda_k}{2} \quad (1.9)$$

- In the $U(1)_Y$ term, B_μ is the weak hypercharge field.
- In the $SU(2)_L$ term, $W_\mu(x) = (W_\mu^1(x), W_\mu^2(x), W_\mu^3(x))$ are a triplet of four-potentials. $\tau/2$ are the Pauli matrices, generators of the $SU(2)$ transformation.

- In the $SU(3)_C$ term, the gluon (color) field is G_μ . λ_k are the Gell-Man matrices, generators of the $SU(3)$ transformation.

The invariance of the Standard Model under $SU(3)_C \times SU(2)_L \times U(1)_Y$ requires massless fermions and massless force carriers.

1.3 The Higgs Mechanism

To introduce mass into the theory, i.e. to change the propagation of the gauge particles and all the fermions, the physical vacuum cannot have all the symmetries of the Standard Model Lagrangian [2]. The symmetries of the physical vacuum must be spontaneously broken, without affecting gauge invariance in the Lagrangian. The Higgs mechanism proposes the existence of a scalar field, or fields, with nonzero vacuum expectation values, which reduce the gauge symmetries of the physical vacuum from $SU(3)_C \times SU(2)_L \times U(1)_Y$ down to $SU(3)_C \times U(1)_{EM}$.

The Higgs field interacts with the gauge bosons and fermions throughout space, impeding their free propagation. The resulting broken symmetry correctly predicts the mass ratio of the neutral (Z) and charged (W) massive electroweak bosons, and predicts that at least one physical degree of freedom in the Higgs field is a particle degree of freedom, called the Higgs boson. The location of the minimum of the Higgs potential can be constrained from previously measured Standard Model parameters, but the shape of the mass distribution of the Higgs boson must be experimentally measured.

The minimal choice of Higgs field comes from the breaking of $SU(2)_L \times U(1)_Y$ down to $U(1)_{EM}$. The smallest $SU(2)$ multiplet is the doublet. The existence of three massive electroweak bosons leads the Higgs sector to have at least three degrees of

freedom. The minimal single-doublet complex scalar Higgs field is

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^+(x) + i\phi_2^+(x) \\ \phi_1^0(x) + i\phi_2^0(x) \end{pmatrix} \quad (1.10)$$

where ϕ_1^+ , ϕ_2^+ , ϕ_1^0 , and ϕ_2^0 are real (four degrees of freedom). By convention, the nonzero vacuum expectation value is assigned to ϕ_1^0 .

The minimal self-interacting Higgs potential that is invariant under $SU(2)_L \times U(1)_Y$ is given by

$$V(\Phi^\dagger \Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad \mu^2 > 0, \lambda > 0 \quad (1.11)$$

where λ is the coupling strength of the four-point Higgs interaction. The potential energy is minimized at

$$\Phi_{\min} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad \text{where } v = \sqrt{\mu^2/\lambda} \quad (1.12)$$

Choosing a fixed orientation of $\langle \Phi \rangle$ out of a continuous set of possible ground states spontaneously breaks the symmetry of the physical vacuum, as illustrated in Fig 1.2.

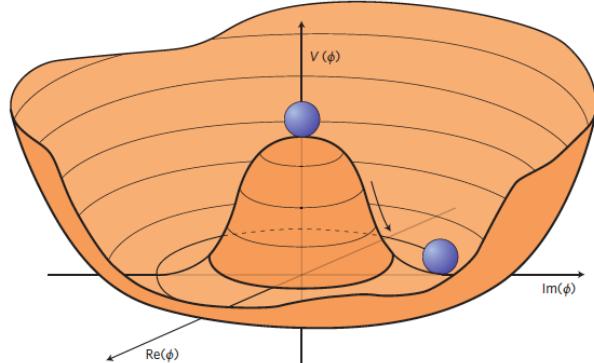


Figure 1.2: An illustration of the Higgs potential [3]. Choosing any of the points at the bottom of the potential breaks spontaneously the rotational $U(1)$ symmetry.

The excitations of the Higgs field with respect to the minimum Φ_{\min} are parameterized by

$$\Phi(x) = \exp(i\boldsymbol{\xi}(x) \cdot \boldsymbol{\tau}) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.13)$$

Three degrees of freedom are coupled directly to the electroweak gauge bosons; this is often referred to as the gauge bosons “eating” the Goldstone bosons to form the longitudinal polarizations of the massive spin-1 boson states. The $H(x)$ excitation is in the radial direction and corresponds to the free particle state of the Higgs boson.

1.4 Two-Higgs Doublet Models

One of the simplest possible extensions to the Standard Model is adding a doublet to the minimal Higgs sector of the Standard Model, which is a $SU(2)_L$ doublet H with hypercharge $Y = +\frac{1}{2}$, denoted here as $H \sim 2_{+1/2}$. These extensions are found in several theories such as supersymmetry. A general 2HDM can be extended with a light scalar (2HDM+S) to obtain a rich set of exotic Higgs decays [4].

The charges of the Higgs fields are chosen to be $H_1 \sim 2_{-1/2}$ and $H_2 \sim 2_{+1/2}$, which acquire vacuum expectation values $v_{1,2}$ which are assumed to be real and aligned [4]. Expanding about the minima yields two complex and four real degrees of freedom:

$$H_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} v_1 + H_{1,R}^0 + iH_{1,I}^0 \\ H_{1,R}^- + iH_{1,I}^- \end{pmatrix} \quad (1.14)$$

$$H_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} H_{2,R}^+ + iH_{2,I}^+ \\ v_2 + H_{2,R}^0 + iH_{2,I}^0 \end{pmatrix} \quad (1.15)$$

The charged scalar and pseudoscalar mass matrices are diagonalized by a rotation angle β , defined as $\tan \beta = v_2/v_1$. One charged (complex) field and one neutral pseudoscalar combination of $H_{1,2,I}^0$ are eaten by the SM gauge bosons after electroweak

symmetry breaking [4]. The other complex field yields two charged mass eigenstates H^\pm , which are assumed to be heavy. The remaining three degrees of freedom yield one neutral pseudoscalar mass eigenstate

$$A = H_{1,I}^0 \sin \beta - H_{2,I}^0 \cos \beta \quad (1.16)$$

and two neutral scalar mass eigenstates (where $-\pi/2 \leq \alpha \leq \pi/2$)

$$\begin{pmatrix} h \\ H^0 \end{pmatrix} = \begin{pmatrix} -\sin \alpha & \cos \alpha \\ \cos \alpha & \sin \alpha \end{pmatrix} \begin{pmatrix} H_{1,R}^0 \\ H_{2,R}^0 \end{pmatrix} \quad (1.17)$$

We assume that the 2HDM is near or in the decoupling limit: $\alpha \rightarrow \pi/2 - \beta$, where the lightest state in the 2HDM is h , which we identify as the 125 GeV Higgs particle [4]. In this limit, the fermion couplings of h become identical to the Standard Model Higgs, while the gauge boson couplings are very close to Standard Model-like for $\tan \beta \gtrsim 5$. All of the properties of h are determined by just two parameters: $\tan \beta$ and α , and the fermion couplings to the two Higgs doublets.

2HDM can be extended by a scalar singlet (2HDM+S) [4]:

$$S = \frac{1}{\sqrt{2}}(S_R + iS_I) \quad (1.18)$$

If this singlet only couples to the Higgs doublets $H_{1,2}$ and has no direct Yukawa couplings, all of its couplings to SM fermions result from mixing with $H_{1,2}$. Under these simple assumptions, exotic Higgs decays $h \rightarrow ss \rightarrow X\bar{X}Y\bar{Y}$ or $h \rightarrow aa \rightarrow X\bar{X}Y\bar{Y}$, and $h \rightarrow aZ \rightarrow X\bar{X}Y\bar{Y}$ are permitted, where $s(a)$ is a (pseudo)scalar mass eigenstate mostly composed of $S_R(S_I)$, and X, Y are Standard Model fermions or gauge bosons. There are two pseudoscalars in the 2HDM+S, and the mostly singlet-like pseudoscalar can be chosen to be the one lighter than the SM-like Higgs. For

$m_a < m_h - m_Z \sim 35$ GeV, the exotic Higgs decay $h \rightarrow Za$ is possible, and for $m_a < m_h/2 \approx 63$ GeV, the exotic Higgs decay $h \rightarrow aa$ is possible.

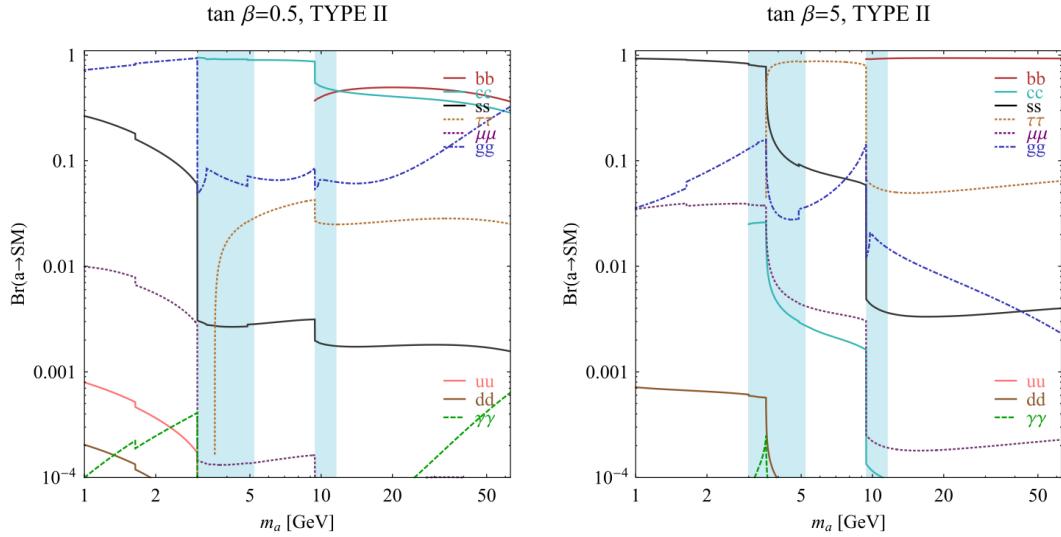


FIG. 7 (color online). Branching ratios of a singletlike pseudoscalar in the 2HDM + S for type-II Yukawa couplings. Decays to quarkonia likely invalidate our simple calculations in the shaded regions.

Figure 1.3: Branching ratios of a singlet-like pseudoscalar in Type II 2HDM+S for $\tan\beta = 0.5$ (*left*) and $\tan\beta = 5$ (*right*) from [4], showing the dependence of the branching ratios on $\tan\beta$, as well as the prominence of the branching ratios to bb and $\tau\tau$, the channels searched for in the analysis presented here.

In 2HDM, and by extension 2HDM+S, there are four types of fermion couplings commonly discussed in the literature that forbid flavour-changing neutral currents at tree level [4]. These are referred to as Type I (all fermions couple to H_2), Type II (MSSM-like, d_R and e_R couple to H_1 , u_R to H_2), Type III (lepton-specific, leptons and quarks couple to H_1 and H_2 respectively) and Type IV (flipped, with u_R , e_R coupling to H_2 and d_R to H_1). The exact branching ratios of the pseudoscalars to Standard Model particles vary depending on the 2HDM+S model and the value of $\tan\beta$ (e.g. Fig. 1.3).

1.5 Two Real Singlet Model

The two real singlet model (TRSM) adds two real singlet degrees of freedom to the Standard Model. These are written as two real singlet fields S and X . Depending on the vacuum expectation values acquired by the scalars, different phases of the model can be realized [5]. To reduce the number of free parameters, two discrete \mathbb{Z}_2 symmetries are introduced. The fields are decomposed as

$$\Phi = \begin{pmatrix} 0 \\ \frac{\phi_h + v}{\sqrt{2}} \end{pmatrix}, S = \frac{\phi_S + v_S}{\sqrt{2}}, X = \frac{\phi_X + v_X}{\sqrt{2}} \quad (1.19)$$

To achieve electroweak-breaking symmetry, $v = v_{SM} \sim 246$ GeV is necessary. If the vacuum expectation values $v_S, v_X \neq 0$ the \mathbb{Z}_2 are spontaneously broken, and the fields $\phi_{h,S,X}$ mix into three physical scalar states. This is called the broken phase and leads to the most interesting collider phenomenology.

The mass eigenstates $h_{1,2,3}$ are related to the fields $\phi_{h,S,X}$ through a 3×3 orthogonal mixing matrix denoted R . The mass eigenstates are assumed to be ordered $M_1 \leq M_2 \leq M_3$. R is parameterized by the three mixing angles θ_{hS} , θ_{hX} , θ_{SX} . The nine parameters of the scalar potential can be expressed in terms of the three physical Higgs masses, the three mixing angles, and the three vacuum expectation values.

After fixing one of the Higgs masses to the mass of the observed Higgs boson, and fixing the Higgs doublet vacuum expectation value to its Standard Model value, there are seven remaining free parameters of the TRSM [5].

In one benchmark scenario of TRSM [5], the heaviest scalar state h_3 is identified with the 125 GeV Higgs, h_{125} , and it can decay asymmetrically $h_{125} \rightarrow h_1 h_2$, which we also denote $h \rightarrow a_1 a_2$ to highlight the similarity with the symmetric decay $h \rightarrow aa$ typically interpreted in 2HDM+S as discussed. The parameter values in TRSM are chosen such that the coupling of h_3 to Standard Model particles are nearly identical to the Standard Model predictions.

In benchmark scenario 1 (benchmark plane 1, or BP1) (Fig. 1.4) [5], the maximal branching ratios for $h_3 \rightarrow h_1 h_2$ reach up to $7 - 8\%$ which translates into a signal rate of around 3 pb. These maximal branching ratios are reached in the intermediate mass state for h_2 , $M_2 \sim 60 - 80$ GeV. For $M_2 < 40$ GeV, although phase space opens up significantly for light decay products, the branching ratio becomes smaller.

If the decay channel $h_2 \rightarrow h_1 h_1$ is kinematically open (i.e. $M_2 > 2M_1$), it is the dominant decay mode leading to a significant rate for the $h_1 h_1 h_1$ final state, in a “cascade” decay. In BP1, $BR(h_2 \rightarrow h_1 h_1) \simeq 100\%$ above the red line in Fig. 1.4. If, in addition, $M_1 \gtrsim 10$ GeV, the h_1 decays dominantly to $b\bar{b}$ leading to a sizable rate for the $b\bar{b}b\bar{b}b\bar{b}$ final state as shown in Fig. 1.4 (*bottom right*).

If the $h_2 \rightarrow h_1 h_1$ decay is kinematically closed (i.e. $M_2 < 2M_1$), both scalars decay directly to Standard Model particles, with branching ratios identical to a Standard Model-like Higgs boson, i.e. with the $b\bar{b}b\bar{b}$ final state dominating, as shown in Fig. 1.4 (*bottom left*), while at smaller masses, combinations with τ leptons and eventually final states with charm quarks and muons become relevant [5].

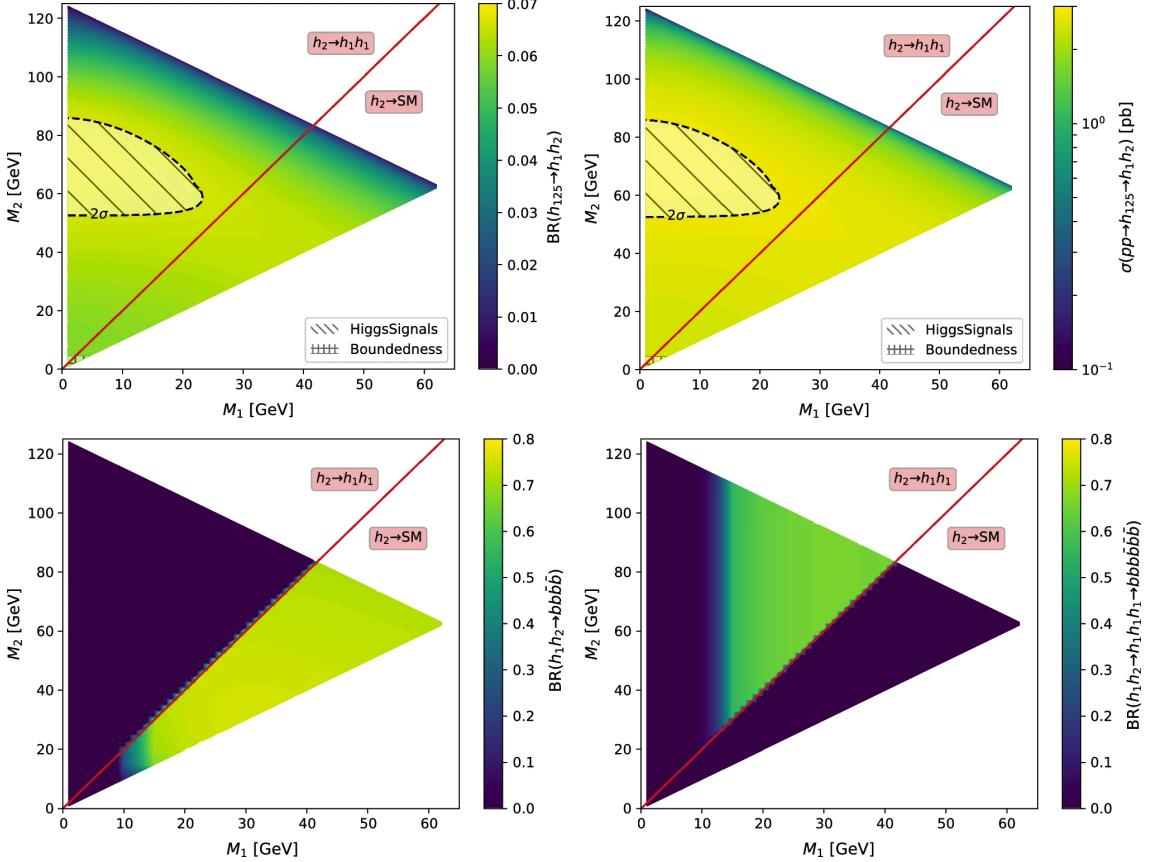


Figure 1.4: Benchmark plane BP1 for benchmark scenario 1 from [5], for the decay signature $h_{125} \rightarrow h_1 h_2$ with $h_{125} \equiv h_3$, defined in the (M_1, M_2) plane. The color code shows $\text{BR}(h_3 \rightarrow h_1 h_2)$ (*top left*) and the 13 TeV LHC signal rate for $pp \rightarrow h_3 \rightarrow h_1 h_2$ (*top right*). The red line separates the region $M_2 > 2M_1$, where $\text{BR}(h_2 \rightarrow h_1 h_1) \sim 100\%$, from the region $M_2 < 2M_1$, where $\text{BR}(h_2 \rightarrow F_{SM}) \sim 100\%$. The *bottom left* and *right* show the branching ratio of the $h_1 h_2$ into (respectively) $b\bar{b} b\bar{b}$, and through a $h_2 \rightarrow h_1 h_1$ cascade to $b\bar{b} b\bar{b} b\bar{b}$. The hatched region indicates where the decay rate slightly exceeds the 2σ upper limit inferred from the LHC Higgs rate measurements, though the region depends on the parameter choices and experimental searches should cover the whole mass range.

Chapter 2

The Large Hadron Collider and the CMS Experiment

2.1 The Large Hadron Collider

The CERN Large Hadron Collider (LHC) is an accelerator complex consisting of a 27-kilometer ring of superconducting magnets with accelerating structures to boost the energy of particles, which collide at a center-of-mass energy of up to 14 TeV. The beams inside the LHC are made to collide at four locations around the accelerator ring, at the locations of four particle detectors: ATLAS, CMS, ALICE, and LHCb. An aerial view of the four major experiments' locations is shown in Fig. 2.1 [6]. ATLAS and CMS are the two general-purpose detectors with broad physics programmes spanning Standard Model measurements and searches for signatures of new physics [7] [8]. The two experiments use different technical solutions and different magnet system designs. ALICE is a general-purpose detector dedicated to measuring LHC heavy-ion collisions, and is designed to address the physics of strongly interacting matter, and the properties of quark-gluon plasma [9]. The LHCb experiment specializes in investigating CP violation through measuring the differences in matter and

antimatter, by using a series of subdetectors to detect mainly forward particles close to the beam direction [10].

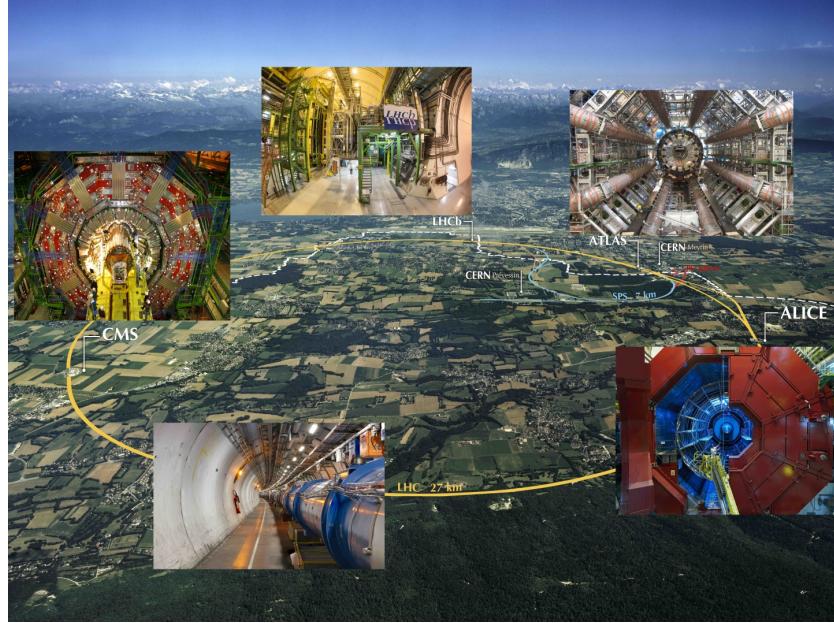


Figure 2.1: Aerial view of the Large Hadron Collider (LHC) spanning the border of France and Switzerland, and the four major experiments located around the ring: CMS (Compact Muon Solenoid), LHCb (LHC beauty), ATLAS (A Toroidal LHC Apparatus), and ALICE (A Large Ion Collider Experiment). [6]

The accelerator complex at CERN is a succession of machines that accelerate particles in stages until they reach their final energy of 6.5 TeV per beam [11] [12]. In Linear accelerator 4 (Linac4), negative hydrogen ions (hydrogen atoms with an additional electron) are accelerated to 160 MeV, and stripped of their two electrons, leaving only protons, before entering the Proton Synchrotron Booster (PSB). These protons are accelerated to 2 GeV, then to 26 GeV in the Proton Synchrotron (PS), and 450 GeV in the Super Proton Synchrotron (SPS). The protons are transferred to the two beam pipes of the LHC, where one beam circulates clockwise and the other counterclockwise. Each LHC ring takes 4 minutes and 20 seconds to fill, and it takes about 20 minutes for the protons to reach their maximum energy. During normal operating conditions, beams circulate for many hours inside the LHC ring.

2.2 Luminosity and pileup

The number of events generated per second by the LHC collisions is given by

$$N_{event} = \mathcal{L} \cdot \sigma_{event} \quad (2.1)$$

where σ_{event} is the cross-section for the event under study, and \mathcal{L} the machine luminosity. The machine luminosity is measured in units of $\text{cm}^{-2} \text{ s}^{-1}$, and depends only on the beam parameters, and can be written for a Gaussian beam distribution as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (2.2)$$

where the parameters are as defined, along with some example typical nominal values in Phase-1 of the LHC [13] [14]:

- N_b is the number of particles per bunch ($N_b \approx 1.15 \times 10^{11}$ protons per bunch)
- n_b is the number of bunches per beam (maximum 2808),
- f_{rev} is the revolution frequency (≈ 11 kHz),
- γ_r is the relativistic gamma factor,
- ϵ_n is the normalized transverse beam emittance (area in a transverse plane occupied by the beam particles),
- β^* is the beta function at the collision point ($\beta^* = 0.55$ m),
- and F is the geometric luminosity reduction factor due to the crossing angle at the interaction points ($F \approx 0.84$ for Phase-1. Note that complete overlap would give $F = 1$).

Peak luminosity at interaction points 1 and 5 reach values of $\sim 1.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, with peak luminosity per bunch crossing reaching $\sim 3.56 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.

Per Eqn. 2.1, the integrated luminosity over time is proportional to the number of events produced, and the size of LHC datasets is commonly presented in terms of integrated luminosity. Collider operation aims to optimize the integrated luminosity. Thus the exploration of rare events in the LHC collisions requires both high beam energies and high beam intensities.

The LHC’s nominal beam luminosities are sufficiently large for multiple proton-proton collisions to occur in the same time window of 25 nanoseconds in which proton bunches collide [15]. These multiple collisions will lead to particle interactions overlapping in the detector. To measure a proton-proton collision, the single collision must be separated from overlapping collisions, which are called “pileup” collisions. A distribution of pileup in the data-taking years 2016-2018 is shown in Fig. 2.2. The pileup is defined as the average number of pp collisions per bunch crossing.

CMS reports an inelastic pp cross section of $\sigma_{\text{inel}} = 68.6$ millibarns at a center-of-mass energy of $\sqrt{s} = 13$ TeV [16], which can be used to estimate pileup as follows:

$$\text{Pileup} = \frac{\mathcal{L} \times \sigma_{\text{inel}}}{n_b \cdot f} \quad (2.3)$$

With the example values above, pileup can be estimated to be ~ 22 .

Thus, higher luminosities create more intense pileup conditions, posing a greater challenge to detector performance and particle reconstruction and identification.

