Search for resonant Higgs boson pair production in the bb$\tau\tau$ decay channel and developments in the reconstruction of High Granularity Calorimeter trigger primitives with the CMS detector at the LHC

Bruno Afonso Fontana Santos Alves

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2024IPPAX098

Thèse de doctorat

Search for resonant Higgs boson pair production in the bbττ decay channel and developments in the reconstruction of High Granularity Calorimeter trigger primitives with the CMS detector at the LHC

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à École polytechnique

École doctorale n°626 Institut Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat : Physique des particules

Thèse présentée et soutenue à Palaiseau, le 12 novembre 2024, par

BRUNO AFONSO FONTANA SANTOS ALVES

Composition du Jury :

Luca Malgeri Directeur de recherche, CERN (Genève)	Président
Aleandro Nisati Directeur de recherche, INFN (Roma)	Rapporteur
Rebeca Gonzalez Suarez Professeure, Université de Uppsala (Uppsala)	Rapporteure
Giovanni Marchiori Directeur de recherche, APC (Paris)	Examinateur
Ilaria Brivio Professeure, INFN (Bologne)	Examinatrice
Andrew Gilbert Professeur, LLR (Palaiseau)	Examinateur
Roberto Salerno Directeur de recherche, LLR (Palaiseau)	Directeur de thèse

Para a Ana, minha casa

Tú, que me lees, ¿estás seguro de entender mi lenguaje?

La biblioteca de Babel

Summary

We perform a search for the resonant production of a pair of Higgs bosons (HH), decaying into a pair of b quarks and a pair of τ leptons, with the CMS experiment at the CERN Large Hadron Collider (LHC). The analysis exploits proton-proton collisions at a center-of-mass energy of 13 TeV, for a total of 138 fb^{-1} collected during the 2016, 2017 and 2018 data-taking years. The gluon-fusion production mode is considered, together with spin-0 and spin-2 hypotheses. This resonant process is strongly motivated by a series of theories able to address current shortcomings of the Standard Model (SM). The decay channel is instead known for its experimental benefits, namely a relatively clean final state signature, balanced by a moderate branching fraction of 7.3%. Additionally, the results of a similar search have been recently reported by the ATLAS Collaboration, where a small tension with the SM was recorded at a resonance mass of 1 TeV. The physics analysis here performed thus aims at confirming or rejecting such an excess. Expected upper limits at a 95% confidence level are set on the production of New Physics signatures, showcasing a compelling improvement over past CMS and ATLAS results. Furthermore, this work tackles a major simplification exploited by resonant searches in High Energy Physics, namely the Narrow Width Approximation (NWA), which assumes that the width of new resonances is negligible when compared to the detector's resolution, ignoring potential interference effects. We show that the current sensitivity level of double Higgs boson analyses is such as to put into question the correctness of the NWA, hinting at the necessity of altogether avoiding such approximation in future HH analyses.

This work is also concerned with sensitivity improvements from a detector perspective. The upcoming High-Luminosity LHC (HL-LHC) will bring large numbers of collisions per proton bunch crossing and extremely high radiation levels, which can only be sustained by a very significant detector upgrade programme within CMS. One of the upgraded sections will be the endcaps, where the novel High Granularity Calorimeter (HGCAL) will be installed. The HGCAL provides ample opportunities for studies and optimizations, and will certainly become a cornerstone of the upcoming CMS HL-LHC phase, providing high spatial and timing resolutions to improve the online and offline reconstruction of physics data. Central to the HGCAL will be the CMS trigger system, which will have to withstand the large rates expected from the HL-LHC. We have specifically developed new algorithms to enable the robust reconstruction of Trigger Primitives, the building blocks of the first level of the online trigger system in CMS. The algorithms include techniques to mitigate the wrongful creation of several energy clusters from single particles, and the computation of calorimetric quantities within a modified coordinate system. These developments are part of a reconstruction framework, implemented from scratch, which also provides a simplified version of HGCAL's geometry. Future efforts will benefit from such tools.

Résumé

Nous presentons une analyse de la production résonante de paires de bosons de Higgs (HH), se désintégrant en quarks b et en leptons τ, avec l'expérience CMS au Grand Collisionneur de Hadrons (LHC) du CERN. L'analyse exploite les $138\,\mathrm{fb}^{-1}$ de données de collisions proton-proton collectés entre 2016 et 2018 à une énergie au centre de masse de 13 TeV. Ce processus résonant est fortement motivé par un grand nombre de théories capables de répondre aux lacunes actuelles du Modèle Standard (SM). Le canal de désintégration étudié présente plusieurs avantages expérimentaux, à savoir une signature de l'état final relativement pure, équilibrée par un rapport d'embranchement modérée de 7.3%. Les résultats d'une étude similaire ont été récemment publiés par la collaboration ATLAS, rapportant une tension avec le SM pour une masse invariante du système HH d'environ 1 TeV. L'analyse effectuée ici vise donc à confirmer ou à rejeter un tel excès. Les limites supérieures attendues, avec un niveau de confiance de 95%, sont fixées sur la production de signatures de nouvelle physique, et représentent une amélioration considérable par rapport aux résultats antérieurs de CMS et d'ATLAS. Par ailleurs, ces travaux s'attaquent à une simplification majeure exploitée par les recherches résonantes en physique des hautes énergies, à savoir l'Approximation de Faible Largeur (NWA), qui suppose que la largeur des nouvelles résonances est négligeable par rapport à la résolution du détecteur, ignorant les effets d'interférence potentiels. Nous montrons que le niveau de sensibilité actuel des analyses HH est tel qu'il remet en question la validité de la NWA, ce qui laisse présager la nécessité d'éviter complètement cette approximation dans des analyses futures.

Les travaux décrits portent également sur l'amélioration de la sensibilité des détecteurs. Le futur LHC à haute luminosité (HL-LHC) apportera un grand nombre de collisions par croisement de paquets de protons et des niveaux de radiation extrêmement élevés, qui ne pourront être soutenus que par un programme très important de mise à niveau des détecteurs au sein de CMS. L'une des sections modernisées sera celle des bouchons, où le nouveau calorimètre à haute granularité (HGCAL) sera installé. Le HGCAL offre de nombreuses possibilités d'études et d'optimisations, et deviendra certainement une pierre angulaire de la prochaine phase HL-LHC de CMS, en fournissant des résolutions spatiales et temporelles élevées pour améliorer la reconstruction en ligne et hors ligne des données de physique. Le système de déclenchement de CMS, qui devra supporter les flux de données importants attendus du HL-LHC, sera critique pour le HGCAL. Nous avons spécifiquement développé de nouveaux algorithmes pour permettre la reconstruction robuste des primitives de déclenchement, les éléments constitutifs du premier niveau du système de déclenchement en ligne de CMS. Ces algorithmes comprennent des techniques permettant d'atténuer la création erronée de plusieurs groupes d'énergie à partir de particules individuelles, et le calcul de quantités calorimétriques dans un système de coordonnées modifié. Ces développements font partie d'outils de reconstruction, mis en œuvre à partir de zéro, qui fournissent également une version simplifiée de la géométrie de HGCAL. Les efforts futurs bénéficieront de ces outils.

Acknowledgments

I start by thanking Luca, Aleandro, Rebeca, Giovanni, Ilaria and Andrew for taking part in my PhD jury committee, briefly setting their many responsibilities aside to come to the École and evaluate my work. I would especially like to acknowledge Rebeca and Aleandro for agreeing to read my Thesis and write an evaluation report, without which I would not have been able to obtain my PhD. On top, the discussion with Aleandro at CERN on the month that preceded the defense was very enjoyable and useful, improving the correctness and clarity of the manuscript.

Thanks to the lab's structure and funding, and with the support of the lab's secretariat personnel (special thanks to Alimata, Laetitia and Marie-Thérèse), I was able to attend some conferences at various locations, to take part in CMS shifts and to spend three months at CERN, where I could get closer to my analysis team. These events surely made me feel more comfortable in such environments, and will surely help in the future, regardless of what I end up doing. It was also thanks to the active environment in the lab that I could join multiple "Fête de la Science" events, which were really rewarding, and in the same vein to take part in some occasional guided lab tours to either the general public or young students. I should also mention the very positive experience that teaching a course to undergraduates was. I had contact with some of the best and worst things a teacher has to endure, and it for sure enriched my perspective on education. I occasionally had students that seemed to know more than myself about the subject being taught, in what really was a humbling experience.

I wish to express my gratitude to the senior people at the lab, which in one way or another provided support when needed. In particular, I would like to thank Roberto and Yves for expressing interest in my coming to the lab, and for ensuring a place would be there for me even if my application to the PhD programme turned out to be unsuccessful. I also thank Roberto for the informal chats we had over the past three years, especially at CERN, where I could get a glimpse of how easy it was for him to discuss any HEP-related topic (except for the "tricky" ones), and where he could always add some dynamic viewing angle to an otherwise static concept. Additionally, it was thanks to him that I got the completely unexpected "Bourse Monge", which provided an enormous help. I would also like to thank Florian which, despite not being my supervisor, provided the ground for some pleasant Physics-related discussions, supported my conference travels, and had the time to discuss future career options, providing a much needed confidence boost. Together with Florian and Roberto, Adinda was also extremely helpful during my Thesis rehearsal and ensuing discussion, and helped improving the final defense presentation. I would finally like to thank Jean-Baptiste, for all the guidance and help with the project on HGCAL, and for having an infinite patience with all my questions. He carefully read and commented on all my detector-related presentations, as well as the third Chapter of this Thesis.

