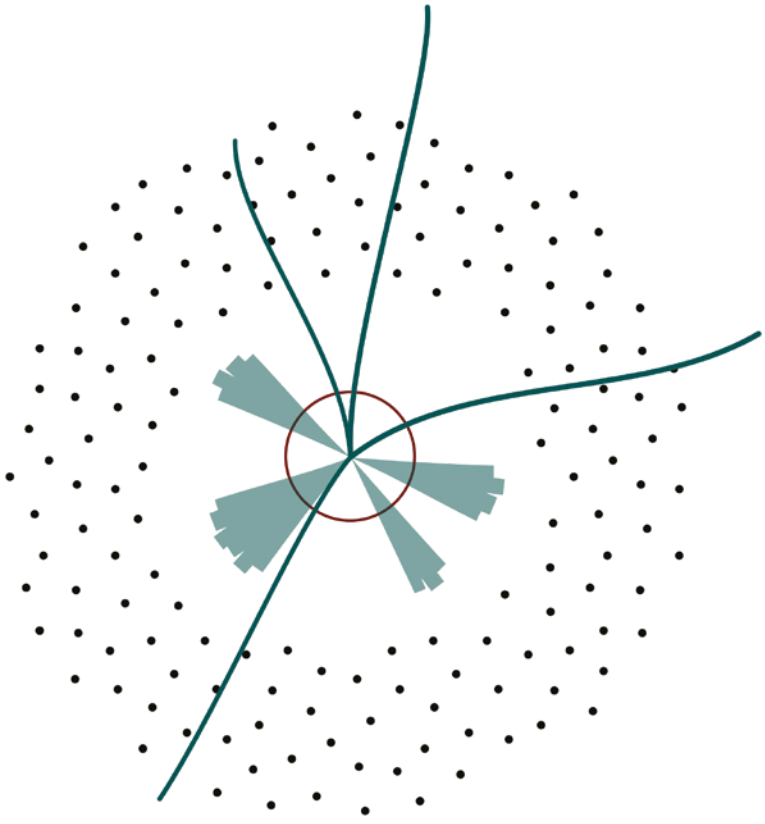
RIGA TECHNICAL UNIVERSITY

Antra Gaile

MEASUREMENT OF THE HIGGS BOSON PAIR PRODUCTION IN THE WWZZ DECAY CHANNEL AND DEVELOPMENT OF MTD CONTROL AND SAFETY SYSTEMS IN THE CMS EXPERIMENT

Summary of the Doctoral Thesis



RTU Press
Riga 2026

RIGA TECHNICAL UNIVERSITY

Faculty of Natural Sciences and Technology
Institute of Particle Physics and Accelerator Technologies

ANTRA GAILE
Doctoral Student of **Riga Technical University's** and **University of Latvia's** Study Programme "Particle Physics and Accelerator Technologies"

MEASUREMENT OF THE HIGGS BOSON PAIR PRODUCTION IN THE WWZZ DECAY CHANNEL AND DEVELOPMENT OF MTD CONTROL AND SAFETY SYSTEMS IN THE CMS EXPERIMENT

Summary of the Doctoral Thesis

Scientific supervisors:
Associate Professor Dr. phys.
TONI ŠĆULAC
Dr. phys.
FRANK GLEGE
Associate Professor Dr. phys.
KĀRLIS DREIMANIS

RTU Press
Riga 2026

Gaile, A. Measurement of the Higgs Boson Pair Production in the WWZZ Decay Channel and Development of MTD Control and Safety Systems in the CMS Experiment. Summary of the Doctoral Thesis. Riga: 2026. – 50 p.
Published in accordance with the decision of RTU IPPAT Promotion Council "RTU P-03" of 26 January 2026, Minutes No. HEP/1/2026.

The research and scientific mobility was supported by the following projects:

- State Research Programme (SRP) "High-energy physics and accelerator technologies" VPP-IZM-CERN-2020/1-0002; VPP-IZM-CERN-2022/1-0001
- COST Action CA22130 – COmprehensive Multiboson Experiment-Theory Action (COMETA)

Cover image by Antra Gaile.
https://doi.org/10.7250/9789934373022
ISBN 978-9934-37-302-2 (pdf)

DOCTORAL THESIS PROPOSED TO RIGA TECHNICAL UNIVERSITY FOR PROMOTION TO THE SCIENTIFIC DEGREE OF DOCTOR OF SCIENCE

To be granted the scientific degree of Doctor of Science (PhD), the present Doctoral Thesis has been submitted for defence at the open meeting of RTU Promotion Council "P-03" on June 19, 2026 at 15:00 at Riga Technical University, Paula Valdena Street, auditorium 271.

OFFICIAL REVIEWERS

Dr. phys. Giulia Di Gregorio
IJCLab, France

Associate professor Dr. phys. Ilaria Brivio
University of Bologna, Italy

Dr. phys. Rhiannon Susan Smith-Jones
University of Sheffield, United Kingdom

DECLARATION OF ACADEMIC INTEGRITY

I hereby declare that the Doctoral Thesis submitted for review to Riga Technical University for the promotion to the scientific degree of Doctor of Science (PhD) is my own. I confirm that this Doctoral Thesis has not been submitted to any other university for the promotion to a scientific degree.

Antra Gaile ...../signature/
Date .....

The Doctoral Thesis has been written in English. It consists of eight chapters, including Introduction and Conclusions, 74 figures, and 20 tables; the total number of pages is 169. The Bibliography contains 157 titles.

ABSTRACT

The continued exploration of the Higgs sector is central for the high energy physics community, it has also been set as one of the main missions for the Large Hadron Collider (LHC) physics program. More than ten years ago the discovery of Higgs boson production was claimed, and in around ten years the discovery of Higgs boson pair (HH) production could be declared. Measuring HH production not only tests one of the last unverified parameters of the Standard Model of particle physics (SM) but also provides a window into the mechanism that is believed to have shaped the early Universe. This Thesis presents the design and implementation of a world-first $HH \to WW^*ZZ^* \to 4l$ analysis. A dataset of 200 fb^{-1} collected with the Compact Muon Solenoid (CMS) detector across five data-taking years has been analysed. These results have allowed to set an observed (expected) 95 % confidence limit (CL) upper limit of 142.3 (176.8) times the SM value of the HH cross section. Findings showcase a proof-of-concept study of how an explicit inclusion of the $HH \to WW^*ZZ^*$ decay channel could improve the combined sensitivity for the HH production in future studies. Additionally, work has been presented on the detector control system (DCS) of a test-stand for one of the new CMS subdetectors, the Minimum ionising particle Timing Detector (MTD), built for the next era of the experiment. The goal of the MTD is to provide precision timing information for charged particles traversing it, thus ensuring flawless operation of the CMS detector. The methods and technical developments presented in this Thesis contribute to the long-term endeavour of ensuring that the CMS experiment remains at the forefront of precision Higgs physics in the years to come.

Contents

Acronyms

ATLAS – A Toroidal LHC ApparatuS . . . 7, 12, 14, 15, 17, 43
BDT – boosted decision tree . . . 30, 33–36, 38, 44
BR – branching ratio . . . 14, 27, 30, 40
BTL – barrel timing layer . . . 19, 23
CERN – European Organization for Nuclear Research . . . 16, 23, 25, 43
CL – confidence level . . . 14, 15, 37, 39, 40, 42, 43
CMS – Compact Muon Solenoid 7–10, 12, 14–20, 22, 24, 25, 27, 28, 30, 33, 37–40, 42, 43
CR – control region . . . 36–39, 42
DCS – detector control system . . . 8, 9, 18, 19, 23–26, 43, 44
DSS – detector safety system . . . 8, 18, 19, 26, 43
ECAL – electromagnetic calorimeter . . . 17, 21, 31
EW – electroweak . . . 7, 12
FSM – Finite State Machine . . . 26
ggF – gluon-gluon fusion . . . 13
HCAL – hadronic calorimeter . . . 17, 21, 31
HH – Higgs boson pair . . . 7–9, 12–15, 27, 33, 35, 37–44
HL-LHC – High-Luminosity LHC . . . 7, 9, 16, 19, 41–44
ID – identifier . . . 21, 31, 32, 37, 44
IP – impact parameter . . . 31, 32
LHC – Large Hadron Collider . . . 7, 8, 16, 17, 23, 43
LO – leading order . . . 21, 28
MC – Monte Carlo . . . 21, 28, 29, 34, 36, 40, 41
MET – missing transverse energy . . . 20, 21, 32
MTD – Minimum ionising particle Timing Detector . . . 7–9, 19, 23–26, 43, 44
MTRS – massive temperature readout system . . . 24, 25
NLO – next-to-leading order . . . 21, 28
PDF – parton distribution function . . . 13, 22, 30
PF – particle flow . . . 20, 22, 31
PLC – programmable logical controller . . . 24, 25
POG – physics object group . . . 32
pp – proton–proton . . . 7, 8, 16, 43
PU – pileup . . . 16, 19, 22, 30–32
SM – Standard Model of particle physics . . . 7, 9, 12–15, 43
SR – signal region . . . 30, 33–40, 42
TIF – tracker integration facility . . . 23–25
WP – working point . . . 30–32

1. INTRODUCTION

Particle physics studies the sub-atomic world of fundamental particles and their interactions. It questions what the Universe is made of and tries to answer the unanswered questions of Nature. Currently, the Standard Model of particle physics (SM) is the most complete theoretical framework that describes elementary particles and their interactions. It has been extensively tested across decades by various experiments in laboratories around the globe and has provided astonishing agreement between the theory and experiment.

The discovery of the Higgs boson in 2012 by the ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) collaborations using data from two Large Hadron Collider (LHC) flagship experiments completed the particle content of the SM [1], [2]. Moreover, this observation confirmed the existence of the field responsible for the electroweak (EW) symmetry breaking and the generation of particle masses. Since the discovery, a great effort has been made studying the properties of the Higgs boson, including its couplings to other particles and to itself. This allows to probe the fundamental structure of the Higgs potential and search for possible deviations from the SM.

Single-Higgs production processes are well measured and understood; moreover, they are consistent with the SM predictions. The Higgs boson pair (HH) production, on the other hand, is yet in the search phase due to its rarity. The HH production provides experimental sensitivity to the Higgs self-coupling, thus shedding light on the shape of the Higgs potential, the stability of the vacuum, and whether the Higgs field alone can explain EW symmetry breaking or additional fields are needed. As the Higgs self-coupling strength is fixed by the SM, any deviations would be indications of new physics.

The measurement of the HH production is equally as challenging as it is exciting. It has roughly three orders of magnitude smaller production cross section at LHC energies than the single-Higgs production. Thus, extracting a possible HH signal signature from the abundant background processes requires advanced event triggering, physics object reconstruction, event selection, and categorisation techniques that have to exploit the full capabilities of the CMS detector. Moreover, robust and reliable analysis methods need to be developed in preparation for the future data-taking periods.

In the near future, the LHC will undergo a major upgrade to increase the number of particle collisions, thus entering into its Phase-II, the High-Luminosity LHC (HL-LHC) [3]. The increase in collisions and the collected data will allow for more detailed studies of the fundamental particles and rare phenomena, such as the HH production. To fully exploit the physics potential of the HL-LHC, the CMS detector is undergoing major upgrades to sustain higher data rates. A completely new subdetector is being built – the Minimum ionising particle Timing Detector (MTD) [4]. It will provide CMS with crucial timing information, allowing to unravel the complex aftermath of particle collisions.

This Thesis is dedicated to the study of HH production in its $HH \to WW^*ZZ^* \to 4l$ decay channel using proton–proton (pp) collision data collected with the CMS detector.

This channel has never been explicitly targeted, and thus this analysis is unique not only to the LHC, but in general to the whole particle physics community. Here, one Higgs boson decays into a pair of W bosons and the other into a pair of Z bosons, resulting in four leptons in the final state. For both vector boson species, one of them is off-shell in terms of its mass, thus denoted with a star. To tackle this channel, the Thesis focusses on the development of the full event selection and categorisation, background estimation, and signal extraction. It builds the foundation for the future $HH \to WW^*ZZ^*$ analysis and contributes to the ongoing Compact Muon Solenoid (CMS) effort to measure the Higgs boson self-coupling. Beyond this physics analysis, a crucial part of this work involves contributions to the detector control system (DCS) and detector safety system (DSS) for one of MTD's test-stands. The flawless operation of the MTD in the Phase-II era will be crucial for an improved measurement of the HH production. For simplicity, throughout this Thesis natural units of $\varepsilon_0 = \hbar = c = 1$ are used, unless explicitly stated.

1.1. Aim of the Doctoral Thesis

The aim of this Thesis is to study the production of HH in pp collisions at the LHC, using data collected with the CMS detector over a span of five years. It includes designing the analysis framework, performing data analysis and setting an upper limit on the production cross section, via the $HH \to WW^*ZZ^* \to 4l$ decay channel. Additionally, the Thesis develops the DCS and DSS for a MTD test-stand that is to become the DCS prototype.

1.2. Tasks of the Doctoral Thesis

1. Review the theoretical motivation for researching HH production, the novelty of using the $HH \to WW^*ZZ^*$ multilepton decay channel.

2. Describe the experimental environment of the LHC and CMS, emphasizing the subsystems and physics object reconstruction relevant to the $HH \to WW^*ZZ^*$ multilepton study.

3. Design and validate the event selection, categorisation, and signal extraction for the $HH \to WW^*ZZ^* \to 4l$ analysis, perform the data analysis and measure the upper limit of the HH cross section.

4. Review the necessary requirements for the MTD test-stand's DCS and contribute to its development.

1.3. Thesis Statements to be Defended

1. The analysis of HH production provides experimental sensitivity to the structure of the Higgs potential and enables to set constraints on the self-coupling parameter within the current experimental reach.

2. The results obtained from the $HH \to WW^*ZZ^* \to 4l$ channel analysis are compatible with the SM predictions within the current statistical and systematic uncertainties.

3. The developed analysis framework is scaleable and readily extendable to larger data sets, providing a solid foundation for future HH production studies at the HL-LHC.

4. A DCS for the MTD test-stand is designed, implemented, and commissioned. The system meets the operational requirements and serves as a functional prototype for the final MTD DCS.

1.4. Scientific Novelty

1. This Thesis presents the first ever measurement of the HH production cross section in the $HH \to WW^*ZZ^* \to 4l$ decay channel using the CMS data.

2. The created design, implementation and optimisation of the analysis strategy for the $HH \to WW^*ZZ^* \to 4l$ channel has allowed to reach sensitivity of the HH production cross section below 200 times the SM value.

3. This Thesis provides insight in the novel DCS for the MTD test-stand.

1.5. Practical Significance

1. This Thesis represents the first feasibility and sensitivity study of the $HH \to WW^*ZZ^*$ four-lepton decay channel at CMS, establishing the methodological framework required for future analyses, and possible combinations.

2. The developed tools and methods for event categorisation, object identification, and calibration and methods for background suppression, signal extraction, systematic uncertainty treatment are directly applicable to broader searches for multiboson, multilepton and exotic signatures.

3. This Thesis encapsulates the needed steps for building a DCS that ensures safe and effective detector operation. The built DCS has been vital for a successful development, assembly and commissioning of the MTD.