2.3 The High-Luminosity LHC

The High-Luminosity LHC (HL-LHC) is a major upgrade of the LHC scheduled to take place in the late 2020s, that will increase the instantaneous luminosity by a factor of five beyond the original design value, and the integrated luminosity by a factor of ten [17]. This will be accomplished through accelerator technological advances: for instance, reduction of the interaction point β^* from 0.55 m down to 0.15

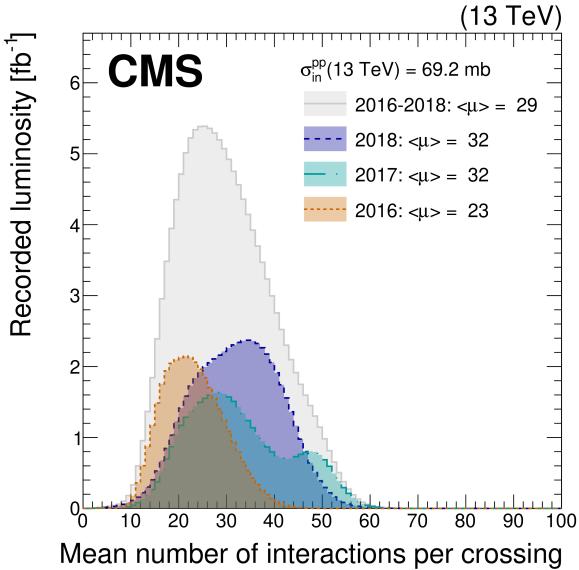


Figure 2.2: Distribution of the mean number of inelastic collisions per bunch crossing (pileup) in data [15], for proton-proton collisions in 2016 (*dotted orange*), 2017 (*dotted light blue*), 2018 (*dotted dark blue*), and integrated over 2016-2018 (*solid grey*). A cross-section of inelastic proton-proton collisions of 69.2 mbarns is assumed. In the running conditions of the High-Luminosity LHC, pileup will reach unprecedented levels of up to 200 per bunch crossing [17].

m by installation of new final-focusing magnets, and improvements in the geometric luminosity loss factor $F \approx 1$ through the installation of crab cavities that optimize the orientation of colliding bunches. A further discussion of the HL-LHC upgrades for the CMS detector follows in Chapter 4.

2.4 The CMS Detector

The Compact Muon Solenoid (CMS) experiment was conceived to study proton-proton and lead-lead collisions at a center-of-mass energy of 14 TeV (5.5 TeV nucleon-nucleon) and at luminosities up to $10^{34}\text{ cm}^{-2}\text{ s}^{-1}$ ($10^{27}\text{ cm}^{-2}\text{ s}^{-1}$) [18] [19]. Starting from the beam interaction region at the center of the CMS detector, particles first pass through a silicon pixel and strip tracker, in which charged-particle trajectories (tracks) and origins (vertices) are reconstructed from signals (hits) in the sensitive

layers. The tracker is immersed in a high-magnetic-field superconducting solenoid that bends the trajectories of charged particles, allowing the measurement of their electric charge and momenta. Electrons and photons are then absorbed in an electromagnetic calorimeter (ECAL) comprised of lead-tungstate scintillating-crystals. The corresponding electromagnetic showers are detected as clusters of energy recording in neighboring cells, from which the direction and energy of the particles can be determined. Charged and neutral hadrons may initiate a hadronic shower in the ECAL as well, which is then fully absorbed in the hadron calorimeter (HCAL). The resulting clusters are used to estimate their direction and energies. Muons and neutrinos pass through the calorimeters with little to no interactions. Neutrinos escaped undetected; muons produce hits in additional gas-ionization chamber muon detectors housed in the iron yoke of the flux-return. A sketch of example particle interactions in a transverse slice of the CMS detector is shown in Fig. 2.3. The collision data is recorded with the use of the Level-1 (L1) trigger (discussed separately in 3.1), high-level trigger (HLT), and data acquisition systems ensuring high efficiency in selecting physics events of interest.

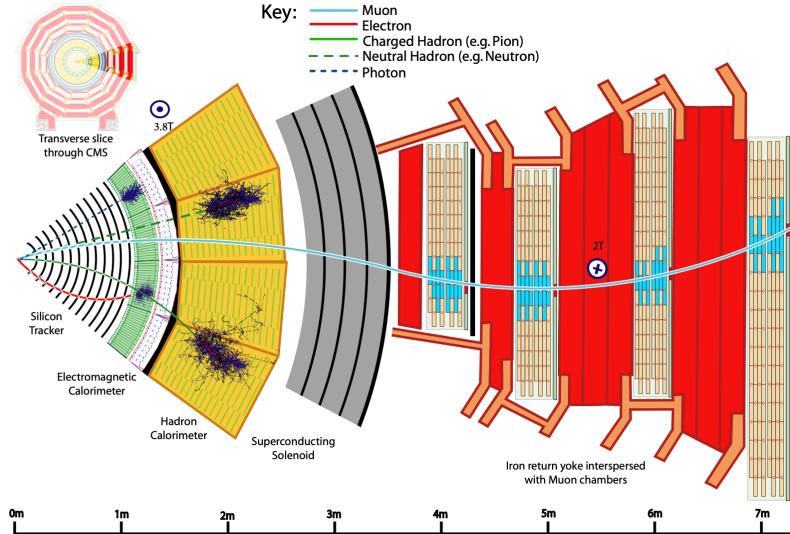


Figure 2.3: Sketch of particle trajectories of muons, electrons, charged and neutral hadrons, and photons in a transverse cross-section of the CMS detector [19].

CMS uses a right-handed coordinate system [18]. The origin is centered at the nominal collision point inside the experiment. The x axis points towards the center of the LHC, and the y axis points vertically upwards. The z axis points along the beam direction. The azimuthal angle, ϕ , is measured from the x axis in the x - y plane, and the radial coordinate in this plane is denoted by r . The polar angle, θ , is measured from the z axis. The pseudorapidity, η , is defined as $\eta = -\ln \tan(\theta/2)$. The momentum and energy transverse to the beam direction, denoted by p_T and E_T respectively, are computed from the x and y components. The momentum imbalance in the transverse plane is called the missing transverse momentum, and its magnitude is denoted by E_T^{miss} .

2.5 Sub-detectors of CMS

This section details the sub-detectors of CMS that operate to identify and precisely measure muons, electrons, photons, and jets over a large energy range.

2.5.1 Inner tracking system

The CMS Tracker performs robust tracking and detailed vertex reconstruction in the 4 T magnetic field of the superconducting solenoidal magnet. The primary sensors used in the tracker are p^+ on n -bulk devices, which allow high voltage operation and are radiation-resistant [20] [21]. The active envelope of the CMS Tracker extends to a radius of 115 cm, over a length of approximately 270 cm on each side of the interaction point [20]. Charged particles in the region $|\eta| \lesssim 1.6$ benefit from the full momentum measurement precision. In this region, a charged particle with p_T of 1000 GeV has a sagitta of $\sim 195 \mu\text{m}$. The Tracker acceptance extends further to $|\eta| = 2.5$, with a reduced radius of approximately 50 cm.

The high magnetic field of CMS causes low p_T charged particles to travel in helical

trajectories with small radii. The majority of events contain particles with a steeply falling p_T spectrum, resulting in a track density which rapidly decreases at higher radii.

A schematic view of the current Phase-1 CMS tracker [22], including the pixel detector, is shown in Fig. 2.4. The Phase-1 pixel detector consists of three barrel layers (BPIX) at radii of 4.4 cm, 7.3 cm, and 10.2 cm, and two forward/backward disks (FPIX) at longitudinal positions of ± 34.5 cm and ± 46.5 cm, and extending in radius from about 6 cm to 15 cm. These pixelated detectors produce 3D measurements along the paths of charged particles with single hit resolutions between 10-20 μm .

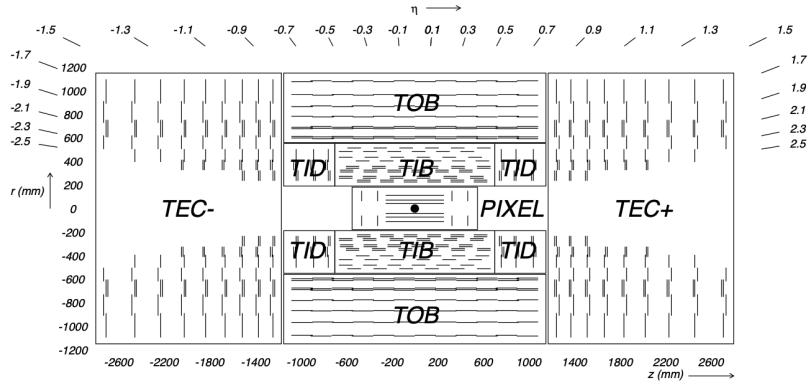


Figure 2.4: Cross section of the current Phase-1 CMS tracker [22]. Each line represents a detector module. Double lines indicate back-to-back modules which deliver two-dimensional (stereo) hits in the strip tracker.

After the pixel and on their way out of the tracker, particles pass through the silicon strip tracker which reaches out to a radius of 130 cm (Fig. 2.4). The sensor elements in the strip tracker are single-sided $p\text{-on-}n$ type silicon micro-strip sensors [18]. The silicon strip detector consists of four inner barrel (TIB) layers assembled in shells, with two inner endcaps (TID), each composed of three small discs. The outer barrel (TOB) consists of six concentric layers. Two endcaps (TEC) close off the tracker on either end.

2.5.2 ECAL

The electromagnetic calorimeter (ECAL) of CMS measures electromagnetic energy deposits with high granularity. One of the driving criteria in the design was the capability of detecting the Standard Model Higgs boson decay to two photons (in fact, the channel in which the 125 GeV Higgs boson was discovered at CMS). ECAL is a hermetic homogeneous calorimeter comprised of 61,200 lead tungstate (PbWO_4) crystals mounted in the central barrel, with 7,324 crystals in each of the two endcaps [18]. A preshower detector is located in front of the endcap crystals. Avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum phototriodes (VPTs) in the endcaps.

The design of the ECAL is driven by the behaviour of high-energy electrons, which predominantly lose energy in matter via bremsstrahlung, and high-energy photons by e^+e^- pair production. The characteristic amount of matter traversed for these interactions is the radiation length X^0 , usually measured in units of g cm^{-2} . The radiation length is also the mean distance over which a high-energy electron loses all but $1/e$ of its energy via bremsstrahlung [23]. Thus high granularity in η and ϕ , and the length of the ECAL crystals, is designed to capture the shower of e/γ produced by electrons and photons.

The barrel part of the ECAL (EB) covers the pseudorapidity range $|\eta| < 1.479$ [18]. The barrel granularity is 360-fold in ϕ and (2×85) -fold in η . The crystal cross-section corresponds to approximately 0.0174×0.0174 in $\eta - \phi$ or $22 \times 22 \text{ mm}^2$ at the front face of the crystal, and $26 \times 26 \text{ mm}^2$ at the rear face. The crystal length is 230 mm, corresponding to $25.8 X_0$.

The ECAL read-out acquires the signals of the photodetectors [18]. At each bunch crossing, digital sums representing the energy deposit in a trigger tower, comprising 5×5 crystals in $\eta \times \phi$, are generated and sent to the Level-1 trigger system (detailed in Section 3.1).

2.5.3 HCAL

The hadronic calorimeter (HCAL) of CMS measures hadronic energy, which is key to characterizing the presence of apparent missing transverse energy which could arise from hadron jets and neutrinos or exotic particles [18]. A schematic of the components of HCAL are shown in Fig. 2.5. The HCAL barrel (HB) and endcaps (HE) are located outside of the tracker and the ECAL, spanning a radius of 1.77 m (outer extent of ECAL) up to 2.95 m (inner extent of the magnet coil). An outer hadron calorimeter (HO) is placed outside the solenoid to complement the barrel calorimeter. Beyond $|\eta| = 3$, the forward hadron calorimeter (HF) at 11.2 m from the interaction point extend the pseudorapidity coverage to $|\eta| = 5.2$.

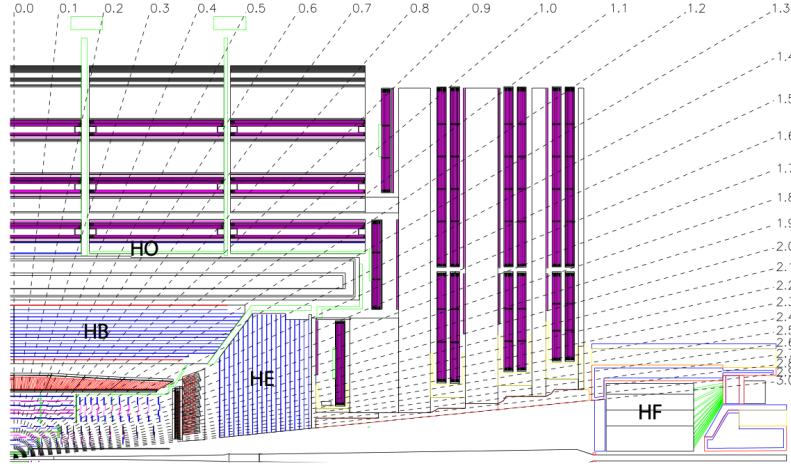


Figure 2.5: Longitudinal view of the CMS detector showing the hadron calorimeter barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeters from [18].

The HB is a sampling calorimeter covering the pseudorapidity range $|\eta| < 1.3$ [18]. It consists of 36 identical azimuthal wedges which form two half-barrels (HB+ and HB-), with a segmentation of $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$. The HE covers pseudorapidity $1.3 < |\eta| < 3$. The HB and endcap HE calorimeters are sampling calorimeters which use brass as the absorber and plastic scintillator as the active material. Light from the plastic scintillator is wavelength-shifted and captured in optic fibers which are

read out by front-end electronics [24].

In the central pseudorapidity region, the combined stopping power of EB plus the HB is insufficient to contain hadron showers [18]. To ensure adequate sampling depth, the hadron calorimeter is extended with a tail catcher, the HO. The size and position of the tiles are designed to roughly map the layers of the HB to make towers with the same granularity of 0.087×0.087 in η and ϕ . HO uses the same active material as the HB and HE calorimeters, but uses the steel return yoke and magnet material of CMS as absorbers [24].

The HF is a Cherenkov calorimeter based on a steel absorber and quartz fibers which run longitudinally through the absorber and collect Cherenkov light, primarily from the electromagnetic component of showers developed in the calorimeter [24]. Photomultiplier tubes are used to collect light from the quartz fibers. The HF is designed to survive in the harsh radiation conditions and high particle flux of the forward region. On average, 760 GeV per proton-proton interaction is deposited into the two forward calorimeters, compared to only 100 GeV for the rest of the detector [18]. Furthermore, this energy has a pronounced maximum at the highest rapidities.

2.5.4 Muon detectors

The CMS muon system is designed to have the capability of reconstructing the momentum and charge of muons over the kinematic range of the LHC, since muons are a powerful handle on signatures of interesting processes over the high background rate of the LHC [18]. For instance, the decay of the Standard Model Higgs boson into ZZ , which in turn decay to 4 leptons, can be reconstructed with high 4-particle mass resolution if all the leptons are muons, since muons are less affected than electrons by radiative losses in the tracker material.

The muon system consists of a cylindrical barrel section and two planar endcap regions [18]. The barrel muon detector consists of drift tube (DT) chambers covering

the pseudorapidity region $|\eta| < 1.2$ (Fig. 2.6). The DTs can be used as tracking detectors due to the barrel region's characteristic low neutron-induced backgrounds, low muon rate, and relatively uniform 4T magnetic field contained in the steel yoke.

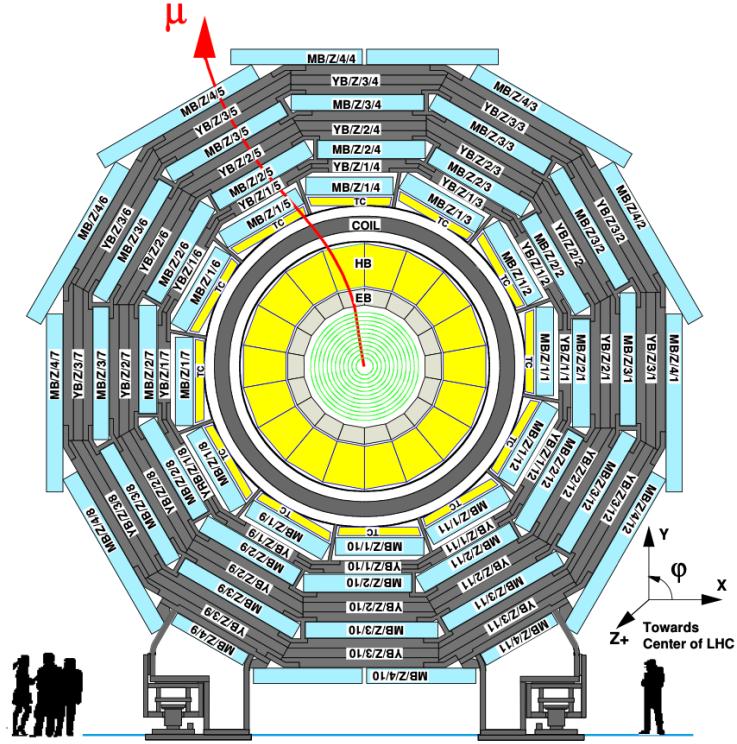


Figure 2.6: Layout of the CMS barrel muon drift tube (DT) chambers in one of the five wheels from [18]. The DTs are organized in 12 sectors of the yoke barrel (YB). In each of the 12 sectors of the yoke, there are 4 muon chambers per wheel (MB1, MB2, MB3, and MB4).

In the two endcap regions, the muon rates and background levels are high and the magnetic field is large and non-uniform [18]. Here, the muon system uses cathode strip chambers (CSCs) to identify muons between $0.9 < |\eta| < 2.4$. The cathode strips of each chamber run radially outwards and provide a precision measurement in the $r - \phi$ bending plane. The anode wires run approximately perpendicular to the strips and are read out in order to measure η and the beam-crossing time of a muon.

In addition to the DT and CSC, a dedicated trigger system consisting of resistive

plate chambers (RPCs) in the barrel and endcap regions provide a fast, independent, and highly-segmented trigger with a sharp p_T threshold over a large portion of the pseudorapidity range ($|\eta| < 1.6$) of the muon system [18]. RPCs have good time resolution but coarser position resolution compared to the DTs or CSCs. The RPCs also play a role in resolving ambiguities in reconstructing tracks from multiple hits in a chamber.

2.5.5 The High-Level Trigger

The HLT is implemented in software running on a large computer farm of fast commercial processors [25] [26]. The algorithms in HLT have access to full data from all CMS sub-detectors, including the tracker, with full granularity and resolution. The HLT reconstruction software is similar to what is used offline for full CMS data analysis. As a result, the HLT can calculate quantities with a resolution comparable to the final detector resolution, compared to the L1 Trigger. The HLT performs more computationally-intensive algorithms, such as combining tau-jet candidates in the calorimeter with high- p_T stubs in the tracker, to form a hadronic tau trigger. The maximum HLT input rate from the L1 Trigger is 100 kHz, and the HLT output rate is approximately 100 Hz.

The HLT contains trigger paths, each corresponding to a dedicated trigger [27]. A path consists of several steps implemented as software modules. Each HLT trigger path must be seeded by one or more L1 trigger bits: the first module always looks for a L1 seed, consisting of L1 bit(s) and L1 object(s). Each module performs a well-defined task such as unpacking (raw to digitized quantities), reconstruction of physics objects (electrons, muons, jet, missing transverse energy, etc.), making intermediate decisions that trigger more detailed reconstruction modules, and calculating the final decision for the trigger path. If an intermediate filter decision is negative, the rest of the path is not executed, and the trigger rejects the event.

2.5.6 Particle reconstruction

To build a description of the physics objects present in the particle collision, the basic elements from the detector layers (tracks and clusters of energy) are correlated to identify each particle in the final state. Measurements from different sub-detectors are combined to reconstruct the particle properties. This approach is called particle-flow (PF) reconstruction [19]. Key to the success of the PF reconstruction is the fine spatial granularity of the detector layers. Coarse-grained detectors can cause the signals from different particles to merge, especially within jets. However, if the subdetectors are sufficiently segmented to separate individual particles, it becomes possible to produce a global event description that identifies all physics objects with high efficiencies and resolution.

2.5.7 Data storage and computational infrastructure

The LHC generates over 15 petabytes (15 million gigabytes) of data every year, necessitating a flexible computing system that can be accessed by researchers working at the four main LHC experiments: ALICE, ATLAS, CMS, and LHCb. The Worldwide LHC Computing Grid (WLCG) [28] is a global collaboration of computer centers that links thousands of computers and storage systems in over 170 centers across 41 countries. These centers are arranged in “tiers”, and provide near real-time access to users processing, analyzing, and storing LHC data. One of the final stages of data analysis at LHC experiments is large-scale data processing taking place over distributing computing, for instance, with the use of Condor [29], a distributed, scalable, flexible batch processing system which accepts a computing job, allocates a resource to it, executes it, and returns the result back to a user transparently.

Chapter 3

The CMS Phase-1 Level-1 Trigger

3.1 The Phase-1 Level-1 Trigger

The design performance of the LHC corresponds to an instantaneous luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ with a 25 ns bunch crossing rate, giving an average pile-up (number of simultaneous events) of 25 per bunch crossing. The large number of minimum bias events per bunch crossing, combined with the small cross-sections of possible physics discovery signatures, necessitates a sophisticated event selection system for filtering this large event rate, as it is impossible to save all events. This data filtering system is implemented by CMS in two stages. The first stage is the Level-1 (L1) Trigger, which is deployed in custom electronic hardware systems and is responsible for reducing the event rate to around 100 kHz. The second stage is the High-Level Trigger (HLT) which is described in Section 2.5.5. This section describes the Phase-1 configuration of the Level-1 Trigger.

The L1 Trigger data flow of Phase-1 is shown in Fig. 3.1, with organization into the L1 calorimeter trigger, the L1 muon trigger, and the L1 global trigger.

The L1 calorimeter trigger begins with trigger tower energy sums formed by the ECAL, HCAL, and HF Trigger Primitive Generator (TPG) circuits from the indi-

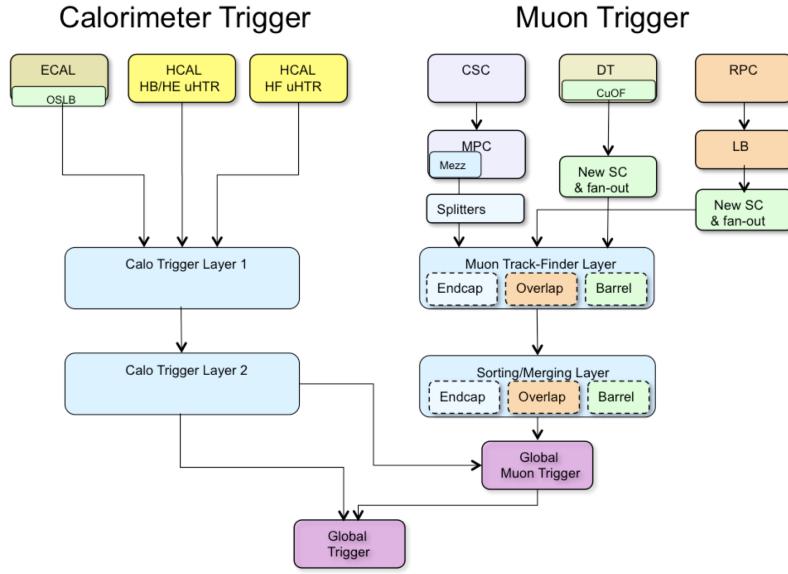


Figure 3.1: Dataflow for the Phase-1 Level-1 Trigger [CITE], which is implemented in custom hardware and is responsible for reducing the event rate from the LHC bunch crossing frequency of 400 MHz (bunch crossings every 25 ns) to a maximum rate of 100 kHz. In Phase-1, the Level-1 Trigger has access to information from the calorimeter and muon detectors.

vidual calorimeter cell energies. In the original configuration, the ECAL energies were accompanied by a bit indicating the transverse extent of the electromagnetic energy deposits, and the HCAL energies were accompanied by a bit indicating the presence of minimum ionizing energy [CITE]. Between Long Shutdowns 1 and 2 (LS1 and LS2), HF was upgraded to provide finer granularity information to the trigger, and the HCAL barrel and endcap front-end electronics were upgraded to provide high-precision timing information and depth segmentation information.

In the original design of the L1 calorimeter trigger, the trigger primitives are processed by the Regional Calorimeter Trigger (RCT, upgraded to Calo Layer 1 after LS2) which finds isolated and non-isolated electron/photon candidates. (At this stage, electrons/photons candidates are treated together since they cannot be definitively distinguished at this stage due to lack of tracking information in the L1 trigger.) The Global Calorimeter Trigger (GCT, upgraded to Calo Layer 2 after LS2) sorts

further the candidate electrons/photons, finds jets (classified as central, forward, and tau) using the E_T sums and performs calibration of the clustered jet energies, and calculates global quantities such as missing E_T . It sends the top four candidates of each type to the global trigger (GT).

Each of the L1 muon triggers has its own trigger logic. The RPC strips are connected to a Pattern Comparator Trigger (PACT), which forms trigger segments that are used to build tracks and calculate p_T . The RPC logic also provides some hit data to the CSC trigger system to resolve ambiguities caused by two muons in the same CSC. The CSCs form local charged tracks (LCTs) from the cathode strips, which are combined with the anode wire information. LCTs are combined into full muon tracks and assigned p_T values.

The Global Muon Trigger (GMT) sorts the RPC, DT, and CSC muon tracks, converts these tracks to the same η , ϕ , and p_T scale, and validates the muon sign. It improves the trigger efficiency by merging muon candidates that were detected in two complementary sub-systems (i.e. DT+RPC, or CSC+RPC). The GMT also contains logic to correlate the found muon tracks with an $\eta - \phi$ grid of quiet calorimeter towers to determine if the muons are isolated, as well as logic to remove duplicate candidates originating in the overlap regions from both DT and CSC systems. The final collection of muons are sorted based on their initial quality, correlation, and p_T , and the top four muons are sent to the Global Trigger.