In addition, I would like to acknowledge Pascal and Giovanni for attending my "mini-soutenance". Their kind words were important, at a time where things were rather stressful.

With the frequent comings and goings from CERN, it was really nice to find Fabian there every time, and to share some nice moments. It is a tad bit disheartening that that period has come to an end, but I have little doubt that what comes next in his life will turn out great, and I will definitely pay a visit once in a while.

In the three months I was at CERN, I was nicely welcomed by the famous house of Prégnin, and I would like to thank in particular Andrea, Angela, Costanza, Fabio, Gabriella, Marta, and Michelangelo, where I could taste some rare delicacies and precious treats.

Concerning my work on including interference effects in resonant HH processes, I must thank Rainer and Alexandra (abbreviated Xanda, not Xana!) for the guidance and comments. The project really was a wonderful addition to my PhD.

Regarding the analysis work, which was made within an international team, I would like to thank Davide, Elvira, Johan, Konstantin, Marcel, Simona, and Valeria for providing help and insights when I was confused, which often happened. A special mention goes to Tobias, with which was a pleasure to share impressions, ideas and food with, and which always impressed me by his quick understanding and overall optimism.

A central and essential part of my PhD time was shared with all the junior people at the lab. I would like to thank the entire young community, associated to all the groups, thinking about the meals we shared, and the nice "get-together moment" we had at the last "Biennale". My time was unevenly spread across people, mostly due to different lines of work and contract timelines, but my thanks extend to everyone. I would like to specifically mention Andrés, Andrew, Andro, Axel, Bharadwaj, Florian, Giacomo, Isaac (pronounced A-sick), Lida, Martina, Matthieu, Nico, Théo, Trisha, Uttya and all the nice people from the new generation that I still did not have the time to properly know. A special word goes to some of the people I spent the most time with, which shared so many nice moments, and helped during the negative ones. Sharing the office with Geliang and Shamik during three years was the perfect setting I could have wished for. Unlike other offices, which were *so loud and unprofessional*, we kept an ambiance of pure dedication to the greatest scientific cause. I specifically enjoyed some discussions with Shamik in somewhat late hours, which provided the needed distraction from systematics and scale factors (or, god forbid, both). Speaking about distractions, it was a pleasure to meet Ana (Šćulac, detta Sculacca) and her relentless positivity. I am looking forward to spent some time in the most beautiful city in the world. I have also learned to appreciate the shy kindness of Alessandra, and the remarkable resilience on sticking to the academic path, despite some less positive outcomes. I wish her the best of luck in the next few years, hoping she will quickly obtain what she so clearly desires. As for Marco, despite much initial desperation, everything is now a rebase away from fixing, and that conveys impressive power. On top, he is the first person I witnessed going from zero to hero in just one year, and I still see much potential waiting to be unleashed. The moments shared outside work were also truly enjoyable, I hope with many yet to come. Jona has proven to be an example to follow in what refers to dedication, perseverance and ambition. I find it truly remarkable that one can dedicate so much energy to one's work, while somehow still managing to have other interests. He surely gave me some additional motivation to perform, when I would have otherwise been less focused. I also thank him and Theresa for letting me stay at their place in Geneva, in what would have otherwise been quite a depressing stay at the CERN hostel. Into what concerns Alessandro, I truly admire his ability to perfectly balance work with personal life, a goal which I would have otherwise thought to be impossible to achieve during a PhD in our field. Together with his natural

spontaneity, his easy-going manners made me very confortable during the time we spent together, and I am sure I am not alone in this. Both Jona's and Alessandro's Theses were an important source of inspiration for my own Thesis, and I am grateful to both of them for sharing all the material needed to facilitate my writing process. I am utterly impressed by Elena's mental resilience and will to succeed, despite the various setbacks that occurred along the way, most notably her Apéro serving skills. It definitely helped to have someone which could understand what I was going through, releasing the burden by just the right amount, and with which I could discuss the specifics of the analysis work, from PhD student to PhD student. It was a pitty we could not share the same office, as I would have benefited from the clarity with which she frames problems, and the determination she brings to solving them. My last words go to Louis, which managed to answer (most of) my questions while we shared the same space, and which, despite having absolutely no obligations in that regard, continued doing so after leaving the lab, serving as a real supervisor, and dedicating countless hours to my scientific insecurities. He also made sure this Thesis was, as much as possible, scientifically correct and complete, providing innumerous comments that improved it by a large margin. Discussing Science with him was and continues to be an ego-shattering experience, and imagining my analysis work without his support is an undeniable challenge. In my honest opinion, Louis embodies what any researcher should strive for, namely a selfless, true and pure passion for the subject at hand. It was an immense pleasure.

To conclude, it is always nice to remember the people that wait for us back home. I would like to mention my old friends Manuel and Vítor, which are always willing to meet and share pleasant stories whenever I go back, despite the many years that have passed since we were all together in the same place. To my parents, a strongly felt thank you, for raising me without dogma, for never forcing a specific path, for encouraging the one I am slowly crafting, and for understanding why I cannot always be there at their side. To my brother, Mauro, which is a constant source of pride, for having fought against disease and suffering like a true champion, for being an intelligent and curious person, and for coming out on top after so many setbacks. At last, to Ana, the one that erases my deepest sorrows, which soothes me when I am down, which supports me when I am shaky, which balances me when I am inordinate, and provides a never-ending spark that makes my life truly worth it.

Contents

Introduction

More than ten years have passed since the observation of the Higgs boson by the CMS and ATLAS Collaborations at the CERN Large Hadron Collider (LHC) [1–3]. What seemed the first of a series of discoveries bound to reshape the field of High Energy Physics, surprisingly turned into an impressive statement of the solid foundations upon which the Standard Model (SM) of Particle Physics is built. Despite its success, the SM cannot be the ultimate answer to all mysteries the Universe still conceals; multiple theoretical puzzles and cosmological observations show otherwise. The seed for the next big discovery might lie where the SM was completed; the Higgs boson can still hide many surprises, and represent the gateway to a new understanding of fundamental interactions. One century ago, no one thought that ordinary matter was about five times less common than a mysterious "dark" form of matter. Before the 50's, every scientist would have sworn that Nature follows a left-right symmetry. Revolutions are found where one least expects them, and the next one might be waiting just around the corner. In the meantime, as physicists, we ought to study everything at our disposal. The High Energy Physics community has access to some of the most complex machinery ever built, representing unmissable opportunities to measure every conceivable physical quantity. This Thesis proposes to pursue such goals by providing contributions on detector and analysis work, to be added to the stack of human knowledge which will inevitably lift the veil on Nature's next revelation.

Are there particles decaying to a pair of Higgs bosons? This apparently simple question has been the source of innumerous studies in the last few years, of ever increasing complexity. The measurement of such processes would provide a clear evidence for physics Beyond the Standard Model (BSM). With a measurement of the resonant production of Higgs boson pairs (HH), this Thesis aims at providing the most precise answer ever obtained to that question, using the $\mathrm{bb}\tau\tau$ decay channel, and exploiting CMS Run 2 proton-proton data obtained at a center-of-massy energy of 13 TeV. Results drawn from searches for New Physics (NP) always benefit the scientific community. Even if no excess is found, the parameter phase-space allowing NP is reduced, forcing theoretical physicists to come up with innovative solutions to the still unresolved inconsistencies. The results here expounded should also take advantage from a statistical combination with similar analyses in other decay channels, to provide the most performant results possible, demonstrating the necessity of joining efforts to push fundamental laws to the limit. This work also serves as a starting point for future endeavours, much like past results provided the inspiration for the manuscript you are now reading. The analysis work is planned to be publicly presented at the "59th Rencontres de Moriond 2025" conference. During the three years of my PhD, I have given several presentations on Higgs physics at multiple CMS internal meetings, in a LHC-wide workshop [4], and in the "31st International Workshop on Deep Inelastic Scattering", leading to a publication in *Proceedings of Science* [5]. An unrelated talk on "Techniques for SMEFT fit" was also given at the "QCD@LHC 2022" conference [6].

The work in the previous paragraph adopts a widespread simplification, namely the usage of the Narrow Width Approximation (NWA), which states that the widths of the probed BSM resonances are much smaller than the resolution of our detectors, thus having no effect on measured quantities. In practice, the signal models against which the data is compared are defined with a negligible width. Recent sensitivity improvements have questioned whether the veracity of such claim still holds. In this Thesis, I exploit the simplest extension of the SM, the singlet model, to probe the relevant phase-space, compare it with current state-of-the-art HH upper limits, and conclude on whether the NWA is still valid, or if the interference between the resonances and the SM background can no longer be ignored. For completeness, singlet model projections at higher luminosity values are produced. This represents the first time the impact of the NWA is assessed at CMS. The work was included as part of CMS review paper title "Searches for Higgs boson production through decays of heavy resonances", which has been accepted for publication at *Physics Reports* [7].