1.6. Approbation of PhD Thesis in Scopus and Web of Science Indexed Articles

During the doctoral study period, the author of the Thesis co-authored 262 publications as a member of the CMS collaboration. Of these, 46 are publications of the CMS Higgs Physics Analysis Group where the author is an active member. Statistics retrieved on 26.01.2026 from https://inspirehep.net/authors/2643592 (Scopus ID of the author: 57214117093, ORCID: https://orcid.org/0000-0003-1350-3523).

Additionally, these six documents can be highlighted, two of which are internal to the CMS collaboration. They are internally reviewed and approved by experts, conveners, and the whole CMS collaboration because they contain valuable, technical information, critical for knowledge transfer and the eventual inclusion in further CMS publications.

[1] *CMS internal note (parts in Chapter 4 in the Thesis)*
A. Gaile, N. Strautnieks, A. Petkovic, A. Sculac, T. Sculac, E. Pajuste, K. Dreimanis. "Low-p_T electron selection efficiency studies". Technical report, CERN, Geneva, 2022, CMS AN-21/165. This note resulted in: CMS Collaboration. "Low-p_T electron Electron ID scale factors from CMS in proton-proton collisions at $\sqrt{s} = 13$ TeV using J/ψ events". CMS Detector Performance Summary, CERN, Geneva, 2023, CMS-DP-2023-081.

[2] *Open access paper (parts in Chapter 3 in the Thesis)*
CMS MIP Timing Detector group, including **A. Gaile**. "Optimization of LYSO crystals and SiPM parameters for the CMS MIP timing detector", JINST 19, 12, P12020, 2024,
DOI:10.1088/1748-0221/19/12/P12020.

[3] *CMS internal note (Chapter 6 in the Thesis)*
A. Gaile, A. Sculac, R. Jiang, C. Charlot, T. Sculac, K. Dreimanis. "Search for non-resonant Higgs boson pair production in WWZZ decay mode with full Run 2 and early Run 3 dataset". Technical report, CERN, Geneva, 2024, CMS AN-24/245.

[4] *Proceedings in the conference LHCP2025 (parts in Chapter 2 in the Thesis)*
A. Gaile. "Non-resonant HH and Higgs self-coupling measurements by CMS". At the Thirteenth Annual Large Hadron Collider Physics conference, Taiwan, 2025, DOI:10.22323/1.499.0105.

[5] *Open access paper (parts in Chapters 3 and 5 in the Thesis)*
CMS BTL group, including **A. Gaile**. "The CMS barrel timing layer: test beam confirmation of module timing performance", Nuclear Instruments and Methods Section A, 1081, 170823, 2026, DOI:10.1016/j.nima.2025.170823.

1.7. Dissemination in International Scientific Conferences

[1] A. Gaile. "Non-resonant HH and higgs-self couplings measurements by CMS". Presentation at the The Thirteenth Annual Large Hadron Collider Physics (LHCP2025) conference, May 2025,
URL: https://indico.cern.ch/event/1419878/contributions/6442906/.

[2] A. Gaile. "Study of Non-Resonant HH Production in WWZZ Decay Mode at CMS Experiment". Presentation at the 4th CERN Baltic Conference, October 2024, URL: https://indico.cern.ch/event/1416853/contributions/6178181/.

[3] A. Gaile. "Detector control and safety system for MIP Timing Detector". Poster at the 3rd CERN Baltic Conference, October 2023,
URL: https://indico.cern.ch/event/1288731/contributions/5600676/.

[4] A. Gaile. "Run 3 electrons, new studies/ recipes – J/ψ studies". Presentation at the HZZ Workshop, June 2023,
URL: https://indico.cern.ch/event/1233959/contributions/5310341/.

[5] A. Gaile. "Development of the MTD Control and Safety Systems at CMS". Presentation at the 2nd CERN Baltic Conference, October 2022,
URL: https://indico.cern.ch/event/1147717/contributions/5069252/.

[6] A. Gaile, N. Strautnieks, J. Proskurins. "Improving low-p_T electron reconstruction for Higgs physics at CMS". Presentation at the 1st CERN Baltic Conference, June 2021, URL:
https://indico.cern.ch/event/970609/contributions/4415897/.

1.8. Use of Artificial Intelligence Tools in the Doctoral Thesis

The summary of the Thesis was prepared using the `Tildes Birojs 2022` machine translation tool. The tool was used from desktop software, therefore, only the standard machine translation model is available, without enhancements offered by artificial intelligence or large language models. The summary was translated from English into Latvian and used as the first translation layer for all chapters.

2. MOTIVATION FOR HIGGS BOSON PAIR PRODUCTION STUDIES

On the 4th of July, 2012, the ATLAS and CMS collaborations jointly announced the discovery of a scalar boson with a mass of about 125 GeV [1], [2]. This boson, compatible with the Higgs boson predicted by the Brout-Englert-Higgs mechanism, is an integral part of the SM as it can explain EW symmetry breaking. An ongoing effort is the measurement of Higgs boson self-coupling which can further shed light on the EW symmetry breaking. The most direct way to probe the Higgs self-coupling is to study the experimentally challenging HH production.

The SM, built as a quantum field theory based on gauge symmetries, remains the best theoretical framework to describe the sub-atomical world of fundamental particles and their interactions. It can precisely predict the properties of the particles contained within it, as well as three of the four fundamental forces, as can be seen in a schematic representation in Fig. 2.1. Importantly for this Thesis, the SM starts with a unified definition of the EW interaction.

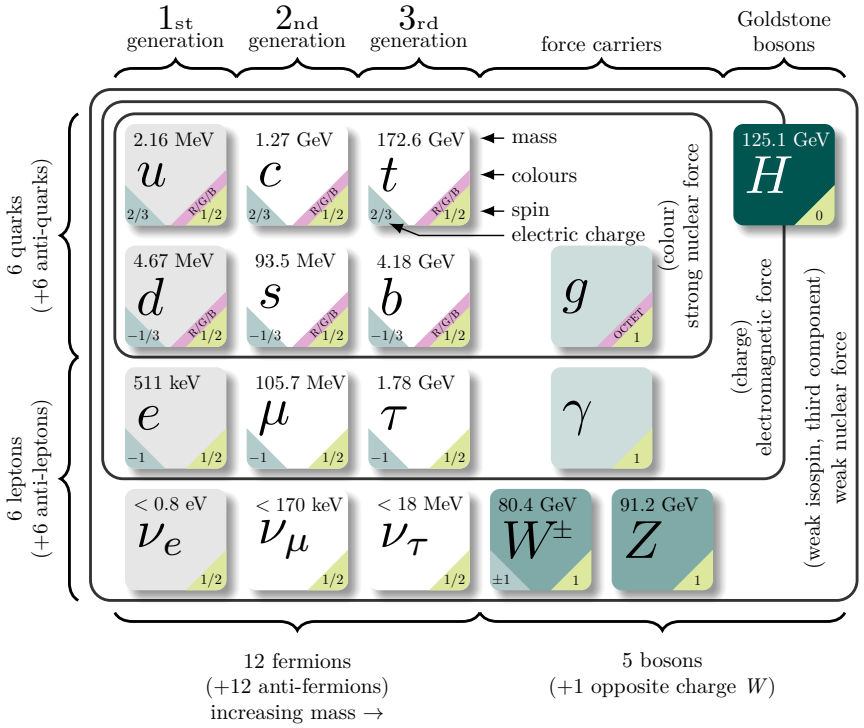


Fig. 2.1. The fundamental particles of the SM, adapted from [5], indicate the mass, colour, spin and charge numbers as well as classification within generations and possible interactions. Corresponding antimatter particles are not shown.

As detailed in the full-length Thesis, there are no explicit mass terms for bosons in the SM Lagrangian, as mass terms break the gauge invariance. To solve this, the Brout-Englert-Higgs mechanism was proposed in 1964 by several independent groups [6], [7], [8], [9]. The mechanism introduces a complex scalar quantum field, the Higgs field, spontaneously acquiring its minimum value other than zero, as a consequence separating electromagnetic and weak interactions and leading to weak gauge bosons gaining mass.

The shape of the Higgs potential is determined by the tri-linear and quadric-linear coupling constants of Higgs boson self-coupling. Thus, studying interactions that involve vertices with three or four Higgs bosons, e.g. HH or Higgs boson triplet production, provides information of the potential's shape. In the SM, the tri-linear Higgs boson self-coupling, measurable in HH production, λ, is a ratio of two constants – the mass of the Higgs boson, m_H, and the vacuum expectation value, v, as shown in Eq. (2.1). Any deviation from unity could indicate new physics.

$$\lambda_{SM} = \frac{m_H^2}{2v^2} = 0.129 \tag{2.1}$$

2.1. Higgs Pair Production Mechanisms

The dominant production mode for HHs is through gluon-gluon fusion (ggF), and it has two contributing processes, as shown schematically in Fig. 2.2. In one of the processes, two Higgs bosons are produced from a quark loop ("box" diagram, see Fig. 2.2 b), while in the other, a virtual Higgs boson is produced that decays into two Higgs bosons ("triangle" diagram, see Fig. 2.2 a).

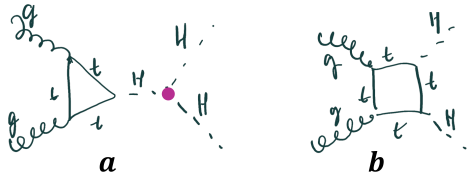


Fig. 2.2. Feynman diagrams of the leading HH production mode: a – ggF via the triangle loop, b – ggF via the box loop; Higgs self-coupling, λ, is marked with a pink dot.

Both diagrams interact destructively, leading to a cross section around three orders of magnitude lower than the cross section for single-Higgs production. At $\sqrt{s} = 13.6$ TeV is estimated as:

$$\sigma_{\mathrm{HH}}^{\mathrm{ggF}} = 34.13^{+6\ \%}_{-23\ \%}(m_t + \mathrm{scale}) \pm 2.3\ \%(\alpha_s + \mathrm{PDF})\ \mathrm{fb}.$$

Further description of the calculations and the considered theoretical uncertainties (the strong coupling constant, α_s, parton distribution function (PDF), top quark mass, m_t, etc.) are found in [10].

2.2. HH Decay Channels and Current Experimental Searches

The small production cross section of HHs poses a significant experimental challenge, while the multitude of possible decay channels further increases the complexity of the search. The branching ratios (BRs) for some of the most common decays are shown in Fig. 2.3. By nature, there is no single "golden channel" with a large BR and a high signal to background ratio. By exploring more decay channels, the sensitivity for HH production

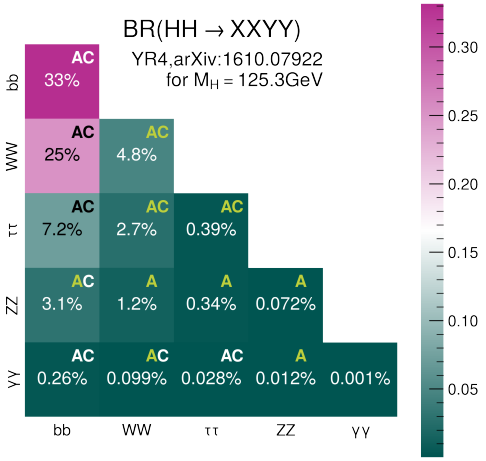


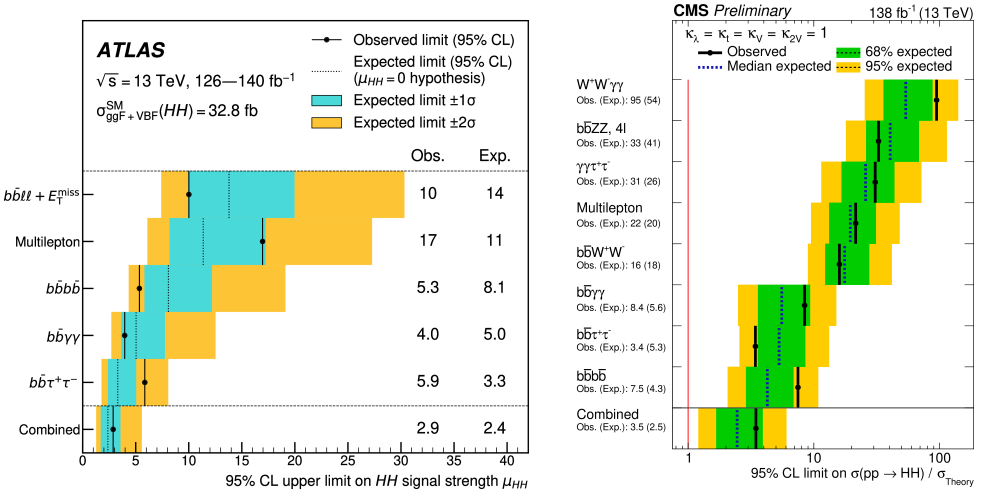
Fig. 2.3. BRs for the most popular HH search channels. Channels marked with "A" correspond to those analysed by the ATLAS collaboration, while "C" marks the ones analysed by the CMS collaboration. Green letter indicates inclusion in multilepton analysis.

measurement can be increased. While it is insufficient, each analysis sets an upper limit on the HH production cross section, or more often the HH production signal strength, μ, as defined in Eq. (2.2), where σ is the measured HH production cross section and $\sigma_{\text{SM (Theory)}}$ the cross section value predicted by the SM (Theory). In addition, several searches can provide restrictions on various other Higgs couplings or new physics scenarios.

$$\mu = \mu_{HH} = \frac{\sigma}{\sigma_{\text{SM}}} = \frac{\sigma}{\sigma_{\text{Theory}}} \tag{2.2}$$

The best multi-channel examples are the HH multilepton analysis where events are selected and categorised based on the final states, more precisely the multiplicity of leptons and hadronically decaying taus, τ_h. Such analysis has been performed by both collaborations. The analysis performed by the CMS collaboration explicitly targets only three decay channels (see Fig. 2.3), categorising them into seven search categories where lepton multiplicity ranges from zero (targeting four τ_h) to four [11]. Overall, this analysis has put a 95 % confidence level (CL) observed (expected) upper limit of 21 (19) on the signal strength of the HH production.

The analysis performed by the ATLAS collaboration includes events from ten decay modes, categorised in nine search categories, with lepton multiplicity ranging from two to four [12]. Overall, this combination has put a 95 % CL observed (expected) upper limit of 17 (11) on the signal strength of the HH production. ATLAS's multilepton search puts a slightly tighter limit while combining more search categories, which CMS has covered in other analyses.



(a) By the ATLAS Collaboration [13] (b) By the CMS Collaboration [14]

Fig. 2.4. Observed and expected 95 % CL upper limits on the signal strength for inclusive HH production.

Up to date, the most complete combination performed by the ATLAS collaboration [13] has set an observed 95 % CL upper limit of 2.9 times the SM value, while 2.4 was expected. The most complete HH production combination done by the CMS collaboration [14] has set an upper observed limit of 95 % CL of 3.5 times the SM value while the expected limit is 2.5. For both combinations, the observed and expected upper limits per analysis channel are shown in Fig. 2.4. Additionally, a combination has been performed also using both collaboration data [15].