Information from the GCT and GT are sent to the Global Trigger (GT), which makes the Level-1 Accept (L1A) decision to either discard or accept the bunch crossing. This is accomplished by sorting ranked trigger objects that are accompanied by positional information in η and ϕ , permitting the trigger to applying criteria with thresholds that can vary based on the location of the trigger objects, and/or to require trigger objects to be close to or opposite from each other. The GT L1A decision arrives at the detector front end with a $3.8\ \mu\text{s}$ latency after the interaction at a rate

which is required to be less than 100 kHz, and triggers a full readout of the detector for further processing.

Chapter 4

The Phase-2 Upgrade of CMS

4.1 High-Luminosity LHC and CMS

In order to sustain and extend the LHC’s physics discovery program and maintain operability for a decade or more, the LHC is undergoing a major upgrade to the High-Luminosity LHC (HL-LHC). In its final configuration, the HL-LHC will deliver a peak luminosity of $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, potentially leading to total integrated luminosity of 4000 fb^{-1} after ten years of operations, scheduled to begin in 2027 [30]. This integrated luminosity is about ten times the predicted luminosity reach of the LHC in its initial configuration. To maximize the discovery potential of this unprecedented amount of data, the CMS detector is undergoing Phase-2 upgrades in order to perform high-precision measurements and searches for physics beyond the Standard Model in the intense running conditions of the HL-LHC.

4.2 The Phase-2 Level-1 Trigger

To achieve the goals of the HL-LHC program and to ensure the collection of information-rich datasets in the HL-LHC, the Phase-2 upgrade of the CMS Level-1 Trigger [30] must be upgraded in conjunction with the CMS sub-detectors and their readouts, to

maintain physics selectivity. The HL-LHC will produce an intense hadronic environment corresponding to 200 simultaneous collisions per beam crossing, necessitating comprehensive upgrades of the trigger system outlined below.

To profit from the extended coverage and increased granularity of the upgraded CMS detector, the latency of the L1 trigger system (time available to produce a L1 Accept signal) will be increased significantly from $3.8 \mu\text{s}$ to $12.5 \mu\text{s}$, with an increased maximum output bandwidth of 750 kHz [30]. With the increased latency, in addition to information from calorimeters and muon detectors (as in the Phase-1 system), information from the new tracker and high-granularity endcap calorimeter can also be included at L1 for the first time. This is illustrated in the functional diagram of the architecture of the Phase-2 trigger system in Fig. 4.1.

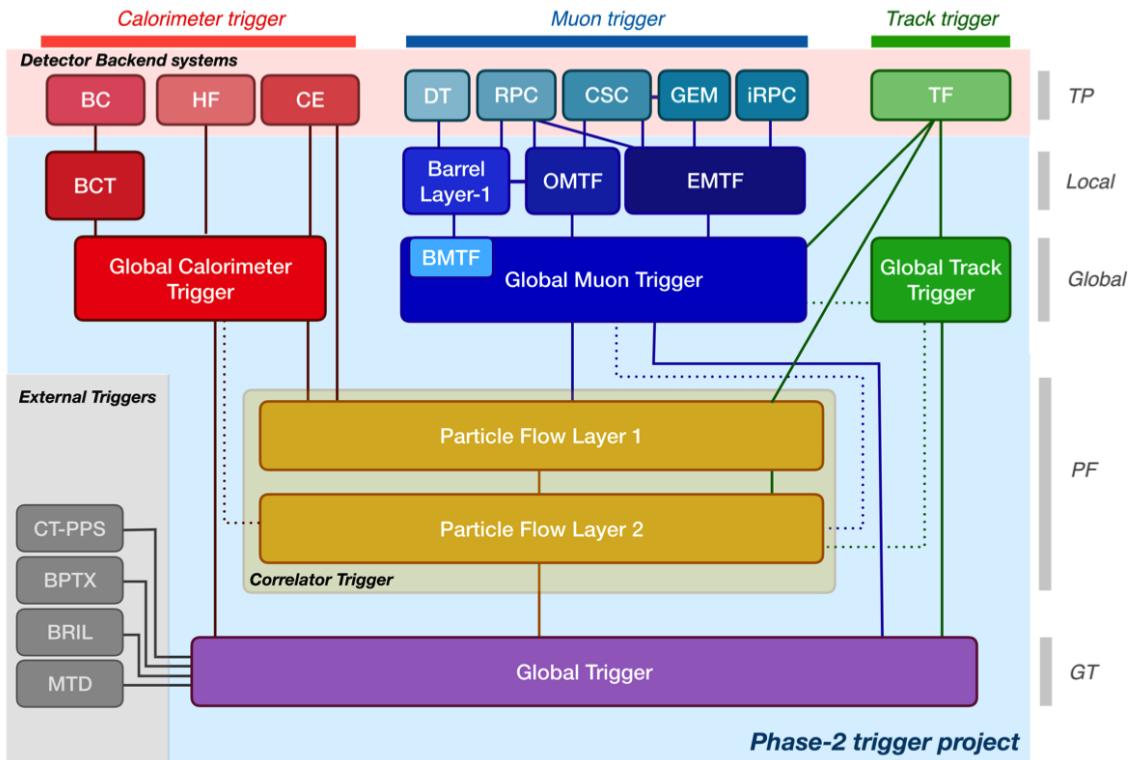


Figure 4.1: Functional diagram of the CMS L1 Phase-2 upgraded trigger design [30], showing the four trigger paths: calorimeter, muon, track, and Particle Flow. For the first time, tracking information will be available as early as the L1 Trigger.

The key feature of the Phase-2 L1 Trigger is the introduction of a correlator layer,

where algorithms produce higher-level trigger objects by combining information from sub-detectors, with a selectivity approaching that of offline reconstruction in the HLT [30]. Four independent data processing paths (grouped together in Fig. 4.1) are implemented: tracking, calorimetry, muon systems, and particle-flow techniques:

- **Calorimeter Trigger path:** (*red*, Fig. 4.1) A barrel calorimeter trigger (BCT) and the HGCAL backend are used to produce high-granularity information from the calorimeters to produce high-resolution clusters and identification variables used for later processing. Outputs from the BCT, HGCAL, and the HF are sent to a global calorimeter trigger (GCT), where calorimeter-only objects such as e/γ candidates, hadronically decaying tau lepton candidates, jets, and energy sums are built.
- **Track Trigger path:** (*green*, Fig. 4.1) Tracks from the Outer Tracker are reconstructed in the track finder (TF) processors as part of the detector backend. A global track trigger (GTT) will reconstruct the primary vertices of the event, along with tracker-only based objects, such as jets and missing transverse momentum.
- **Muon Trigger path:** (*blue*, Fig. 4.1) Trigger primitives are processed by muon track finder algorithms, again separated into the barrel (barrel muon track finder, BMTF), overlap (overlap muon track finder, OMTF), and endcap (endcap muon track finder, EMTF). Standalone muons and stubs containing information such as position, bend angle, and timing, as well as L1 tracks, are sent to the global muon trigger (GMT).
- **Particle-Flow Trigger path:** (*yellow*, Fig. 4.1) The correlator trigger (CT) aims to approach the performance of offline Particle Flow, and is implemented in two layers. “Layer-1” produces the particle-flow candidates from matching

calorimeter clusters and tracks. “Layer 2” builds and sorts final trigger objects and applies additional identification and isolation criteria.

The outputs from the above trigger paths are combined in the Global Trigger (GT) (*purple*, Fig. 4.1), which calculates the final trigger decision (Level-1 Accept), transmitting it to the Trigger Control and Distribution System (TCDS), which distributes it to the detector backend systems, initiating the readout to the DAQ. The GT also provides the interface to external triggers (*grey*, Fig. 4.1), such as triggers for the precision proton spectrometer (PPS), beam position and timing monitors (BPTX), and luminosity and beam monitoring (BRIL) detectors [30]. The design of the Phase-2 Level-1 Trigger allows for future inclusion of triggering information, for instance information about minimum ionizing particles (MIPs) from the MIP Timing Detector (MTD) [31].

4.3 Standalone Barrel Calorimeter electron/photon reconstruction

The reconstruction and identification of electrons and photons (e/γ) begin with the trigger primitives of the barrel ECAL and HCAL detectors and endcap HGCAL calorimeters, covering the pseudorapidity region $|\eta| < 3$. The barrel and endcap regions of the detector are intrinsically different enough to warrant different approaches to e/γ reconstruction. This work focuses on the Standalone Calorimeter e/γ reconstruction taking place in the barrel (Fig. 4.2).

4.3.1 Phase-2 geometry of the ECAL Barrel trigger

In Phase-2, the upgrade of both on-detector and off-detector electronics for the barrel calorimeters trigger primitive generator (TPG) will stream single crystal data from

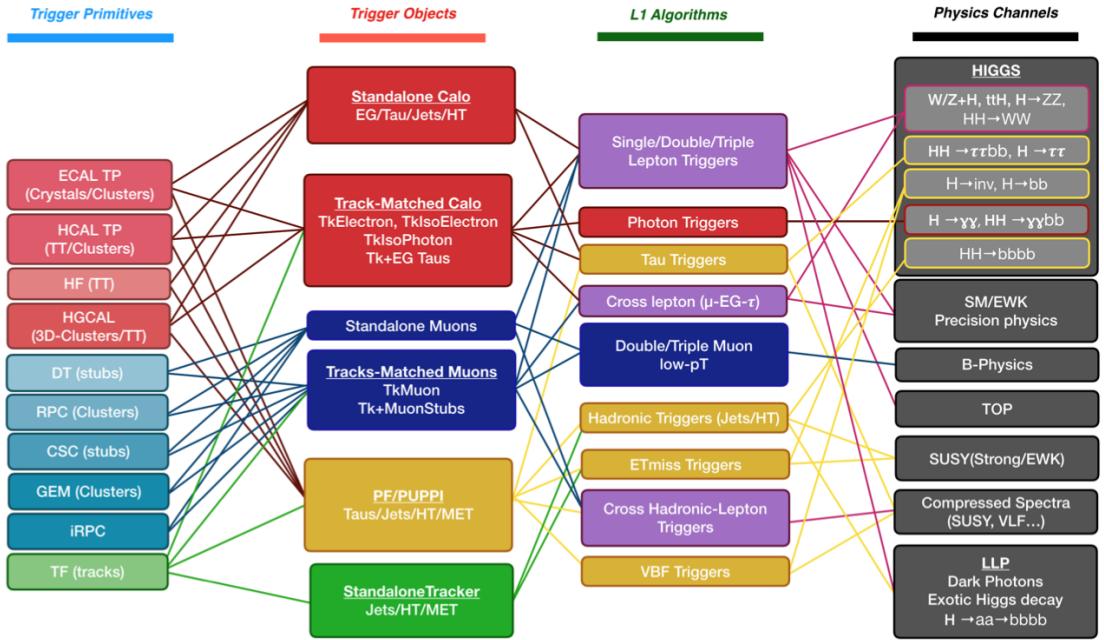


Figure 4.2: Summary of the links between the trigger primitives (*first column*), the trigger objects (*second column*), the Level-1 algorithms used in the menu (*3rd column*), and the physics channels (*4th column*), from [30], where a full description of the Phase-2 L1 algorithms can be found. This work focuses on developments for the Standalone Calorimeter electron and photon ("EG") reconstruction algorithm.

the on-detector to the backend electronics, in contrast to the lower-granularity output of the Phase-1 ECAL TPG that is restricted to providing trigger tower sums of 5×5 crystals [30]. A schematic representation of the geometry of the ECAL barrel in the Regional Calorimeter Trigger (RCT) is shown in Fig. 4.3. The barrel is spanned by 36 RCT cards, each spanning 17×4 towers in $\eta \times \phi$. Each RCT card is subdivided into five “regions” as shown in Fig. 4.4. After initial clustering and processing, the outputs of the RCT card are sent to the Global Calorimeter (GCT) trigger, which is processed in three cards as shown in Fig. 4.5.

4.3.2 Phase-2 electron/photon reconstruction algorithm

The standalone barrel algorithm for reconstructing and identifying electrons and photons in the Phase-2 Level-1 Trigger takes as input the digitized response of each crystal

of the barrel ECAL, with a granularity 0.0175×0.0175 in $\eta \times \phi$, which is 25 times higher than the input to the Phase-1 trigger, which consisted of trigger towers with a granularity of 0.0875×0.0875 . In HCAL the tower size of 0.0875×0.0875 is unchanged. The trigger algorithm is designed to closely reproduce the algorithm used in the offline reconstruction, with limitations and simplifications due to trigger latency.

In the RCT, an initial requirement of $p_T > 0.5$ GeV is imposed on the input trigger primitives (i.e. energies from the ECAL crystals and HCAL towers) to reject contribution from pileup. In one of the regions inside a RCT card (Fig. 4.4), the crystal containing the highest energy deposit is identified as the seed crystal, as shown in Fig. 4.6. The energy in the crystals in a window of size 3×5 in $\eta \times \phi$ around the seed cluster is added into a cluster. The energy is considered “clustered”. The process is repeated with the remaining “unclustered” energy, until up to four clusters are produced in the region.

To improve e/γ identification and to reduce background contributions, identification and reconstruction algorithms are implemented at this stage:

- Shower shape: The energy deposit sums around the seed crystal is computed in windows of size 2×5 and 5×5 (Fig. 4.6, *dashed lines*), with true e/γ clusters tending to produce showers that deposit most of their energy in a 2×5 region.
- Bremsstrahlung recovery: e/γ tend to spread in the ϕ direction due to charged particles being bent by the magnetic field of the CMS solenoid. If sufficient energy comparable to the core 3×5 cluster is found in the adjacent 3×5 windows (Fig. 4.6, *shaded yellow*), the energy is added to the core cluster and no longer considered unclustered energy.

After parallel processing in the regions, the clusters in a RCT card are stitched together if they are located directly along the borders of a region (Fig. 4.3). The remaining unclustered ECAL energy is summed into ECAL towers.

From each RCT card, the twelve highest-energy clusters, as well as any remaining unclustered energy, are sent to the GCT. Since each GCT card has information from sixteen RCT cards (Fig. 4.5), final stitching across the boundaries of the RCT cards is performed. One more identification algorithm is performed at this stage:

- Isolation: One handle to reject backgrounds from e.g. pileup, comes from the tendency for background to be spread more uniformly across a large area in the detector, whereas genuine e/γ are expected to produce showers concentrated in the 3×5 crystal window. The energy sum in a large window of 7×7 in towers is computed and used to reject background.

The performance of the standalone barrel e/γ algorithm in Phase-2 conditions is summarized in the efficiency and rates. The efficiencies are measured with a simulated Monte Carlo sample containing electrons. The rates are measured with a simulated minimum bias sample intended to closely mimic generic proton-proton collisions in the CMS detector. The performance of the Phase-2 emulator discussed in this work, which closely mimics the firmware logic and uses fixed-precision integers, is shown to be comparable to the previous emulator which used floats and idealized logic.

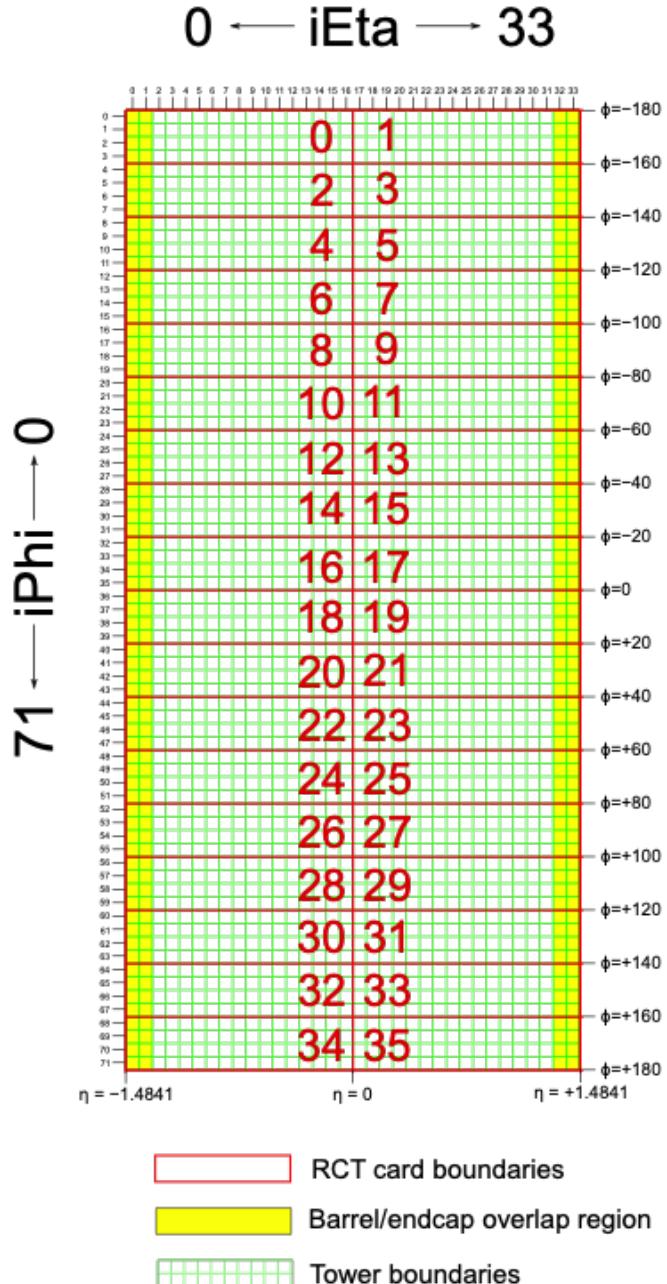


Figure 4.3: Schematic of the geometry of the Phase-2 ECAL barrel in the Regional Calorimeter Trigger (RCT), showing the division of the barrel region into 36 Regional Calorimeter Trigger (RCT) cards (*red*). Each card spans 17×4 towers in $\eta \times \phi$ (*green*), and each tower is 5×5 in single crystals in $\eta \times \phi$. Towers in the overlap region (*shaded yellow*) are read out to both the barrel and endcap.

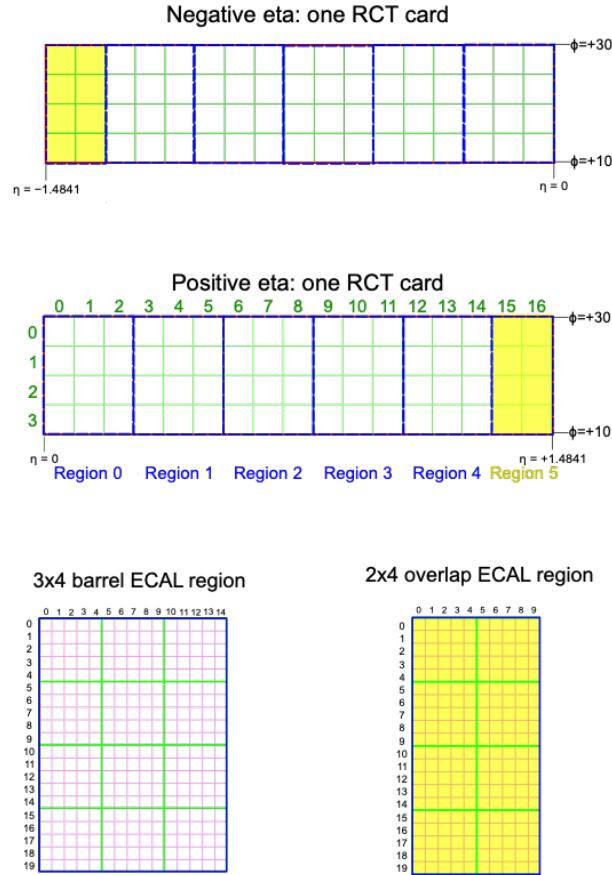


Figure 4.4: Schematic of two example RCT cards in the negative eta (*top*) and positive eta (*center*) regions of the ECAL barrel. Each RCT card is divided into five regions: four regions are of size 3×4 towers in $\eta \times \phi$ (*bottom left*), and a fifth smaller overlap region of size 2×4 towers (*bottom right*). Each tower is 5×5 ($\eta \times \phi$) in crystals.

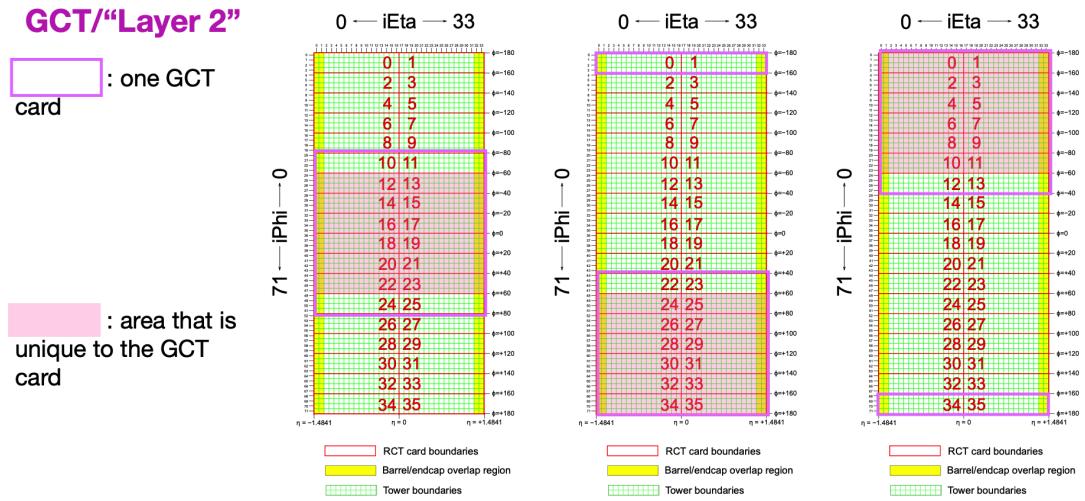


Figure 4.5: Schematic of the Phase-2 ECAL barrel in the Global Calorimeter Trigger (GCT), which will process the outputs of the Regional Calorimeter Trigger (RCT) in three cards (*magenta highlights*). Each card in the GCT processes the equivalent of sixteen RCT cards, with the center twelve being unique to that GCT card (*shaded pink*), and the remaining four processed in overlap with the other GCT cards.

3x4 barrel ECAL region

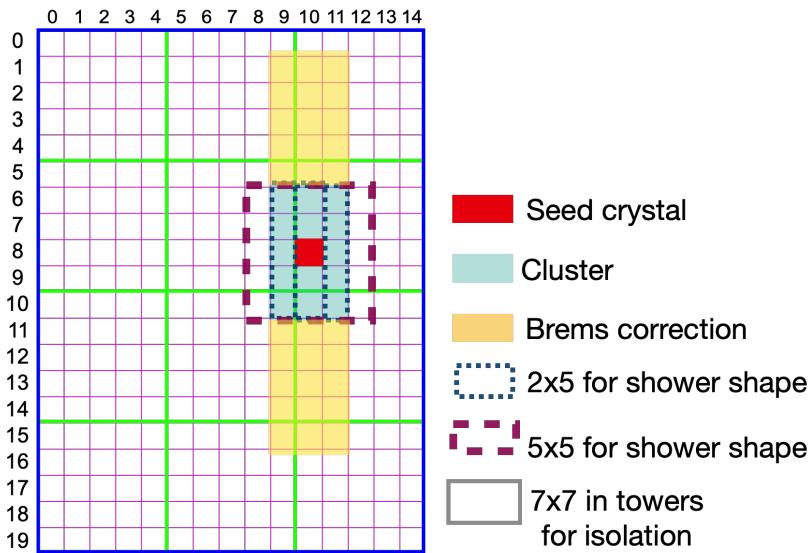


Figure 4.6: Illustration of an example electron/photon (e/γ) cluster in the Phase-2 Level-1 Trigger standalone barrel e/γ reconstruction, in a region of 15×20 crystals (3×4 towers). Each small pink square is one crystal, the highest-granularity ECAL trigger primitives available to the L1 Trigger in Phase-2. The core cluster consists of the energy sum in a 3×5 window of crystals, (*shaded light blue*) centered around the seed crystal (*red*). Bremsstrahlung corrections are checked in the adjacent 3×5 windows in the ϕ direction (*shaded light yellow*). The relative energies in windows of size 2×5 and 5×5 in crystals (*dashed dark blue and dark red*) are used to compute shower shape variables to identify true e/γ objects. Lastly, an isolation sum is computed in a window of size 7×7 in towers (not shown in figure).

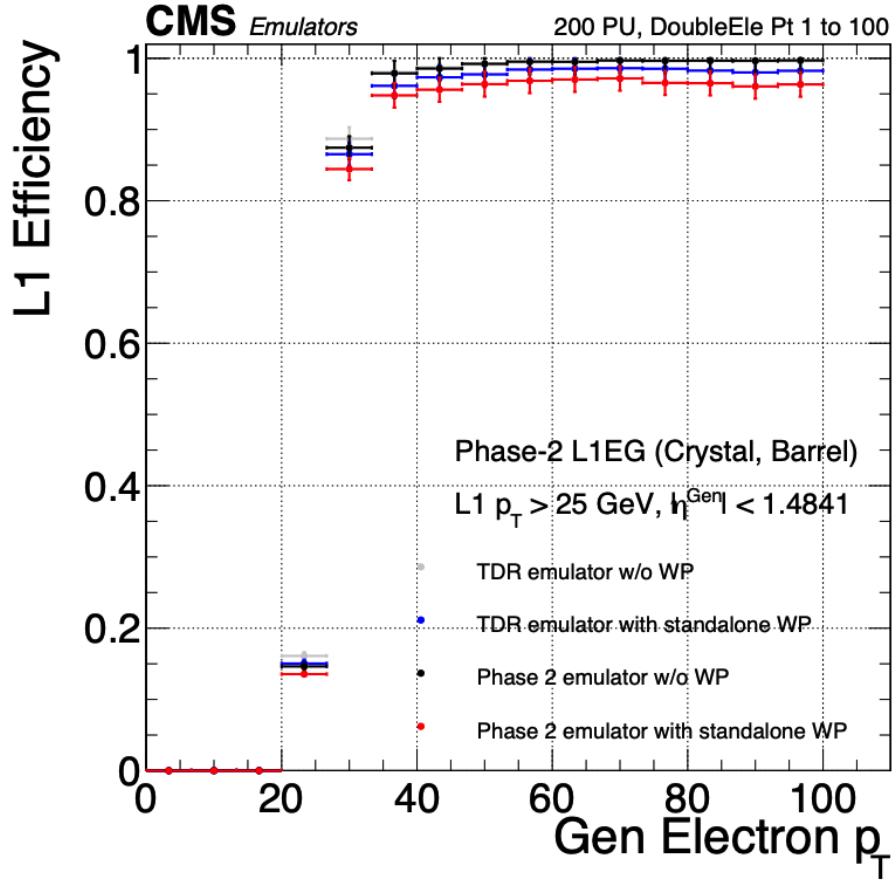


Figure 4.7: Efficiency of the standalone barrel e/γ reconstruction, measured in a simulated sample of electrons, as a function of the true electron’s transverse momentum p_T . The performance of the previous, idealized algorithm as shown in the 2021 Phase-2 TDR [30] with and without the isolation and shower shape discrimination variables (“standalone working point/ WP”) (*dark blue, grey*). The Phase-2 emulator discussed in this work with and without the same working point (*black, red*) is shown to have comparable performance.