One should never forget that all the interesting studies above are fully dependent on the workings of accelerators and detectors, made possible by the joint effort of thousands of physicists and engineers. From the point of view of someone doing a physics analysis, I think it is quite humbling to realize that a poor analysis leads to a weak result, but a poor detector leads to no result at all. As experimental apparatuses get more and more complex, contributions must come from all sides to establish a sound hardware and software infrastructure, ultimately leading to an improved reconstruction. I am naturally considering the future High Luminosity LHC (HL-LHC), an accelerator meant to push the boundaries of what the LHC has so far provided, leading to a plethora of possible improvements of the measurements performed by related experiments, including Higgs physics. The CMS detector will be forced to undergo a major upgrade, at all levels, to preserve its capabilities, and possibly extend them, while facing the harshest particle environment ever sustained by any collider experiment. Particularly interesting in the context of this work is the High Granularity Calorimeter (HGCAL), one of the most ambitious projects to have ever been undertaken at CERN. It consists of an endcap calorimeter with unprecedented transversal and longitudinal granularities, designed with a mixed silicon and scintillator technology, and capable of tolerating the expected extreme radiation levels. I have developed algorithms within a new reconstruction framework, thus contributing to the reconstruction of Trigger Primitives (TPs) in HGCAL. TPs constitute a foundational element of the first level of the online CMS trigger, for it to make an informed decision on which events to keep, among the colossal data rates provided by the HL-LHC. The work consists of providing a convenient TP reconstruction framework to be used by several people, and the development of mitigation techniques to address inefficiencies in the clustering stage. The framework also includes a simplified version of HGCAL's geometry. The work has been presented at the "26th International Conference on Computing in High Energy and Nuclear Physics (CHEP)" conference, leading to a publication at *EPJ Web of Conferences* [8]. The developed software further lead to a citable presentation at the "PyHEP 2023 Workshop" [9].

Chapter 1 begins with a portray of the theoretical, phenomenological and experimental landscape surrounding Higgs boson physics. The underlying SM fundamentals are provided, together with its noticeable shortcomings, unambiguously stating the clear motivations to pursue studies on the Higgs boson, and especially on processes involving Higgs boson pairs. I also contextualize the resonant $bb\tau\tau$ channel within the profusion of both single and double Higgs boson phenomena, which have been measured since the

discovery of the Higgs boson, 12 years ago. These discoveries would not have been possible without the CMS detector at the LHC, presented in Chapter 2. Besides the detector itself, I provide a succinct description of its trigger system and offline reconstruction, giving special emphasis to the Particle Flow algorithm, so that every physics object used throughout the Thesis is well defined. The processing of data and Monte Carlo samples within the CMS software is also covered. Chapter 3 contains my first original contribution, connected to the HL-LHC and the new HGCAL calorimeter that CMS will exploit. There, contributions to the reconstruction of Trigger Primitives in the L1 trigger of CMS are exhaustively described. The analysis of CMS datasets is then depicted, specifically the search for resonant HH production in the bb$\tau\tau$ channel. I detail the analysis' triggering, selection and correction steps in Chapter 4, and proceed to obtain the results in Chapter 5, where a description of the final discriminant, systematic uncertainties and statistical methodology is included, and where future prospects are delineated. The Thesis concludes in Chapter 6, where the study on interference effects is presented. There, we can deduce the unfolding of a paradigm shift concerning Higgs boson resonant searches.

1 Higgs Boson Pair Production in the Standard Model and Beyond

The SM of Particle Physics is undoubtedly one of the most successful theoretical frameworks in all of Science. Despite colossal experimental efforts, including the most sophisticated particle accelerators and detectors ever built, together with the involvement of thousands of scientists and engineers, most of the hundreds of petabytes of data so far produced match SM predictions, in an inspiring demonstration of scientific accomplishment. And yet, we know the SM to be incomplete, from the surprising behavior of rotation curves of galaxies, to a mysterious form of energy permeating the entire Universe, or the apparent absence of a charge-mirrored world.

Soon after the Higgs boson was discovered by the A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS) Collaborations at the CERN LHC [1–3], studies of its properties followed, in what seemed a natural and quick way to observe New Physics (NP). While some of these properties still remain under-constrained by the experiments, none of the highly anticipated new resonances have been discovered, such as the ones predicted by Supersymmetry (SUSY). NP could unfortunately lie at such high energies to substantiate the dreaded "particle desert" scenario: a vast region between TeV and Beyond the Standard Model (BSM) scales where no resonances exist. Future discoveries would then become impossible in the foreseeable future. Notwithstanding, such fears might prove wrong, as a single unexpected discovery would overturn the current state of affairs. After discovering the top quark [10, 11], physicists had to wait 17 years for the Higgs boson; the next momentous discovery could very well lie around the corner.

The production of Higgs boson pairs (HH) arguably represents one of the best probes to observe the unfolding of new phenomena, especially since it is still mostly unexplored. Alongside the ability to search for BSM resonances, the interaction of two Higgs bosons should enable the determination of one of the fundamental SM parameters still left to be measured: the Higgs boson self-coupling. This quantity is profoundly connected to the internals of electroweak symmetry breaking (EWSB) and the stability of the electroweak (EW) vacuum, and its measurement could lead to a revolution in the not-so-distant future.

This Chapter introduces the SM, giving special emphasis to the Higgs particle. The first Section sets the tone, by detailing the underlying mechanisms of the SM, including a description of the Higgs mechanism. Next, we devote our attention to Higgs phenomenology, most notably HH production and decay. We then present the reasons that lead us to believe something lies beyond the current theory, specifically discussing resonant and nonresonant HH analyses, just like the one presented in this Thesis. We conclude by summarizing the experimental status of HH physics, highlighting both direct and indirect searches, and a combination of the two.

1.1 The Standard Model of Particle Physics

Little did Thomson know, in 1897, that the observed deflection of a stream of charged particles by a magnet would be the seed for the gradual development of the theory explaining the interactions of all known fundamental particles. Following the discovery of the electron, Rutherford famously demonstrated the structure of the atom, and soon after Bohr managed to develop a mathematical framework predicting the spectrum of Hydrogen. Discoveries followed at great pace: Chadwick's neutrons, Planck's and Einstein's quanta. Theoretical and experimental developments continued for half a century, facing frequent obstacles, but eventually reaching a surprisingly effective framework, capable of describing wildly diverse natural phenomena: from the core of an atom to interstellar space. This overarching theory is known as the SM of Particle Physics, which serves as the foundation of modern high-energy physics (HEP).

This Section provides a standard description of the SM, covering its major aspects, focusing on the items most relevant for later sections. The building blocks of the SM are listed, hopefully providing a terse but clear picture of what is currently known at the most fundamental level. It immediately follows a description of the far-reaching relationship between abstract symmetries and measurable interactions, and its impact on the development of the SM as we know it. Inconsistencies are elegantly solved with the Higgs mechanism, discussed in Section 1.1.3; a revolutionary idea defining the modern understanding of Particle Physics. There, elementary particles acquire mass, and a new scalar is introduced, the Higgs boson, a central piece of the experimental programme of current high-energy collider experiments. We finalize this first Section by presenting a concise account of the phenomenology and experimental status of single Higgs physics: the quantities worth measuring, and the results such measurements have so far yielded.

1.1.1 The Building Blocks of the SM

Out of the four known fundamental forces, three are described by the SM: the electromagnetic (EM), the weak and the strong forces. Electromagnetism is responsible for most phenomena we can readily perceive, such as friction, elasticity or drag, the color of the sky and how it changes during sunsets, from impressive lightning storms to high-speed Maglev trains. The weak force is related to particle decays, and is fundamental for the nuclear fusion reactions that power the Sun, while the strong force is what keeps atoms and quarks together. Gravity is not integrated within the SM, but it is by far the weakest of all forces, and can be safely ignored at the subatomic scales here discussed.