To fill in an uncharted territory of HH searches, this Thesis and the accompanying CMS physics analysis summary presents a novel $HH \to WW^*ZZ^* \to 4l$ search, an integral part of the $HH \to WW^*ZZ^*$ search. Later, it could be integrated in the whole CMS HH multilepton analysis. The decay channel of $HH \to WW^*ZZ^*$ is included in the ATLAS multilepton's analysis, however, not its final state of four leptons [12], while CMS's multilepton [11] four-lepton search category targets the $HH \to WW^*WW^*$ decay. Thus, motivating the $HH \to WW^*ZZ^* \to 4l$ search.

3. EXPERIMENTAL ENVIRONMENT

Each physics theory must be complemented by experimental data that validate or reject it. Fundamental theories describing elementary particles and their interactions have extremely high requirements for the experimental environment to test them. Therefore, to push the boundaries and advance this field, scientists need to collaborate, working together for decades to construct the required experimental apparatus. European Organization for Nuclear Research (CERN) (from the French *Conseil Européen pour la Recherche Nucléaire*) is a prime example of the necessary experimental environment, providing an accelerator complex, ground for experimental apparatus, and a fruitful community for growing science. All data analysed in this Thesis have been recorded at CERN, with the CMS detector, using LHC's pp collisions.

3.1. The Large Hadron Collider

Currently, CERN's flagship accelerator is the LHC, which at the time of writing of this Thesis is the most powerful particle collider on Earth. However, the whole accelerator complex at CERN is powerful and versatile, and is not only used to fill the LHC but also to provide protons or other ions to a multitude of experiments.

The concept of having a pp collider in the Large Electron-Positron Collider tunnel was already discussed in 1984, while operations began in the 2010's. Several data-taking Runs spanning multiple years have been conducted and are foreseen to continue until the 2040s. It is a particle-particle collider with two counter-rotating beams in two separate rings housed in a shared twin bore magnet and mostly collides pp beams. The LHC was designed to create pp collisions of centre-of-mass, $\sqrt{s}$, energies up to 14 TeV. Amongst other important characteristics for the machine is the luminosity that characterises the rate of event production, as well as the pileup (PU) that characterises simultaneous interactions. Both parameters have been increasing and have exceeded the initial design value. For more details, see the full-length Thesis.

During the operational span of the machine, it has and will continue to evolve. In the late 2020's, an extensive upgrade is foreseen for the LHC, the accelerator complex, and for the detectors [3]. This major upgrade will mark the transition from LHC Phase-I to the Phase-II or the HL-LHC era. The upgrade will allow for Higgs boson property measurements with unprecedented precision and increased potential for the search of new physics, as the total number of the collisions are expected to increase by a factor of five with respect to the Phase-I LHC. The preliminary design study expects to record an unprecedented 3000 fb^{-1} in about a decade.

3.2. CMS Experiment

The CMS is one of LHC's multi-purpose detectors, and its name highlights its main design features [16]. It is *compact* due to its dense design – although the CMS detector is significantly smaller in volume compared to the ATLAS detector, it is almost twice as heavy. It has an extensive *muon* system comprising around 80 % of its total volume. And finally, at the heart of CMS sits a ≈4 T superconducting *solenoid* magnet [17], providing bending of charged particles. The detector has a cylindrical shape and azimuthal symmetry with a diameter of 15 m and length along the beam axis of around 21 m, completely surrounding the interaction point. A cutaway diagram is shown in Fig. 3.1 illustrating the subdetectors, the beamline and a person for scale. In the diagram, the central part (barrel) and the two forward regions (endcaps) are clearly distinguishable.

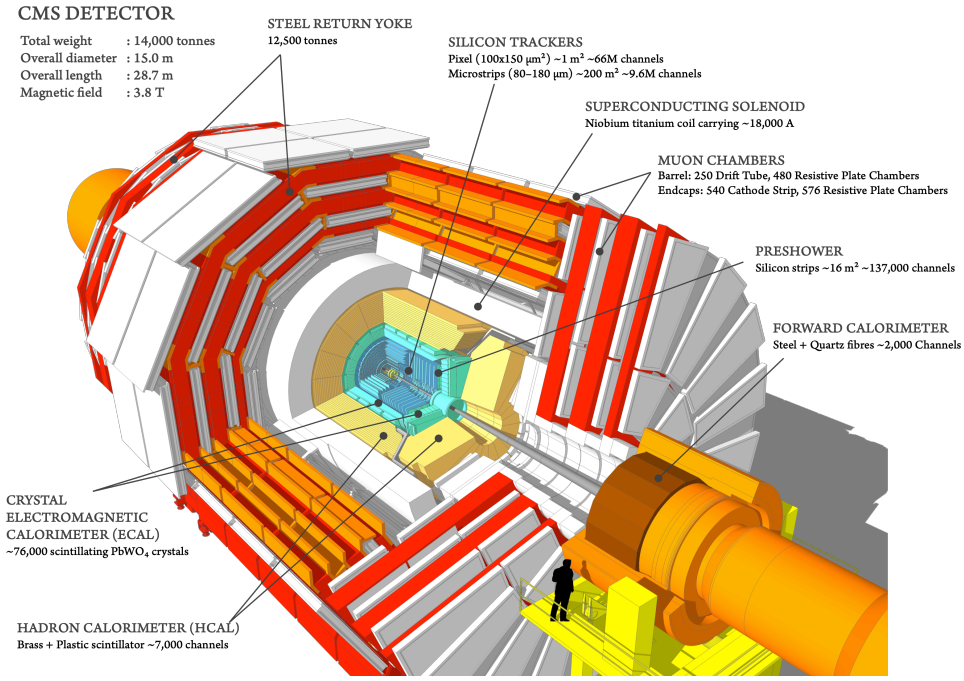


Fig. 3.1. Cutaway diagram of CMS detector [18].

The CMS consists of multiple concentric subdetectors with high-granularity and good time resolution, each specialised to detect and measure various particles and their properties, which together can reconstruct what has happened in the collision. The silicon tracker [19], denoted in cyan, traces the paths and momenta of charged particles, as well as measures the position of the collision vertices. Between the tracker and the solenoid is the electromagnetic calorimeter (ECAL) [20] and hadronic calorimeter (HCAL) [21], which measure the energy deposited by electrons, photons, and hadrons. Both calorime-

ters are shown in yellow in Fig. 3.1. Finally, outside the solenoid is the muon system [22] which tracks muon paths and measures their momenta. System-wise, the CMS would be powerless without a dedicated trigger, and the data acquisition system [23]. For more Thesis-specific details, see the full-length Thesis.

Having the right detecting material alone is not sufficient to gather high quality physics data. We need to be able to control the detector in an optimal, reliable, and safe manner, starting from "turning it on/off" to optimising its performance. This is the main goal of the DCS, while the DSS is responsible for protecting the detector and its equipment from various hazards. In the CMS experiment, these two systems are separated while they often share data from the same probes. The scope of the DCS reaches beyond just the control of CMS; it also communicates with external systems, such as the accelerator, and provides a graphical user interface. A well-designed DCS ensures optimal functioning of the system.

3.2.1. The Coordinate System

A right-handed coordinate system is used in the CMS to characterise both its layout and the reconstructed particles. Due to CMS's shape, spherical coordinates are used. A schematic illustration of the standard coordinate system at the CMS detector is shown in Fig. 3.2. To combat the unknown initial momentum of the proton constituents, mo-

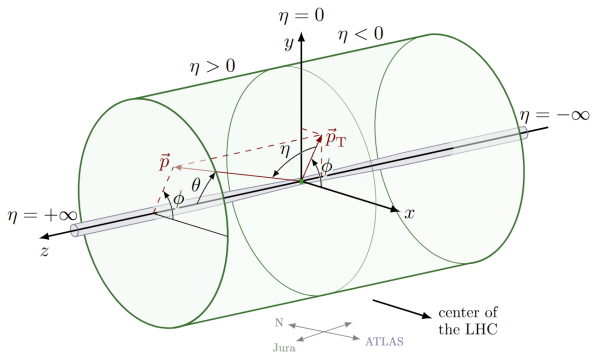


Fig. 3.2. Schematic illustration of the standard coordinate system at the CMS detector, adapted from [24].

mentum and energy transverse to the beam direction (in the xy plane), respectively, p_T and E_T, are used to characterise the final state particles as the initial momentum in the transverse plane is zero. As accustomed at colliders, the angle from the transverse plane of a particle is defined as pseudorapidity, η. Lastly, it is convenient to describe a Lorentz invariant quantity, the angular distance, ΔR, in terms of the distances in $\eta - \phi$ with Eq. (3.1).

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.1}$$

3.2.2. Phase-II Upgrade and the MIP Timing Detector

Just as the accelerator machine will enter its Phase-II, all detectors will also undergo severe and ambitious upgrades to get the most out of the HL-LHC and the opportunities it will bring. To tackle the challenges of the increased PU of 200, the CMS detector will have a higher geometric coverage [25], [26], higher resolution and granularity [27], [28]. Moreover, it will use precision timing and support higher data rates to analyse its data. Existing detectors will evolve, new ones will be built (are being built), and several improvements in the read-out systems, such as the Level-1 trigger [29], are also expected.

Outside the tracker, a completely new detector will provide a totally new capability for the CMS – precision timing for charged particles [4]. The MTD is expected to provide timing resolution for single charged tracks of 30 ps to 40 ps resolution at the start of operation, degrading due to radiation damage to 60 ps or 70 ps by its end-of-life. The MTD will help with particle identification, both indirectly by increasing the efficiency of PU track removal, helping with isolated lepton identification, and directly because the layers will allow for time-of-flight measurements thus allowing for improved charged hadron identification. The MTD will consist of two systems, shown in Fig. 3.3 – the barrel

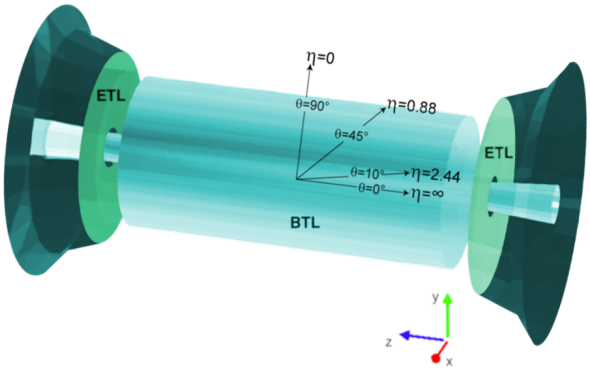


Fig. 3.3. A schematic view of MTD's geometry, showing its barrel and endcaps part [4].

timing layer (BTL), a single-layer cylindrical detector covering $|\eta| < 1.48$, and the endcap timing layer, a two-disc system covering $1.6 < |\eta| < 3.0$. To mitigate the dark current and increase the signal over background ratio, the detector will be cooled to around $-30\,^{\circ}\mathrm{C}$ using a two-phase CO_2 flow. Due to the geometry, both systems are experiencing different levels of irradiation and have converged on using different sensors and other solutions to deal with possible complications. The endcap layer is expected to be serviced in-between running periods, while the BTL is likely to lack such a luxury. Therefore, the design and operation solutions have to be thoroughly thought out. An effective DCS and DSS will be absolutely crucial for effective operations.

4. PHYSICS OBJECT RECONSTRUCTION AND IDENTIFICATION

As particles have traversed the CMS detector, leaving their signatures in different layers, and the trigger system has selected this event as potentially interesting, the whole detector information can be used to interpret these signatures as "physics objects". Low-level objects such as tracks and energy clusters are correlated to form high-level objects such as electrons or muons, which in turn can be grouped to build even more complex objects, such as jets. This chapter will focus on electrons, muons, jets and missing transverse energy (MET) (carried away by neutrinos) as they are important for the $HH \to WW^*ZZ^*$ analysis; however, with data collected by the CMS detector it is possible to reconstruct also other physics objects, e.g. photons.

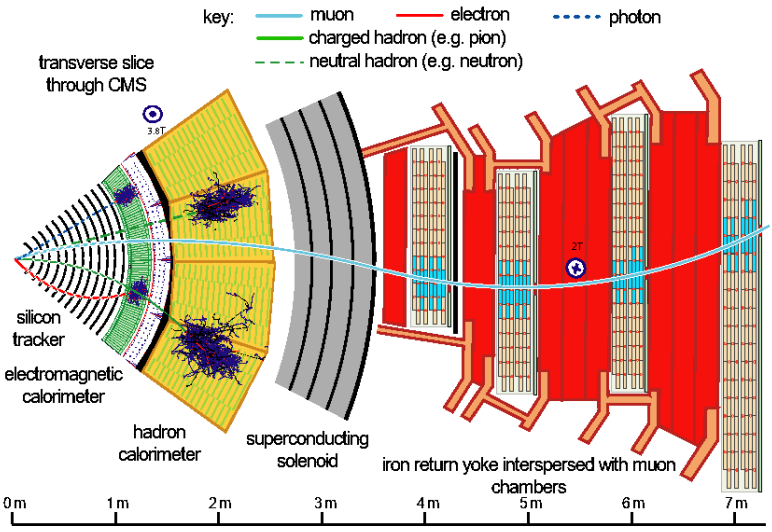


Fig. 4.1. Several particle interactions and typical signatures in a transverse slice of the CMS detector [30].

In Fig. 4.1, a slice of the transverse cross section of the CMS detector is shown, highlighting various physics objects and their typical signatures. The CMS uses the particle flow (PF) approach [30] that correlates the signal of all subdetectors to create an improved reconstruction of events. First, it reconstructs charged particle tracks, sequentially masking the used signal to avoid double counting. Second, it clusters energy deposits in the calorimeters. Third, a linking algorithm tests any of these combinations for a possible physics object candidate. The algorithm starts with reconstructing muons, electrons, photons, then hadrons, and finally accounts for any MET. This holistic approach has enabled CMS to deliver such great results.

Muons traverse the entire detector and often escape it altogether; however, as charged

particles they leave signal in the traversed silicon tracker layers and the muon chambers. These signals or hits can be grouped to form tracks that mark their trajectory and are used to extrapolate to their origin or vertex. Muon candidates can be built starting from the tracker hits or the muon system's hits [31]. Overall, muon identification and reconstruction are higher than 99 %, and dedicated identifiers (IDs) can increase this efficiency even more.

Electrons leave hits in the tracker and are absorbed in the ECAL, thus the main information for their reconstruction is provided by these two systems. Electron reconstruction can be ECAL-based or tracker-based [32], associating tracks or energy deposits of the other system. Attention is paid to account for the energy lost through bremsstrahlung and photon conversion [33].

Both charged and neutral hadrons lose part of their energy in the ECAL and are absorbed in the HCAL. In hadron colliders, it is common to talk about jets – hadronised quarks and gluons, manifesting in cone-shaped objects, inclusively measured by the calorimeters [34]. Several methods exist to reconstruct jets [30], [35]. The tagging of jets originating from, e.g., a b quark needs information from the tracker, as hadrons containing b quarks tend to have a longer lifetime, thus producing a more displaced secondary vertex. Finally, neutrinos escape the detector undetected. However, their energy can be estimated by the MET, which is obtained by vectorially summing the transverse momenta of the reconstructed objects.

4.1. Event Simulation with Monte Carlo Generators

Monte Carlo (MC) simulations are the backbone for event selection and optimisation, as well as the development of various methods. These simulations reproduce the whole chain of particle collision, starting from parton level interactions until a complete detector response. Thus, allowing for a quantitative comparison between theory with controlled assumptions and the beautifully random and real experiment.