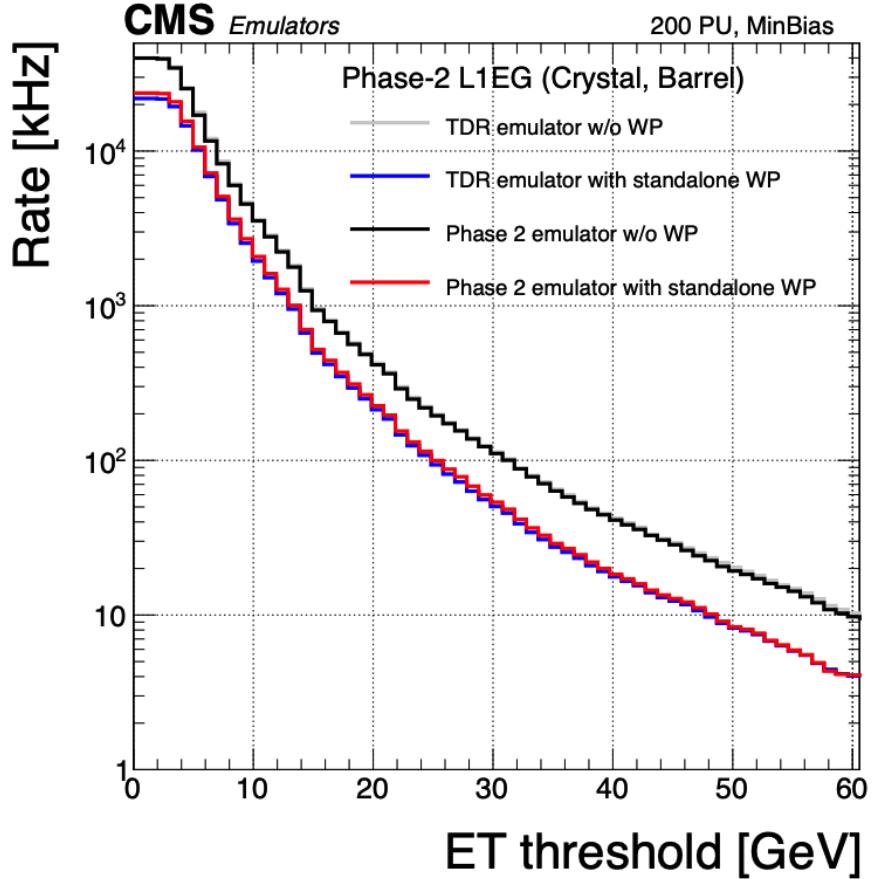


Figure 4.8: Rates of the standalone barrel e/γ reconstruction, evaluated on a minimum bias sample, measured as a function of the minimum energy (E_T) required of the reconstructed e/γ object in each event. The performance of the previous, idealized algorithm as shown in the 2021 Phase-2 TDR [30] with and without the isolation and shower shape discrimination variables (“standalone working point/ WP”) (dark blue, grey). The Phase-2 emulator discussed in this work with and without the same working point (black, red) is shown to have comparable performance.

Chapter 5

Event reconstruction

We review the properties of the particles most pertinent to the analyses presented in this work (taus, muons, electrons, jets, and b-jets), their signatures in the CMS detector, and dedicated reconstruction techniques used at CMS.

5.1 Taus

5.1.1 Decay modes

The lifetime of τ leptons is short enough that they decay before reaching the detector elements. In two thirds of the cases, τ leptons decay hadronically, typically into one or three charged mesons (predominantly π^+ , π^-), often accompanied by neutral pions (that decay $\pi^0 \rightarrow \gamma\gamma$), and a ν_τ . These hadronic decays are denoted τ_h . In the remainder of the decays, the tau decays to the lighter leptons (electron or muon), termed leptonic decays. In all cases, at least one neutrino is produced, resulting in missing transverse energy in the CMS detector. The tau's largest decay branching ratios (proportional to probability of decay) are listed below [23]:

- 17.8% decay to $e^- \bar{\nu}_e \nu_\tau$
- 17.4% decay to $\mu^- \bar{\nu}_\mu \nu_\tau$

- 25.5% decay to $\pi^-\pi^0\nu_\tau$ (ρ^- resonance at 770 MeV)
- 10.8% decay to $\pi^-\nu_\tau$
- 9.3% decay to $\pi^-\pi^0\pi^0\nu_\tau$ (a_1^- resonance at 1200 MeV)
- 9.0% decay to $\pi^-\pi^-\pi^+\nu_\tau$ (a_1^- resonance at 1200 MeV)

The neutrinos escape undetected from the CMS detector and are not considered in the reconstruction. Charged hadrons leave tracks in the tracking detector before being absorbed in the hadronic calorimeter; in CMS tau reconstruction terminology, they are often called “prongs”, i.e. the dominant τ_h decay modes are termed “1 prong” (π^\pm), “1 prong + $\pi^0(s)$ ”, and “3-prong”. Neutral pions decay to two photons which lose their energy in the electromagnetic calorimeter. Taus that decay to electrons and muons, are typically triggered on and reconstructed as electrons and muons respectively.

5.1.2 Hadron plus strips (HPS) reconstruction

At CMS, hadronically decaying tau leptons are reconstructed with the hadron plus strips (HPS) algorithm [32] [33]. The HPS algorithm capitalizes on photon conversions in the CMS tracker material, which originate from the neutral pion (π^0) decaying to two photons. The bending of electron/positron tracks due to the CMS solenoid magnetic field leads to a spread of the neutral pions’ calorimeter signatures in the ϕ direction. This motivates the reconstruction of photons in “strips”: objects that are built out of PF photons and electrons. The strip reconstruction starts with centering a strip on the most energetic electromagnetic particle in a PF jet. Among other electromagnetic particles located in a window of size $\Delta\eta = 0.05$ and $\Delta\phi = 0.20$ around the strip center, the most energetic one is associated with the strip and its momentum is added to the strip momentum. This is repeated iteratively until no

further particles can be associated. Lastly, strips satisfying a requirement of $p_T^{\text{strip}} > 1$ GeV are combined with charged hadrons to reconstruct individual τ_h decay modes, where h stands for both π and K :

- *Single hadron:* $h^- \nu_\tau$ and $h^- \pi^0 \nu_\tau$ decay modes, in which the neutral pions have too little energy to be reconstructed as strips.
- *One hadron + one strip:* $h^- \pi^0 \nu_\tau$ decay modes, where the photons from the π^0 decay are close together in the calorimeter.
- *One hadron + two strips:* $h^- \pi^0 \nu_\tau$ decay modes, where the photons from the π^0 decay are well separated.
- *Three hadrons:* $h^- h^+ h^- \nu_\tau$ decay modes. The three charged hadrons are required to originate from the same secondary vertex.

The $h^- \pi^0 \pi^0 \nu_\tau$ and $h^- h^+ h^- \pi^0 \nu_\tau$ decay modes do not have their own treatment are reconstructed with the above topologies.

In the HPS algorithm, the direction of the reconstructed tau momentum \vec{p}^{τ_h} is required to fall within a distance of $\Delta R = 0.1$ from the original PF jet. All charged hadrons and strips are required to be contained within a cone of size $\Delta R = (2.8 \text{ GeV})/p_T^{\tau_h}$, from the τ_h as reconstructed by the HPS.

All charged hadrons are assumed to be pions, and they are required to be consistent with the masses of the intermediate meson resonances (if applicable), with the following allowed windows for candidates: 50-200 MeV for π^0 , 0.3-1.3 GeV for ρ , and 0.8-1.5 GeV for a_1 . If the τ_h decay is compatible with more than one hypothesis, the one giving the highest $p_T^{\tau_h}$ is chosen. Lastly, an isolation requirement is applied: aside from the τ_h decay products, no charged hadrons or photons can be present within an isolation cone of size $\Delta R = 0.5$ around the direction of the τ_h . The outputs of the HPS algorithm are the reconstructed decay mode and the visible four-momentum (i.e. the four-momenta of all decay products excluding the neutrinos).

5.1.3 DeepTau

The identification of τ_h candidates in CMS has historically been divided into separate discriminators against jets, electrons, and muons. Discriminators versus jets and electrons use information from derived quantities, such as the p_T sum of particles near the τ_h axis. Building on the previous multivariate analysis (MVA) classifier [34] based on a boosted decision tree (BDT), DeepTau is a more recent classifier based on a deep neural network (DNN) that simultaneously discriminates against jets, electrons, and muons. The DNN uses a combination of high-level inputs, similar to previous algorithms, and also uses convolutional layers in $\eta\text{-}\phi$ space to process information from all reconstructed particles near the τ_h axis. Convolutional layers are based on the principle that an image can be processed independently of its position.

The final DeepTau discriminators against jets, muons, and electrons are given by

$$D_\alpha(y) = \frac{y_\tau}{y_\tau + y_\alpha} \quad (5.1)$$

where y_τ (y_α) are estimates of the probabilities for the τ_h candidate to come from a genuine τ_h (jet, μ , e). Working points for each discriminator with different τ_h identification efficiencies are defined for D_e , D_μ , and D_{jet} , for usage in physics analyses and derivation of data-to-simulation corrections [35].

5.2 Muons

Muons are identified with requirements on the quality of the track reconstruction and on the number of measurements in the tracker and the muon systems [36]. In the standard CMS reconstruction, tracks are first reconstructed independently in the inner tracker (tracker track) and in the muon system (standalone-muon track). Next, these tracks are processed in two different methods.

The first is Global Muon reconstruction (outside-in) [36], which fits combined hits from the tracker track and standalone-muon track, using the Kalman-filter technique. At large transverse momenta, $p_T \gtrsim 200$ GeV, the global-muon fit can improve the momentum resolution compared to the tracker-only fit.

The second is Tracker Muon reconstruction (inside-out) [36], which starts with tracker tracks with $p_T > 0.5$ GeV and total momentum $p_T > 2.5$ GeV. These tracks are extrapolated outwards to the muon system and matched to detector segments there, taking into account the magnetic field, expected energy losses, and multiple Coulomb scattering in the detector material. Tracker Muon reconstruction is more efficient than the Global Muon reconstruction at low momenta, $p \lesssim 5$ GeV, because it only requires a single muon segment in the muon system, whereas Global Muon reconstruction typically requires segments in at least two muon stations.

To further suppress fake muons from decay in flight, isolation cuts are used. A relative isolation variable is defined to quantify the energy flow of particles near the muon trajectory. A relative isolation is defined similarly for muons and electrons:

$$I^\ell \equiv \frac{\sum_{\text{charged}} p_T + \max\left(0, \sum_{\text{neutral}} p_T - \frac{1}{2} \sum_{\text{charged, PU}} p_T\right)}{p_T^\ell} \quad (5.2)$$

where $\sum_{\text{charged}} p_T$ is the scalar sum of the p_T of the charged particles originating from the primary vertex and located in a cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4(0.3)$ centered on the direction of the muon (electron). The sum $\sum_{\text{neutral}} p_T$ is the equivalent for neutral particles. The sum $\sum_{\text{charged, PU}} p_T$ is the scalar sum of the p_T of the charged hadrons in the cone originating from pileup vertices. The factor 1/2 comes from simulation estimations, which find that the ratio of neutral to charged hadron production in the hadronization process of inelastic pp collisions is 1/2. Thus the subtracted term is intended to subtract contribution from pileup, from the neutral particle contribution to the isolation sum. Finally, this is divided by the lepton

transverse momentum, p_T^ℓ .

5.3 Electrons

Electrons are reconstructed by associating a track reconstructed in the silicon tracking detector with a cluster of energy in the ECAL. Performance is maximized via a combination of a stand-alone approach and the complementary global particle-flow approach [37].

In the stand-alone approach, the electron energy, which is typically spread over several crystals of the ECAL, is clustered with the “hybrid” algorithm in the barrel and the “multi- 5×5 ” in the endcaps [37]. The hybrid algorithm collects energy in a small window in η and an extended window in ϕ . It identifies a seed crystal, and adds arrays of 5×1 crystals in $\eta \times \phi$ in a range of $N = 17$ crystals in both directions of ϕ , if their energies exceed a minimum threshold, thus forming a supercluster (SC). In the endcap, crystals are not arranged in an $\eta \times \phi$ geometry; instead clusters are build around seed crystals in clusters of 5×5 crystals that can partly overlap. Nearby clusters are grouped into a supercluster, and energy is recovered from associated deposits in the preshower.

In the PF reconstruction [37], PF clusters are reconstructed by aggregating around a seed all contiguous crystals with energies two standard deviations above the electronic noise observed at the beginning of a data-taking run. The energy of a given crystal can be shared among two or more clusters.

The electron track reconstruction is performed in two ways [37]: the ECAL-based seeding, which begins with the SC energy and positioning, and the tracker-based seeding (part of the PF reconstruction algorithm), which uses tracks reconstructed from the general algorithm for charged particles, extrapolated towards the ECAL and matched to an SC. Kalman filter (KF) tracks with a small number of hits or that are

not well-fitted, are re-fitted with a dedicated Gaussian sum Filter (GSF).

A global identification variable [37] is defined using a multivariate analysis (MVA) technique that combines information on track observables (kinematics, quality of the KF track and GSF track), the electron PF cluster observables (shape and pattern), and the association between the two (geometric and kinematic observables). For electrons seeded only through the tracker-based approach, a weak selection is applied on this MVA variable. For electrons seeded through both approaches, a logical OR is taken.

Electron isolation, i.e. the presence of energy deposits near the electron trajectory, is a separate key handle in rejecting significant background. Compared to isolated electrons, electrons from misidentified jets or genuine electrons within a jet resulting from semileptonic decays of b or c quarks tend to have significant energy deposits near the primary trajectory [37]. Offline analyses benefit from the PF technique for defining isolation, which sums the PF candidates reconstructed located within a specified isolation cone around the electron candidate, as in Eqn. 5.2.

5.4 Jets

The vast majority of processes of interest at the LHC contains quarks or gluons in the final state, but these particles cannot be observed directly. In a process called hadronization, they fragment into spatially-grouped collections of particles called jets, which can be detected in the tracking and calorimeter systems. Hadronization and the subsequent decays of unstable hadrons can produce hundreds of nearby particles in the CMS detector. Jets are reconstructed by the PF algorithm (PF jets), or from the sum of the ECAL and HCAL energies deposited in the calorimeter towers (Calo jets). In PF jets, typically used in offline analyses, jets are built using the anti- k_T (AK) clustering algorithm [38]. The anti- k_T algorithm iterates over particle pairs and

finds the two that are closest in a distance measure d , and determines whether to combine them:

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta_{ij}^2}{R^2}, \text{ combine when } d_{ij} < p_{T,i}^{-2}; \text{ stop when } d_{ij} > p_{T,i}^{-2} \quad (5.3)$$

where $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ and $p_{T,i}$, η_i , ϕ_i are the transverse momentum, rapidity, and azimuthal angle of particle i . The power -2 means that higher-momentum particles are clustered first, leading to jets that tend to be centered on the hardest (highest p_T) particle.

There are several methods to remove contributions of pileup collisions from jet clustering [39]:

- Charged hadron subtraction (CHS), which removes all charged hadron candidates associated with a track that is not associated with the primary vertex.
- PileUp Per Particle Identification (PUPPI), which weighs input particles based on their likelihood of arising from pileup. QCD particles tend to have a collinear structure, compared to soft diffuse radiation coming from pileup. The local shape for charged pileup, used as a proxy for all pileup particles, is used on an event-by-event basis to calculate a weight for each particle. PUPPI is deployed in Run-2 and is more performant than CHS in high pileup scenarios.

5.5 B jets

Jets that arise from bottom-quark hadronization (b jets) have overwhelming background from processes involving jets from gluons (g) and light-flavour quarks (u, d, s), and from c -quark fragmentation. The ability to identify b jets, or b -tagging, exploits the b hadrons' relatively large masses, long lifetimes, and daughter particles with hard momentum spectra [38].

The impact parameter (IP) of a track is the 3-dimensional distance between the track and the primary vertex (PV) at the point of closest approach. The IP is positive if the track originates from the decay of particles travelling along the jet axis. The resolution of the IP depends on the p_T and η of the track, motivating the use of the impact parameter significance S_{IP} (ratio of the IP to its estimated uncertainty) as an observable [38].

Because of the large but finite lifetimes of the b hadrons, b hadrons tend to travel a short distance before decaying at a secondary vertex (SV), which can be measured and reconstructed separately from the primary vertex due to the excellent position resolution of the pixel detector [38]. Previous b-tagging algorithms (e.g. CSV, cMVAv2, and DeepCSV) have capitalized on variables such as the presence of a SV, the flight distance and direction (computed from the vector between the PV and the SV), and kinematics of the system of associated secondary tracks (e.g. track multiplicity, mass, and energy).

The DeepJet (formerly known as DeepFlavour) algorithm [40] is a deep-neural-network multi-classification algorithm, which uses 16 properties of up to 25 charged and 6 properties of 25 neutral particle-flow jet constituents, as well as 17 properties from up to 4 secondary vertices associate with the jet. Compared to the previous classifying algorithm DeepCSV, DeepJet has been demonstrated to have higher efficiency with lower misidentification probability in Phase-1 data [41].

Chapter 6

Datasets and Monte Carlo samples

6.1 Datasets used

The $h \rightarrow aa \rightarrow 2b2\tau$ analysis (CMS CADI line HIG-22-007) is based on proton-proton collision data at a center-of-mass energy of 13 TeV collected in full Run-2 (2016–18) with the CMS detector. The data analyzed corresponds to a total integrated luminosity of 138 fb^{-1} (36.33 fb^{-1} for 2016, 41.53 fb^{-1} for 2017, and 59.74 fb^{-1} for 2018) [42] [43] [44].

Data collected with the single muon trigger is used for the $\mu\tau_h$ channel. For the $e\tau_h$ channel, data collected with the single electron trigger is used; and for the $e\mu$ channel, data collected with the electron + muon trigger is used. A more in-depth discussion of the triggers used follows in a later section.

A full list of samples used can be found in the full documentation [45] [46].

6.2 Monte Carlo samples

Modeling and computing observables originating from arbitrary physics processes at the tree level and at next-to-leading order (NLO) is performed by Monte Carlo (MC) event generators, such as Powheg and MadGraph5_amCNLO [47] [48]. The informa-

tion generated, e.g. the computation of the differential cross sections and kinematics of the final state particles, is saved in a compressed file and used to generate MC samples that are used in physics analyses. The samples are digitized using GEANT4 [49], a platform used at the LHC and other facilities to comprehensively simulate the passage of particles through matter. The digitized samples are passed through the same detector reconstruction as real data events collected in the detector.

The samples for modeling the signal ($h \rightarrow aa \rightarrow 2b2\tau$ and $h \rightarrow a_1a_2$) in the 2HDM+S and TRSM are generated at tree-level, for a range of masses of the light neutral scalar a . For $h \rightarrow aa$, the mass hypotheses for the a range from $m_a = (12 \text{ GeV}, 62.5 \text{ GeV})$. For $h \rightarrow a_1a_2$, the mass hypotheses for the two light scalars span combinations of m_{a1} , m_{a2} ranging from $(12 \text{ GeV}, 62.5 \text{ GeV})$ for the two scalars.

6.3 Embedded samples

An embedding technique is used to estimate background from the Standard Model Z boson decaying to $\tau\tau$, from data with minimal simulation input [50]. In data events selected to have $Z \rightarrow \mu\mu$ decays, all energy deposits of the recorded muons are removed from the event, and replaced with simulated tau leptons with the same kinematic properties as the removed muons. This results in a hybrid data format containing information from both observed and simulated events. The advantage of the Embedded samples is that the portions of the event that are difficult to model and describe (e.g. the underlying event or production of additional jets) are directly taken from data.

Chapter 7

Event selection

7.1 General procedure for all channels

For the search for $h \rightarrow aa \rightarrow bb\tau\tau$, three final states of the $\tau\tau$ system are considered: $\mu\tau_h$, $e\tau_h$, and $e\mu$. The $\tau_h\tau_h$ final state is not considered because signal events in the $\tau_h\tau_h$ channel would typically produce hadronic taus with momenta below data-taking trigger thresholds.

In all three final states, events are required to have at least one b-tag jet passing the medium working point of the DeepFlavour tagger, with $p_T > 20$ GeV, and $|\eta| < 2.4$. A second b-tag jet is not required because such a requirement would reduce signal acceptance by 80% compared to only requiring one b-tag jet.

Events in MC samples are sorted into one of the three $\tau\tau$ channels if they pass the following trigger requirements and requirements on the offline reconstructed objects in the event, first checking the HLT paths for the $\mu\tau_h$ channel, then $e\tau_h$, and finally $e\mu$. The two leading leptons (e.g. muon and hadronic tau for the $\mu\tau_h$ channel) that were determined to have originated from the $\tau\tau$ decay, are called the $\tau\tau$ “legs” and are respectively subscripted 1 and 2 in this work. For events in data and embedded samples, the HLT paths requirements for the corresponding channel are checked.

After sorting events by HLT paths and identifying the leading tau legs in the offline reconstructed objects, the p_T of the offline objects is checked against the online trigger thresholds. Trigger matching is also performed, which checks the correspondence between each offline reconstructed object used in the analysis (e.g. a muon), and a trigger object in the HLT (e.g. a HLT muon). An offline object is considered to be matched, if it corresponds to a trigger object of the same object type, with $\Delta R < 0.5$. This matched trigger object is also required to pass the filter(s) of the HLT trigger.

Further cuts are made on the offline objects in each channel to obtain the signal region, or other data regions used to perform data-driven background estimations. These are detailed below.

7.2 Event selection in the $\mu\tau_h$ channel

In all three years, a single muon trigger is used if the muon has sufficiently high p_T , otherwise a dilepton $\mu\tau_h$ cross-trigger is used (Tables 7.1, 7.2, and 7.3). For data taken in 2017-2018 (2016), the logical OR of the single muon triggers with online p_T thresholds 24 and 27 (23) GeV is used, with the corresponding offline muon required to have with p_T 1 GeV above the online threshold. For data taken in 2017-2018 (2016), a dilepton $\mu + \tau_h$ cross-trigger with p_T thresholds of 20 (19) and 27 (20) GeV for the muon and tau respectively, is used. The τ_h is required to have $|\eta| < 2.3$ if the single trigger is fired, $|\eta| < 2.1$.

The muon and τ_h are required to have opposite charge and be separated by $\Delta R > 0.4$. The muon is required to have $|\eta| < 2.4$, and the τ_h is required to have $|\eta| < 2.3$ unless a cross-trigger is required, in which case we require $|\eta| < 2.1$ as discussed above.

The muon is required to pass the medium identification (ID) working point [51], which is defined by the Muon POG as a loose muon (i.e. a Particle Flow muon that

is either a global or a tracker muon - see Section 5.2) with additional requirements on track quality and muon quality. This identification criteria is designed to be highly efficiently for prompt muons and for muons from heavy quark decays. In addition to the ID, for prompt muons it is recommended to apply cuts on the impact parameter [51]: we apply $|\Delta(z)| < 0.2$ and $|\Delta(xy)| < 0.045$.

In addition, a cut is applied on the muon relative isolation (defined in Section 5.2), to be less than 0.15 in a cone size of $\Delta R = 0.4$, which corresponds to the Tight Particle Flow isolation requirement [51].

The τ_h is required to pass a cut on its impact parameter of $|\Delta(z)| < 0.2$. The τ_h is also required to pass the VLoose (Very Loose) DeepTau working point vs. electron, the Tight DeepTau working point vs. muons, and the VVVLoose and Medium DeepTau working point vs. jets. Events with taus reconstructed in two of the decay modes (labeled 5 and 6) are rejected, since these decay modes are meant to recover 3-prong taus, but are only recommended for use in analyses where the benefits in final significance outweigh the resulting increase in background [52].

For the estimation of the background from jets faking τ_h , which is described in Section 9.7, anti-isolated events are selected, by requiring events to pass all the selections described above, except failing the Medium DeepTau working point vs. jets.

7.3 Event selection in the $e\tau_h$ channel

The HLT trigger paths for the $e\tau_h$ channel are summarized in Tables 7.1, 7.2, and 7.3. Similarly to the $\mu\tau_h$ channel, a single electron trigger is used if the electron has sufficiently high p_T in 2018 and 2017. For data taken in 2018 (2017), the OR of the single electron triggers with online p_T thresholds at 32 and 35 (27 and 32) GeV are used, with the corresponding offline electrons required to have p_T greater than 33 (28) GeV. A $e + \tau_h$ cross-trigger is used for electrons with lower offline p_T between

25 and 33 GeV (25 and 28 GeV). For the 2016 dataset, there is no cross trigger but only a single electron trigger with online p_T threshold at 25 GeV, which is used if the offline electron has p_T greater than 26 GeV.