The three forces are represented by a $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ local gauge symmetry, which is obeyed by the full SM Lagrangian. The $\mathrm{SU(3)_C}$ group is generated by the mediators of the strong force, the *gluons*, described by Quantum Chromodynamics (QCD). The $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry, instead, encompasses both the weak and EM forces, respectively mediated by the $\mathrm{W^+}$, $\mathrm{W^-}$ and Z particles, and by the photon γ. Particles are categorized according to their spin, being called *bosons* when possessing integer spin, or *fermions* when the spin is half-integer. All mediators are bosons, and 12 exist in the SM: 8 gluons plus the four mediators just mentioned. In contrast, fermions, which comprise all ordinary matter and are subdivided into *quarks* and *leptons*, exchange the mediators. In total, 24 fermionic fields are defined in the SM, where half correspond to particles identical to the ones described but with opposite internal quantum numbers, i.e. an-

tiparticles. Fermions are split across three generations, where each comprises either two leptons with electric charges -1 and 0, or two quarks with charges $2/3$ and $-1/3$:

$$\text{leptons: } \begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}, \quad \text{quarks: } \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}. \tag{1.1}$$

Each generation is arranged in a doublet, for reasons made clear when discussing the Weinberg-Salam model, in Section 1.1.2. The e, μ and τ symbols for charged leptons refer to, respectively, electrons, muons and tau leptons. They are identical except for their masses, which span multiple orders of magnitude: $m_e \approx 511\,\text{keV}$ [12], $m_\mu \approx 105.7\,\text{MeV}$ [12] and $m_\tau \approx 1.8\,\text{GeV}$ [13]. Contrary to electrons and muons, tau leptons decay too fast for observation at a collider experiment; we thus measure only their decay products. The symbols $\nu_{e/\mu/\tau}$ refer instead to neutrinos and their three flavors, associated to the three leptons we just introduced. Neutrinos do not have an electric or color charge, and thus interact solely via the weak force. Their tiny interaction cross sections make them extremely hard to detect. The most precise measurements of their near-zero masses set them at values below 1 eV, with their sum constrained to $\lesssim 0.15\,\text{eV}$ [13]. It is however important to note that the three masses do not correspond to the three flavors, as they mix via the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [14], leading to neutrino oscillations [15]. All leptons interact via the EM and weak forces, but not strongly. On the contrary, quarks "feel" the three forces in the SM, and are the internal constituents of *hadrons*. Groups of two quarks are called *mesons* and groups of three quarks are called *baryons*. Baryons include protons and neutrons, which are made of up and down quarks in different proportions. Hadrons with even more quarks have been observed, but are all unstable. Quarks exist in 6 flavors, as represented in Eq. (1.1): up, down, charm, strange, top and bottom, all with different masses. Due to the QCD property of confinement, quarks cannot exist independently, and must be associated to other quarks, forming color-neutral bound states. This is the basis of *hadronization*, described in Section 2.4 when discussing jets, a type of object measured in collider experiments. Only the top quark cannot be part of hadrons, as it decays faster ($\tau_t \approx 0.5 \times 10^{-24}\,\text{s}$ [16]) than the time required to form a strong bound state. Finally, thanks to a mechanism detailed in Section 1.1.3, the scalar Higgs boson completes the collection of SM particles, representing a crucial ingredient for the Thesis here discussed. It is the only scalar particle in the SM, and is responsible for the masses of all particles. The pieces of the SM are illustrated in Fig. 1.1, where antiparticles are not represented for simplicity.

1.1.2 From Symmetries to Interactions

There has always been a deep connection between the symmetries in the Universe and the laws that dictate its behavior. Geometrical patterns had been observed from ancient times; Greeks for instance believed the motion of objects had to follow symmetrical trajectories, despite frequent discrepancies with observed data. Newton instead transferred the idea of symmetries to the equations describing the motion, rather than focusing on the solutions of those same equations. In 1915, the synergy between symmetries and Nature became firmly established in Noether's famous theorem [18], stating that for each continuous symmetry a conservation law ensues or that, conversely, every conservation law reflects an underlying symmetry. Energy conservation follows from the invariance of the laws of Physics under translations in time, the invariance under space translations

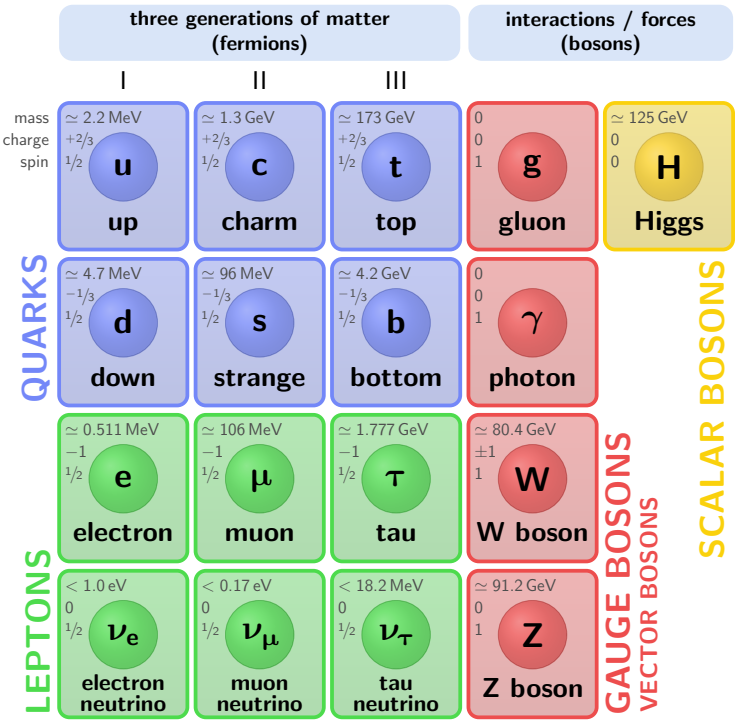

Figure 1.1: Schematic illustration of the SM particle content, organized according to the particle's type and generation, where blue stands for quarks, green for leptons and red for the force carriers. The Higgs boson, in yellow, is the only scalar particle in the SM. The mass, charge and spin are listed for each particle. Antiparticles are not explicitly represented for simplicity. Courtesy of Izaak Neutelings [17].

leads to momentum conservation, and so forth. In this Section, the profound ideas that cement the SM are described, covering the surprising way in which symmetries lead to all known forces and particle interactions [19, 20].

Yang-Mills theory

Interactions are dictated by symmetry principles. For example, Einstein used the invariance under general coordinate transformations to establish the general theory of relativity in 1915. Initially taken at a global level, i.e. without depending on space-time points $x^\mu \equiv x$, in 1953 Yang and Mills promoted the invariances to the *local* realm. They started from the Dirac Lagrangian for two elementary half-integer spin fields ψ:

$$\mathcal{L} = i\bar{\psi}(x)\partial\!\!\!/\psi(x)\,, \quad \partial\!\!\!/ \equiv \gamma^\mu\partial_\mu\,, \tag{1.2}$$

where ψ is a two-component column vector, or spinor, and γ^μ are the Dirac γ matrices. For simplicity, we here consider massless fields (we do not introduce the $-m\bar{\psi}\psi$ term), without loss of generality. The original Yang-Mills theory *required* the fields to be invariant under a local non-Abelian SU(2) transformation:

$$\psi(x) \to e^{-ig\boldsymbol{\sigma}.\boldsymbol{\lambda}(x)}\psi(x)\,; \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.3}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices, $\boldsymbol{\lambda}(x)$ is a vector depending on space-time, and g is a coupling constant analogous to the electric charge. Without further changes, the Lagrangian in Eq. (1.2) is not invariant under the local SU(2) transformation, since the derivative picks up an extra term:

$$\partial_\mu \psi(x) \to e^{-ig\boldsymbol{\sigma}\cdot\boldsymbol{\lambda}(x)}[\partial_\mu - ig\boldsymbol{\sigma}\,\partial_\mu \boldsymbol{\lambda}(x)]\psi(x)\,. \tag{1.4}$$

If we insist on the concept of local invariance, something has to be added to the Lagrangian in order to cancel the new term. This can be achieved by replacing the standard derivative ∂_μ by the *covariant derivative*:

$$D_\mu \equiv \partial_\mu + ig\boldsymbol{\sigma}\,\boldsymbol{A}_\mu\,, \tag{1.5}$$

where a new set of *gauge*[1] fields $\boldsymbol{A}_\mu = (A^1_\mu, A^2_\mu, A^3_\mu)$ had to be introduced. It is highly non-trivial to find the transformation rule for $\boldsymbol{A}_\mu$ that restores local SU(2) invariance to the Lagrangian. Keeping only the first-order terms for the case of small $\boldsymbol{\lambda}$, the solution is:

$$\boldsymbol{A}_\mu \to \boldsymbol{A}_\mu + \partial_\mu \boldsymbol{\lambda} - 2g(\boldsymbol{\lambda} \times \boldsymbol{A}_\mu)\,. \tag{1.6}$$

Finally, without forgetting the additional kinematic term brought by the new field:

$$\mathcal{L}_{\text{Kin}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, \quad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]\,, \tag{1.7}$$

the new, local invariant Lagrangian becomes:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}(x)\not{D}\psi(x)\,, \quad \not{D} \equiv \gamma^\mu D_\mu\,, \tag{1.8}$$

where one should notice that the new gauge fields A_μ couple to the particle fields ψ via the covariant derivative, introducing interactions that become relevant in the context of the SM. Since a term of the form $A_k A^k$ is not invariant under the gauge transformation in Eq. (1.6), the new fields must be massless to honor the symmetry in Eq. (1.8). Crucially, it is the insistence on local gauge invariance which leads to three new vector gauge fields.

The original Yang-Mills theory turned out to be of little use, on account of starting from the assumption that there are two elementary half-integer spin particles with exactly the same mass, which were never observed. However, the same idea could be generalized to many other non-Abelian symmetry groups, and ended up representing the basis of the SU(3)_C color symmetry for strong interactions and of the isospin-hypercharge $\text{SU(2)}_\text{L} \otimes \text{U(1)}_\text{Y}$ symmetry for weak interactions.