Figure 4.2 summarises the main steps in data production, showing both the paths taken by the experimental data and the MC. The generation of MC events begins with the simulation of the hard scattering process of the incoming partons (quarks or gluons) from each proton that produces outgoing partons. The probability of the given interaction is computed from matrix elements and can be done with various precision. The leading order (LO) calculates basic tree-level interactions, the next-to-leading order (NLO) includes one-loop correction, and each "Next-to" adds another level of correction. Generally, higher order corrections improve the agreement between theory and data; however, they are more computationally heavy.

A multitude of dedicated generators are available and customisable for various calculations. Examples are `MadGraph5 amc@NLO` [37], `Pythia` [38], `Powheg box` [39]. Generators

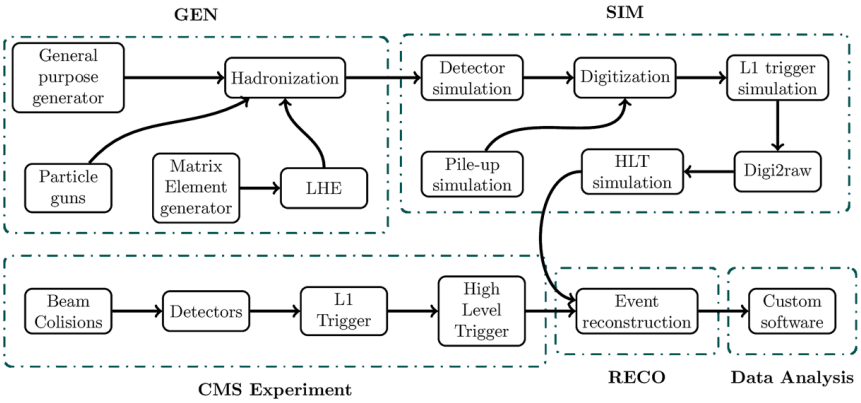


Fig. 4.2. The main steps in the production chain of simulated data [36], following event generation (GEN), simulation (SIM), and reconstruction (RECO).

need specific input data. As the momenta of quarks and gluons inside the proton are not constant over time, they can be described by PDFs, which characterises the probability of finding a parton with a certain momentum at a given scale.

After the hard scattering, the outgoing coloured partons undergo parton showering, including the initial and final state radiation. This is a crucial step often performed by `Pythia` or `Herwig++` [40]. Generators have various hadronisation models which can either use string or cluster modelling. At this level, the particles are called "generator-level" or even more poetically as "truth-level" particles, as afterwards they meet the "reality". Alongside the primary hard interaction, additional or overlapping events can take place, such as simultaneous soft scatterings or simply beam remnants. As PU contaminates object reconstruction, understanding its modelling is crucial.

The next processing step involves understanding how such an event would be seen in the detector. This is done by `Geant4` [41] passing the event through a model of the CMS detector. `Geant4` simulates the passage of particles through matter, modelling energy loss through ionisation, bremsstrahlung, and showering. It also tracks the particle as it traverses various layers through the magnetic field of the CMS, and can simulate detector-specific properties, such as the resolution and response time, and shape for various sensors. After the modelling, the response is digitalised by introducing electronic noise and other detector response limitations to mimic the raw detector output. Finally, the "raw detector response" is fed through the standard CMS reconstruction chain just as the real data. This means emulating triggers, reconstructing tracks, clustering calorimeter energy deposits, employing PF and identifying objects. The beauty of simulated data is that the generator-level or the truth-level information will always be available, thus allowing to validate or simply to compare object reconstruction, event categorisation, as well as to dig into the origin of the particle and much more.

5. BUILDING A DETECTOR CONTROL SYSTEM FOR THE MTD TEST-STAND

Not only each detector, but also each test system needs a control system. In this chapter the extensive work carried out as the main priority during the first half of the Thesis studies is summarised. It centres around the prototype control system which was built to accommodate one of the MTD test-stands. The goal of the project was not only to facilitate data-taking vital for a successful development, assembly and commissioning of the BTL, but also to create a DCS prototype. For more details on the basic concepts and tools used for building a DCS at the LHC experiments, see the full-length Thesis.

5.1. MTD Test-Stand and its Prototype DCS

Testing and validating is a natural process for any detector system. Each test-stand can investigate individual components or the interlink amongst many. The MTD test-stand in question is a test-stand for a single BTL tray and is located in the tracker integration facility (TIF) at CERN. It was initially created for cooling tests, but has evolved over time to hold a wide variety of tests. A "tray" is the largest subcomponent of the barrel part of the MTD, as 72 trays create the whole BTL.

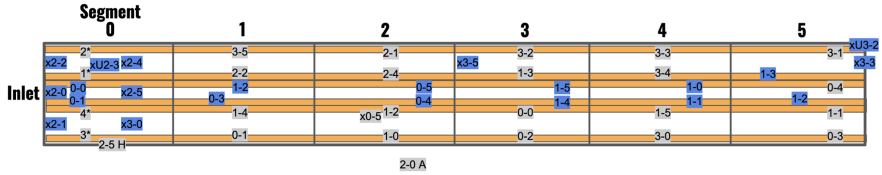


Fig. 5.1. A schematic representation of the BTL tray cooling test-stand, indicating the segment numbering, gas inlet, and noting microstrip heaters in yellow. Gray and blue boxes represent a possible layout of various sensors.

This test-stand is a full-size prototype, and a schematic representation of it is shown in Fig. 5.1. It is approximately 40 cm wide, approximately 250 cm long and around 40 mm thick, logically it is sectioned into six segments. It is made of an aluminium plate with etched pathways for steel pipes installed for the two-phased CO_2 cooling. It has an insulating cover ensuring an enclosed volume that can be flushed with dry air. In the first iteration of the test-stand, it consisted of only the tray with its cooling network. In the second step, load simulating resistive strip heaters were placed on each segment (six per tray, as can be seen in Fig. 5.1, each corresponding to one readout unit). In the third iteration, the simulating heaters were removed from one segment and replaced with real sensor modules. More precisely, the replacement mimicked one half of the readout unit, consisting of a mock-up concentrator card and a front-end board with pre-series SiPM

(silicon photomultipliers) and LYSO crystals (a Cerium doped Lutetium-based scintillation crystal). All heaters and modules (and their thermoelectric coolers) were powered by laboratory power supplies, allowing for easy testing. A relay system was attached to cut the power to these supplies if the situation required it. The main sensing within the cold volume is done with temperature, humidity, and pressure sensors. The majority of temperature sensors are connected via the massive temperature readout system (MTRS), while other sensors are connected through a programmable logical controller (PLC).

Several functional, performance, and safety requirements were set for this DCS. Regarding the functional requirements, the very first and most important ones imposed by the team performing the tests were the monitoring and data logging. In addition, an alarm system was needed to inform if the connection with any of the hardware devices was lost for a prolonged period, exceeding the maximum latency. A functionality put in a "wish-list" is the control of the power supplies and the modules. However, this has not been possible in the current test-stand scenarios and is reserved for future evolutions as described in Section 5.2. Regarding performance requirements, the DCS must be running 24/7 and be able to self-recover after unlikely situations of lost power. The safety criteria set for the system are rather loose and require an operator for the safe supervision of the states. No hardware or software redundancy is created for the DCS. As control functions are limited, safety measures are not fail-safe and mainly are in the form of alarms as mentioned above. The integration of a relay system allowed to cut the power to the power supplies, however, this is done in a harsh way and does not allow for soft shut down. However, for the initial tests, this was sufficient.

The MTD TIF DCS project was built following the CMS production project standards, with its initial version being available in May 2022. Its graphical interface was made available for all `MTD_Operator`s via the `CMS Terminal server`. The project uses a set of JCOP framework components [42], as well as many from the CMS framework. Finally, within this Thesis, a set of MTD specific components was created, as summarised in Table 5.1. The building of this prototype project has been a team effort, initially led and always supervised by the field's experts. For the author of this Thesis, it was a great learning experience, and her contribution to the project evolved over time. It started with creating the graphical user interface panels for various components wherever it was needed. It encapsulated both the graphical design, designing and optimising the functionality, and reusing or creating new scripts and libraries to support the said functionalities. The only component for which she did not create the graphical interface was the `CMS_MTD_ILK`, as it is the latest addition to the set of components, already delivered by the following expert. For the design of this component, the author of this Thesis took on a supervisory role, as it was the onboarding project of the following MTD DCS expert.

With evolving knowledge of how a functional component has to be designed and packaged, the author of this Thesis followed the full creation cycle to create the `CMS_MTD_DIP` and `CMS_MTD_RDB_SMOOTHING` components. The first component accesses the CO_2 plant's

Table 5.1

Functional components created for the MTD TIF DCS project

Component	Description
`CMS_MTD_TIF_BASE`	Ensures connexion to the configuration and conditions databases
`CMS_MTD_GENERAL`	Hosts supervisory panel as well as the heartbeat panels for hardware, indicating the connexion and latency to the PLC and MTRS
`CMS_MTD_AC`	Handles the access control, people with role of `MTD_Operator` have access rights to the project, while `MTD_Expert` have rights to alter the project
`CMS_MTD_PLC`	Ensures the MTD specific PLC configurations to be installed
`CMS_MTD_MTRS`	Ensures the MTD specific MTRS configurations
`CMS_MTD_DIP`	Accesses published information of the CO_2 plant, such as the flow-rate, temperature
`CMS_MTD_RDB_SMOOTHING`	Allows to record data without any smoothing applied for intensive testing over short periods
`CMS_MTD_ILK`	Provides a software level interlock, allowing to use MTRS probes for opening the relays

published data, the environment data via the data interchange protocol [43], that extends the distributed information management, a protocol created at CERN [44]. The creation of this package allowed to deepen the understanding of creating a new data point type, data point creation and loading to the configuration database, appending a manager, and of course, creating a functional graphical user interface. All performed by the author of this Thesis with quality control provided by the CMS Central DCS team. The second component allows to alter the smoothing of the data archiving for various probes. The work on this component provided a great learning opportunity for creating dedicated scripts that use data from various data points, designing a new manager, and new data point types.

Other maintenance and upkeep tasks were also performed throughout the project development carried out by the author of this Thesis. These included the updates of graphical interfaces, the inclusion of new data points and their types, as well as the update of the project from `WinCC-OA 3.16` to `WinCC-OA 3.19`, ensuring compatibility for all components. This work allowed to practise working with various centralised tools, and practice the DCS building philosophy adapted within the CMS collaboration.

5.2. Prospects of the Project

The initial DCS prototype met the necessary goals set by the team working on the test-stand. However, it is foreseen that the project will evolve as more functionality becomes

available. The current DCS project will be used to further develop three important aspects for controlling the MTD. They are expanded below, indicating the introduced functionalities. Moreover, it could evolve into the first full DCS prototype towards the final detector. And even if this DCS prototype never evolves into the final DCS, it has served as a good testing ground for many features, design choices and functionalities that could benefit any further system.

Firstly, the controlling sequence needs to be designed and tested. This scenario becomes available upon the arrival of dedicated power supplies. From the test-stand point of view, this will provide the opportunity to test the component powering as intended in the final detector. Various tests will become available – finding the most optimal powering sequence, designing the ramping up and down of the voltage, evaluating the necessary power in various conditions, and estimating any losses, as well as optimising the placement of the power cables. DCS-wise, it will be possible to work with the designated components, implementing the power supplies as data points, having the ability to control and monitor them via the DCS. This will allow to build the first prototypes of the Finite State Machine (FSM), design and test out basic functionality with actions, such as "turning on", etc. This will allow to extensively test a set of DSS aspects. This scenario could be investigated further by populating the full tray with sensor modules, thus analysing collective behaviour, defining annealing, standby and data-taking regimes.

Secondly, the communication with the Data acquisition system needs to be tested. This second scenario becomes available upon the collaboration between both team, and upon creating a unified test-stand for investigating the interplay between the DCS and the physics data acquisition. Here, both from the test-stand and the DCS point of view, the main gain will be streamlining the intercommunication between both systems and finding any possible bottlenecks. This will allow to define clear boundaries of each system's responsibility, test and improve the efficiency of their communication.

Finally, to create the first iteration of the DCS for the final detector, the supertray collective behaviour needs to be researched. Thus, the third scenario is expanding this test-stand to a supertray unit. Test-stand-wise, this allows to study and investigate the collective behaviour of a set of six trays, explore the mechanical solutions, communication interfaces and similar. As an example, such a system will provide ample room for tests regarding the uniformity of cooling. From the DCS point of view, this setup will allow to test the scalability of the project, and will provide ample ground for testing various aspects of the FSM, as well as alarm design. As an example, parts of the supertray could be forced to exit the desired running ranges, forcing the system to act.

6. BUILDING BLOCKS OF THE NON-RESONANT HIGGS BOSON PAIR PRODUCTION IN WWZZ DECAY CHANNEL ANALYSIS

The primary focus during the final years of the author's doctoral studies was the design and development of a new CMS analysis focused on the non-resonant HH production in the $HH \to WW^*ZZ^* \to 4l$ decay channel. As there was no previous CMS analysis of this channel, the analysis workflow had to be designed, optimised, tested, and finally used to measure the HH production cross section. This chapter is a short glimpse into the design process; however, more details are presented in the full length of this Thesis.

6.1. Motivation to Target the WWZZ Decay Channel

The $HH \to WW^*ZZ^*$ channel has a BR of only around 1 %. Neither of the vector bosons is detectable in our detector; however, several of their decays are. Figure 6.1 shows all the possible decays of $HH \to WW^*ZZ^*$. For the initial analysis, seven categories were created based on lepton multiplicity as summarised in Table 6.1, aiming at final states with charged leptons. The initial focus was on the three- and four-lepton categories. Of these, the three-lepton category with a partial dataset is developed in a parallel analysis, see [45], while the four-lepton category is detailed in this Thesis.

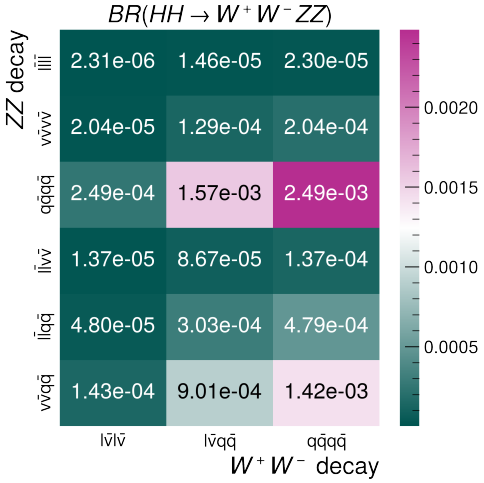


Fig. 6.1. Possible decays of HHs through the WWZZ decay channel.