The electron and τ_h are required to have opposite charge and be separated by $\Delta R > 0.4$. The electron is required to be within $|\eta| < 2.3$ when no cross trigger is used, and $|\eta| < 2.1$ when the cross trigger is fired. The τ_h is required to have $|\eta| < 2.3$ if no cross trigger is fired, and have $|\eta| < 2.1$ if the cross trigger is fired.

The electron is required to have a relative isolation (same definition as in Section 5.2) of less than 0.1 in a cone size of $\Delta R = 0.3$, which is the standard recommended cone size giving minimal pileup dependence and reduced probability of other objects overlapping with the cone. The isolation quantity used includes an “effective area” (EA) correction to remove the effect of pileup in the barrel and endcap parts of the detector [53].

The electron is also required to pass cuts on its impact parameter of $|\Delta(z)| < 0.2$ and $|\Delta(xy)| < 0.045$. It is also required to pass the non-isolated MVA working point corresponding to 90% efficiency. The electron’s number of missing hits, which are gaps in its trajectory through the inner tracker [53], must be less than or equal to 1. The electron must pass a conversion veto, which rejects electrons coming from photon conversions in the tracker, which should instead be reconstructed as part of the photon [53].

The impact parameter cut for the τ_h is $|\Delta(z)| < 0.2$. In contrast to the $\mu\tau_h$ event selection, the vs. electron and vs. muon DeepTau working points are flipped, to reject muons faking the τ_h leg. The τ_h is required to pass the Tight DeepTau working point vs. electrons, the VLoose DeepTau working point vs. muons, and the Medium DeepTau working point vs. jets.

As in the $\mu\tau_h$ channel, for the estimation of the background from jets faking τ_h , which is described in Section 9.7, anti-isolated events are selected, by requiring events

to pass all the selections described above, except failing the Medium DeepTau working point vs. jets.

7.4 Event selection in the $e\mu$ channel

The HLT trigger paths for the $e\mu$ channel are summarized in Tables 7.1, 7.2, and 7.3. Events are selected with the logical OR of two $e + \mu$ cross triggers, where either the electron or muon can have larger p_T : (1) leading electron, where the electron has online $p_T > 23$ GeV and muon has online $p_T > 8$ GeV, or (2) leading muon, where electron has online $p_T > 12$ GeV and muon has online $p_T > 23$ GeV.

The leading and sub-leading leptons are required to have an offline p_T greater than 1 GeV above the online threshold (i.e. $p_T > 24$ GeV). If the sub-leading lepton is the electron, the offline p_T threshold is 1 GeV above the online threshold ($p_T > 13$ GeV), but if it is a muon, the offline p_T threshold is required to be at least 5 GeV greater than the online threshold (i.e. $p_T > 13$ GeV). This is because of poor data and simulation agreement for low- p_T muons with p_T between 9 GeV and 13 GeV, and the higher probability of mis-identifying jets as muons at lower p_T . With no effect on the expected limits, the offline p_T threshold for muons is raised to 13 GeV instead of 9 GeV, even though it may lead to loss in signal acceptance. Both the electron and muon are required to be have $|\eta| < 2.4$.

The electron and muon are required to have opposite charge and be separated by $\Delta R > 0.3$ (note the decreased separation requirement compared to the other two channels). The electron is required to pass the non-isolated MVA identification working point corresponding to 90% efficiency, and to have a relative isolation less than 0.1 for a cone size of $\Delta R = 0.3$ with the EA pileup subtraction correction. The electron must have one or fewer missing hits and pass the conversion veto (both described previously in Section 7.3).

2016 $\mu\tau_h$ trigger paths	
Notes	HLT Path
	HLT_IsoMu22_v
	HLT_IsoMu22_eta2p1_v
	HLT_IsoTkMu22_v
	HLT_IsoTkMu22_eta2p1_v
	HLT_IsoMu19_eta2p1_LooseIsoPFTau20_v
	HLT_IsoMu19_eta2p1_LooseIsoPFTau20_SingleL1_v
2016 $e\tau_h$ trigger paths	
Notes	HLT Path
	HLT_Ele25_eta2p1_WPTight_Gsf_v
2016 $e\mu$ trigger paths	
Notes	HLT Path
runs B-F and MC	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v
runs B-F and MC	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v
runs G-H	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
runs G-H	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v

Table 7.1: High-Level Trigger (HLT) paths used to select data and simulation events in 2016 for the three $\tau\tau$ channels.

The muon is required to pass the medium identification working point (described earlier in 7.2), and to have a relative isolation less than 0.15 for a cone size of $\Delta R = 0.4$. The muon impact parameter is required to have $|\Delta(z)| > 0.2$ and $|\Delta(xy)| < 0.045$.

For the QCD multijet background estimation described in Section 9.8, the same-sign region is selected by requiring all the above selections, except the legs are required to have the same electric charge rather than opposite.

7.5 Extra lepton vetoes in all channels

Events containing a third lepton (electron or muon) that is neither of the leading $\tau\tau$ legs are rejected, and events with di-muons and di-electrons are vetoed, with criteria taken from the Standard Model $H \rightarrow \tau\tau$ working group [54].

The event is vetoed if a third electron is found with the following properties: $p_T > 10$ GeV, $|\eta| < 2.5$, impact parameter $|\Delta(z)| < 0.2$ and $|\Delta(xy)| < 0.045$, passing non-isolation MVA identification with 90% efficiency, conversion veto, ≤ 1 missing

2017 $\mu\tau_h$ trigger paths	
Notes	HLT Path
	HLT_IsoMu24_v
	HLT_IsoMu27_v
	HLT_IsoMu20_eta2p1_LooseChargedIso_PFTau27_eta2p1_CrossL1_v
2017 $e\tau_h$ trigger paths	
Notes	HLT Path
	HLT_Ele32_WPTight_Gsf_v
	HLT_Ele35_WPTight_Gsf_v
	HLT_Ele24_eta2p1_WPTight_Gsf_Loose_ChargedIsoPFTau30_eta2p1_CrossL1_v
2017 $e\mu$ trigger paths	
Notes	HLT Path
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v

Table 7.2: High-Level Trigger (HLT) paths used to select data and simulation events in 2017 for the three $\tau\tau$ channels.

2018 $\mu\tau_h$ trigger paths	
Notes	HLT Path
	HLT_IsoMu24_v
	HLT_IsoMu27_v
only data run < 317509	HLT_IsoMu20_eta2p1_ (contd.) LooseChargedIsoPFTauHPS27_eta2p1_CrossL1_v
MC and data run \geq 317509	HLT_IsoMu20_eta2p1_ (contd.) LooseChargedIsoPFTauHPS27_eta2p1_TightID_CrossL1_v
2018 $e\tau_h$ trigger paths	
Notes	HLT Path
	HLT_Ele32_WPTight_Gsf_v
	HLT_Ele35_WPTight_Gsf_v
only data run < 317509	HLT_Ele24_eta2p1_WPTight_Gsf_ (contd.) LooseChargedIsoPFTauHPS30_eta2p1_CrossL1_v
MC and data run \geq 317509	HLT_Ele24_eta2p1_WPTight_Gsf_ (contd.) LooseChargedIsoPFTauHPS30_eta2p1_TightID_CrossL1_v
2018 $e\mu$ trigger paths	
Notes	HLT Path
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v

Table 7.3: High-Level Trigger (HLT) paths used to select data and simulation events in 2018 for the three $\tau\tau$ channels. In 2018 a HLT trigger path using the hadron plus strips (HPS) tau reconstruction algorithm became available.

hits, and relative isolation < 0.3 with cone size $\Delta R = 0.3$. The event is also vetoed if a third muon is found with the following properties: $p_T > 10$ GeV, $|\eta| < 2.4$, impact parameter $|\Delta(z)| < 0.2$ and $|\Delta(xy)| < 0.045$, medium ID, and isolation < 0.3 with cone size $\Delta R = 0.4$.

A di-muon veto is applied, which rejects events containing a pair of muons with opposite charge and separation of $\Delta R > 0.15$, that both pass the following selections: $p_T > 15$ GeV, $|\eta| < 2.4$, flag for global muons, flag for tracker muon, flag for Particle Flow muon, $|\Delta(z)| < 0.2$, $|\Delta(xy)| < 0.045$, and isolation < 0.3 with cone size $\Delta R = 0.4$.

A similar di-electron veto is applied to reject events containing a pair of electrons with opposite charge and separation of $\Delta R > 0.15$, that both pass the following selections: $p_T > 15$ GeV, $|\eta| < 2.5$, a dedicated electron ID (cut-based) for vetoing third leptons, $|\Delta(z)| < 0.2$, $|\Delta(xy)| < 0.045$, with pileup-corrected relative isolation < 0.3 with cone size $\Delta R = 0.3$.

These vetoes on extra leptons also ensure orthogonality of events to analyses such as the $bb\mu\mu$ final state, whose results are combined with this $bb\tau\tau$ final state as described in Section 13.2.

Chapter 8

Scale factors and corrections

As the modeling of the CMS detector and underlying physics processes are imperfect, scale factors and corrections are applied to Monte Carlo events and in some cases, the embedded samples, to improve agreement with measurements in data. For the standard reconstruction physics objects in this work, central recommendations provided by dedicated physics object groups (POGs) in CMS are used.

Uncertainties in scale factors and corrections are also sources of systematic errors in the analysis, detailed in Chapter 11. Systematic uncertainties in the tau, muon, and electron energy scales can shift the p_T of the leptons up or down, causing a migration of number of events that pass the offline p_T thresholds described in the previous section.

8.1 Tau energy scale

An energy scale is applied to the transverse momentum p_T and mass of the hadronic tau τ_h in the $\mu\tau_h$ and $e\tau_h$ channels, to correct for a deviation of the average reconstructed τ_h energy from the generator-level energy of the visible τ_h decay products. These correction factors are derived centrally [34], by fitting to events in $e\tau_h$ and $\mu\tau_h$ final states in Z/γ^* events separately for the h^\pm , $h^\pm\pi^0$, and $h^\pm h^\mp h^\pm$ decays. The

values used are shown in Table 8.1.

When applying the energy scale to the τ_h , the 4-momentum of the missing transverse energy (MET) is adjusted such that the total 4-momenta of the τ_h and the MET remains unchanged [52].

Tau energy scale factor				
Decay mode	2018	2017	2016 pre-VFP	2016 post-VFP
0	0.991 ± 0.008	0.986 ± 0.009	0.987 ± 0.01	0.993 ± 0.009
1	1.004 ± 0.006	0.999 ± 0.006	0.998 ± 0.006	0.991 ± 0.007
10	0.998 ± 0.007	0.999 ± 0.007	0.984 ± 0.008	1.001 ± 0.007
11	1.004 ± 0.009	0.996 ± 0.01	0.999 ± 0.011	0.997 ± 0.016

Table 8.1: Energy scales applied to genuine hadronic tau decays τ_h by data-taking year/era and decay mode, along with systematic errors.

8.2 Muon energy scale

An energy scale is applied to the p_T and mass of genuine muons from τ decays in the $e\mu$ and $\mu\tau_h$ channels [55]. The applied values are the same for MC and embedded samples and are shown in Table 8.2. Following the SM $H \rightarrow \tau\tau$ analysis, Rochester corrections are not applied, and instead prescriptions from [56] are followed.

Muon energy scale factor	
Eta range	Value for all years
$ \eta \in [0.0, 1.2)$	1.0 ± 0.004
$ \eta \in [1.2, 2.1)$	1.0 ± 0.009
$ \eta \in [2.1, 2.4)$	1.0 ± 0.027

Table 8.2: Energy scales and systematic errors applied to genuine muons. The values are the same for MC and embedded for all years [54] [56].

8.3 Electron energy scale

Corrections to the electron energy scale are applied to genuine e from τ decays, and are binned in two dimensions by electron p_T and η for barrel vs. endcap [57]. The

scale factors are binned in p_T and η for MC samples: e.g. values for 2018 are shown in Fig. 8.1 from [58]. For embedded samples the electron energy scale is taken as only binned in η (Table 8.3).

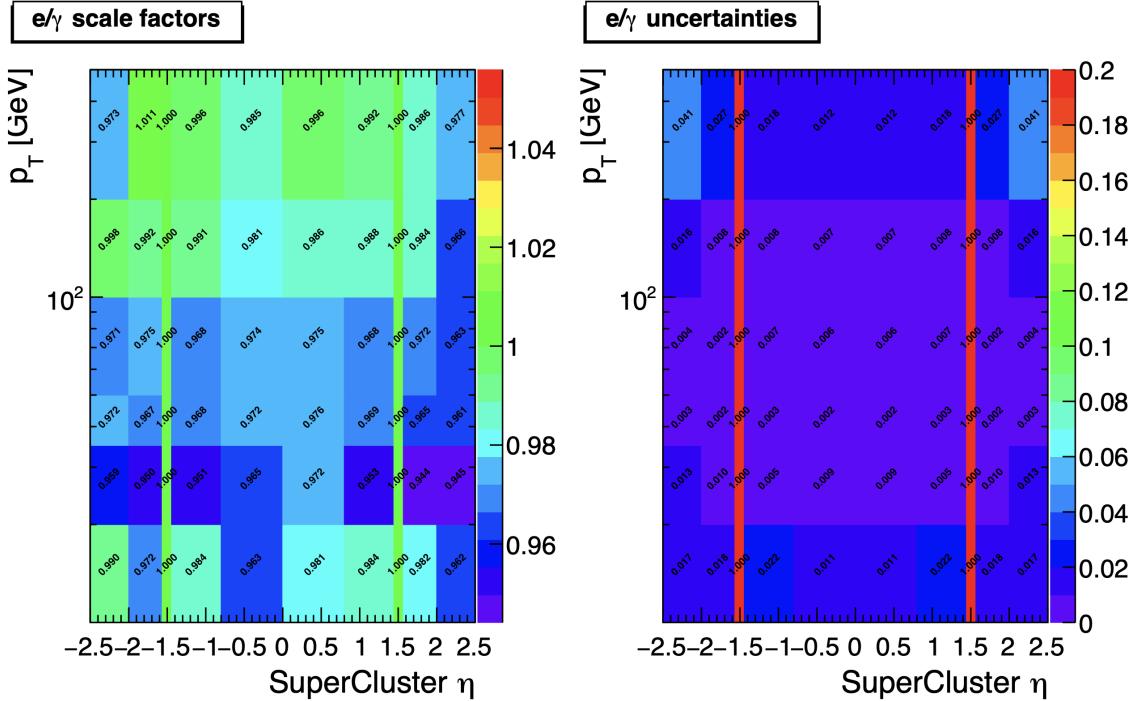


Figure 8.1: Electron/photon energy scale factors (*left*) and corresponding uncertainties (*right*) binned in the electron η and p_T , for the data-taking year 2018 [58].

Electron energy scale factor for embedded samples			
Eta range	2018	2017	2016
$ \eta \in [0.0, 1.479)$	0.973 ± 0.005	0.986 ± 0.009	0.9976 ± 0.0050
$ \eta \in [1.479, 2.4)$	0.980 ± 0.0125	0.887 ± 0.0125	0.993 ± 0.0125

Table 8.3: Energy scales and systematic errors applied to electrons in embedded samples, binned in the electron η , by data-taking year [59] [60] [61].

8.4 τ_h identification efficiency

The τ_h identification efficiency can differ in data and MC [52]. Recommended corrections are provided by the Tau POG, and we use the medium DeepTau vs. jet working

point values. The identification efficiency is measured in $Z \rightarrow \tau\tau$ events in the $\mu\tau_h$ final state, and is binned in p_T due to clear p_T dependence of the DeepTau ID.

Tau ID efficiency for DeepTau Medium vs. jet WP in 2018						
p_T (GeV)	< 20	(20, 25]	(25, 30]	(30, 35]	(35, 40]	(40, 500]
Central value	0	0.945	0.946	0.916	0.921	1.005
Up value	0	1.001	0.981	0.946	0.950	1.035
Down value	0	0.888	0.981	0.883	0.893	0.953

Table 8.4: Tau ID efficiency for the DeepTau vs. jet medium working point, with central, up, and down values for 2018, binned in the tau p_T [52].

8.5 $e \rightarrow \tau_h$ and $\mu \rightarrow \tau_h$ energy scales and misidentification efficiency

Muons faking τ_h receive a nominal energy scale correction of 1.0, with an uncertainty of ± 0.1 .

Corrections on identification efficiencies are applied to genuine electrons and muons misidentified as τ to account for differences in data and MC.

The specific values depend on the vs. electron and vs. muon discriminator working points used. For misidentified $e \rightarrow \tau_h$, the scale factors are split into barrel and endcap regions, dictated by the ECAL detector geometry.

For misidentified $\mu \rightarrow \tau_h$, the scale factors are split into different $|\eta|$ regions, determined by the CMS muon and tracker detector geometries.

8.6 Electron and muon ID, isolation, and tracking

Scale factors to correct for differences between data and MC in the identification, isolation, and tracking of the electrons and muons are taken from the Standard Model $H \rightarrow \tau\tau$ analysis [62], due to the similarities in event and object selection.

8.7 Recoil corrections

In proton-proton collisions, W and Z bosons are predominantly produced through quark-antiquark annihilation. Higher-order processes can induce radiated quarks or gluons that recoil against the boson, imparting a non-zero transverse momentum to the boson [63]. Recoil corrections accounting for this effect are applied to samples with W+jets, Z+jets, and Higgs bosons [54]. The corrections are performed on the vectorial difference between the measured missing transverse momentum and the total transverse momentum of neutrinos originating from the decay of the W, Z, or Higgs boson. This vector is projected onto the axes parallel and orthogonal to the boson p_T . This vector, and the resulting correction to use, is measured in $Z \rightarrow \mu\mu$ events, since these events have leptonic recoil that do not contain neutrinos, allowing the 4-vector of the Z boson to be measured precisely. The corrections are binned in generator-level p_T of the parent boson and also the number of jets in the event.

8.8 Drell-Yan corrections

The Z boson transverse momentum distribution disagrees between leading-order (LO) simulations and data in a $Z \rightarrow \mu\mu$ control region with at least one b-tag jet [64]. Per-event weights derived by the 2016 data-only version of this analysis [64] are applied to $Z \rightarrow \tau\tau/\ell\ell$ events, as a function of the generator-level Z boson p_T to provide better matching of MC to data.

8.9 Pileup reweighting

Reweighting is performed to rescale MC events to account for differences between MC and data, in the distribution of the pileup (number of additional proton-proton interactions per bunch crossing). A tool for calculating the pileup reweighting for the

MC samples used is provided centrally by the Luminosity POG [65].

8.10 Trigger efficiencies

Scale factors are applied to correct for differences in trigger efficiencies between MC and data, and embedded samples and data. They are taken from the Standard Model $H \rightarrow \tau\tau$ working group which uses similar trigger paths [54].

8.11 Top p_T spectrum reweighing

In Run-1 and Run-2 it was observed that the p_T spectra of top quarks in $t\bar{t}$ data was significantly softer than those predicted by MC simulations [66]. Possible sources of this discrepancy are higher order QCD and/or electroweak corrections, and non-resonant production of $t\bar{t}$ -like final states. To account for this, corrections derived from Run-2 data by the Top Physics Analysis Group (PAG) are applied to the p_T of the top and anti-top quarks in MC simulations, computed as a function of their generator-level p_T [66].

8.12 B-tagging efficiency

In order to predict correct b-tagging discriminant distributions and event yields in data, the weight of selected MC events is reweighed according to recommendations by the BTV POG [67]. The reweighing depends on the jet p_T , η , and the b-tagging discriminant. In this method, there is no migration of events from one b-tag multiplicity bin to another.

8.13 Jet energy resolution and jet energy smearing

Calibration of jet energies, i.e. ensuring that the energy and momentum of the reconstructed jet matches that of the quark/gluon-initiated jet, is a challenging task due to time-dependent changes in the detector response and calibration and high pileup [68] [69]. Jet calibration is done via jet energy corrections (JECs) applied to the p_T of jets in MC samples, accounting successively for the effects of pileup, uniformity of the detector response, and residual data-simulation jet energy scale differences [70]. Typical jet energy resolutions reported at $\sqrt{s} = 8$ TeV in the central rapidities are 15-20% at 30 GeV and about 10% at 100 GeV [68]. Jet energy corrections are also propagated to the missing transverse energy.

Measurements show that the jet energy resolution (JER) in data is worse than in simulation, and so the jets in MC need to be smeared to describe the data. JER corrections are applied after JEC on MC simulations, and adjust the width of the p_T distribution based on pileup, jet size, and jet flavour [71]. Tools for applying JEC and JER are provided centrally by the JER Corrections group.

Chapter 9

Background estimation

This section describes methods used to estimate sources of background from Standard Model processes in the search for $h \rightarrow aa \rightarrow bb\tau\tau$. Similar background estimation methods are being used for the $h \rightarrow a_1a_2$ analysis. The background contributions directly taken from MC are described first, followed by backgrounds estimated from data-driven methods to produce sufficient statistics in the signal region.

9.1 Z+jets

A major source of background for $\tau\tau$ analyses is the Drell-Yan (DY) process (Z+jets). The Z boson decays to $\tau\tau/\mu\mu/ee$ with equal probability of 3.4% each, with the dominant decay modes being to hadrons (around 70%) and neutrinos (invisible) (20%) [23].

The Drell-Yan contribution with genuine taus, $Z \rightarrow \tau\tau$, is estimated using embedded samples, described in Section 6.3. To avoid double-counting between embedded and MC samples, in all MC samples, events with legs that originated from genuine τ are discarded.

The other decays of the Z, $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$, are estimated from MC simulation, and are hereafter referred to as simply the Drell-Yan background. These MC samples

are generated to leading order (LO) with different numbers of jets (jet multiplicity) in the matrix element: Z+1 jet, Z+2jets, Z+3 jets, Z+4 jets, and inclusive Z+jets. The cross-sections of the samples with ≥ 1 jets are normalized to next-to-NLO (NNLO) in QCD.

For the inclusive Drell-Yan sample, two samples are used with different thresholds for the di-lepton invariant mass ($m_{\ell\ell}$) at the generator level: one with $m_{\ell\ell} > 50$ GeV and the other with $10 < m_{\ell\ell} < 50$.

9.2 W+jets

The dominant W boson decay modes are to hadrons (67.4%), $e + \nu_e$ (10.7%), $\mu + \nu_\mu$ (10.6%), and $\tau + \nu_\tau$ (11.4%) [23]. The W+jets background is estimated from MC simulation. Similarly to the Z+jets, the W+jets samples are generated with different jet multiplicities in the matrix element. LO samples are used for greater statistics and are normalized to NNLO cross sections.

9.3 $t\bar{t}$ + jets

In hadron collisions, top quarks are produced singly with the weak interaction, or in pairs via the strong interaction, with interference between these leading-order processes possible in higher orders of the perturbation theory. The top quark is the heaviest fermion in the Standard Model and has a short lifetime ($\sim 10^{-25}$ s), decaying without hadronization into a bottom quark and a W boson [23], with the decay modes of the W boson as listed in the previous section. With two top quarks, the final states of the two resulting W bosons can be described as fully leptonic, semileptonic, and fully hadronic. These three final states are modeled separately with MC simulation in 2018 and 2017, while for 2016 the sample used is inclusive.

9.4 Single top

There are three main production modes of the single top in pp collisions [72]: the exchange of a virtual W boson (t channel), the production and decay of a virtual W boson (s channel), and the associated production of a top quark and W boson (tW , or W-associated) channel. As the s channel process is rare and only 3% of the total production, the dominant production mode of the t -channel and the tW production are considered and modeled with MC.

9.5 Diboson

In pp collisions, the production of dibosons (pairs of electroweak gauge bosons, i.e. WW, WZ, and ZZ) is dominated by quark-antiquark annihilation, with a small contribution from gluon-gluon interaction [73]. MC is used to model the pair production and decays of VV to $2\ell 2\nu$, WZ to $2q 2\ell$ and $3\ell\nu$, and ZZ to 4ℓ and $2q 2\ell$ (q being quarks and ℓ being leptons).

9.6 Standard Model Higgs

MC is used to simulate backgrounds from major production modes of the Standard Model 125 GeV Higgs boson: gluon-gluon fusion (ggH), vector boson fusion (VBF), associated production with a W or Z (WH, ZH), and associated production with a top pair (ttH) (see Fig. 9.1 for leading-order diagrams). For these production modes, samples with the Higgs decaying to $\tau\tau$ or to WW are used. Samples made with higher-order diagrams for WH and ZH that include the production of a jet, with the Higgs decaying to WW, are also used.

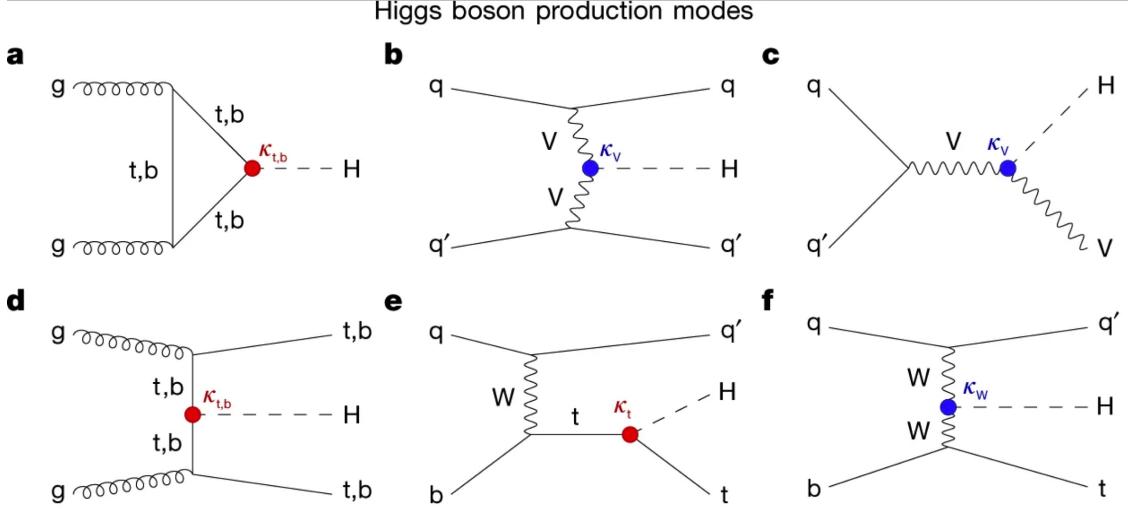


Figure 9.1: Leading-order Feynman diagrams of Higgs production from [74], in ggH (a) and vector boson fusion (VBF; b), associated production with a W or Z (V) boson (VH; c), associated production with a top or bottom quark pair (ttH or bbH); d, and associated production with a single top quark (tH; e, f).