Chromodynamics

According to the SM, each quark flavour comes in three colors: red, blue and green. We can express its free Lagrangian exactly as in Eq. (1.2), but where now the fields are a three-component column vector $\bar{\psi} \equiv (\bar{\psi}_r\, \bar{\psi}_b\, \bar{\psi}_g)$ for the three colors. The Lagrangian naturally exhibits a global SU(3) symmetry, which we again require to be local. Under SU(3)_C, the fields transform as follows (compare to Eq. (1.3)):

[1] The term *gauge* refers to the regulation of redundant degrees of freedom. It can be thought of as a historical misnomer for the word *phase* [19].

$$\psi(x) \to e^{-ig\boldsymbol{\phi}.\boldsymbol{\theta}}\psi(x) \tag{1.9}$$

where $\boldsymbol{\phi}$ represents the eight generators of $\mathrm{SU(3)_C}$, also know as Gell-Mann matrices, and g symbolizes the strength of the interaction. As in Eq. (1.4), a covariant derivative is introduced to satisfy the local invariance request, which this time will include 8 gauge fields instead of three. The gauge fields also transform in a more complex way, due to the larger number of generators:

$$\boldsymbol{A}_\mu \to \boldsymbol{A}_\mu + \partial_\mu\boldsymbol{\phi} - 2g(\boldsymbol{\phi}\times\boldsymbol{A}_\mu)\,, \quad (\boldsymbol{\phi}\times\boldsymbol{A}_\mu)^i = \sum_{j,k=1}^{8} f^{ijk}\phi^j A_\mu^k\,, \tag{1.10}$$

where the f^{ijk} symbols denote the structure constants of $\mathrm{SU(3)_C}$, defined based on the following commutation rules:

$$\left[\frac{\phi^i}{2}, \frac{\phi^j}{2}\right] = if^{ijk}\frac{\phi^k}{2}\,. \tag{1.11}$$

Adding the gluon kinetic term, we obtain the complete QCD Lagrangian for massless fields:

$$\mathcal{L}_{\mathrm{QCD}} = i\bar{\psi}\partial\!\!\!/\psi - \frac{1}{16\pi}\boldsymbol{F}^{\mu\nu}\boldsymbol{F}_{\mu\nu} - g\bar{\psi}\gamma^\mu\boldsymbol{\lambda}\psi\boldsymbol{A}_\mu\,, \tag{1.12}$$

where:

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\mu A^i_\nu + 2gf^{ijk}A^j_\mu A^k_\nu\,. \tag{1.13}$$

The Lagrangian is invariant under local $\mathrm{SU(3)_C}$, describing three fields interacting with eight massless vector fields, the gluons. Contrary to the Yang-Mills theory discussed before, scientists believe Eq. (1.12) actually reproduces reality, describing the strong interaction. The strength of the interaction is usually redefined as $\alpha_\mathrm{S} = g^2/4\pi$, the strong coupling. Notice that, once more, the gluons must be massless for the local invariance to hold. The second term in Eq. (1.12) will lead to multiple combinations in A^i_μ, originating cubic and quartic interactions between gluons.

The Weinberg-Salam Model

In 1956, and against most expectations, Wu observed that parity was maximally violated in beta decays of Co^{60} [21]. This motivated the exploration of chiral theories, where left and right handed components behave differently under applied symmetries. Just as Quantum Electrodynamics (QED) had be obtained by imposing a U(1) EM local gauge invariance, it was postulated that the EM and weak forces could be united into a single force by imposing invariance under a $\mathrm{SU(2)_L}\otimes\mathrm{U(1)_Y}$ local transformation. Under the latter, fermion fields of left chirality are represented as doublets, as shown in Eq. (1.1), while right chirality fermions are singlets. Dirac spinors can be decomposed into left- and right-handed chiral components L and R:

$$\Psi_\mathrm{L} = \frac{1}{2}(1+\gamma^5)\begin{pmatrix}\psi\\ \psi'\end{pmatrix} = \begin{pmatrix}\psi_\mathrm{L}\\ \psi'_\mathrm{L}\end{pmatrix}, \quad \psi_\mathrm{R} = \frac{1}{2}(1-\gamma^5)\psi\,, \quad \psi'_\mathrm{R} = \frac{1}{2}(1-\gamma^5)\psi'\,, \tag{1.14}$$

with the γ^5 matrix defined as:

$$\begin{pmatrix} 0_{2\times2} & \mathbb{I}_{2\times2} \\ \mathbb{I}_{2\times2} & 0_{2\times2} \end{pmatrix} . \tag{1.15}$$

With the notation above, the Lagrangian can be written as:

$$\mathcal{L} = \bar{\Psi}_{\mathrm{L}}(i\not{D})\Psi_{\mathrm{L}} + i\bar{\psi}_{\mathrm{R}}\not{D}\psi_{\mathrm{R}} + i\bar{\psi}'_{\mathrm{R}}\not{D}\psi'_{\mathrm{R}} , \tag{1.16}$$

where it becomes clear the generators of $\mathrm{SU}(2)_{\mathrm{L}}$ are zero for the right singlets, and thus only the left chirality fields interact with the W gauge bosons. This corresponds to Wu's observation mentioned above, that Nature does not treat left- and right-handed particles equally. If we follow the Yang-Mills approach as in Eq. (1.4), but for the EW $\mathrm{SU}(2)_{\mathrm{L}} \otimes \mathrm{U}(1)_{\mathrm{Y}}$ transformation instead, the covariant derivative becomes:

$$D_\mu \equiv \partial_\mu + ig\boldsymbol{TW}_\mu - ig'\frac{\mathcal{Y}}{2}B_\mu \tag{1.17}$$

with $T_i \equiv \sigma_i/2$ representing SU(2) generators given by Eq. (1.3) for the left fields, and $\mathcal{Y}$ standing for the generator of the $\mathrm{U}(1)_{\mathrm{Y}}$, the hypercharge, which is divided by two for convention. The couplings are represented by g and g'. Rewriting Eq. (1.18) we obtain:

$$\mathcal{L}_{\mathrm{EW}} = \mathcal{L}_{\mathrm{kin}} + \mathcal{L}_{\mathrm{CC}} + \mathcal{L}_{\mathrm{NC}} , \tag{1.18}$$

where the first term is purely kinematical, and interactions are encoded in the neutral and charged current terms, NC and CC, respectively. The kinematic term describes the propagation of fermions and gauge bosons through space-time, and is given by:

$$\mathcal{L}_{\mathrm{kin}} = \bar{\Psi}_{\mathrm{L}}(i\not{\partial})\Psi_{\mathrm{L}} + \bar{\psi}_{\mathrm{R}}(i\not{\partial})\psi_{\mathrm{R}} + \bar{\psi}'_{\mathrm{R}}(i\not{\partial})\psi'_{\mathrm{R}} - \frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} , \tag{1.19}$$

where the field strength tensors are given by:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + gf^{ijk}W_{\mu j}W_{\nu k} , \quad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu . \tag{1.20}$$

By linearly combining the two gauge bosons as:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp W^2_\mu\right) , \tag{1.21}$$

we can create a term representing charged currents, forming two bosons corresponding to the observed SM positive and negative W fields:

$$\begin{aligned} \mathcal{L}_{\mathrm{CC}} &= \frac{g}{\sqrt{2}}\left[W^+_\mu \bar{\Psi}_{\mathrm{L}}\gamma^\mu\sigma^+\Psi_{\mathrm{L}} + W^-_\mu \bar{\Psi}_{\mathrm{L}}\gamma^\mu\sigma^-\Psi_{\mathrm{L}}\right] \\ &= \frac{g}{\sqrt{2}}\left[W^+\left(\bar{\psi}_{\mathrm{L}}\gamma^\mu\psi'_{\mathrm{L}}\right) + W^-\left(\bar{\psi}'_{\mathrm{L}}\gamma^\mu\psi_{\mathrm{L}}\right)\right] , \end{aligned} \tag{1.22}$$

Finally, the term referring to neutral interactions is given by:

$$\mathcal{L}_{\mathrm{NC}} = \frac{g}{\sqrt{2}}W^3_\mu\left[\bar{\psi}_{\mathrm{L}}\gamma^\mu\psi_{\mathrm{L}} - \bar{\psi}'_{\mathrm{L}}\gamma^\mu\psi'_{\mathrm{L}}\right] + \frac{g'}{\sqrt{2}}\mathcal{Y}B_\mu\bar{\Psi}\gamma^\mu\Psi . \tag{1.23}$$

If a rotation with the Weinberg angle θ_W is performed in the space of the two neutral gauge fields W_μ^3 and B_μ, one can recover the two gauge bosons observed in nature:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} . \tag{1.24}$$

where A_μ refers to the (massless) photon and Z_μ to the Z boson. Indeed, by replacing the new fields back in Eq. (1.23), we get:

$$\mathcal{L}_{NC} = \overline{\Psi}\gamma^\mu \left(g \sin\theta_W I_3 + g' \cos\theta_W \frac{\mathcal{Y}}{2} \right) \Psi A_\mu + \overline{\Psi}\gamma^\mu \left(g \cos\theta_W I_3 - g' \sin\theta_W \frac{\mathcal{Y}}{2} \right) \Psi Z_\mu \tag{1.25}$$