Due to the rarity of $HH \to WW^*ZZ^* \to 4l$, this search suffers from various background sources, both irreducible and reducible. They include the production of vector bosons with additional jets, vector boson production in association with top quark pair,

Table 6.1

Requirements for the initial design of analysis categories, targeted vector boson decays, number of opposite electric charge sign (OS) and same flavour (SF) or opposite flavour (OF) lepton pairs, other set requirements, as well as initial estimates for signal ad background yields for 200 fb^{-1}

	2.1	2.2	3	4.1	4.2	5	6
WW decay	$l\nu l\nu$	$qqqq$	$l\nu qq$	$l\nu l\nu$	$qqqq$	$l\nu qq$	$l\nu l\nu$
ZZ decay	$qqqq$	$llqq$	$llqq$	$llqq$	$llll$	$llll$	$llll$
OSOF pairs	1	0	0	1	0	0	0
OSSF pairs	0	1	1	1	2	2	2
$m_{ll} > 12$ GeV	1	1	1	2	2	2	2
$m_{ll} \in (25, 100)$ GeV	0	0	1	1	1	1	1
$m_{ll} \in (80, 100)$ GeV	0	0	0	0	1	1	1
$n_{\text{jets}} \geq$	2	2	3	2	2	0	0
$E_T \geq$ X GeV	30	0	20	30	0	20	20
signal yield	0.605	1.486	0.915	0.083	0.093	0.029	0.004
background yield	3042891	15184261	907191	20270	8618	1077	43

single and pair production of top quarks, as well as diboson and triboson production, and other even rarer processes. Additionally, single-Higgs boson production is considered as a background source. Overall, physics object choice and event categorisation should disfavour events coming from backgrounds, with a moderate impact on the signal efficiency.

6.2. Data Samples

The analysed data have been collected with the CMS detector during the full Run 2 (2016, 2017, and 2018) and the early Run 3 (2022 and 2023). These periods correspond to around 200 fb^{-1}. The remnants of $HH \to WW^*ZZ^* \to 4l$ are looked for in single and double lepton primary datasets, and only data certified with the Golden standard are used in this work. More details are provided in the full-length Thesis.

Due to drastic changes in the detector, the 2016, 2022, and 2023 data-taking periods have been split into two eras per year that require dedicated data analysis or handling. `2016pre` and `2016post` cover 20.2 fb^{-1} and 16.6 fb^{-1}, respectively [46]. `2022` and `2022EE` cover 8.08 fb^{-1} and 26.67 fb^{-1}, respectively [47]. `2023` and `2023BPix`, cover 18.41 fb^{-1} and 9.45 fb^{-1}, respectively [47]. Despite the challenges, the data collected by the CMS detector remain state-of-the-art.

6.3. Monte Carlo Simulations and their Reweighing

This Thesis has used LO and NLO (or even higher when possible) MC samples generated by `MadGraph5_amc@nlo` [37], `Powheg v1.0` or `Powheg v2.0` [39], [49], [50], [51], [52],

Table 6.2

List of used MC samples; † has an extra restriction on $4 < m_{ll} < 50$ GeV, and †† has an extra restriction on $m_{ll} > 50$ GeV; "N/A" indicates the sample was not available in the given Run; abbreviations used: PW(2) – Powheg (v2), MG5 – MadGraph5, FXFX – Frederix–Frixione merging scheme [48], MS – MadSpin, JHU – JHUGen

Process	Decay modes	Generators	$\sigma^{\text{Run 2}}$, pb	$\sigma^{\text{Run 3}}$, pb
ggFHH	$HH \to VV^*VV^*$	PW	1.87e-03	2.07e-03
ggFHH	$HH \to bb\tau\tau$	PW+MG5@LO	2.46e-03	2.46e-03
ggFHH	$HH \to bbZZ^*$	PW	4.92e-06	4.92e-06
ggFHH	$HH \to \tau\tau VV^*$	PW	9.42e-04	1.04e-03
ggFHH	$HH \to \tau\tau\tau\tau$	PW	1.20e-04	1.33e-04
ggFH	$H \to ZZ \to 4l$	PW/PW2+JHUGen	0.013	0.014
VBFH	$H \to ZZ \to 4l$	PW/PW2+JHUGen	1.04e-03	1.12e-03
W^-H	$H \to ZZ \to 4l$	PW/PW2+MiNLO+JHUGen	1.46e-04	1.56e-04
W^+H	$H \to ZZ \to 4l$	PW/PW2+MiNLO+JHUGen	2.31e-04	2.44e-04
ZH	$H \to ZZ \to 4l$	PW/PW2+MiNLO+JHUGen	6.25e-04	7.75e-04
$t\bar{t}H$	$H \to ZZ \to 4l$	PW/PW2+JHUGen	1.39e-04	1.56e-04
$b\bar{b}H$	$H \to ZZ \to 4l$	PW/PW2+JHUGen	1.34e-04	1.45e-04
$t\bar{t}WH$	all	MG5@LO	1.14e-03	N/A
$t\bar{t}ZH$	all	MG5@LO	1.13e-03	N/A
$Z/\gamma^* + j$	$2l$	MG5@NLO+FXFX	360	386
$t\bar{t} + j$	$2l2\nu$	PW/MG5@NLO+FXFX	88.3	95.5
t or $\bar{t}[tWll]$	all	MG5@LO	0.011	N/A
t or $\bar{t}[tZq]$	all	MG5@LO	0.076	N/A
tW^-	$2l2\nu$	PW	N/A	36.0
$\bar{t}W^+$	$2l2\nu$	PW	N/A	36.0
$t\bar{t}\gamma + j$	all	MG5@NLO+FXFX(+MS)	3.76	4.22
$t\bar{t}\gamma$	all	MG5@LO	1.50	N/A
$WZ\gamma$	$W \to l\nu,\ Z \to 2q$	MG5@NLO	0.043	0.084
$t\bar{t}W + j$	$W \to l\nu$	MG5@NLO+FXFX+MS	0.216	N/A
$t\bar{t}WW$	all	MG5@LO(+MS)	7.00e-03	8.20e-03
$t\bar{t}WZ$	all	MG5@LO	2.45e-03	2.71e-03
$t\bar{t}ZZ$	all	MG5@LO(+MS)	2.45e-03	1.58e-03
$t\bar{t}Z$	$Z \to 2l$	MG5@NLO	0.075	N/A
$t\bar{t}Z$ †	$Z \to 2l$	MG5@NLO	N/A	0.086
$t\bar{t}Z$ ††	$Z \to 2l$	MG5@NLO	N/A	0.039
WW	$2l2\nu$	PW	11.1	11.8
WZ	$3l\nu$	PW+MG5@NLO+FXFX	5.21	4.92
ZZ	$4l$	PW	1.32	1.39
WWW	all	MG5@NLO(+MS)	0.216	0.233
WWZ	all	MG5@NLO	0.171	0.185
WZZ	all	MG5@NLO	0.057	0.062
ZZZ	all	MG5@NLO	0.015	0.016
$t\bar{t}t\bar{t}$	all	MG5@NLO	8.21e-03	9.65e-03
$WWZZ$	all	MG5@LO	4.44e-04	4.91e-04

while hadronisation has always been done using `Pythia` [38]. Several resonances are decayed with `MadSpin` [53], and `JHUGen` [54], and extended with `MiNLO` [55]. A short summary is shown in Table 6.2. Regarding `Pythia`, the most comprehensive and modern event "tune" developed for CMS is used: CP5 [56]. It uses `NNPDF3.1 NNLO` [57] PDF set. Cross sections are from [58], [59], [60], [61]. More information is given in the full-length Thesis.

Low-level inaccuracies are corrected using per-event weights. Each event gets an assigned weight as in Eq. (6.1):

$$w_{event} = \frac{L_{int} \cdot \sigma \cdot \mathcal{BR} \cdot w_{PU} \cdot SF_{leptons} \cdot SF_{jets} \cdot w_{L1} \cdot w_{generator}}{\sum w_{generator}}, \tag{6.1}$$

where L_{int} is the integrated luminosity, σ is the cross section of the simulated sample and $\mathcal{BR}$ is the BR of the simulated events, w_{PU} is the PU weight, SF are the scale factors, w_{L1} is the Level-1 trigger pre-firing [62], and finally, $w_{generator}$ is the weight assigned by the event generator. In general, a scale factor is the efficiency ratio of, e.g. lepton identification in simulations and in the data, as minor modelling imperfections always exist. More details can be found in [63], and the full-length Thesis.

6.4. Designing the Signal Region

Events must pass several stages of selection before being considered in the signal region (SR) of this analysis. The very first step is the trigger selection, afterwards object selection is performed, then events are categorised based on several object kinematics or candidate parameters, and finally a boosted decision tree (BDT) threshold is imposed. Each stage aims to target the signal while rejecting background processes as much as possible.

Firstly, events are selected if tight working point (WP) single lepton, or ((very) very) loose WP double or triple lepton High-level trigger paths have been triggered. The OR combination provides trigger efficiency of around 98 % for the signal events. When selecting events from several primary datasets, care is taken to not double count them.

6.4.1. Lepton Selection and Calibration

After the event has passed the trigger selection, the physics objects have to be selected. This Thesis is not explicitly targeting taus or neutrinos; therefore, only three main physics objects are considered – muons, electrons, and jets. This being a completely new study, gave the opportunity to design the desired selection by launching several studies to find a clear separation between signal and background leptons and jets. For more details, see the full-length Thesis, as here only a summary will be provided.

The final requirement set is summarised in Table 6.3 for electron selection and in

Table 6.3

Electron selection

Observable	Description	Value
p_T any	transverse momentum	> 10 GeV
$\|\eta\|$	abs. pseudorapidity	< 2.5
$\|d_{xy}\|$	abs. IP in xy plane	< 0.05 cm
d_z	IP along the beam line	< 0.1 cm
d/σ_d	3D IP	< 8
I_r	isolation criteria	$< 0.4 \cdot p_T$
$\sigma_{i\eta i\eta}$	electron supercluster shape	$\{0.011/0.030\}$
H/E	H (E) energy deposited in the HCAL (ECAL)	< 0.1
1/E-1/p	E – cluster energy, p – track momentum	> -0.04
conv.	photon conversion rejection	True
miss.	missing inner hits in tracker	$= 0$
ID	`mvaFall17V2noIso_WP[L/90]` ID	$\geq$ WP-loose/WP-90
ID	`MVA ttH`	$\geq$ WP-loose

Table 6.4

Muon selection

Observable	Description	Value
p_T any	transverse momentum	> 10 GeV
$\|\eta\|$	abs. pseudorapidity	< 2.4
$\|d_{xy}\|$	abs. IP in xy plane	< 0.05 cm
d_z	IP along the beam line	< 0.1 cm
d/σ_d	3D IP	< 8
I_r	isolation criteria	$< 0.4 \cdot p_T$
ID	PF muon ID	$\geq$WP-medium
ID	`MVA ttH`	$\geq$WP-loose

Table 6.5

Jet selection, (* – requirements that are applied for Run 2 period only)

Observable	Description	Value
p_T	transverse momentum	> 25 GeV
$\|\eta\|$	pseudorapidity	< 2.4
`btagDeepFlavB`	b-tagger	$<$ WP-loose
jet ID	jet quality ID	True
jet PU ID*	PU rejection	True
overlap	rejected if an ID'd lepton in $\Delta R < 0.4$	True

Table 6.4 for muon selection. Several studies focused on investigating various quality cuts – lepton isolation criteria, thresholds on the impact parameters (IPs), numerous characteristics for deposited energy, as well as other criteria. Both electrons and muons have to pass ID criteria developed by the respective physics object group (POG). For electrons with $p_T > 10$ GeV, the loose WP corresponds to around 98 % signal efficiency, while WP90 has only 90 % signal efficiency. For muons with $p_T > 20$ GeV, the chosen ID ensures 99.5 % signal efficiency. To combat the $t\bar{t}$ contribution, an additional ID, `MVA ttH`, is applied for both types of leptons. Its WP-loose provided the highest statistical significance in this search. More details are found in the full-length Thesis.

Moreover, the kinematic differences between signal and background leptons can be exploited. Even though one of the vector bosons always is off-shell, thus producing low-p_T leptons, signal leptons are expected to be slightly boosted due to the Higgs decays. At the same time, many background processes have leptons with a softer p_T spectrum overall. Thus, $p_T > 25$ GeV is imposed on the leading lepton and $p_T > 15$ GeV on the sub-leading lepton, giving a significant overall improvement.

6.4.2. Jet and Missing Transverse Energy Selection and Calibration

Due to the hadronic decays of vector bosons, jets are equally important as leptons for this analysis. Ideally, the category 4.1 would expect two jets, while category 4.2 would expect four, while in the decay channel where $W \to \nu l$, a part of the energy is carried away by neutrinos, thus MET is expected. For both jets and MET, a loose selection is chosen to increase signal acceptance.

Table 6.5 summarises the final selection for jets. Jets have to pass the tight jet ID, which has signal efficiency and background rejection of more than 98 %. Problematic jet regions and jets that overlap leptons are avoided. For Run 2 jets with $p_T < 50$ GeV, an additional PU ID is applied to reject PU jets. This ID has a signal efficiency larger than 95 % in the entire detector region regardless of the average PU. The `BTV POG`'s (b-tagging and vertexing POG) recommended `DeepJet` [64] tagger was used for jet flavour tagging. Its WP-loose provides around 10 % mis-tagging efficiency for light jets. As always, the `JetMET POG` (Jet and missing transverse energy POG) recommendations for jet corrections and energy resolution are followed.

6.4.3. Event Categorisation

If the event satisfies the trigger selection described in Section 6.2 and four leptons pass the criteria described in Section 6.4.1, the next step is the event categorisation. Three $HH \to WW^*ZZ^*$ decays can produce four leptons. However, in the case where $Z \to \nu\nu$, a significant amount of the decay channel's energy is carried away by neutrinos. Thus,

the event selection has been developed with two other decay channels in mind, creating categories 4.1 and 4.2. The event is sequentially tested for requirements shown in Fig. 6.2, forming the pre-selection phase.

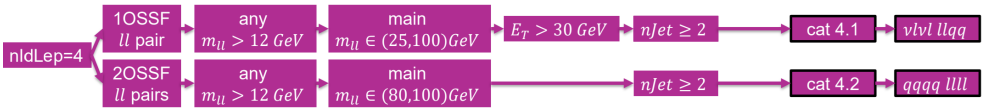


Fig. 6.2. Sequential event categorisation for four-lepton categories.

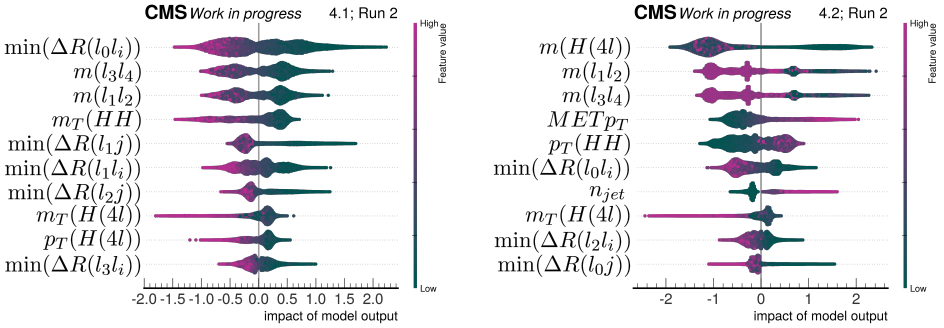
Initial studies used a looser lepton invariant mass window of $m_{ll} \in (60, 120)$ GeV; however, tightening it to $m_{ll} \in (80, 100)$ GeV provided a medium improvement in statistical significance. This will make it easier for future combinations, as the four-lepton category in CMS multilepton analysis [11] requires an inverted mass window. Requiring exactly four jets in the category 4.2 proved to be inefficient. However, requiring the event to have at least two non-b-tagged jets gave the highest statistical significance and enough training data for the BDT model. For category 4.1, the targeting of two jets proved to be effective and was kept as the main requirement.