9.7 Jet faking τ_h for the $\mu\tau_h$ and $e\tau_h$ channels

Events with a jet mis-reconstructed as the hadronic tau leg τ_h are a major source of background in the $\mu\tau_h$ and $e\tau_h$ channels. The main processes contributing to jet $\rightarrow \tau_h$ events are QCD multijet, W+jets, and $t\bar{t}$ production. These events are estimated using a data-driven method adapted from past analyses [62] [64]. This background includes contributions from W+jets, QCD multijets, and $t\bar{t}$ +jets. To estimate this background, a sideband region is constructed, where events are required to pass all baseline $\mu\tau_h/e\tau_h$ selection criteria, but fail the τ_h isolation criteria. The events in this sideband region are reweighted with a factor $f/(1 - f)$, where f is the probability for a jet to be misidentified as a τ_h . The jet $\rightarrow \tau_h$ background is the anti-isolated, reweighted MC and embedded events subtracted from the anti-isolated, reweighted data events.

The fake factor is measured in $Z \rightarrow \mu\mu + \text{jets}$ events in data in the $\mu\mu\tau_h$ final state, as any reconstructed τ_h in these events must originate from a jet. The two muons are required to be isolated (< 0.15), have opposite electric charge, and have

an invariant mass between 76 and 106 GeV (close to the Z mass). These events are selected with a double muon trigger, with the leading muon having offline $p_T > 20$ GeV and the subleading muon $p_T > 10$ GeV. Simulated diboson (ZZ and WZ) events are subtracted to avoid contamination from events with real τ_h . The denominator of the fake rate corresponds to fake taus passing the VVVLoose working point of the discriminator vs. jets, while the numerator corresponds to those passing the Medium working point, i.e. $f = N_{\text{jet passing tight}} / N_{\text{jet passing loose}}$.

f is measured as a function of the τ_h transverse momentum and is 8% - 10% in each of the data-taking years. f is derived separately for the $\mu\tau_h$ and $e\tau_h$ channels because the channels use different anti-lepton identification working points.

9.8 QCD multijet background for the $e\mu$ channel

In the $e\mu$ channel, events with jets faking electrons or muons originating from QCD multijet, is estimated from data events with the same baseline selection as in the signal region, except with same-signed (SS) charged $e + \mu$, ensuring orthogonality with the signal region which requires opposite-sign (OS) $e\mu$ pairs. All same-sign MC events (both events with real and fake $e + \mu$) are subtracted from same-sign data events to remove contamination from other backgrounds. i.e. $\text{QCD}_{\text{SS}} = \text{Data}_{\text{SS}} - \text{MC}_{\text{SS}}$.

Three scale factors are applied to the QCD_{SS} events to compute the QCD multijet background [64] [45]:

- *OS-to-SS scale factor:* This scales the SS QCD to the OS region, and is measured from an orthogonal region with an isolated electron and an anti-isolated muon. Only the muon is chosen to be anti-isolated because this scale factor was observed to depend more strongly on electron isolation than that of the muon. This scale factor is treated as a function of the ΔR separation of the trajectories

of the electron and muon, and is measured separately for events with 0 jets, 1, jet, and greater than 1 jet.

- *2D closure correction for the lepton p_T :* This factor accounts for subleading dependencies of the first scale factor on the p_T of the two leptons. A 2D weight is derived in a similar fashion, as a ratio of QCD_{OS} events to QCD_{SS} events, but parameterized by both electron and muon p_T , where the SS events have the previous scale factor applied.
- *Isolation correction for the muon:* The third and final factor is an isolation correction, which is a bias correction to account for the fact that the fake factor was determined for less-isolated muons. This factor is obtained as the ratio of the OS-to-SS scale factors measured in two other control regions: (1) events where the electron is anti-isolated ($0.15 < \text{iso} < 0.5$) and the muon is isolated, and (2) events where both leptons are anti-isolated.

Chapter 10

Event categorization

10.1 B-tag jet multiplicity

The increased statistics of the full Run-2 dataset enables the separation of events into events with exactly 1 b-tag jet and events with greater than 1 b-tag jet. Further event categorization is performed with deep neural networks (DNNs) described below. The DNNs are used only for separating events into signal and control regions in the 1 b-tag and 2 b-tag jets scenarios. The final results are extracted from the statistical fitting to the mass of the $\tau\tau$, $m_{\tau\tau}$.

10.2 DNN-based event categorization

A brief overview of the DNN-based event categorization is given below with a focus on the physics aspects, with full details of the machine learning training in [45] and associated documentation.

Training samples

Neural networks for event categorization are trained for each of the $\mu\tau_h$, $e\tau_h$, and $e\mu$ channels, for 1 and 2 b-tag jets, giving $3 \times 2 = 6$ networks in total. In the training,

the signal is taken to be all of the possible pseudoscalar mass m_a hypotheses together. The backgrounds for each DNN are taken to be a representative combination of the three major backgrounds: $Z \rightarrow \tau\tau$, $t\bar{t} + \text{jets}$, and fake backgrounds. The proportions of each background for each channel and b-tag jet multiplicity are taken from the yields in the $m_{\tau\tau}$ distribution. For instance, in the $\mu\tau_h$ 1 b-tag jet category, the composition of the background for training is 17.4% from $Z \rightarrow \tau\tau$, 42.4% from $t\bar{t} + \text{jets}$, and 40.2% fakes.

Input variables

The input variables capture the key differences between the signal and the background:

- Transverse momentum p_T of the electron and muon in the $e\tau_h$ and $\mu\tau_h$ channels, where the signal tends to have a softer p_T spectrum (lower energy) than the background.
- p_T of the b-tag jet(s). The signal sample b-tag jet(s) tend to have softer p_T .
- Invariant masses of the various objects ($\tau\tau$ legs and the b-tag jet(s)), which tend to be smaller for the signal samples.
- The angular separation ΔR between pairs of the objects, where signal samples peak at smaller ΔR values.
- The transverse mass between the missing transverse energy p_T^{miss} and each of the four objects [64], defined as

$$m_T(\ell, p_T^{\text{miss}}) \equiv \sqrt{2p_T^\ell \cdot p_T^{\text{miss}}[1 - \cos(\Delta\phi)]} \quad (10.1)$$

where p_T^ℓ is the transverse momentum of the object ℓ , and $\Delta\phi$ is the difference in azimuthal angle between the object and the p_T^{miss} . Events from $t\bar{t} + \text{jets}$ and

jets faking τ_h backgrounds have larger p_T^{miss} resulting in larger transverse mass values compared to the signal, which tends to have smaller p_T^{miss} that is also more aligned with the lepton legs.

- The variable D_ζ [64], defined as

$$D_\zeta \equiv p_\zeta - 0.85p_\zeta^{\text{vis}} \quad (10.2)$$

where the ζ axis is the bisector of the transverse directions of the visible τ decay products. p_ζ is the component of the p_T^{miss} along the ζ axis, and p_ζ^{vis} is the sum of the components of the lepton p_T along the same axis. This variable captures the fact that in signal the p_T^{miss} is small and approximately aligned with the $\tau\tau$. In contrast, the $Z \rightarrow \tau\tau$ background tends towards large D_ζ values because the p_T^{miss} is collinear to the $\tau\tau$, and the $t\bar{t}$ +jets events tend to have small D_ζ due to a large p_T^{miss} not aligned with the $\tau\tau$.

- For events with 2 b-tag jets, one additional variable is defined to capture the difference in the invariant mass of the bb and the $\tau\tau$:

$$\Delta m_{a_1} \equiv (m_{bb} - m_{\tau\tau})/m_{\tau\tau} \quad (10.3)$$

This variable peaks at zero for the $h \rightarrow aa \rightarrow 2b2\tau$ signal.

Categorization using the DNN score

After training, events in data, MC, and embedded are evaluated with the six DNNs and assigned a raw score between 0 and 1 (background-like or signal-like). In order to flatten the distribution of the score and define score thresholds for categorizing events, the raw output scores are transformed with the function $\tilde{p}(n) = \text{arctanh}(p \times \tanh(n))/n$ where n is a positive integer. The thresholds of the DNN score used for

signal/control region definition are determined using scans that optimize the signal sensitivity and are shown in Tables 10.1 and 10.2.

	1bNN $\tilde{p}(n = 1.5)$			
	SR1	SR2	SR3	CR
$\mu\tau_h$ 2018	> 0.98	$\in [0.95, 0.98]$	$\in [0.90, 0.95]$	< 0.90
$\mu\tau_h$ 2017	> 0.97	$\in [0.94, 0.97]$	$\in [0.90, 0.94]$	< 0.90
$\mu\tau_h$ 2016	> 0.97	$\in [0.94, 0.97]$	$\in [0.89, 0.94]$	< 0.89
	1bNN $\tilde{p}(n = 1.5)$			
	SR1	SR2	SR3	CR
$e\tau_h$ 2018	> 0.97	$\in [0.945, 0.97]$	$\in [0.90, 0.945]$	< 0.90
$e\tau_h$ 2017	> 0.985	$\in [0.965, 0.985]$	$\in [0.93, 0.965]$	< 0.93
$e\tau_h$ 2016	> 0.985	$\in [0.965, 0.985]$	$\in [0.93, 0.965]$	< 0.93
	1bNN $\tilde{p}(n = 2.5)$			
	SR1	SR2	SR3	CR
$e\mu$ 2018	> 0.99	$\in [0.95, 0.99]$	$\in [0.85, 0.95]$	< 0.85
$e\mu$ 2017	> 0.985	$\in [0.95, 0.985]$	$\in [0.85, 0.95]$	< 0.85
$e\mu$ 2016	> 0.99	$\in [0.95, 0.99]$	$\in [0.85, 0.95]$	< 0.85

Table 10.1: Event categorization based on DNN scores for events with exactly 1 b-tag jet (1bNN), for the three $\tau\tau$ channels and three eras.

	2bNN $\tilde{p}(n = 1.5)$		
	SR1	SR2	CR
$\mu\tau_h$ 2018	> 0.99	$\in [0.96, 0.99]$	< 0.96
$\mu\tau_h$ 2017	> 0.98	$\in [0.94, 0.98]$	< 0.94
$\mu\tau_h$ 2016	> 0.97	$\in [0.93, 0.97]$	< 0.93
	2bNN $\tilde{p}(n = 1.5)$		
	SR1	SR2	CR
$e\tau_h$ 2018	> 0.96	NA	< 0.96
$e\tau_h$ 2017	> 0.985	NA	< 0.985
$e\tau_h$ 2016	> 0.96	NA	< 0.96
	2bNN $\tilde{p}(n = 2.5)$		
	SR1	SR2	CR
$e\mu$ 2018	> 0.98	$\in [0.94, 0.98]$	< 0.94
$e\mu$ 2017	> 0.97	$\in [0.93, 0.97]$	< 0.93
$e\mu$ 2016	> 0.98	$\in [0.94, 0.98]$	< 0.94

Table 10.2: Event categorization based on DNN scores for events with 2 b-tag jets (2bNN), for the three $\tau\tau$ channels and three eras.

Chapter 11

Systematic uncertainties

The handling of systematic uncertainties is separated into normalization uncertainties (those that affect the total yield of a variables' distribution) and shape uncertainties (those that shift the distribution of events). Normalization uncertainties are expressed as multiplicative factors, while shape uncertainties are represented as up and down shifts of a variable's distribution.

Up/down shifts of shape uncertainties can change the number of background events in a distribution. For instance, hadronic taus receive corrections from the nominal tau energy scale, with the nominal, up, and down energy scales provided centrally by CMS. For the $\mu\tau_h$ channel, an event could have a τ_h with p_T just below the offline threshold of 20 GeV (for instance, 19.5 GeV), so in the nominal distribution of $m_{\tau\tau}$ (or any other variable for this channel), the event is excluded. However, when we build our distributions with the tau energy scale “up” shift, the energy of this τ_h may be scaled up to, say, 20.5 GeV, and now the event passes the offline p_T threshold for the single muon trigger, leading to the event's inclusion in the distributions made with the tau energy scale “up” shift.

In evaluating the up and down shifts of a specific source of uncertainty, all other corrections and scale factors are held at their nominal values, and the full chain

of object and event selection and event categorization is performed to obtain the observable distributions. Any “downstream” variables that depend on the shifted variable, e.g. the invariant di-tau mass $m_{\tau\tau}$, must be computed for the nominal case, and then re-computed separately for each up and down shift of the tau legs’ energy scale. The objective of this process is to quantify the effect of a single source of uncertainty on the resulting observable distributions.

11.1 Uncertainties associated with physics objects

Each scale factor and correction described in Ch. 8 has an associated uncertainty. The binning of the uncertainties follows that of the nominal scale factor value.

11.1.1 Uncertainties in the lepton energy scales

The uncertainties in the tau energy scales [52] are binned by the tau decay mode and are taken as shape uncertainties treated as uncorrelated across the tau decay modes and years. Same as with the application of the nominal scale factor, when applying the up or down shifts, the missing transverse energy (p_T^{miss}) of the event is adjusted so that the 4-vector sum of the tau p_T^{miss} is unchanged.

The uncertainties in the muon energy scale [55] are 0.4% for $|\eta| < 1.2$, 0.9% for $1.2 < |\eta| < 2.1$, and 2.7% for $2.1 < |\eta| < 2.4$, and are treated as shape uncertainties, fully uncorrelated between embedded and MC samples.

The uncertainties in the electron energy scale [57] in MC are 0.5% and 0.69% for the barrel and endcap respectively. The uncertainties for the embedded sample are on the order of 0.5% and 1.25% for the barrel and endcap [61].

There are also uncertainties in the energy scales for electrons and muons misidentified as τ_h . The uncertainty for muons misidentified as τ_h is 1% [52]. For electrons misidentified as τ_h , the uncertainty is binned in barrel/endcap η and by 1-prong and

1-prong + π_0 decays. The probability for e/μ faking a 3-prong decay mode is much lower.

11.1.2 Uncertainties from other lepton corrections

Uncertainties associated with the τ_h identification efficiencies are treated as shapes, uncorrelated across the seven p_T bins and years. The shape uncertainties in the embedded samples are taken as 50% correlated with those of the MC samples.

The uncertainties on electron and muon identification efficiencies are taken as normalization uncertainties of 2% each, with a 50% correlation between embedded and MC samples.

In the $e\tau_h$ channel, there is an additional uncertainty for the vs. jet discrimination efficiency [52], because the analysis uses a looser anti-lepton working point (VLoose WP) than the working points used in the measurement of the efficiency (namely, VLoose WP vs e, and Tight WP vs mu). For nominal $\tau_h p_T < 100$ GeV, an additional uncertainty of 3% (5%) is used in MC (embedded), and for high p_T an uncertainty of 15% is used for both.

The uncertainties in trigger efficiencies are taken as shapes [52]. In the $e\tau_h$ and $\mu\tau_h$ channels, there are uncertainties for the single and cross lepton triggers, and in the $e\mu$ channel there is one uncertainty each for the two $e + \mu$ triggers, and one combined uncertainty since their trigger phase spaces are not mutually exclusive.

11.1.3 Uncertainties from jet energy scale and resolution

The jet energy scale uncertainties are taken as shape uncertainties: there are eleven in total, with seven correlated across years (labeled “Year” below) and the remainder uncorrelated across years. They affect the b-tag jet p_T and mass, and hence the missing transverse energy p_T^{miss} . The shifts are propagated through the b-tagging scale factor calculation and b-tag jet counting.

The uncertainties in the jet energy correction and resolution [68] [75] are as follows:

- *Absolute, AbsoluteYear*: flat absolute scale uncertainties.
- *BBEC1, BBEC1Year*: for sub-detector regions, with barrel “BB” in $|\eta| < 1.3$ and endcap region 1 “EC1”: $1.3 < |\eta| < 2.5$.
- *EC2, EC2 year*: for sub-detector regions, with endcap region 2 “EC2” in $2.5 < |\eta| < 3.0$.
- *HF, HF year*: for sub-detector regions, with hadron forward “HF” in $|\eta| > 3$.
- *FlavorQCD*: for uncertainty in jet flavour (uds/c/b-quark and gluon) estimates based on comparing Pythia and Herwig (different MC generator) predictions.
- *RelativeBal*: account for difference between log-linear fits of the two methods used to study the jet energy response: MPF (missing transverse momentum projection fraction) and p_T balance.
- *RelativeSample*: account for η -dependent uncertainty due to a difference between relative residuals, observed with dijet and Z+jets in Run D of 2018 data.
- *JetResolution*: uncertainty in the jet energy resolution.

11.1.4 Uncertainties from b-tagging scale factors

The b-tagging scale factor has its own set of associated uncertainties (not to be confused with shifts in the b-tagging scale factor due to the propagation of the jet energy scale uncertainties described in the previous section 11.1.3). They are:

- *hf*: contamination from heavy flavour (b+c) jets in the light flavour region.
- *hfstats1, hfstats2*: linear and quadratic statistical fluctuations from b jets.
- *lf*: contamination from light flavour (udsg+c jets) in the heavy flavour region.

- $lfstats1$, $lfstats2$: linear and quadratic statistical fluctuations from udsg jets.
- $cferr$, $cferr2$: uncertainty for charm jets.

The variations for “lf, hf, hfstats1/2, lfstats1/2” are applied to both b and udsg jets.

For c-flavour jets, only “cferr1/2” is applied.

11.1.5 Uncertainties from MET

Samples where recoil corrections were applied ($Z+jets$, $W+jets$, and Standard Model Higgs, as described in Section 8) have uncertainties from the response and resolution of the hadronic recoil against the leptonic system. These are each binned in jet multiplicity.

11.2 Uncertainties associated with samples used

Normalization uncertainties related to the samples used are:

- *Cross-section uncertainties*: $\sigma(t\bar{t})$: 4.2%, $\sigma(\text{diboson})$: 5%, $\sigma(\text{single top})$: 5%, $\sigma(\text{ggH})$: 3.2%, $\sigma(\text{qQH})$: 2.1%, $\sigma(\text{WH})$: 1.9%, $\sigma(\text{ZH})$: 1.3%, $\sigma(\text{ttH})$: 3.6%
- *Uncertainties in QCD renormalization scale*: QCD scale(qqH): +0.43%-0.33%, QCD scale(WH): +0.5%-0.7%, QCD scale(ttH): +5.8%-9.2%
- *Branching ratio uncertainties*: $\text{BR}(\text{H} \rightarrow \tau\tau)$: 1.8%, and $\text{BR}(\text{H} \rightarrow \text{WW})$: 1.5%.
- *Normalization uncertainties*: 2% for Drell-Yan, 4\$ for embedded, 20% pre-fit for the QCD multijet background in the $e\mu$ channel, 20% pre-fit for the jet faking background.

The $t\bar{t}$ process has additional acceptance uncertainties from QCD scale variation and parton shower uncertainties [76]. Parton shower uncertainties originate from the modeling of perturbative and non-perturbative QCD effects handled in parton

shower MC generators. The scale variations are determined from the envelope of the 6 provided shapes due to variations in the factorization scale, renormalization scale, and their combined variation [76].

The Z p_T reweighting uncertainty in Drell-Yan samples is taken to be 10% of the nominal value, taken as a shape uncertainty.

The fake rate uncertainties are taken as shape uncertainties. For the weight applied to scale up anti-isolated events in cross-trigger regions, 20% of the nominal weight is taken as a shape uncertainty.

11.3 Other uncertainties

A 3.6% yield uncertainty in the signal is used to cover uncertainties in the parton distribution functions, α_s (fine structure constant), and QCD scale.

Normalization uncertainties from luminosity are applied to all MC samples, divided into those uncorrelated across years, those correlated between 2017 and 2018, and one for 2018 [65].

11.4 Pulls and impacts

The top impacts and pulls computed for the combination of all channels and years is shown in Fig. 11.1. The top impacts are related to uncertainty in the signal sample and cross-section of the $t\bar{t}$ cross-section, and also the yields of the jet faking τ_h background, which is a major background in all channels and expected to be constrained due to the yield uncertainty which is taken to be 20% pre-fit.

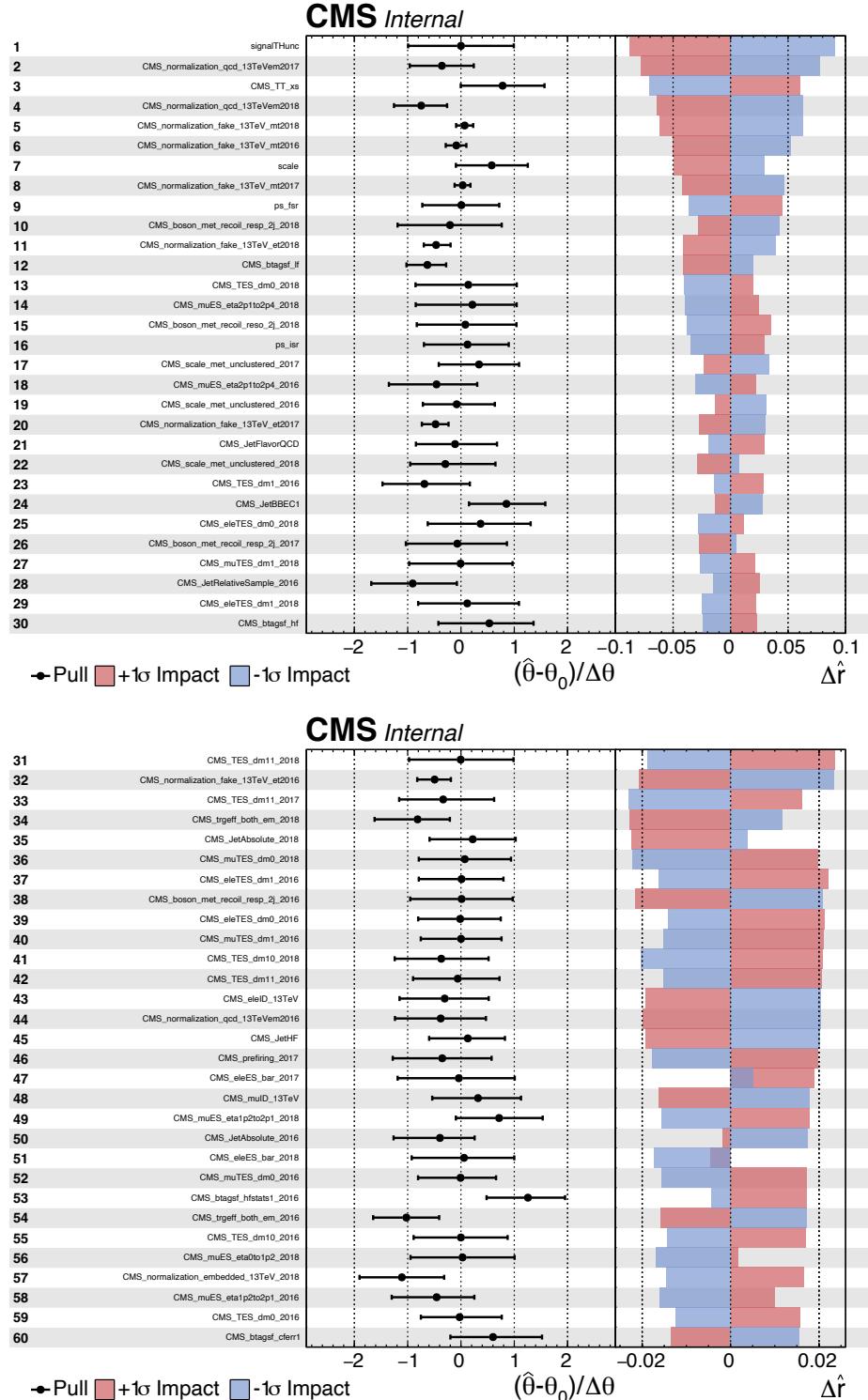


Figure 11.1: Top sixty impacts for the combination of all channels and years [77].

Chapter 12

Signal extraction from the di-tau mass

12.1 Estimation of the $\tau\tau$ mass

The final signal extraction is done to the total $\tau\tau$ mass, which is estimated from the visible $\tau\tau$ mass using the FastMTT algorithm [78]. FastMTT is based on the SVFit algorithm, originally developed for the Standard Model $H \rightarrow \tau\tau$ analysis [79]. Both the SVFit algorithms, and the FastMTT algorithm, are described below, to give a complete picture of how tau decays are parametrized.

To specify a hadronic τ decay, six parameters are needed [79]: the polar and azimuthal angles of the visible decay product system in the τ rest frame, the three boost parameters from the τ rest frame to the laboratory frame, and the invariant mass m_{vis} of the visible decay products. For a leptonic τ decay, two neutrinos are produced, and a seventh parameter, the invariant mass of the two-neutrino system, is necessary. The unknown parameters are constrained by four observables that are the components of the four-momentum of the system formed by the visible decay products of the τ lepton, measured in the laboratory frame. The remaining unconstrained

parameters for hadronic and leptonic τ decays are thus:

- The fraction of the τ energy in the laboratory frame carried by the visible decay products,
- ϕ , the azimuthal angle of the τ direction in the laboratory frame,
- $m_{\nu\nu}$, the invariant mass of the two-neutrino system in leptonic τ decays (for hadronic τ decays, $m_{\nu\nu}$ is set to 0).

E_x^{miss} and E_y^{miss} , the x and y components of the missing transverse energy E_T^{miss} provide two further constraints.