The first term can be interpreted as the current that couples the photon field A_μ to the fermions Ψ, by setting the electric charge as:

$$g \sin\theta_W I_3 + g' \cos\theta_W \frac{\mathcal{Y}}{2} = eQ . \tag{1.26}$$

The hypercharge $\mathcal{Y}$ multiplies the coupling g', and thus leaves space from some arbitrariness in its value. With the appropriate choice of $\mathcal{Y}(\bar{\Psi}_L) = -1$, and considering the quantum numbers of the leptons in the up and down components of the doublets in Eq. (1.1), a set of equations can be obtained:

$$\begin{cases} 0 = (g/2)\sin\theta_W - (g'/2)\cos\theta_W \\ -e = -(g/2)\sin\theta_W - (g'/2)\cos\theta_W \end{cases} , \tag{1.27}$$

so that:

$$g \sin\theta_W = g' \cos\theta_W = e , \tag{1.28}$$

and the Gell-Mann–Nishijima relation becomes apparent:

$$Q = I_3 + \frac{\mathcal{Y}}{2} . \tag{1.29}$$

The remaining hypercharge values can be computed from Eq. (1.29), leading to $\mathcal{Y}(\psi_R) = 0$ and $\mathcal{Y}(\psi'_R) = -2$. The neutral current term can be finally rewritten as follows:

$$\mathcal{L}_{NC} = e\bar{\Psi}\gamma^\mu Q \Psi A_\mu + e\bar{\Psi}\gamma^\mu \frac{I_3 - Q\sin^2\theta_W}{\cos\theta_W \sin\theta_W} \Psi Z_\mu . \tag{1.30}$$

The charged current interactions couple the left spinors and are mediated by two W bosons, while the interactions arising from the neutral current are mediated by the neutral Z boson and photon. Charged currents can change the particle's flavour, and connect only left-chirality fermions. Once the last two kinematic terms in Eq. (1.19) are developed, a rich structure of gauge boson interactions is revealed. The theory thus predicts cubic and quartic gauge couplings: ZWW, γWW, ZZWW, $\gamma\gamma$WW, γZWW and WWWW.

The model presented in this Section does manage to integrate the EM and weak forces into the same framework, into what represented a resounding success. However, while photons and gluons are massless, vector bosons are known to be massive, but Dirac mass terms are not gauge invariant to the symmetries so far explored. We could add those terms anyways, but unrenormalizable divergences would appear, getting more severe at

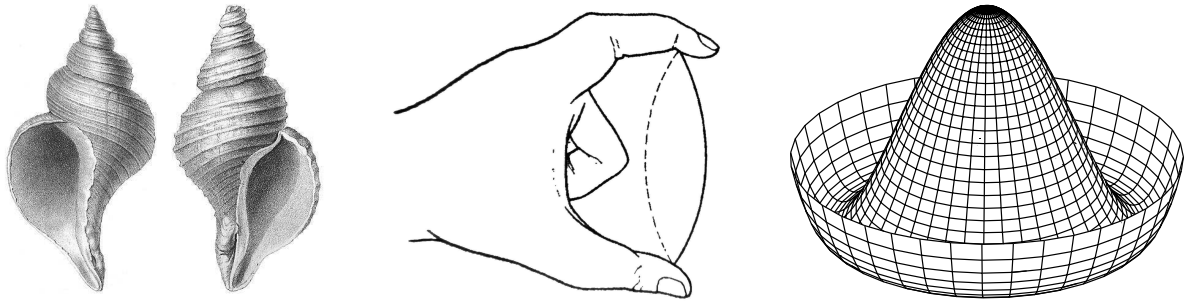

Figure 1.2: (Left) Shells of two species of sea snails, exhibiting similar characteristics with opposite spiral symmetries. (Middle) A plastic strip being compressed along its longitudinal axis conveys the idea of spontaneous symmetry breaking. Taken from [20]. (Right) Shape of the Higgs potential, also known as "mexican hat", which minima break the symmetry of the SM Lagrangian.

each additional loop order, making the theory unpredictive and ultimately meaningless. Fortunately, an alternative procedure was found.

1.1.3 The Higgs Mechanism

The Higgs mechanism revolutionized Particle Physics, and has driven HEP experiments and studies for the past ∼50 years, of which this Thesis is a good illustration. The mechanism is responsible for generating the masses of the W and Z gauge bosons, and for explaining the masses of fermions. It is deeply connected to the stability of the EW vacuum, and thus possibly the fate of our Universe, it can have a role in baryogenesis [22], and its homonymous particle might even be identified as the possible inflaton of the primordial Universe [23]. Despite its central role in modern particle physics, the Higgs mechanism is at its core based on a relatively simple phenomenon, that of symmetry breaking, which is widely present in Nature (see Fig. 1.2, left). A striking example of symmetry breaking takes place in biology [24, 25], where molecules that are central to life have just one out of two possible chiralities: the DNA and RNA are right-handed, while amino acids and proteins are all left-handed. Gravity represent yet another example, breaking 3D space symmetry, and establishing a difference between "up" and "down". We instead speak of *spontaneous* symmetry breaking when the states of a system do not possess the same symmetries of the theory that describes them [26]. The true symmetry of the system is simply hidden by the necessary choice of an asymmetrical lower energy state. We can for instance think of a plastic rod being compressed along its longitudinal axis (Fig. 1.2, middle): before applying a vertical force the system is symmetric under rotations, but after being compressed the ground state necessarily breaks the symmetry[2].

In the context of HEP, the presence of an additional scalar field spontaneously breaks the symmetry of the original EW Lagrangian in Eq. (1.18). A $\mathrm{SU(2)_L}$ doublet with 4 scalar real fields is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}, \tag{1.31}$$

[2]Spontaneous symmetry breaking is also present in bleeding edge diffusion models for content generation [27]!

where ϕ must be a scalar due to space isotropy and homogeneity arguments. To add this term to the SM Lagrangian, we introduce an additional Lagrangian term with standard kinetic and potential terms:

$$\mathcal{L}_{\mathrm{H}} = (D_\mu \phi)^\dagger (D^\mu \phi) - \mathrm{V}_{\mathrm{H}}(\phi^\dagger \phi)\,, \tag{1.32}$$

where D_μ is the covariant derivative introduced in Eq. (1.17). The potential term above can be written as:

$$\mathrm{V}_{\mathrm{H}}(\phi^\dagger \phi) = -\mu^2 \phi^\dagger \phi + \lambda \big(\phi^\dagger \phi\big)^2\,, \quad \text{with} \quad \mu^2\,,\, \lambda > 0\,, \tag{1.33}$$

and is illustrated in Fig. 1.2 (right). The minima of the potential are clearly not located at $\phi = 0$, but satisfy instead $|\phi^2| = \mu^2/2\lambda$, which describes a circle in the complex plane. Once a particular minimum is chosen, the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry is broken, but the ground state remains invariant under the $\mathrm{U(1)_{EM}}$ symmetry, since it stays parallel to ϕ^0. By defining the vacuum expectation value (VEV) as $v = \mu/\sqrt{\lambda}$ and expanding around the true minimum, the field is treated as a fluctuation around the ground state, and can be rewritten as:

$$\phi(x) = \frac{1}{\sqrt{2}} \exp\left[\frac{i\sigma_j \theta^j(x)}{v}\right] \begin{pmatrix} 0 \\ v + \mathrm{H}(x) \end{pmatrix}, \tag{1.34}$$

where the sum across j is implied. Eq. (1.34) is interpreted as the presence of a scalar real massive field H and three massless fields θ^i. The massless fields follow from Goldstone's theorem [28], which states that whenever a continuous symmetry is not apparent in the ground state, or spontaneously broken, a number of massless bosons will be generated equal to the number of the broken generators of the symmetry, three in this case. By applying the *unitary gauge* transformation on top, we can easily see that the θ^i bosons are nonphysical:

$$\phi(x) \to e^{-\frac{i}{v}\sigma_j \theta^j(x)} \phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \mathrm{H}(x) \end{pmatrix}. \tag{1.35}$$

Only one scalar field remains, corresponding to a new scalar boson, the Higgs particle. The Higgs field is the only known scalar field in Nature. Using the definition of covariant derivative in Eq. (1.17) plus the unitary gauge in the Lagrangian Eq. (1.32), we obtain:

$$\begin{aligned} \mathcal{L}_{\mathrm{H}} = {} & \frac{1}{2}\partial^\mu \mathrm{H} \partial_\mu \mathrm{H} - \frac{1}{2}\big(2\lambda v^2\big)\mathrm{H}^2 \\ & + \left[\left(\frac{gv}{2}\right)^2 \mathrm{W}^{\mu+}\mathrm{W}^-_\mu + \frac{(g^2+g'^2)v^2}{8}\mathrm{Z}^\mu \mathrm{Z}_\mu\right]\left(1 + \frac{\mathrm{H}}{v}\right)^2 \\ & + \lambda v \mathrm{H}^3 + \frac{1}{4}\lambda \mathrm{H}^4 - \frac{1}{4}\lambda v^4\,, \end{aligned} \tag{1.36}$$

where the Higgs mass term in the first line establishes a new free parameter of the theory, $m_{\mathrm{H}} = \sqrt{2\lambda}v = \sqrt{2}\mu$, and the second line shows mass terms for the W and Z gauge bosons:

$$m^2_{\mathrm{W}} = \frac{g^2 v^2}{4}\,, \quad m^2_{\mathrm{Z}} = \frac{(g^2+g'^2)v^2}{4}\,. \tag{1.37}$$

These masses have been experimentally measured to be [29–31]:

$$m_{\mathrm{W}} = 80.3692 \pm 0.0133\,\mathrm{GeV}\,, \quad m_{\mathrm{Z}} = 91.1880 \pm 0.0020\,\mathrm{GeV}\,. \tag{1.38}$$

We must here mention the recent and precise CMS measurement of $m_{\mathrm{W}} = 80.3602 \pm 0.0099\,\mathrm{MeV}$ [32], which is not yet part of the PDG combination above.