6.4.4. The BDT Classifier

A dedicated observable with strong background discriminating power is constructed by a BDT, built using the `XGBoost` method [65]. A BDT model is trained per category and per Run. After obtaining the BDT score distribution, a threshold is evaluated to classify an event in the SR. More details are available in the full-length Thesis.

As the first step, the most optimal input data was investigated for the BDT training. Signal-wise, dedicated filtered samples of $HH \to WW^*ZZ^*$ with at least four leptons were used. To increase statistics, eras within a Run were merged as no significant differences were found across them for variable distributions. Similarly, for background, samples within a Run were merged in a controlled way to increase the statistics while keeping the mixture representative. Overall, fewer events are available to train the BDTs for category 4.1 compared to category 4.2; thus, each category has a dedicated BDT model, as a combined one is not powerful enough.

As the second step, the SHAP values [66] were used to find the minimal sufficient set of 25 high-level input features, that encapsulate kinematic quirks characteristic of signal events. The final set for both categories can be seen in the full-length Thesis, while the top 10 highest ranking feature SHAP shapes can be seen in Fig. 6.3 for the Run 2 BDT models. The BDT for category 4.2 strongly benefits from the Higgs boson and the HH candidate features, while the BDT for category 4.1 has learnt to recognise angular variables. Overall, multiple features have strongly polar distributions with relatively high impact.



(a) BDT for category 4.1, Run 2

(b) BDT for category 4.2, Run 2

Fig. 6.3. Top 10 BDT feature SHAP values.

As the third step, the BDT architecture was optimised using the k-fold cross-validation on sets of hyperparameters to find the most optimal set of values giving the lowest average error. These parameters are detailed in the full-length Thesis. To evaluate the performance of BDTs, receiver operating characteristic curve and the area under curve were used as metrics. The 4.1 category's BDTs took fewer epochs to converge, while the 4.2 category's BDTs reached lower loss function values. For all BDTs the area under the curve was greater than 0.9.

Table 6.6

BDT score threshold and relative uncertainties considered; the relative MC uncertainty on signal, S_r, is shown in the third column, the relative MC uncertainty on background, B_r, in fourth, and the relative uncertainty on statistical significance in fifth

Cat	Threshold	S_r, %	B_r, %	$\left(\frac{S}{\sqrt{B}}\right)_r$, %	Year
4.1	0.6	15.23	14.53	16.88	2016
4.1	0.7	14.11	18.94	17.00	2017
4.1	0.6	12.22	25.59	17.70	2018
4.1	0.4	7.02	24.27	14.02	2022
4.1	0.2	5.11	23.34	12.74	2023
4.2	0.8	12.85	3.24	12.95	2016
4.2	0.9	17.83	3.70	17.92	2017
4.2	0.8	13.31	7.37	13.81	2018
4.2	0.4	5.15	17.62	10.21	2022
4.2	0.8	8.01	7.93	8.93	2023

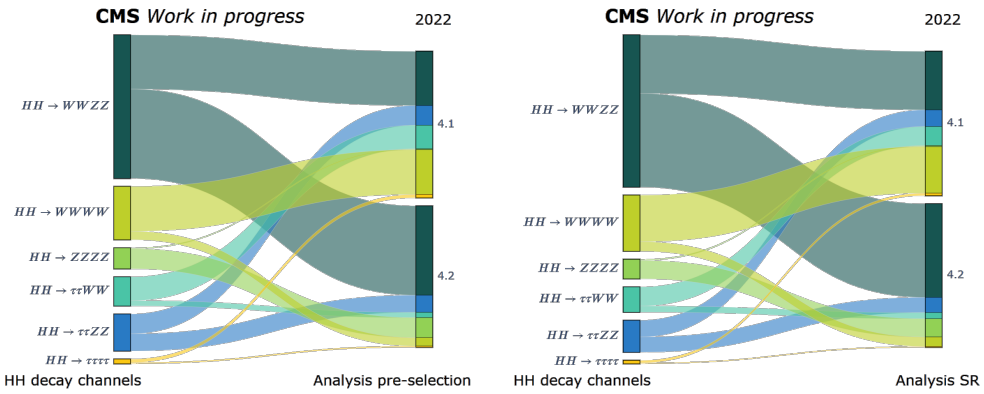
As the fourth step, a suitable threshold was evaluated on the BDT score to classify an event to be in the SR. This threshold signifies that it is more likely to find signal above this threshold than below it. As the MC statistics was the main limiting factor, the threshold was reevaluated in each year, reaching the highest statistical significance with a tamed MC uncertainty. Table 6.6 summarises these values per year.

Finally, the influence of the BDT score threshold is evaluated. As an example, for

the most luminosity-rich era, 2018, the 4.1 BDT reduced the signal contribution by 22 % while reducing the background by around 81 %, and the 4.2 BDT reduced the signal contribution by 26 % while reducing the background by around 97 %. Similar results have been reached in other years. A more detailed description can be found in the full-length Thesis. The contribution of various processes in each category and SR will be shown in Chapter 7.

6.5. Considering Other HH Contributions

Although the $HH \to WW^*ZZ^*$ decay is leading in each SR, other HH decays also contribute, as can be seen in Fig. 6.4. In the 2022 example, $HH \to WW^*ZZ^*$ occupies



(a) Pre-selection (b) Signal region

Fig. 6.4. Relative comparison for several HH decay mode contribution in the pre-selection and SR phase-space, example from the 2022 data-taking year.

around 41 % and 63 % in categories 4.1 and 4.2, respectively. After pre-selection $HH \to WW^*WW^*$ is of similar contribution as $HH \to WW^*ZZ^*$ in category 4.1; however, after the full selection it is reduced to around 32 %. At the same time, in category 4.2, the sub-leading contributions are $HH \to \tau^+\tau^-ZZ^*$ and $HH \to ZZ^*ZZ^*$ contributing in the order of 10 %.

The dedicated BDT has allowed to keep the sensitivity to HH decays while not sacrificing the $HH \to WW^*ZZ^*$ decay channel. When evaluating the contributions in the SRs, the situation is improved as the $HH \to WW^*ZZ^*$ contribution is slightly increased; however, the yield undergoes a significant decrease. These results give a basis for a great question, if the final signal-background separation could be done even more effectively. And moreover, should the $HH \to WW^*ZZ^*$ decay channel be analysed separately or in combination with other multilepton decays.

6.6. Data and MC Agreement

To validate physics modelling and the constructed categories, data and MC agreement is checked in control regions (CRs). Within these regions, the distributions of high-level and low-level variables are compared between data and MC simulations. Two CRs per category are constructed, they are very close to the SR, but

- with inverted BDT score threshold (`CRlow`);
- with inverted jet multiplicity requirement (`CRno`).

Clearly the `CRno` provides more statistics than the `CRlow`; however, this phase-space is further away from the SR's phase-space. Also here, the CRs for the category 4.1 remain to be a mixture of backgrounds, while CRs for category 4.2 are dominated by ZZ production.

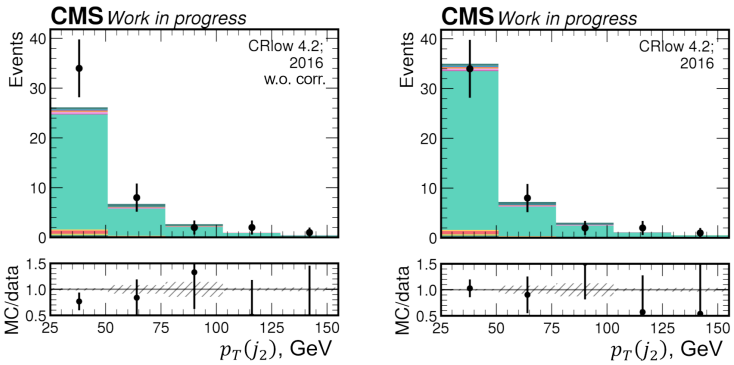


Fig. 6.5. Transverse momentum distributions before and after data-driven ZZ modelling correction for sub-leading jets, example of CRlow 4.2, 2016.

Although an overall good shape agreement was observed across years, within most of 4.2 CRs significant data surplus was noticed, indicating poor MC modelling, most likely coming from ZZ production. Investigation revealed poor high-jet multiplicity modelling. Therefore, a data-driven correction for the ZZ modelling was created based on the transverse momentum of the sub-leading (second) and sub-sub-leading (third) jet (leading in p_T). More details are found in the full-length Thesis. An example of the transverse momentum distributions before and after data-driven ZZ modelling correction for the sub-leading jets are shown in Fig. 6.5. The correction improves the data and MC agreement with slight tension remaining in 2017.

6.7. Uncertainty Estimations

No measurement can be quoted without its uncertainty. Although this measurement is completely statistically limited, it is necessary to account for potential biases and vari-

ations in the measurement. This Thesis considers both theoretical uncertainty sources and experimental ones, of which the greatest influence comes from electron ID efficiency (below 10 %, mainly inflated by low-p_T electrons), pileup (below 5 %), and b-tagging efficiency (below 3 %). For more details, see the full-length Thesis.

6.8. The Measurement

The main goal of this Thesis analysis is to compute an exclusion limit on the HH production signal strength, μ, when approaching it through the $HH \to WW^*ZZ^*$ decay. The limit is calculated using the $\mathrm{CL_S}$ method [67], which is a modified frequentist approach and noted as CL. A statistical model needs to be built.

Statistical analysis based on expected observations is performed with CMS statistical analysis software, `COMBINE` [68]. It allows building the model using expected observations, including nuisance parameters, running statistical tests, and finally providing tools for inspecting the model and the tests performed. Due to the low expected yield, a simple counting experiment is chosen. The test is set up with two SRs and two CRs for constraining the ZZ modelling, resulting in four bins. For easier manipulation of uncertainties, also called nuisances, backgrounds are subdivided into four groups: H for single-Higgs related backgrounds, VV for dibosons, mainly the ZZ, t for top quark related backgrounds, mainly $t\bar{t}$, and other backgrounds.

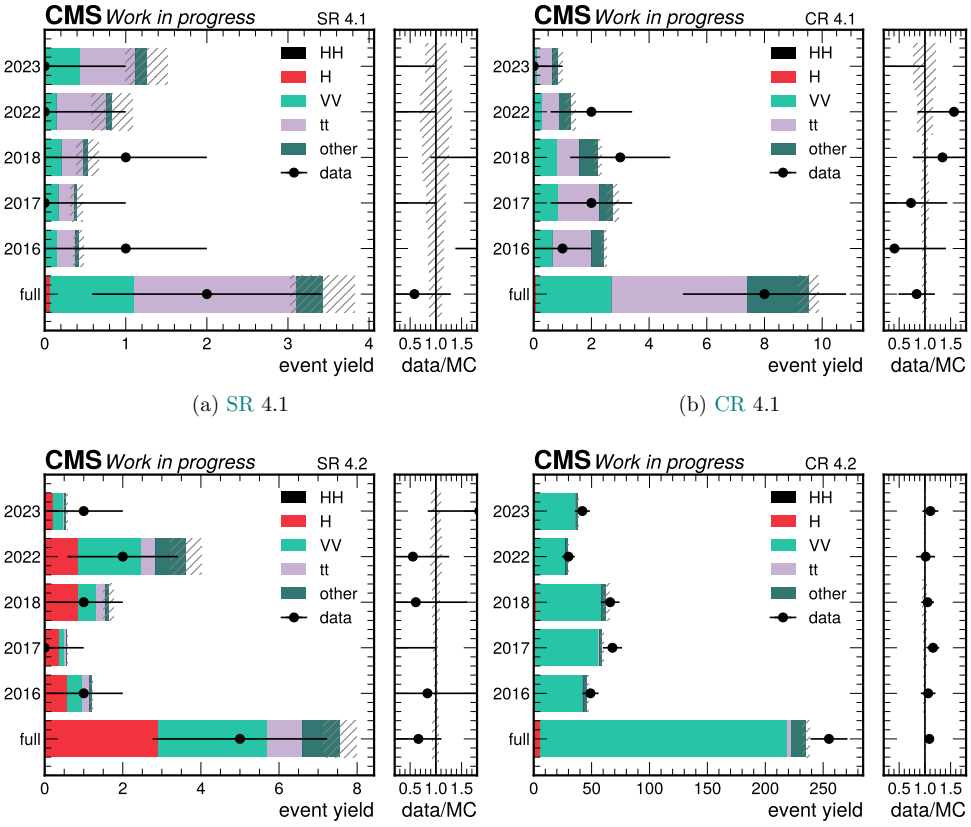
A profile likelihood ratio is used as the test statistic, q_μ. To avoid unwanted exclusion of small values of μ, the simple ratio is replaced by a piecewise function as shown in Eq. (6.2), where $\hat{\mu}$ is the maximum likelihood estimator for μ, and $\hat{\nu}$ is the nuisance parameters that maximise the likelihood function, $\mathcal{L}$. This definition encapsulates that the HH signal strength is physically non-negative.

$$q_\mu = \begin{cases} -2\log\left(\dfrac{\mathcal{L}(\mu, \hat{\nu}(\mu))}{\mathcal{L}(\mu = 0, \hat{\nu}(\mu = 0))}\right) & \hat{\mu} < 0 \\ -2\log\left(\dfrac{\mathcal{L}(\mu, \hat{\nu}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\nu}(\mu))}\right) & 0 < \hat{\mu} < \mu \\ 0 & \mu < \hat{\mu} \end{cases} \tag{6.2}$$

The statistics model is tested by evaluating the nuisance impact on the best fit value by simultaneously fitting them and evaluating any pulls. These fits simply maximise the likelihood, and no test statistic (with any priors) is applied; moreover, it is evaluated on pseudo-data. Overall, no strong pulls or abnormal influence of the systematics can be seen, thus rendering the model viable. For visual representation and more details on the model, see the full-length Thesis.

7. RESULTS AND PERSPECTIVE

As the building blocks of the $HH \to WW^*ZZ^* \to 4l$ analysis have been summarised, it is time to evaluate the results. To reiterate, the goal of this analysis is to set an upper limit on the HH signal strength, using data collected with the CMS experiment in the full Run 2 and early Run 3 (2022 and 2023) data-taking periods, amounting to 200 fb^{-1}. A comparison of expected and observed event yields per category is shown in Fig. 7.1. As the BDT threshold varies in each year, the expected yield is not proportional to the integrated luminosity.



(a) SR 4.1 (b) CR 4.1

(c) SR 4.2 (d) CR 4.2

Fig. 7.1. Expected and observed events per year, per SR and CRlow.

Overall, good agreement is seen between the data and the simulations. Less than five events are expected in SR 4.1; however, only two have been observed. In CR 4.1, around eight events are expected and eight are observed. Similarly in category 4.2, less than eight events were expected and five were observed in the SR, while in the CR, around

220 were expected but over 250 were observed. The leading background in SR 4.1 and in CR 4.1 comes from the $t\bar{t}$ group, while the leading backgrounds in SR 4.2 come from the single-Higgs group, mainly ggH, followed by the diboson background, mainly ZZ. The diboson background group is the dominant in CR 4.2.