12.1.1 Original SVFit ‘‘standalone’’: maximum likelihood

In one of the original versions of SVFit, called ‘‘standalone’’ SVFit [79], a maximum likelihood fit method is used to reconstruct the mass $m_{\tau\tau}$ by combining the measured observables E_x^{miss} and E_y^{miss} with a likelihood model that includes terms for the τ decay kinematics and the E_T^{miss} resolution [79]. The likelihood function $f(\vec{z}, \vec{y}, \vec{a}_1 \vec{a}_2)$ of the parameters $\vec{z} = (E_x^{\text{miss}}, E_y^{\text{miss}})$ in an event is constructed, where the remaining parameters are the kinematics of the two τ decays, denoted $\vec{a}_1 = (x_1, \phi_1, m_{\nu\nu,1})$ and $\vec{a}_2 = (x_2, \phi_2, m_{\nu\nu,2})$, and the four-momenta of the visible decay products with the measured values $\vec{y} = (p_1^{\text{vis}}, p_2^{\text{vis}})$.

The likelihood f is the product of three likelihood functions. The first two likelihood functions model the decay parameters \vec{a}_1 and \vec{a}_2 of the two τ leptons. For leptonic decays, the likelihood function is modeled using matrix elements for τ decays, and integrated over the allowed phase space $0 \leq x \leq 1$ and $0 \leq m_{\nu\nu} \leq m_\tau \sqrt{1-x}$. For hadronic τ decays, a model based on the two-body phase space is used and integrated over $m_{\text{vis}}^2/m_{\tau\tau}^2 \leq x \leq 1$. The third likelihood function quantifies the compatibility of a τ decay hypothesis with the reconstructed \vec{E}_T^{miss} in an event, assuming the neutrinos are the only source of missing transverse energy. The expected \vec{E}_T^{miss} resolution

is represented by a covariant matrix, estimated on an event-by-event basis using a significance algorithm [80].

12.1.2 “Classic SVFit” with matrix element

Classic SVFit is an improved algorithm of the original “standalone” SVFit using the formalism of the matrix element (ME) method [78]. In the ME method, an estimate for the unknown model parameter Θ (here, the mass $m_{\tau\tau}$) is obtained by maximizing the probability density \mathcal{P} . The key ingredients of the probability density are the squared modulus of the matrix element $|\mathcal{M}(\mathbf{p}, \Theta)|^2$ and the transfer function $W(\mathbf{y}|\mathbf{p})$ (probability density to observe the measured observables \mathbf{y} given the phase space point \mathbf{p}). The best estimate $m_{\tau\tau}$ is obtained by computing the probability density \mathcal{P} for a range of mass hypotheses and finding the value of $m_{\tau\tau}$ that maximizes \mathcal{P} .

Distributions illustrating the performance of the classic matrix element SVFit algorithm are shown in Fig. 12.1 from [78], showing the di-tau mass after and before application of SVFit to recover energy lost to neutrinos. The SVFit algorithm is found to improve the sensitivity of the Standard Model $H \rightarrow \tau\tau$ analysis performed by CMS by about 30%, compared to performing the same analysis using only the visible mass m_{vis} .

12.1.3 FastMTT: optimized SVFit

FastMTT [81] is a further simplification to the matrix element method of Classic SVFit which has comparable performance but is about 100 times faster. FastMTT drops the matrix element component of the computation without significant impact on the final mass resolution, and simplifies the computation of the transfer functions. The opening angle of the τ decay products with respect to the initial τ momenta approaches 0 for τ with high $\gamma = E_\tau/m_\tau$, with typical τ decays from the Z boson decays already satisfying this condition. In this collinear approximation, the dimensionality

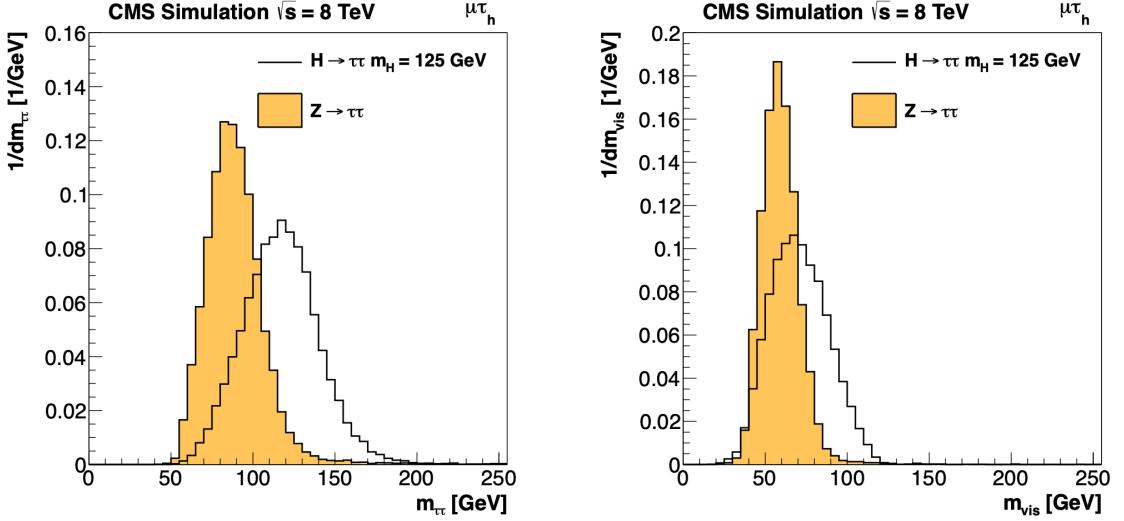


Figure 12.1: Distributions from [78], of $m_{\tau\tau}$ after reconstruction with the original SVFit algorithm (*left*), and before SVFit with only the visible tau decay products (*right*), for $H \rightarrow \tau\tau$ signal events of mass $m_H = 125$ GeV (black line) and the $Z/\gamma^* \rightarrow \tau\tau$ background (orange, solid), in the decay channel $\tau\tau \rightarrow \mu\tau_h$.

of the transfer function can be reduced in the computation of FastMTT, while still yielding similar results to Classic SVFit [81].

12.2 Signal extraction

A binned maximum likelihood fit is performed on the $m_{\tau\tau}$ distribution with the systematic uncertainties described in Chapter 11. Event categories described in Section 10, in all three channels, are included in a simultaneous fit.

Limits and confidence intervals are obtained using the modified frequentist CL_s approach, with an asymptotic approximation to the distribution of the profile likelihood test statistic.

Chapter 13

Results

13.1 Results from $bb\tau\tau$

The postfit final $m_{\tau\tau}$ distributions in the signal region (SR) and control regions (CRs) for 1 and 2 b-tag jet multiplicities, are shown for the $\mu\tau_h$ channel in Fig. 13.1, $e\tau_h$ channel in Fig. 13.2, and $e\mu$ channel in Fig. 13.3.

The 95% CL exclusion limits on the signal strength of the branching ratio $\mathcal{B}(h \rightarrow aa \rightarrow bb\tau\tau)$, shown in percentage and normalized to the Standard Model Higgs production cross-section, is shown in Fig. 13.4. The pseudoscalar mass hypotheses m_a between 15 GeV and 60 GeV are searched for in all channels. The $e\mu$ channel is the only channel that can provide sensitivity to the $m_a = 12$ GeV mass point, as its cut on the ΔR between the two τ legs is the smallest, which increases the signal acceptance for low mass signals.

No excess of events above the Standard Model expectations is observed. Combined expected (observed) limits between 1.4 and 5.6% (1.7 and 7.6%) are set for the pseudoscalar mass range between 12 and 60 GeV. The best sensitivity is attained at intermediate mass points, since the analysis targets a resolved signature: at low mass points, objects have a boosted signature, and at high mass points, the $m_{\tau\tau}$

distributions in signal have larger overlap with background distributions.

The branching fraction of the two pseudoscalars depends on the 2HDM+S model type, the pseudoscalar mass m_a , and the ratio of the two vacuum expectation values $\tan \beta$. In Type I models, the branching fraction is independent of $\tan \beta$, while in Type III and IV models, it is a function of m_a and $\tan \beta$. Using the predicted values of $B(a \rightarrow ff)$, the observed limits can be used to constrain the $\tan \beta$ vs. m_a phase space.

13.2 Combination procedure with $bb\mu\mu$ final state

A combination of the results from $bb\tau\tau$ is performed with the analysis for the $h \rightarrow aa \rightarrow bb\mu\mu$ final state [82]. While the $bb\mu\mu$ final state has a small branching ratio, it provides competitive results due to the excellent di-muon resolution measured by CMS. The $bb\mu\mu$ analysis uses an unbinned fit to the data using the di-muon mass $m_{\mu\mu}$ distribution.

A combination is possible because the two final states use event selections that are orthogonal to each other, as the $bb\tau\tau$ channels veto events with extra leptons. Several systematic uncertainties are treated as correlated: the integrated luminosity normalization, the b-tagging scale factor, the scale factors related to muon reconstruction, identification, and trigger efficiencies, the inefficiency in the ECAL trigger readout, and the theoretical uncertainties related to signal modeling.

The dimuon and ditau mass distributions ($m_{\mu\mu}$ and $m_{\tau\tau}$) are compared to the data in a combined maximum likelihood fit to derive upper limits on $B(h \rightarrow aa)$.

13.3 Results from combination

The observed limits at 95% CL on $B(h \rightarrow aa)$ for different 2HDM+S scenarios are shown for the $bb\mu\mu$ analysis in Fig. 13.5, the $bb\tau\tau$ analysis in Fig. 13.6, and the

combined analyses in Fig. 13.7. For 2HDM+S Types II, III, and IV, exclusion limits are set as a function of $\tan \beta$ and m_a , and are shown in Fig. 13.8.

The most stringent constraints are observed for the Type III model, which predicts combined branching fractions between 0.47 and 0.42 for $\tan \beta = 2.0$ and values of m_a between 15 and 60 GeV, while the 95% CL upper limits are between 0.08 and 0.03. For the Type IV model, the predicted branching fractions from theory are between 0.26 and 0.20 for $\tan \beta = 0.6$ for values of m_a between 15 and 60 GeV, and the 95% CL observed upper limits are between 0.12 and 0.05.

The combined analyses' results are also shown in summary plots alongside the other current CMS results for 2HDM+S limits: for type I in Fig. 13.9, for type II with $\tan \beta = 2.0$ in Fig. 13.10, and for type III with $\tan \beta = 2.0$ in Fig. 13.11. In other scenarios, e.g. type III with $\tan \beta = 5.0$, more stringent limits are set by analyses in other final states, $\mu\mu\tau\tau$ in this case. Other summary plots can be found at [83].

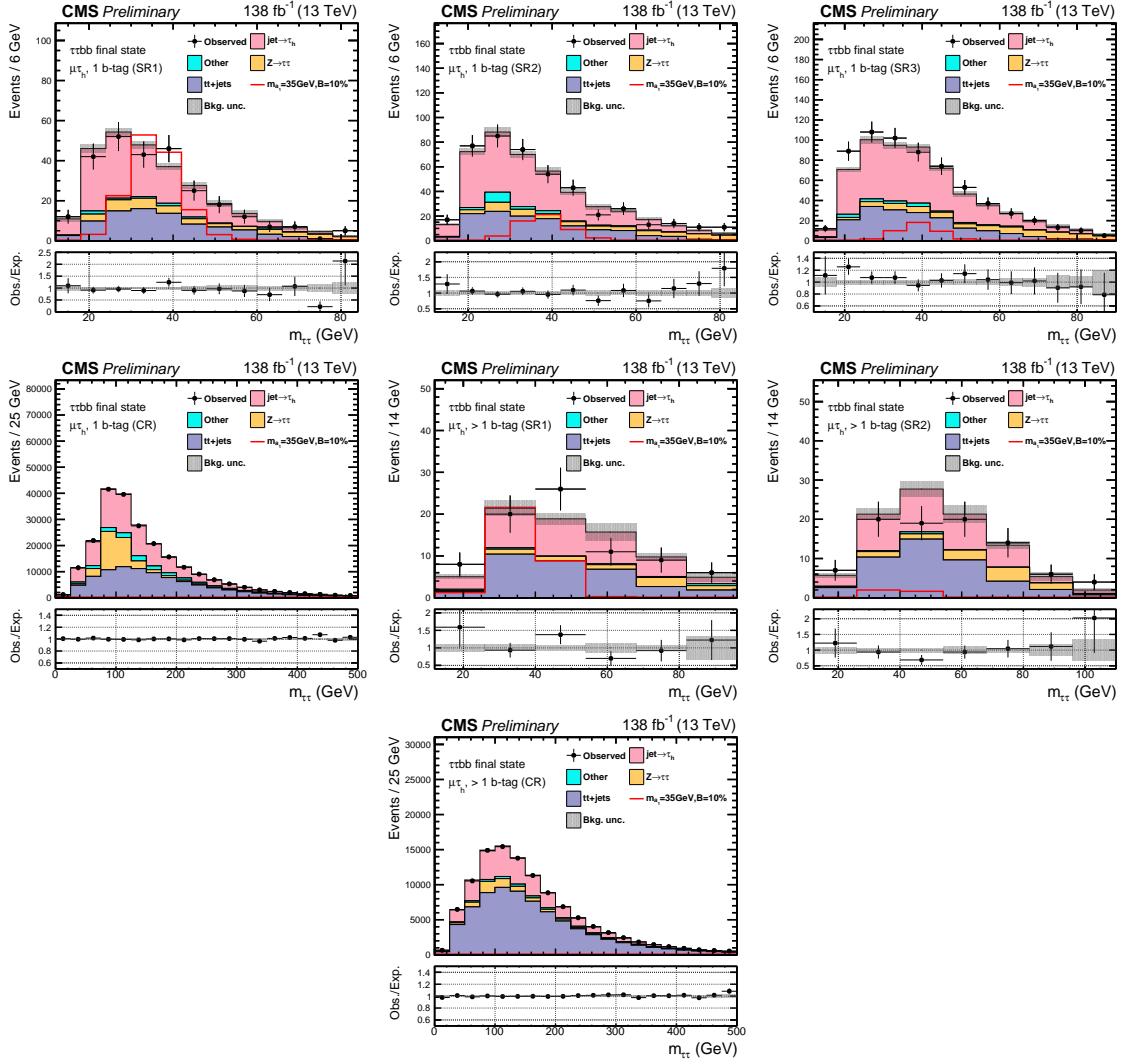


Figure 13.1: Postfit final $m_{\tau\tau}$ distributions in the $\mu\tau_h$ channel [77]. Statistical and systematic uncertainties are included. *Top row:* 1 b-tag jet categories: three signal regions (SR1, SR2, SR3). *Middle row, left to right:* 1 b-tag jet categories: control region (CR), and 2 b-tag jet categories: two signal regions (SR1, SR2). *Bottom:* 2 b-tag jet categories: control region (CR).

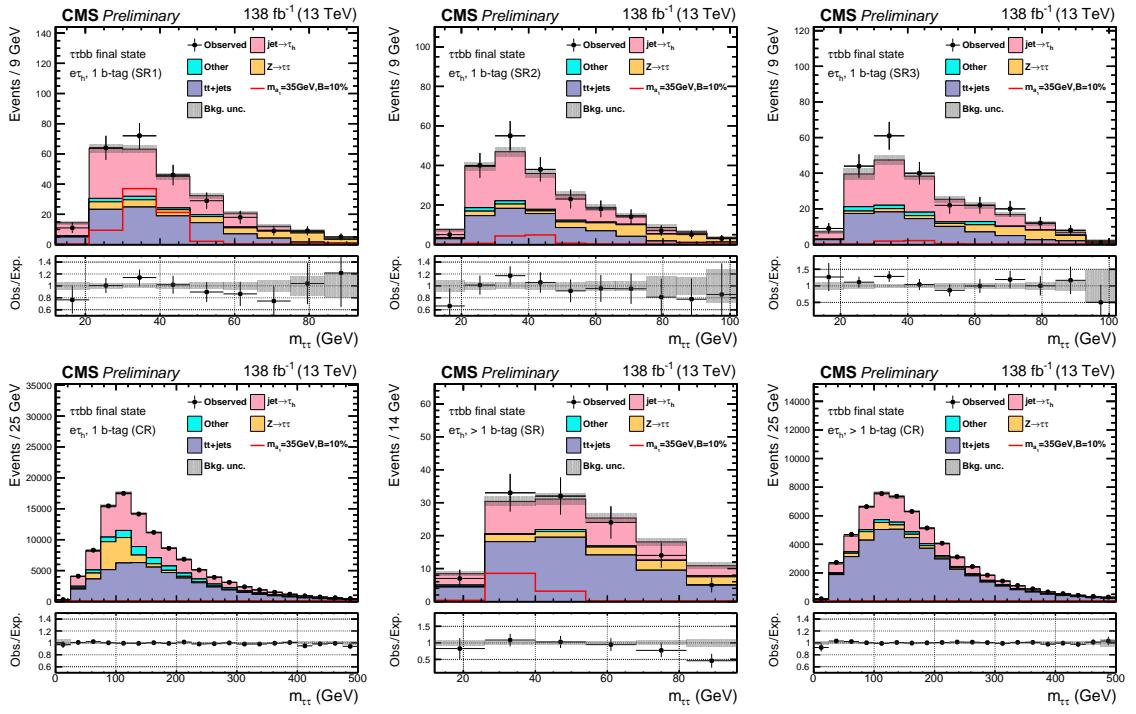


Figure 13.2: Postfit final $m_{\tau\tau}$ distributions in the $e\tau_h$ channel [77]. Statistical and systematic uncertainties are included. *Top row:* 1 b-tag jet categories: three signal regions (SR1, SR2, SR3). *Bottom row, left to right:* 1-btag jet categories: control region (CR), and 2 b-tag jet categories: signal region (SR) and control region (CR).

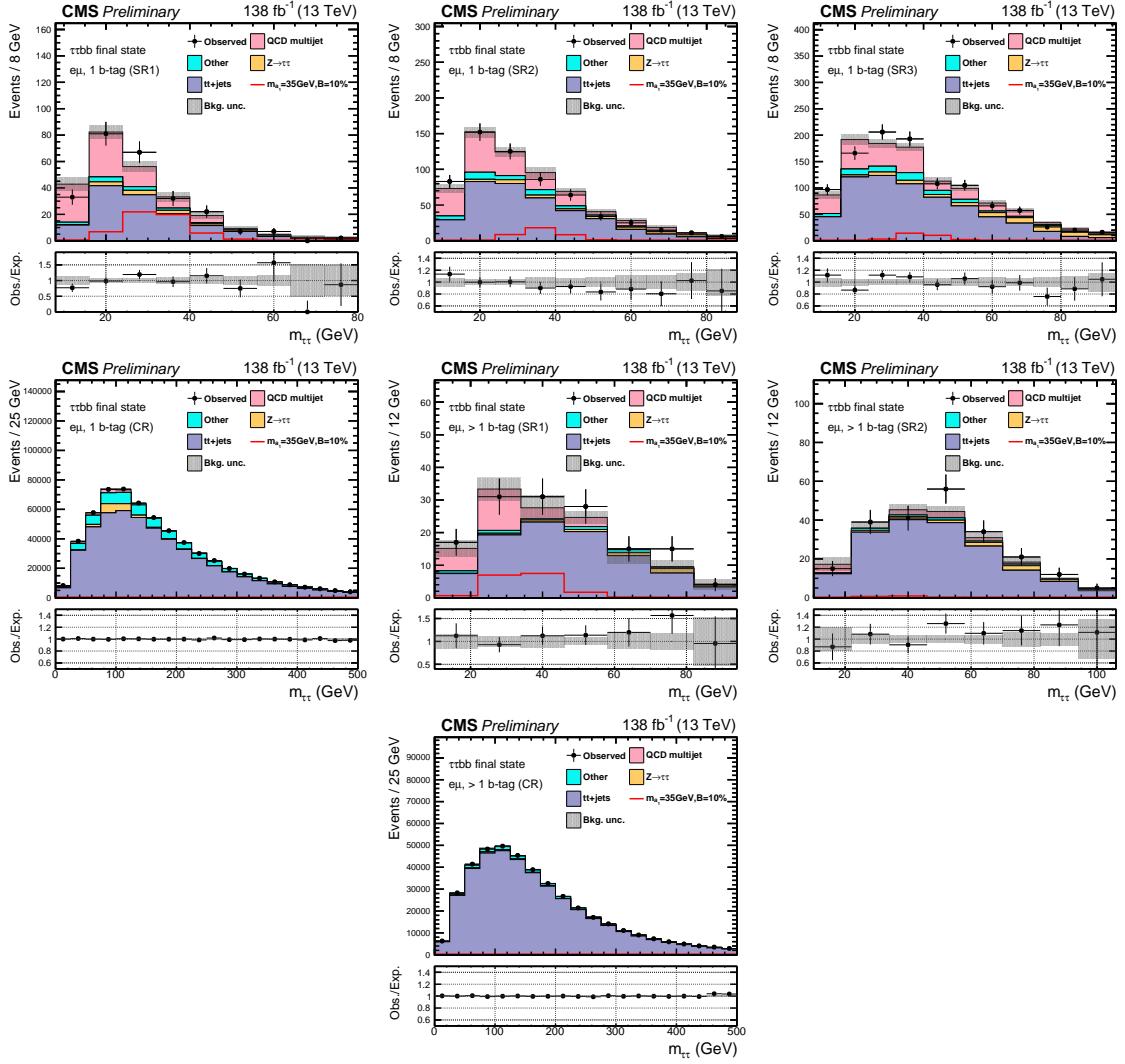


Figure 13.3: Postfit final $m_{\tau\tau}$ distributions in the $e\mu$ channel [77]. Statistical and systematic uncertainties are included. *Top row:* 1 b-tag jet categories: three signal regions (SR1, SR2, and SR3). *Middle row, left to right:* 1 b-tag jet categories: control region (CR), and 2 b-tag jet categories: signal regions (SR1 and SR2). *Bottom:* 2 b-tag jet categories: control region (CR).

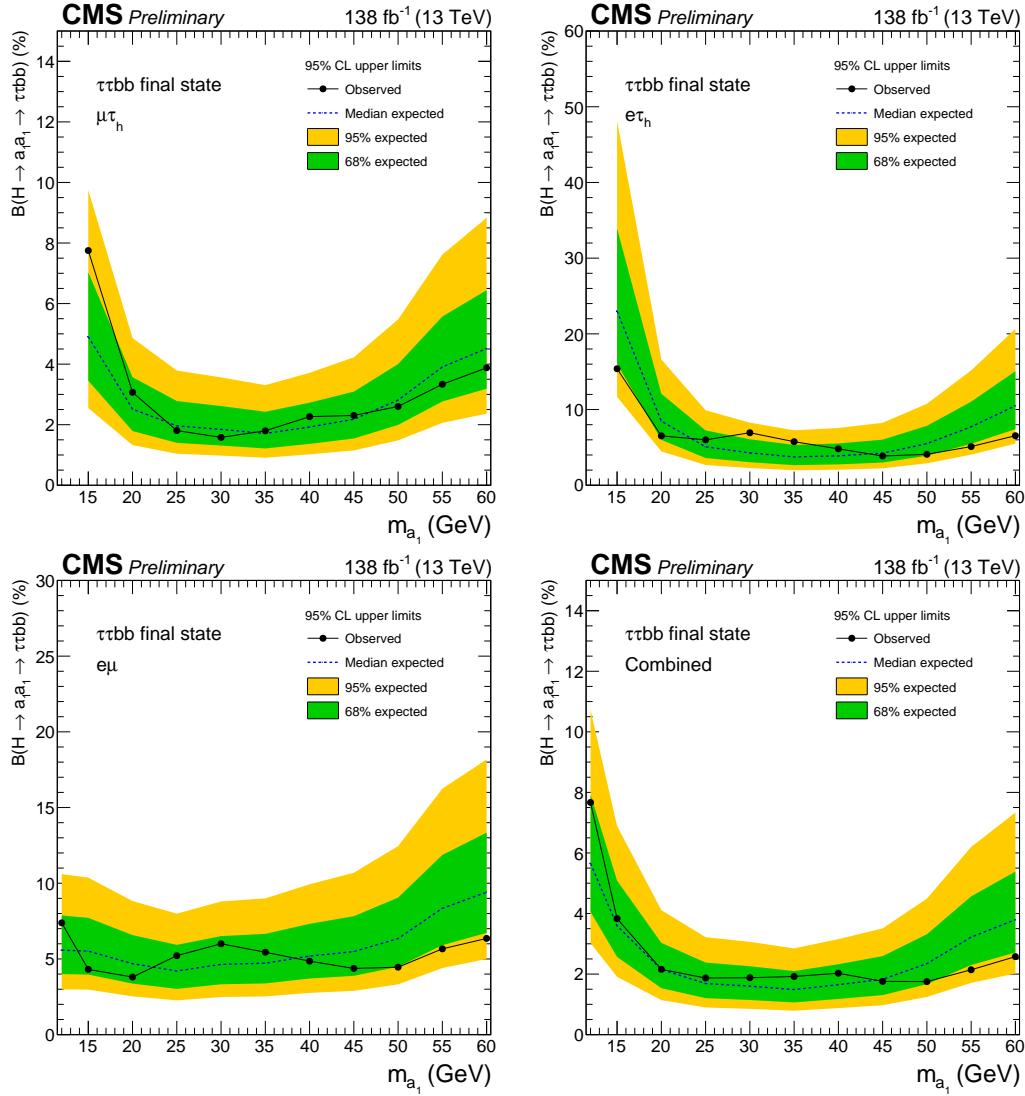


Figure 13.4: 95% CL exclusion limits on $B(h \rightarrow aa \rightarrow bb\tau\tau)$ in %, for the combination of all years by channel (*top left*: $\mu\tau_h$ channel, *top right*: $e\tau_h$ channel, and *Bottom left*: $e\mu$ channel) and the combination of all channels (*bottom right*) [77].