The unitary gauge effectively removed the Goldstone bosons θ^i, which were absorbed as additional degrees of freedom of the weak gauge bosons, corresponding to their longitudinal polarizations. We can finally introduce the Higgs potential $\mathrm{V_H}$ as it is usually presented:

$$\mathrm{V_H(H)} = \frac{1}{2} m_{\mathrm{H}}^2 \mathrm{H}^2 + \lambda_{\mathrm{HHH}} v \mathrm{H}^3 + \frac{\lambda_{\mathrm{HHHH}}}{4} \mathrm{H}^4\,, \tag{1.39}$$

with the self-couplings defined as:

$$\lambda_{\mathrm{HHH}} = \frac{G_F}{\sqrt{2}} m_{\mathrm{H}}^2 = \frac{m_{\mathrm{H}}^2}{2v^2} \simeq 0.13\,, \quad \lambda_{\mathrm{HHH}} = \lambda_{\mathrm{HHHH}}\,. \tag{1.40}$$

We can conclude that the Higgs self-couplings are fully determined in the SM, given the measurements of m_{H} and v, the latter fixed by Fermi's constant G_F, which is in turn measured via muon decays. From an experimental point of view, the measurement of λ_{HHH} thus provides a unique opportunity to perform a SM closure test, independently from the gauge interactions of the EW sector.

The Yukawa couplings

The Higgs mechanism represents an economical way to provide mass to the weak bosons in the theory. It can also be shown that the new massive bosons preserve renormalizability [33]. However, a second feat is yet to be achieved: justify the existence of fermion masses. To do so, Yukawa (renormalizable) $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ invariant terms are added to the SM Lagrangian, coupling left to right chiral fields:

$$\mathcal{L}_{\mathrm{Yukawa}} = -y_{f'}\left(\bar{\Psi}_{\mathrm{L}} \phi \psi'_{\mathrm{R}} + \bar{\psi}'_{\mathrm{R}} \phi^\dagger \Psi_L\right) - y_f \left(\bar{\Psi}_{\mathrm{L}} \tilde{\phi} \psi_{\mathrm{R}} + \bar{\psi}_{\mathrm{R}} \tilde{\phi}^\dagger \Psi_L\right)\,, \tag{1.41}$$

where ψ and ψ' denote up and down fermions ($I_3 = 1/2$ and $I_3 = -1/2$, respectively, see Eq. (1.14)), and:

$$\tilde{\phi} = i\sigma_2 \phi^* = \begin{pmatrix} \phi_0^* \\ -\phi_+^* \end{pmatrix} \xmapsto{\text{EWSB}} \frac{1}{\sqrt{2}} \begin{pmatrix} v + \mathrm{H}(x) \\ 0 \end{pmatrix}. \tag{1.42}$$

If we want for instance to generate the electron mass, it suffices to replace the fields with $\bar{\Psi}_{\mathrm{L}} = (\bar{\nu_e}, \bar{e})$ and similarly for the right-handed singlets. Note that fermion mass terms of the form $m\bar{\psi}\psi$ are forbidden by gauge invariance, and it is the addition of the Higgs doublet that changes the scenario.

After EWSB, using Eq. (1.35) and Eq. (1.42), and summing over fermion up and down types, Eq. (1.41) is simplified to:

$$\mathcal{L}_{\mathrm{Yukawa}} = -\sum_f m_f \left(\bar{\psi}_{\mathrm{L}} \psi_{\mathrm{R}} + \bar{\psi}_{\mathrm{R}} \psi_{\mathrm{L}}\right)\left(1 + \frac{\mathrm{H}}{v}\right)\,, \tag{1.43}$$

with new $m_f = y_f v/\sqrt{2}$ mass terms for the fermions. Very importantly, we have simplified the calculations: Eq. (1.41) can be expressed as a series of mass terms only after diagonalizing the original Yukawa matrices that multiply the fields. The process of diagonalization implies that the lepton fields are not mass eigenstates; they have to be redefined as in $\psi' \to U\psi$. When the redefinition is propagated back to the full EW Lagrangian, only the term containing charged currents is modified:

$$\mathcal{L}_{\mathrm{CC}} = \frac{e}{\sqrt{2}\sin\theta_{\mathrm{W}}} W_\mu^+ V_{\mathrm{CKM}}\, \bar{\psi}_{\mathrm{L}} \gamma^\mu \psi'_{\mathrm{L}} + \mathrm{h.c.}\,, \tag{1.44}$$

where the last term refers to the hermitian conjugate to simplify the expression, and V_{CKM} is the Cabibbo-Kobayashi-Maskawa (CKM) matrix, receiving contributions from the two redefined fields ψ_{L}. The CKM matrix describes the mixing between all quark generations, and makes possible the existence of flavour-changing charged currents in the SM. This is not the case for leptons, where couplings to charged W's only happen within each generation; a decay like $e^- \to \nu_\mu W^-$ is not allowed. The values of the CKM matrix are shown in the following:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.45}$$

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97435 \pm 0.00016 & 0.22501 \pm 0.000658 & 0.003732^{+0.000090}_{-0.000085} \\ 0.22487 \pm 0.00068 & 0.97349 \pm 0.00016 & 0.04183^{+0.00079}_{-0.00069} \\ 0.00858^{+0.00019}_{-0.00017} & 0.04111^{+0.00077}_{-0.00068} & 0.999118^{+0.000029}_{-0.000034} \end{pmatrix}, \tag{1.46}$$

where the experimentally measured [13, p. 12] magnitudes of the CKM elements represent the probability amplitude for a quark i to transform into a quark j. The matrix also includes an additional phase, which will be further discussed in Section 1.3.1. Since the mass diagonalization does not affect neutral currents, neutral processes changing the flavour of fermions are forbidden at first-order in the SM. The fact that the matrix is relatively close of being diagonal implies a preference for couplings within the same generation. This is for instance why top quarks decay virtually always to b quarks, and other decays channels can be neglected.

The SM therefore explains fermion masses by introducing 12 *ad hoc* Yukawa coupling terms between the Higgs field and all fermion fields, breaking the symmetry between the three lepton families. At the same time, the origin or fundamental reason for such terms remains unknown, with the masses being free parameters of the theory. The Higgs mechanism does however predict a proportionality between its couplings and the masses of all fermions and gauge bosons (recall Eq. (1.36)). This has been experimentally verified by the CMS and ATLAS Collaboration, and is shown in Fig. 1.3. The Higgs field also has an enormous impact in other fronts. As we will see in Section 1.2, there is a strong link between the λ_{HHH} and λ_{HHHH} self-couplings and the stability of the vacuum, and the Higgs boson is also speculated to be connected to inflation, dark matter production, and baryogenesis. The last item is also discussed in Section 6.1.

Despite strongly motivated, and serving as the main motivation for building the Large Hadron Collider (LHC), all the above remained uncertain until very recently. Additionally, since the mass of the new boson was a free parameter of the theory, there was no definitive guarantee that a Higgs particle could be produced in a HEP experiment. However, some considerations on unitarity and perturbative behavior conveyed some op-