Table 7.1

Observed and expected yields for various processes in the SRs, combined across all years, equalling to around 200 fb^{-1}

Process	SR 4.1, events	SR 4.2, events
HH $\rightarrow$ 2W2Z ($\times$100)	0.798 ± 0.080	1.136 ± 0.089
HH $\rightarrow$ 4W ($\times$100)	0.752 ± 0.074	0.117 ± 0.028
HH $\rightarrow$ 4Z ($\times$100)	0.002 ± 0.002	0.156 ± 0.035
HH $\rightarrow$ 2W2τ ($\times$100)	0.412 ± 0.059	0.070 ± 0.026
HH $\rightarrow$ 2Z2τ ($\times$100)	0.183 ± 0.038	0.238 ± 0.048
HH $\rightarrow$ 4τ ($\times$100)	0.046 ± 0.006	0.006 ± 0.002
HH $\rightarrow$ 2b2Z ($\times$100)	-	0.043 ± 0.002
total expected signal ($\times$100)	2.19 ± 0.13	1.76 ± 0.11
Higgs boson backgrounds	0.06 ± 0.00	2.90 ± 0.03
dibosons, mainly ZZ	0.97 ± 0.10	2.26 ± 0.06
$t\bar{t}$ (and t) backgrounds	2.77 ± 0.40	0.91 ± 0.23
other backgrounds	0.40 ± 0.02	0.96 ± 0.34
total expected background	4.54 ± 0.42	7.66 ± 0.42
data	2	5

The total expected yields for the signal and background processes in the SRs across all five data-taking years are given in Table 7.1. It can be seen that the signal contribution is around 200 times lower than the background contribution in SR 4.1 and around 400 times lower in SR 4.2. The $t\bar{t}$ group has a higher statistical uncertainty because there are fewer events available for evaluation.

The expected and observed upper limits on the HH production signal strength in each category per Run are shown in Fig. 7.2. SR 4.1 provides a lower upper limit mainly due to the relatively lower background rates, and similarly, Run 2 is more sensitive due to the higher integrated luminosity. Not shown in the figure, but the tightest limit is set using data from the luminosity richest period, the 2018 data-taking period. Altogether, **the combination of both categories in both Runs allows to set an observed (expected) 95 % CL upper limit of 142.3 (176.8) on the HH signal strength**. The observed limit is lower than expected, as on average less data has been observed in the SRs than expected. However, it is well within the one standard deviation of the expected value.

It is not possible to do a direct comparison with other or older results as this is a novel analysis. Moreover, at the time of writing of this Thesis summary, not many analyses have combined data from the full Run 2 and early Run 3. The closest comparison can be made with the full Run 2 multilepton analyses performed by the CMS collaboration [11].

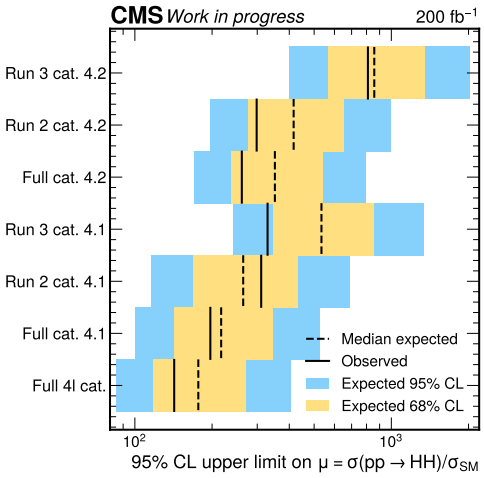


Fig. 7.2. Expected and observed upper limits on the signal strength modifier μ per four lepton category in each period and combined.

More specifically, its four lepton category, which targets $WWWW \to \nu l \nu l\ \nu l \nu l$. For integrated luminosity of 138 fb^{-1}, this channel has set an observed (expected) 95 % CL upper limit of 56.7 (76.8) on the HH signal strength. Thus rendering the result achieved by the $HH \to WW^*ZZ^* \to 4l$ channel with significantly smaller BR impressive. Clearly, it does not aim to become the leading HH decay channel; however, it is a complementing one for the combination.

To estimate the statistical constraints in the model, the measurement is repeated without including any systematics ("stat.-only"). Table 7.2 summarises the limits calculated with the "stat.-only" model and the full model ("stat. and syst."). The comparison of these limits quantifies the degradation of sensitivity as a result of the systematic effects, and it can be seen that the influence is less than a half percent.

Table 7.2

Comparison of the expected and observed 95 % CL limit on the μ set from the $HH \to WW^*ZZ^* \to 4l$ analysis, using full CMS Run 2 and early Run 3 dataset

Scenario	Expected limit	Observed limit
stat.-only	169.5	135.8
stat. and syst.	176.8	142.3

Although the measurement is statistically dominated, it is useful to understand which are the dominating systematics and how to mitigate them, so that when more statistics are available, systematics will not be the bottleneck of this analysis. In Fig. 7.3, the observed impact of nuisances is shown for the fitted signal strength. Here, the greatest influence comes from the MC statistics, more specifically, SR 4.1 in 2018, followed by the contribution of electron efficiency and the ZZ normalisation. It clearly outlines the

insufficient MC statistics. Overall, post-fit values are not shifted far away from their priors, and within uncertainties overlap (filled points). Also, pulls are moderate and condensed around the central value. When investigating nuisance impacts, no priori has been applied, and a negative, unphysical signal strength value is reached. However, it is not unusual when performing calculations near a physical boundary. More details are in the full-length Thesis. The negative value is consistent with the observed data deficit.

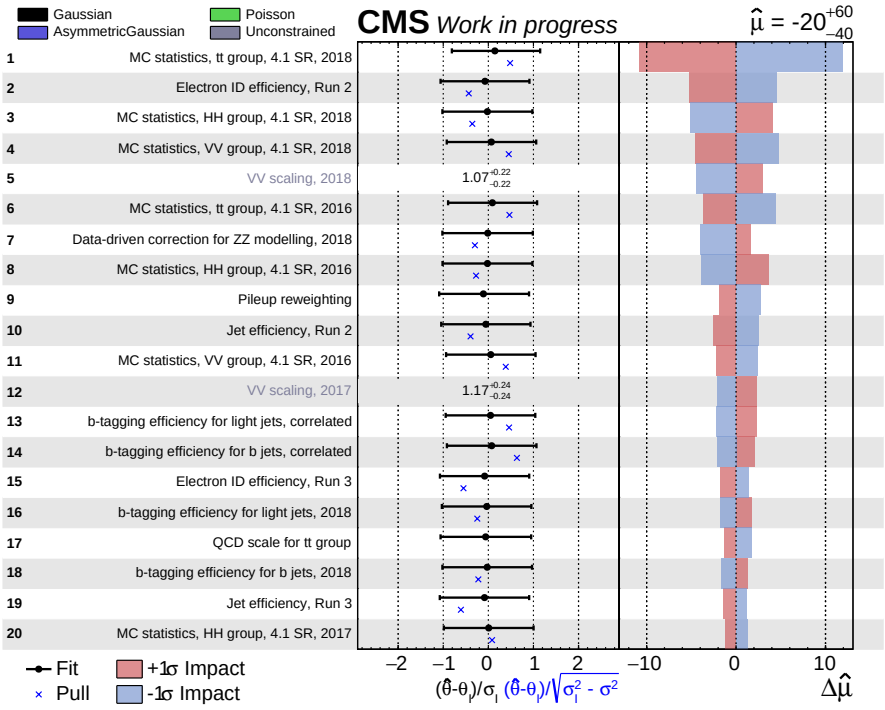


Fig. 7.3. The observed ranking of the top 20 nuisance parameters by their impact on the signal strength modifier μ, for the 4l categories.

7.1. Projection to HL-LHC and Combination with Other Channels

As the upcoming HL-LHC is expected to deliver an integrated luminosity up to 3 ab^{-1} at a centre-of-mass energy of at least 13.6 TeV, it is interesting to see how the calculated limit would be affected. It is well established that such a major increase in data will allow precision measurements of rare processes such as the HH production. A simple extrapolation to 3 ab^{-1} is employed, scaling the signal and background yields with luminosity and assuming the current efficiencies and systematic uncertainties. The total yields in both

Table 7.3

Expected yields for 3 ab^{-1} in the signal and control regions considered for the extrapolation

Process	SR 4.1, events	SR 4.2, events	CRlow 4.1, events	CRlow 4.2, events
signal (×100)	34.05 ± 1.95	26.55 ± 1.65	7.96 ± 1.83	12.14 ± 2.63
background	68.1 ± 6.3	114.9 ± 6.3	122.4 ± 15.9	3298 ± 395

CRs and SRs are shown in Table 7.3, as they are used in the calculation. The extrapolation allows to set a median 95 % CL expected upper limit of 51.5 on the HH signal strength. Even with such an increase in data, the set limit is weaker than the one set by the CMS multilepton analysis, using only Run 2 data [11]. Thus, strengthening the idea that this channel could be included in the HH search combination.

As it had been intended from the very beginning that the $HH \to WW^*ZZ^* \to 4l$ channel is combined with other $HH \to WW^*ZZ^*$ channels to improve the sensitivity of this analysis, a HL-LHC estimate is given for this combination. With full Run 2 data, the three-lepton category has been able to set an observed (expected) 95 % CL upper limit of 410 (240) on the HH signal strength [45]. For the HL-LHC estimate, this result will be extrapolated based on the increase in integrated luminosity. Moreover, since both the five- and six-lepton categories have not yet matured to output a result, their initial sensitivity (Table 6.1) will be extrapolated with the assumption that the sensitivity is inversely proportional to the signal strength. Finally, all categories are combined.

Table 7.4

The estimate of the expected upper limit on the signal strength μ at 3 ab^{-1}

Category	Cat. 2	Cat. 3	Cat. 4	Cat. 5	Cat. 6	Combined
limit	3494	46.5	51.5	2038	2953	34.5

Following these assumptions, for 3 ab^{-1} the combination of all the mentioned WW^*ZZ^* categories, gives an estimate of the mean value of the 95 % CL upper limit of 34 on the HH signal strength. A breakdown of the limit per category is shown in Table 7.4. Clearly, this approach ignores any correlated systematics, but is powerful enough to give a "back of the envelope" estimate that motivates future studies. Also, the used sensitivities for other $HH \to WW^*ZZ^*$ decays are slightly outdated, and significant improvement is expected. Moreover, a possible assimilation in the CMS multilepton analysis could allow for an even greater combined improvement and a complete reduction of the overlap.

8. CONCLUSIONS AND OUTLOOK

Within this Thesis, extensive work was carried out on the DCS and DSS for a MTD's test-stand, as well as a novel physics analysis was designed and performed in the four-lepton channels of the $HH \to WW^*ZZ^*$ decay channel. The set tasks for the Doctoral Thesis have been accomplished, theses defended, and the work outlined in this Thesis has allowed to form the following conclusions.

1. The review of literature summarised in Chapter 2 has shown that the analysis of HH production provides a direct probe of the Higgs potential and allows to experimentally constrain the Higgs boson self-coupling parameter.

2. In recent years the HH production has become an increasingly active area of research within the LHC program; it is especially highlighted by the HH production combination articles published by both the CMS and ATLAS collaboration.

3. A dedicated $HH \to WW^*ZZ^*$ decay channel analysis in the four-lepton decay channels as part of its categories is a strong and complementing contender for the HH analysis community, moreover, previously, this decay channel has not been explicitly targeted and optimised for.

4. As summarised in Chapter 3 and Chapter 4, the CERN accelerator complex with the LHC provides high-energy and high-luminosity pp collisions, and the multi-purpose CMS detector allows to reconstruct and identify many final states of these collisions, thus making it the perfect experimental setup for the $HH \to WW^*ZZ^* \to 4l$ decay channel analysis.

5. A novel $HH \to WW^*ZZ^* \to 4l$ decay channel analysis has been developed, and Chapter 6 encapsulates all the necessary building blocks of it. Several optimisation steps are shown for the designing of the framework, categories and object selection.

6. The performed analysis of $HH \to WW^*ZZ^* \to 4l$ decay channel has allowed to set an observed (expected) 95 % CL upper limit of 142.3 (176.8) times the SM value of the HH cross section. This is in agreement with the SM expectations, given the current experimental uncertainties.

7. The developed analysis framework is applicable for future $HH \to WW^*ZZ^* \to 4l$ and more generally, the $HH \to WW^*ZZ^*$ decay channel studies with increased amount of data, motivating such an analysis for the HL-LHC era studies.

8. Work detailed in Chapter 5 shows how a DCS has been built for the MTD test-stand. It has been built following the CMS production guidelines and has met all the set requirements. Moreover, it serves as a prototype for the final MTD DCS.

8.1. Outlook on the DCS Work

Now, as the HL-LHC is getting closer and closer, the work on the DCS becomes increasingly relevant. Section 5.2 already detailed three development scenarios for the built DCS prototype. The successful continuation of the project is crucial not only for the MTD test-stand, but also for the whole progression towards the final detector's DCS. It will need to continue to keep the needs of the experimental team, as well as follow the recommended production guidelines, to shape into a self-sufficient, reliable, and robust system. Judging from previous experience, attaining such qualities is well within reach.

8.2. Outlook on the HH Analysis

The timing of this Thesis allowed to analyse the full Run 2 dataset and only a partial Run 3 dataset. Now, as the Run 3 is near its end, it is clear to see that when reanalysed, this analysis would be dominated by the Run 3 data. Moreover, the integrated luminosity would be more than doubled, ensuring a great playground for any HH production analysis.

Run 3 has brought many technical improvements in particle reconstruction and identification that this analysis could benefit from. By choosing the new b-tagger [47], higher b-tagging efficiency and lower mis-tagging rates would be reached. This would allow to relax the jet transverse momentum requirements, thus recovering more signal while not sacrificing precision. Similarly, attention would need to be given to the lepton identification. With the increase in collected luminosity, the detector material has dealt with significant radiation damage. Thus, the analysis would benefit from using a dedicated ID trained with a high signal efficiency, also for the low-p_T electrons. The sensitivity of this analysis heavily depends on it.

Finally, it would be truly interesting to open two more aspects of the current analysis. The first would be the inclusion of a dedicated new physics scenario study. It would be interesting to train a separate BDT to evaluate the upper limits on the HH production in these scenarios. Second, it would be interesting to explicitly target other HH production modes.

Altogether, the foreseen data-taking periods ensure exciting years ahead for the HH production analyses and analysers. It will truly be a collaborative effort, as claiming the discovery of the HH production is in closer reach when working together, when combining as many decay channels and analysis as possible, thus showcasing what the high energy particle physics field is all about. Although the $HH \to WW^*ZZ^*$ decay channel will not play a central role in this search, it will be a complementing contributor, pulling its own weight. I am excited to see what the next years will bring and I will celebrate together with the community the discovery of HH production when the time comes.

ACKNOWLEDGEMENTS

I wish I could say that these formative years have been nothing but frolicking through meadows and greeting sunrises on mountain tops. If I had to do this again, I would change so many, many things (oh, hindsight, you silly thing), mainly being more present and cherishing the beautiful, beautiful souls around me throughout this journey. I must admit, even this thank you list is shortened, so feel free to check out the full-length Thesis!