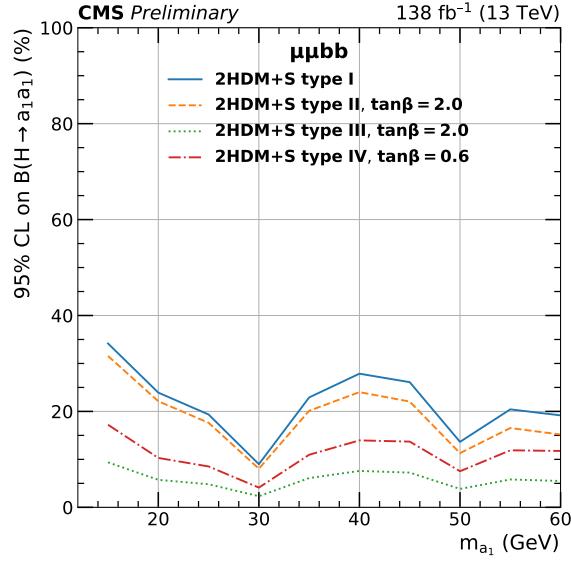


Figure 13.5: Observed 95% CL upper limits on $B(h \rightarrow aa_1)$ in %, for the $bb\mu\mu$ channel using the full Run 2 integrated luminosity of 138 fb^{-1} for different 2HDM+S models. Limits on the Type I 2HDM+S model do not depend on $\tan\beta$ [77].

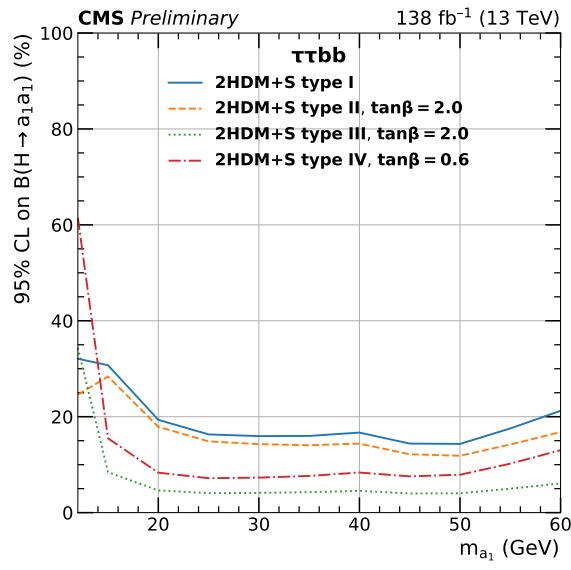


Figure 13.6: Observed 95% CL upper limits on $B(h \rightarrow aa_1)$ in %, for the $bb\tau\tau$ channel using the full Run 2 integrated luminosity of 138 fb^{-1} for different 2HDM+S models. Limits on the Type I 2HDM+S model do not depend on $\tan\beta$ [77].

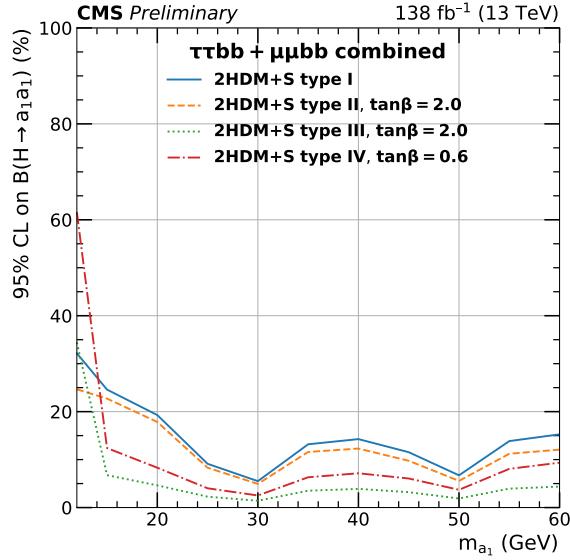


Figure 13.7: Observed 95% CL upper limits on $\mathcal{B}(h \rightarrow aa)$ in %, for the combination of $bb\mu\mu$ and $bb\tau\tau$ channels using the full Run 2 integrated luminosity of 138 fb^{-1} for different 2HDM+S models. Limits on the Type I 2HDM+S model do not depend on $\tan \beta$ [77].

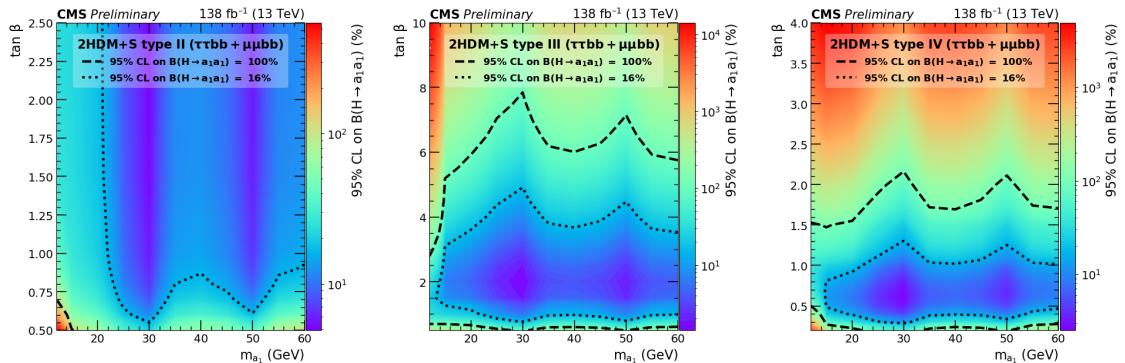


Figure 13.8: Observed 95% CL upper limits on $\mathcal{B}(h \rightarrow aa)$ in %, for the combination of $bb\mu\mu$ and $bb\tau\tau$ channels using the full Run 2 integrated luminosity of 138 fb^{-1} for Type II (left), Type III (middle), and Type IV (right) 2HDM+S in the $\tan \beta$ vs. m_a phase space. The contours corresponding to branching ratios of 100% and 34% are drawn using dashed lines, 34% corresponding to the previous upper limit on Higgs to BSM particle decays during Run 1. Linear extrapolation has been used between different points on the figures [77].

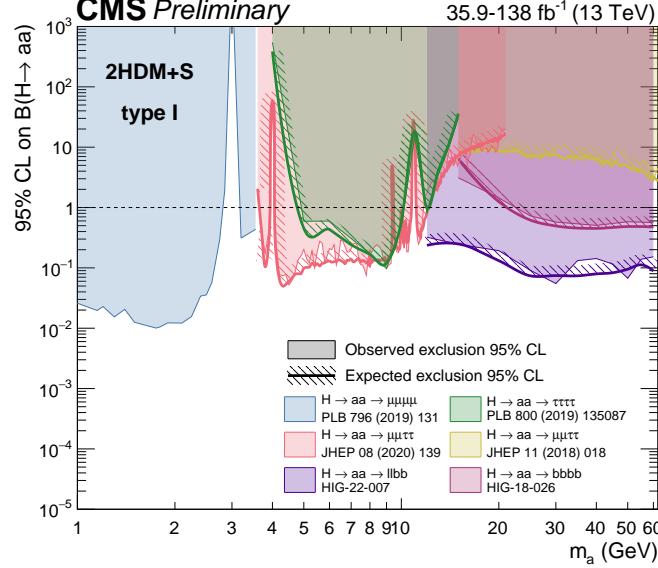


Figure 13.9: 95% limits on $\frac{\sigma(h)}{\sigma_{\text{SM}}} \times \mathcal{B}(h \rightarrow aa)$ in the 2HDM+S type-I scenario for exotic Higgs decay searches, performed with data collected at 13 TeV. Branching fractions of the pseudoscalar to Standard Model particles are computed following prescriptions in [84]. The combined result for $h \rightarrow aa \rightarrow \ell\ell bb$ [45] is shown in purple, representing the most stringent limits in the probed mass ranges in this scenario.

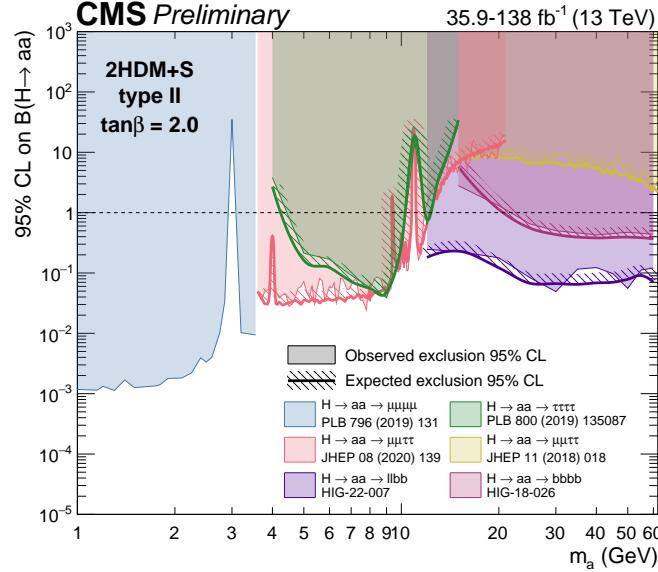


Figure 13.10: 95% limits on $\frac{\sigma(h)}{\sigma_{\text{SM}}} \times \mathcal{B}(h \rightarrow aa)$ in the 2HDM+S type-II scenario with $\tan \beta = 2.0$ for exotic Higgs decay searches, performed with data collected at 13 TeV. Branching fractions of the pseudoscalar to Standard Model particles are computed following prescriptions in [84]. The combined result for $h \rightarrow aa \rightarrow \ell\ell bb$ [45] is shown in purple, representing the most stringent limits in the probed mass ranges in this scenario.

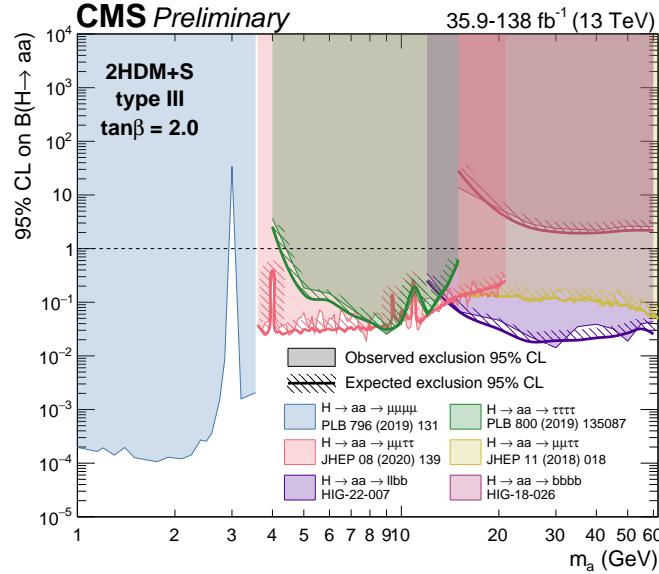


Figure 13.11: 95% limits on $\frac{\sigma(h)}{\sigma_{\text{SM}}} \times \mathcal{B}(h \rightarrow aa)$ in the 2HDM+S type-III scenario with $\tan \beta = 2.0$ for exotic Higgs decay searches, performed with data collected at 13 TeV. Branching fractions of the pseudoscalar to Standard Model particles are computed following prescriptions in [84]. The combined result for $h \rightarrow aa \rightarrow llbb$ [45] is shown in purple, representing the most stringent limits in the probed mass ranges in this scenario.

Bibliography

- [1] Paul H. Frampton. Journeys Beyond the Standard Model. 54(1):52–52. ISSN 0031-9228. doi: 10.1063/1.1349615. URL <https://doi.org/10.1063/1.1349615>. eprint: https://pubs.aip.org/physicstoday/article-pdf/54/1/52/11109432/52_1_online.pdf.
- [2] Christopher G. Tully. *Elementary Particle Physics in a Nutshell*. Princeton University Press, Princeton, 2012. ISBN 9781400839353. doi: doi:10.1515/9781400839353. URL <https://doi.org/10.1515/9781400839353>.
- [3] John Ellis. Higgs Physics. In *2013 European School of High-Energy Physics*, pages 117–168, 2015. doi: 10.5170/CERN-2015-004.117.
- [4] David Curtin, Rouven Essig, Stefania Gori, and Others. Exotic decays of the 125 GeV Higgs boson. *Phys. Rev. D*, 90: 075004, Oct 2014. doi: 10.1103/PhysRevD.90.075004. URL <https://link.aps.org/doi/10.1103/PhysRevD.90.075004>.
- [5] Tania Robens, Tim Stefaniak, and Jonas Wittbrodt. Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios. *Eur. Phys. J. C*, 80(2):151, 2020. doi: 10.1140/epjc/s10052-020-7655-x.
- [6] Antonella Del Rosso. Aerial view of the LHC and the four major experiments. 2017. URL <https://cds.cern.ch/record/2253966>. General Photo.

- [7] ATLAS Collaboration. ATLAS: Detector and physics performance technical design report. Volume 1. 5 1999.
- [8] CMS Collaboration. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006.
- [9] L Musa. Conceptual Design Report for the Upgrade of the ALICE ITS. Technical report, CERN, Geneva, 2012. URL <https://cds.cern.ch/record/1431539>.
- [10] S. Amato et al. LHCb technical proposal: A Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays. 2 1998.
- [11] R. Schmidt. Accelerator physics and technology of the LHC. In *ROXIE: Routine for the Optimizazion of Magnet X-Sections, Inverse Field Calculation and Coil End Design*, pages 7–17, 1998.
- [12] J. Vollaire et al. *Linac4 design report*, volume 6/2020 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 9 2020. ISBN 978-92-9083-579-0, 978-92-9083-580-6. doi: 10.23731/CYRM-2020-006.
- [13] Werner Herr and B Muratori. Concept of luminosity. 2006. doi: 10.5170/CERN-2006-002.361. URL <https://cds.cern.ch/record/941318>.
- [14] Olivier S. Brning and Frank Zimmerman. Parameter space for the LHC luminosity upgrade. ISBN 978-3-95450-115-1. URL <https://accelconf.web.cern.ch/IPAC2012/papers/moppc005.pdf>.
- [15] CMS Collaboration. Pileup mitigation at CMS in 13 TeV data. *JINST*, 15(09):P09018, 2020. doi: 10.1088/1748-0221/15/09/P09018.
- [16] CMS Collaboration. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV. *JHEP*, 07:161, 2018. doi: 10.1007/JHEP07(2018)161.

- [17] CMS Collaboration. High-Luminosity Large Hadron Collider (HL-LHC): Technical design report. 10/2020, 12 2020. doi: 10.23731/CYRM-2020-0010.
- [18] CMS Collaboration. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
- [19] CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017. doi: 10.1088/1748-0221/12/10/P10003.
- [20] V Karimki, M Mannelli, P Siegrist, H Breuker, A Caner, R Castaldi, K Freudenreich, G Hall, R Horisberger, M Huhtinen, and A Cattai. *The CMS tracker system project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997. URL <https://cds.cern.ch/record/368412>.
- [21] The Phase-2 Upgrade of the CMS Tracker. Technical report, CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2272264>.
- [22] CMS Collaboration. CMS Technical Design Report for the Pixel Detector Upgrade. 9 2012. doi: 10.2172/1151650.
- [23] R. L. Workman and Others. Review of particle physics. 2022:083C01. doi: 10.1093/ptep/ptac097.
- [24] CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter. 9 2012. doi: 10.2172/1151651.
- [25] CMS Collaboration. The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger. Technical report, CERN, Geneva, 2021. URL <https://cds.cern.ch/record/2759072>. This is the final version of the document, approved by the LHCC.

- [26] C. Foudas. The CMS Level-1 Trigger at LHC and Super-LHC. In *34th International Conference on High Energy Physics*, 10 2008.
- [27] CMS Software Guide. High Level Trigger (TWiki), 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideHighLevelTrigger>.
- [28] The Worldwide LHC Computing Grid. 2012. URL <https://cds.cern.ch/record/1997398>.
- [29] Douglas Thain, Todd Tannenbaum, and Miron Livny. Distributed computing in practice: the Condor experience. *Concurrency and Computation: Practice and Experience*, 17(2-4):323–356, 2005. doi: <https://doi.org/10.1002/cpe.938>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.938>.
- [30] Alexandre Zabi, Jeffrey Wayne Berryhill, Emmanuelle Perez, and Alexander D. Tapper. The Phase-2 Upgrade of the CMS Level-1 Trigger. 2020.
- [31] Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade. Technical report, CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2296612>.
- [32] CMS Collaboration. Reconstruction and identification of tau lepton decays to hadrons and tau neutrinos at CMS. *Journal of Instrumentation*, 11(01):P01019–P01019, January 2016. ISSN 1748-0221. doi: 10.1088/1748-0221/11/01/p01019. URL <http://dx.doi.org/10.1088/1748-0221/11/01/P01019>.
- [33] CMS Collaboration. Performance of τ -lepton reconstruction and identification in CMS. *Journal of Instrumentation*, 7(01):P01001, jan 2012. doi: 10.1088/1748-0221/7/01/P01001. URL <https://dx.doi.org/10.1088/1748-0221/7/01/P01001>.

- [34] CMS Collaboration. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(10):P10005, 2018. doi: 10.1088/1748-0221/13/10/P10005.
- [35] CMS Collaboration. Identification of hadronic tau lepton decays using a deep neural network. *JINST*, 17:P07023, 2022. doi: 10.1088/1748-0221/17/07/P07023.
- [36] CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation*, 7(10):P10002–P10002, October 2012. ISSN 1748-0221. doi: 10.1088/1748-0221/7/10/p10002. URL <http://dx.doi.org/10.1088/1748-0221/7/10/P10002>.
- [37] CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Journal of Instrumentation*, 10(06):P06005, 2015. doi: 10.1088/1748-0221/10/06/P06005. URL <https://dx.doi.org/10.1088/1748-0221/10/06/P06005>.
- [38] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *Journal of Instrumentation*, 8(04):P04013–P04013, April 2013. ISSN 1748-0221. doi: 10.1088/1748-0221/8/04/p04013. URL <http://dx.doi.org/10.1088/1748-0221/8/04/P04013>.
- [39] CMS Collaboration. Pileup Removal Algorithms. Technical report, CERN, Geneva, 2014. URL <https://cds.cern.ch/record/1751454>.
- [40] CMS Collaboration. CMS Phase 1 heavy flavour identification performance and developments. 2017. URL <https://cds.cern.ch/record/2263802>.
- [41] CMS Collaboration. Performance of the DeepJet b tagging algorithm using 41.9 fb^{-1} of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector. 2018. URL <https://cds.cern.ch/record/2646773>.

- [42] CMS Collaboration. CMS luminosity measurement for the 2016 data-taking period. CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017. URL <https://cds.cern.ch/record/2257069>.
- [43] CMS Collaboration. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018. URL <https://cds.cern.ch/record/2621960>.
- [44] CMS Collaboration. CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019. URL <https://cds.cern.ch/record/2676164>.
- [45] CMS Collaboration. Search for exotic decays of the Higgs boson to a pair of pseudoscalars in the $\mu\mu bb$ and $\tau\tau bb$ final states. 2 2024.
- [46] Search for exotic Higgs boson decays to a pair of pseudoscalars in the $\mu\mu bb$ and $\tau\tau bb$ final states in proton-proton collisions with the CMS experiment. Technical report, CERN, Geneva, 2023. URL <https://cds.cern.ch/record/2853298>.
- [47] J. Alwall, R. Frederix, S. Frixione, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014 (7), July 2014. ISSN 1029-8479. doi: 10.1007/jhep07(2014)079. URL [http://dx.doi.org/10.1007/JHEP07\(2014\)079](http://dx.doi.org/10.1007/JHEP07(2014)079).
- [48] R. Frederix, S. Frixione, V. Hirschi, et al. The automation of next-to-leading order electroweak calculations. *Journal of High Energy Physics*, 2018 (7), July 2018. ISSN 1029-8479. doi: 10.1007/jhep07(2018)185. URL [http://dx.doi.org/10.1007/JHEP07\(2018\)185](http://dx.doi.org/10.1007/JHEP07(2018)185).
- [49] S. Agostinelli, J. Allison, K. Amako, et al. Geant4 - a simulation toolkit.

- 506(3):250–303. ISSN 0168-9002. doi: 10.1016/S0168-9002(03)01368-8. URL <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [50] CMS Collaboration. An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data. *JINST*, 14(06):P06032, 2019. doi: 10.1088/1748-0221/14/06/P06032.
- [51] CMS MUO POG. Muon Physics Object Group: Baseline muon selections for Run-II, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>.
- [52] CMS TAU POG. Tau Physics Object Group: Tau ID Recommendation For Run 2, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendationForRun2>.
- [53] CMS ELE POG. Electron Identification Based on Simple Cuts, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EgammaPublicData>.
- [54] CMS HTT working group. Higgs To Tau Tau Working TWiki for the full Run-2 legacy analysis, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/HiggsToTauTauWorkingLegacyRun2>.
- [55] CMS MUO POG. Muon Physics Object Group: Recommendations, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG>.
- [56] CMS MUO POG. Muon Physics Object Group: Reference guidelines and results for muon momentum scale and resolution in Run II, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceScaleResolRun2>.
- [57] CMS ELE POG. Electron Physics Object Group: Recommendations, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaRunIIRecommendations>.

- [58] CMS ELE POG. Electron Physics Object Group: Recommendations for 2016 to 2018 UL, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaUL2016To2018>.
- [59] CMS TAU Embedding Group. Tau embedding samples using 2016 data, 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/TauTauEmbeddingSamples2016Legacy>.
- [60] CMS TAU Embedding Group. Tau embedding samples using 2017 data, 2024. URL <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauTauEmbeddingSamples2017>.
- [61] CMS TAU Embedding Group. Tau embedding samples using 2018 data, 2024. URL <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauTauEmbeddingSamples2018>.
- [62] CMS Collaboration. Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 83(7):562, 2023. doi: 10.1140/epjc/s10052-023-11452-8.
- [63] V.M. Abazov, B. Abbott, M. Abolins, et al. A novel method for modeling the recoil in W boson events at hadron colliders. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 609(2):250–262, 2009. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2009.08.056>. URL <https://www.sciencedirect.com/science/article/pii/S0168900209016623>.
- [64] CMS Collaboration. Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two b quarks and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 785:462, 2018. doi: 10.1016/j.physletb.2018.08.057.

- [65] CMS LUMI POG. Luminosity Physics Object Group: Recommendations, 2024.
URL <https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM>.
- [66] The modeling of the top quark p_T (TWiki), 2024. URL
<https://twiki.cern.ch/twiki/bin/view/CMS/TopPtReweighting>.
- [67] CMS BTV group. Methods to apply b-tagging efficiency scale factors (TWiki), 2024. URL
<https://twiki.cern.ch/twiki/bin/view/CMS/BTagShapeCalibration>.
- [68] CMS Collaboration. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12(02):P02014, 2017. doi: 10.1088/1748-0221/12/02/P02014.
- [69] Garvita Agarwal. Jet Energy Scale and Resolution Measurements in CMS. *PoS*, ICHEP2022:652, 2022. doi: 10.22323/1.414.0652.
- [70] CMS JERC group. Jet Energy Corrections (TWiki), 2024. URL
<https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC>.
- [71] CMS JERC group. Jet Energy Resolution (TWiki), 2024. URL
<https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution>.
- [72] Jose Enrique Palencia Cortezon. Single top quark production at CMS. Technical report, CERN, Geneva, 2018. URL <https://cds.cern.ch/record/2640578>.
- [73] CMS Collaboration. Measurements of the electroweak diboson production cross sections in proton-proton collisions at $\sqrt{s} = 5.02$ TeV using leptonic decays. *Phys. Rev. Lett.*, 127(19):191801, 2021. doi: 10.1103/PhysRevLett.127.191801.
- [74] CMS Collaboration. A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature*, 607(7917):60–68, 2022. doi: 10.1038/s41586-022-04892-x.

- [75] CMS JERC group. Jet energy scale uncertainty sources (TWiki), 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources>.
- [76] TOP Systematic Uncertainties (Run 2) (TWiki), 2024. URL <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopSystematics>.
- [77] Cécile Caillol, Pallabi Das, Sridhara Dasu, Pieter Everaerts, Stephanie Kwan, Isobel Ojalvo, and Ho-Fung Tsoi. CMS AN-20-213 (internal): Search for an exotic decay of the 125 GeV Higgs boson to light pseudoscalars, with a pair of b jets and a pair of tau leptons in the final state, 2020.
- [78] Lorenzo Bianchini, John Conway, Evan Klose Friis, and Christian Veelken. Reconstruction of the Higgs mass in $H \rightarrow \tau\tau$ Events by Dynamical Likelihood techniques. *Journal of Physics: Conference Series*, 513(2):022035, jun 2014. doi: 10.1088/1742-6596/513/2/022035. URL <https://dx.doi.org/10.1088/1742-6596/513/2/022035>.
- [79] CMS Collaboration. Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons. *JHEP*, 05:104, 2014. doi: 10.1007/JHEP05(2014)104.
- [80] CMS Collaboration. Missing transverse energy performance of the CMS detector. *JINST*, 6:P09001, 2011. doi: 10.1088/1748-0221/6/09/P09001.
- [81] Artur Kalinowski. CMS AN-19-032 (internal): Reconstruction of a τ pair invariant mass with a simplified likelihood scan, 2019.
- [82] Elham Khazaie, Maryam Zeinali, Hamed Bakhshiansohi, and Abideh Jafari. CMS AN-21-058 (internal): Search for exotic decays of the Higgs boson to a pair of new light bosons with two muons and two b jets in the final states at $\sqrt{s} = 13$ TeV, 2021.

- [83] CMS Higgs Physics Analysis Group. Summary of 2HDM+S searches at 13 TeV (Run 2), 2024. URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/Summary2HDMSSRun2>.
- [84] David Curtin, Rouven Essig, Stefania Gori, Prerit Jaiswal, Andrey Katz, Tao Liu, Zhen Liu, David McKeen, Jessie Shelton, Matthew Strassler, Ze’ev Surujon, Brock Tweedie, and Yi-Ming Zhong. Exotic decays of the 125 GeV Higgs boson. 90(7):075004. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.90.075004. URL <http://arxiv.org/abs/1312.4992>.