Enormous thanks go out to my supervisors. I thank you, **Kārli**, with all my heart for being our forever-light house, for your guidance and support, although at times I wished we had five of you. Thank you, **Toni**, for training me to be a statistics police, your enthusiasm and time, even when you were away on your World tour! Thank you, **Frank**, for your calm and encouraging presence, for being the most complete source of information on everything CERN and CMS!

I need to shine some spotlight on the catalysts in my life. Thank you, **CERN**, the most wonderful village, for bringing so many motivated and sparkling people together and into my life. Thank you, **CMS**, for showing what particle physics is about, and how it is to work in a big collaboration. Thank you, **LLR**, for the most physics discussions I've had in a two-week span. Thank you, **DTF**, for adopting the little IPPAT island.

Oh, my comrades in the trenches, thank you for keeping up the morale! Dear **419**, thank you for the lunches! **Normund**, I'm happy we survived the electron task. I had my doubts. Thank you, **Lazar**, for trusting me with ice-skating, thank you, **Dace**, for saying yes to outings, although we did too little of them. **Marija**, thank you for your spark and trust in us all. Keep your momentum, gain new boosts, and I will be happy to see you reach your dreams. I'm sorry we sometimes were too cloudy. **Andri**, thank you for being at the right place at the right time. I (we) needed you so much. Will you ever reckon your impact? Thank you for finding CERN's gems and introducing me to them.

Finally, my nearest and dearest, I couldn't have done it without you. My dearest **Ame**, thank you for being the most loyal friend I've ever had. I'm happy our paths keep crossing; however, I wish they'd be more in sync. My closest family, thank you for your unconditional love. **Ruta** and **Andris**, thank you for reminding me of afternoon cacao, and oh, how you read all your grandkids like open books, thank you! **Dad**, I'm so happy that you are proud of me, and I'm so proud of you slaying your PhD! **Brother**, it has always been so lovely to be on the same wavelength. Thank you for always understanding me! **Mārieši**, thank you for visiting me at CERN. **Vitaut** and **Inita**, thank you for asking about the LHC. **Adrian** and **Aiden**, thank you for all the puzzles and Legos, and the ones yet to come! Thanks to your parents for always warm invitations to visit! **Reini**, thank you for filling my heart with overflowing love, and taking the leap of faith and sampling the Francophone sphere with me. And above all, despite the always annoying, time-warying distances between us, thank you for wanting to go around Lac Leman with me.

References

[1] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.020 (cit. on pp. 7, 12).

[2] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021 (cit. on pp. 7, 12).

[3] CERN. *The High-Luminosity LHC Project. 298th Meeting of Scientific Policy Committee*. Tech. rep. CERN, 2016. URL: https://cds.cern.ch/record/2199189 (cit. on pp. 7, 16).

[4] CMS Collaboration. *A MIP Timing Detector for the CMS Phase-2 Upgrade*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2667167 (cit. on pp. 7, 19).

[5] D. Galbraith and C. Burgard. *Standard model of physics*. https://texample.net/model-physics/. Accessed: 2025-09-01 (cit. on p. 12).

[6] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321 (cit. on p. 13).

[7] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508 (cit. on p. 13).

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585 (cit. on p. 13).

[9] A. A. Migdal and A. M. Polyakov. "Spontaneous Breakdown of Strong Interaction Symmetry and the Absence of Massless Particles". In: *Sov. Phys. JETP* 24 (1967), pp. 91–98 (cit. on p. 13).

[10] B. D. Micco et al. "Higgs boson potential at colliders: Status and perspectives". In: *Reviews in Physics* 5 (2020), p. 100045. ISSN: 2405-4283. DOI: 10.1016/j.revip.2020.100045 (cit. on p. 13).

[11] CMS Collaboration. "Search for Higgs boson pairs decaying to WW*WW*, WW*$\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s}$= 13 TeV". In: *Journal of High Energy Physics* 95 (2023), p. 101801. ISSN: 1029-8479. DOI: 10.1007/JHEP07(2023)095 (cit. on pp. 14, 15, 33, 39, 42).

[12] ATLAS Collaboration. "Search for non-resonant Higgs boson pair production in final states with leptons, taus, and photons in pp collisions at $\sqrt{s}$= 13 TeV with the ATLAS detector". In: *Journal of High Energy Physics* 164 (2024). ISSN: 1029-8479. DOI: 10.1007/JHEP08(2024)164 (cit. on p. 15).

[13] ATLAS Collaboration. "Combination of Searches for Higgs Boson Pair Production in pp Collisions at $\sqrt{s}$= 13 TeV with the ATLAS Detector". In: *Physics Review Letters* 133 (2024), p. 101801. DOI: https://doi.org/10.1103/PhysRevLett.133.101801 (cit. on p. 15).

[14] CMS Collaboration. *Combination of searches for nonresonant Higgs boson pair production in proton-proton collisions at $\sqrt{s}$= 13 TeV*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2917252 (cit. on p. 15).

[15] CMS and ATLAS Collaborations. *Combination of ATLAS and CMS searches for Higgs boson pair production at 13 TeV*. Tech. rep. Geneva: CERN, 2025. URL: https://cds.cern.ch/record/2947521 (cit. on p. 15).

[16] CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004 (cit. on p. 17).

[17] CMS Collaboration. *The CMS magnet project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. DOI: 10.17181/CERN.6ZU0.V4T9 (cit. on p. 17).

[18] T. Sakuma and T. McCauley. "Detector and Event Visualization with SketchUp at the CMS Experiment". In: *Journal of Physics: Conference Series* 513.2 (June 2014), p. 022032. DOI: 10.1088/1742-6596/513/2/022032 (cit. on p. 17).

[19] V. Karimäki et al. *The CMS tracker system project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/368412 (cit. on p. 17).

[20] CMS Collaboration. *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375 (cit. on p. 17).

[21] CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/357153 (cit. on p. 17).

[22] J. G. Layter. *The CMS muon project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814 (cit. on p. 18).

[23] S. Cittolin, A. Rácz, and P. Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical design report. CMS. Geneva: CERN, 2002. URL: https://cds.cern.ch/record/578006 (cit. on p. 18).

[24] I. Neutelings. *Izaak Neutelings' contribution to tikz*. https://tikz.net/author/izaak/. Accessed: 2025-06-19 (cit. on p. 18).

[25] CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.QZ28.FLHW (cit. on p. 19).

[26] CMS Collaboration. *The Phase-2 Upgrade of the CMS Muon Detectors*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.5T9S.VPMI (cit. on p. 19).

[27] CMS Collaboration. *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283187 (cit. on p. 19).

[28] CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.IV8M.1JY2 (cit. on p. 19).

[29] CMS Collaboration. *The Phase-2 Upgrade of the CMS Level-1 Trigger*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2714892 (cit. on p. 19).

[30] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: *Journal of Instrumentation* 12.10 (Sept. 2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003 (cit. on pp. 20, 21).

[31] CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$=13 TeV". In: *Journal of Instrumentation* 13.06 (June 2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015 (cit. on p. 21).

[32] CMS Collaboration. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 16.05 (May 2021), P05014. DOI: 10.1088/1748-0221/16/05/P05014 (cit. on p. 21).

[33] W. Adam et al. "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC". In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (July 2005), N9. DOI: 10.1088/0954-3899/31/9/N01 (cit. on p. 21).

[34] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-kt jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063 (cit. on p. 21).

[35] D. Bertolini et al. "Pileup per particle identification". In: *Journal of High Energy Physics* 59.10 (Oct. 2014), pp. 1029–8479. DOI: 10.1007/JHEP10(2014)059 (cit. on p. 21).

[36] CMS Collaboration. *CMS Monte Carlo production overview*. https://opendata.cern.ch/docs/cms-mc-production-overview. Accessed: 2025-07-07 (cit. on p. 22).

[37] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (July 2014), p. 79. ISSN: 1029-8479. DOI: 10.1007/JHEP07(2014)079 (cit. on pp. 21, 28).

[38] T. Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 2006.05 (May 2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026 (cit. on pp. 21, 30).

[39] P. Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". In: *Journal of High Energy Physics* 2004.11 (Dec. 2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040 (cit. on pp. 21, 28).

[40] J. Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8 (cit. on p. 22).

[41] J. Allison et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (2016), pp. 186–225. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2016.06.125 (cit. on p. 22).

[42] M. Gonzalez-Berges. "The Joint COntrols Project Framework". In: *arXiv: Instrumentation and Detectors* (2003). URL: https://api.semanticscholar.org/CorpusID:18412592 (cit. on p. 24).

[43] B. Copy et al. *Monitoring of CERN's Data Interchange Protocol (DIP) system*. Tech. rep. CERN, 2018, THPHA162. DOI: 10.18429/JACoW-ICALEPCS2017-THPHA162 (cit. on p. 25).

[44] C. Gaspar, M. Dönszelmann, and P. Charpentier. "DIM, a portable, light weight package for information publishing, data transfer and inter-process communication". In: *Computer Physics Communications* 140.1 (2001). CHEP2000, pp. 102–109. ISSN: 0010-4655. DOI: https://doi.org/10.1016/S0010-4655(01)00260-0 (cit. on p. 25).

[45] A. Sculac. "Measurement of electron identification efficiency and the search for Higgs boson pair production in the WWZZ decay channel using the CMS detector at the at the Large Hadron Collider". Theses. Institut Polytechnique de Paris; PMF, University of Zagreb, Nov. 2025. URL: https://theses.hal.science/tel-05540175 (cit. on pp. 27, 42).

[46] CMS Collaboration. "Operation and performance of the CMS silicon strip tracker with proton-proton collisions at the CERN LHC". In: *JINST* 20.08 (2025), P08027. DOI: 10.1088/1748-0221/20/08/P08027 (cit. on p. 28).

[47] CMS Collaboration. *Run 3 commissioning results of heavy-flavor jet tagging at $\sqrt{s}$ =13.6 TeV with CMS data using a modern framework for data processing*. Tech. rep. CERN, 2024. URL: https://cds.cern.ch/record/2898463 (cit. on pp. 28, 44).

[48] R. Frederix and S. Frixione. "Merging meets matching in MC@NLO". In: *Journal of High Energy Physics* 2012.12 (Dec. 2012), p. 61. ISSN: 1029-8479. DOI: 10.1007/JHEP12(2012)061 (cit. on p. 29).

[49] G. Heinrich, J. Lang, and L. Scyboz. "Erratum to: SMEFT predictions for gg → hh at full NLO QCD and truncation uncertainties". In: *Journal of High Energy Physics* 2023.10 (Oct. 2023), p. 86. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2023)086 (cit. on p. 28).

[50] S. Alioli et al. "NLO Higgs boson production via gluon fusion matched with shower in POWHEG". In: *JHEP* 04 (2009), p. 002. DOI: 10.1088/1126-6708/2009/04/002 (cit. on p. 28).

[51] P. Nason and C. Oleari. "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG". In: *JHEP* 02 (2010), p. 037. DOI: 10.1007/JHEP02(2010)037 (cit. on p. 28).

[52] T. Melia et al. "W+W-, WZ and ZZ production in the POWHEG BOX". In: *JHEP* 11 (2011), p. 078. DOI: 10.1007/JHEP11(2011)078 (cit. on p. 28).

[53] P. Artoisenet et al. "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations". In: *Journal of High Energy Physics* 2013.3 (Mar. 2013), p. 15. ISSN: 1029-8479. DOI: 10.1007/JHEP03(2013)015 (cit. on p. 30).

[54] A. V. Gritsan et al. "New features in the JHU generator framework: constraining Higgs boson properties from on-shell and off-shell production". In: *Phys. Rev. D* 102.5 (2020), p. 056022. DOI: 10.1103/PhysRevD.102.056022 (cit. on p. 30).

[55] K. Hamilton, P. Nason, and G. Zanderighi. "MINLO: multi-scale improved NLO". In: *Journal of High Energy Physics* 2012.10 (Sept. 2012), p. 155. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2012)155 (cit. on p. 30).

[56] A. M. Sirunyan et al. "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements". In: *Eur. Phys. J. C* 80 (2020), p. 4. DOI: 10.1140/epjc/s10052-019-7499-4 (cit. on p. 30).

[57] R. D. Ball et al. "Parton distributions from high-precision collider data". In: *The European Physical Journal C* 77.10 (Oct. 2017), p. 663. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5199-5 (cit. on p. 30).

[58] CMS GEN XSEC TASK FORCE. *Cross Section DB Tool*. https://xsecdb-xsdb-official.app.cern.ch/xsdb/. Accessed: 2025-02-18 (cit. on p. 30).

[59] J. Baglio et al. "$gg \to HH$: Combined Uncertainties". In: *Phys. Rev. D* 103 (2021). 8 pages, 2 figures, minor extensions of text, matches published version, p. 056002. DOI: 10.1103/PhysRevD.103.056002 (cit. on p. 30).

[60] B. Mellado Garcia et al. "CERN Report 4: Part I Standard Model Predictions". In: (2016). URL: https://cds.cern.ch/record/2150771 (cit. on p. 30).

[61] Common ATLAS-CMS effort. *NNLO+NNLL top-quark-pair cross sections*. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO. Accessed: 2025-02-17 (cit. on p. 30).

[62] CMS Collaboration. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{S}$=13 TeV". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10017. DOI: 10.1088/1748-0221/15/10/P10017 (cit. on p. 30).

[63] CMS Collaboration. *Low-p_T Electron ID scale factors from CMS in proton-proton collisions at $\sqrt{s}$ =13 TeV using J/ψ events*. Tech. rep. CERN, 2023. URL: https://cds.cern.ch/record/2879279 (cit. on p. 30).

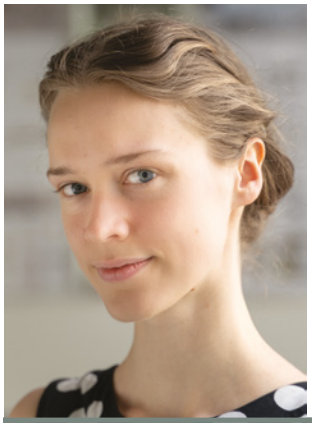
[64] E. Bols et al. "Jet flavour classification using DeepJet". In: *Journal of Instrumentation* 15.12 (Dec. 2020), P12012. DOI: 10.1088/1748-0221/15/12/P12012 (cit. on p. 32).

[65] T. Chen and C. Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN: 9781450342322. DOI: 10.1145/2939672.2939785 (cit. on p. 33).

[66] S. M. Lundberg and S.-I. Lee. "A unified approach to interpreting model predictions". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17. Long Beach, California, USA: Curran Associates Inc., 2017, pp. 4768–4777. ISBN: 9781510860964 (cit. on p. 33).

[67] A. L. Read. "Presentation of search results: the CLs technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313 (cit. on p. 37).

[68] CMS Collaboration. "The CMS Statistical Analysis and Combination Tool: Combine". In: *Computing and Software for Big Science* 8.1 (Nov. 2024), p. 19. ISSN: 2510-2044. DOI: 10.1007/s41781-024-00121-4 (cit. on p. 37).



Antra Gaile was born in 1997. She obtained a Bachelor's degree (2019) and a Master's degree (2021) in Physics from the University of Latvia. She has worked at the MHD Technology Laboratory at the Institute of Physics of the University of Latvia and then at the Institute of Particle Physics and Accelerator Technologies of Riga Technical University. She developed an interest in particle physics during the CERN summer student programme in 2019 and during her two PhD years at CERN.