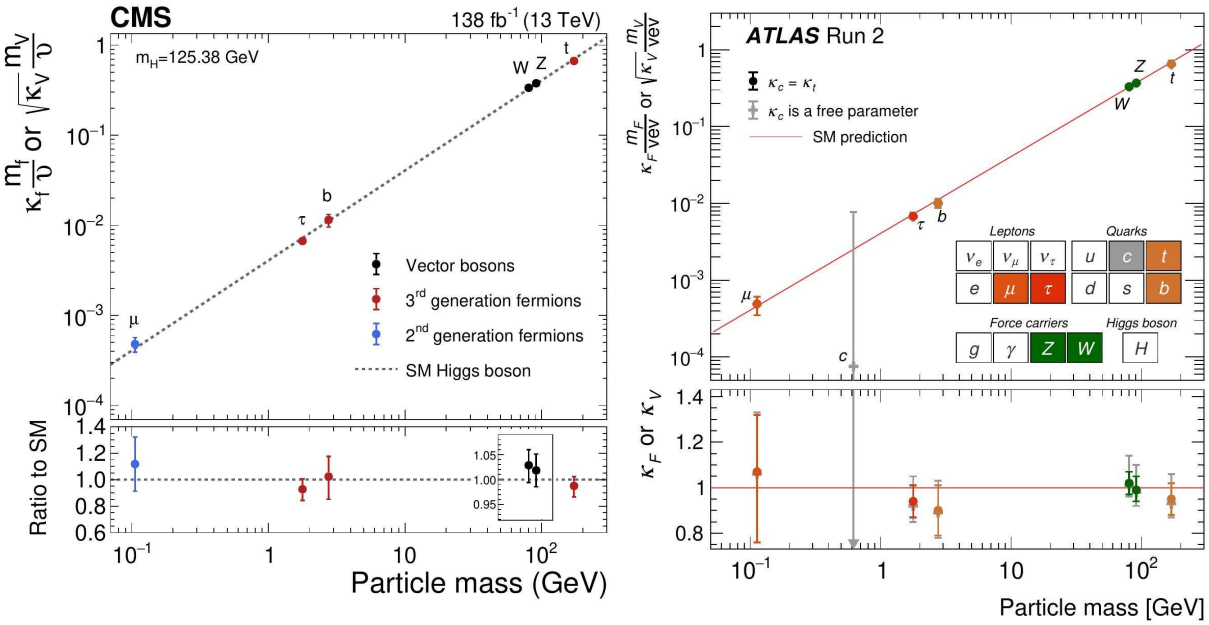

Figure 1.3: The measured coupling modifiers of the Higgs boson to fermions and heavy gauge bosons, as a function of fermion or gauge boson masses, where v stands for the vacuum expectation value of the Higgs field. For gauge bosons, the square root of the coupling modifier is plotted, to keep a linear proportionality to the mass, as predicted in the SM. Measurements used full Run 2 LHC data. Their p-value with respect to the SM prediction is 37.5% (CMS, left) and 65% (ATLAS, right). Taken from [38, 39].

timism, implying upper limits of $m_{\mathrm{H}} \lesssim 700\,\mathrm{GeV}$ [34, 35], and to lower limits driven by vacuum stability arguments, where the nontrivial minima at $v \neq 0$ would disappear if m_{H} was too small [36]. The Higgs boson was finally jointly discovered by the CMS and ATLAS experiments, and was announced to the world on July 2012 [1–3]. All measurements so far suggest that the new particle truly is the predicted SM scalar Higgs boson, with zero charge and a mass slightly above 125 GeV. Interestingly, the value found for m_{H} almost perfectly maximizes the product of the 14 branching ratios of the SM Higgs boson into the known fermion and boson pairs [37]. It is not known if this fact is accidental or if it instead hints at a deeper yet hidden theory, but it certainly increases the richness of Higgs physics to be explored at the LHC and beyond.

1.1.4 Phenomenology and Experimental Status of Higgs Boson Physics

The measurements of Higgs properties at the LHC are in general challenging, given the low cross sections involved. The most sensitive analyses to single Higgs boson processes are the ones featuring the $\mathrm{H} \to \gamma\gamma$ and $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ decay channels. Both are quite rare in the SM, with the di-photon channel not even being allowed at tree level, since the Higgs boson couples to massive particles. They however possess extremely clean photonic or leptonic signatures, with relatively few competing background processes. The two channels have been responsible for increasingly precise measurements of the properties of the Higgs boson, namely its spin-parity $J^P = 0^+$ [40, 41], and its mass, here shown for the $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ channel in Run 1 + Run 2 combinations [42, 43]:

$$\begin{aligned} m_{\mathrm{H}} &= 124.94 \pm 0.17\,(\mathrm{stat.}) \pm 0.03\,(\mathrm{syst.}) \quad (\mathrm{ATLAS})\,, \\ m_{\mathrm{H}} &= 125.08 \pm 0.10\,(\mathrm{stat.}) \pm 0.05\,(\mathrm{syst.}) \quad (\mathrm{CMS})\,. \end{aligned} \tag{1.47}$$

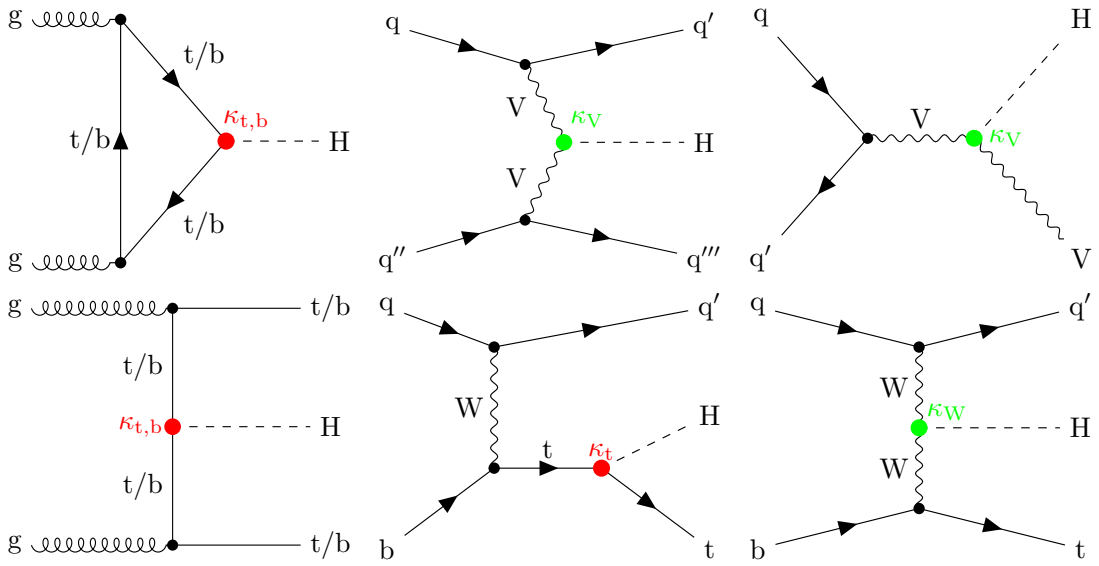

Figure 1.4: Feynman diagrams for the leading Higgs boson production processes. Top row, from left to right: gluon fusion, vector boson fusion and associated production with a W or Z (V) boson. Bottom row, from left to right: associated production with a top or bottom quark pair, associated production with a single top quark, in the t-channel (two diagrams). Yukawa coupling modifiers are highlighted.

We also mention the combination performed by the ATLAS experiment of the two channels above [44], which achieves the best precision so far:

$$m_{\mathrm{H}} = 125.11 \pm 0.09\,(\mathrm{stat.}) \pm 0.06\,(\mathrm{syst.})\,. \tag{1.48}$$

The precision of m_{H} measurements is a very important parameter to consider, on the grounds that m_{H} is a free parameter of the theory, and many observables depend on it.

The width is also an important parameter to characterize the Higgs boson. One can measure it directly, using on-shell processes, but the results have large uncertainties. CMS has for instance measured $\Gamma_{\mathrm{H}} < 60\,\mathrm{MeV}$ at 68% confidence level using the $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ channel [42, 45, 46]. On the other hand, the precision is greatly increased when combining on-shell and off-shell cross section measurements. Despite the small total width, the $\mathrm{H} \to \mathrm{ZZ}^*$ cross section is enhanced due to ZZ and $\mathrm{t\bar{t}}$ threshold effects [47]. Since the on-shell to off-shell cross section ratio is proportional to the width, measurements lead to [42, 48]:

$$\begin{aligned} \Gamma_{\mathrm{H}} &= 4.5^{+3.3}_{-2.5}\,\mathrm{MeV} \quad (\mathrm{ATLAS})\,, \\ \Gamma_{\mathrm{H}} &= 2.9^{+2.3}_{-1.7}\,\mathrm{MeV} \quad (\mathrm{CMS})\,. \end{aligned} \tag{1.49}$$

The Higgs boson is created via a series of production modes, the most important of which are shown in Fig. 1.4 and listed in Table 1.1, where cross sections are provided for a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$, with $m_{\mathrm{H}} = 125.09\,\mathrm{GeV}$. Looking at Fig. 1.5 (left), we can expect a cross section increase for higher energies. By far the most dominant production mode is gluon-gluon fusion (ggF), which as a consequence is also the most targeted by physics analyses at LHC experiments. One order of magnitude below, the vector-boson fusion (VBF) and VH "Higgs-strahlung" associated production modes follow, the latter split into WH and ZH, where WH includes processes with the two

Production mode	Approx. fraction [%]	Cross section [pb]
ggF	87.30	$48.61^{+4.27\%}_{-6.49\%}\,(\text{theory}) \pm 1.85\%\,(\text{PDF})^{+2.59\%}_{-2.62\%}(\alpha_S)$
VBF	6.76	$3.766^{+0.43\%}_{-0.33\%}\,(\text{scale}) \pm 2.1\%\,(\text{PDF}+\alpha_S)$
WH	2.44	$1.358 \pm 0.51\%\,(\text{scale}) \pm 1.35\%\,(\text{PDF}+\alpha_S)$
ZH	1.58	$0.880^{+3.50\%}_{-2.68\%}\,(\text{scale}) \pm 1.65\%\,(\text{PDF}+\alpha_S)$
ttH	0.90	$0.5065^{+5.8\%}_{-9.2\%}$
bbH	0.87	$0.4863^{+20.1\%}_{-23.9\%}$
tH (t-channel)	0.13	$0.07426^{+6.5\%}_{-14.7\%}$

Table 1.1: Relative abundances and cross sections plus corresponding uncertainties for the most significant Higgs boson production modes, at $\sqrt{s} = 13\,\text{TeV}$, for $m_H = 125.09\,\text{GeV}$ [49].

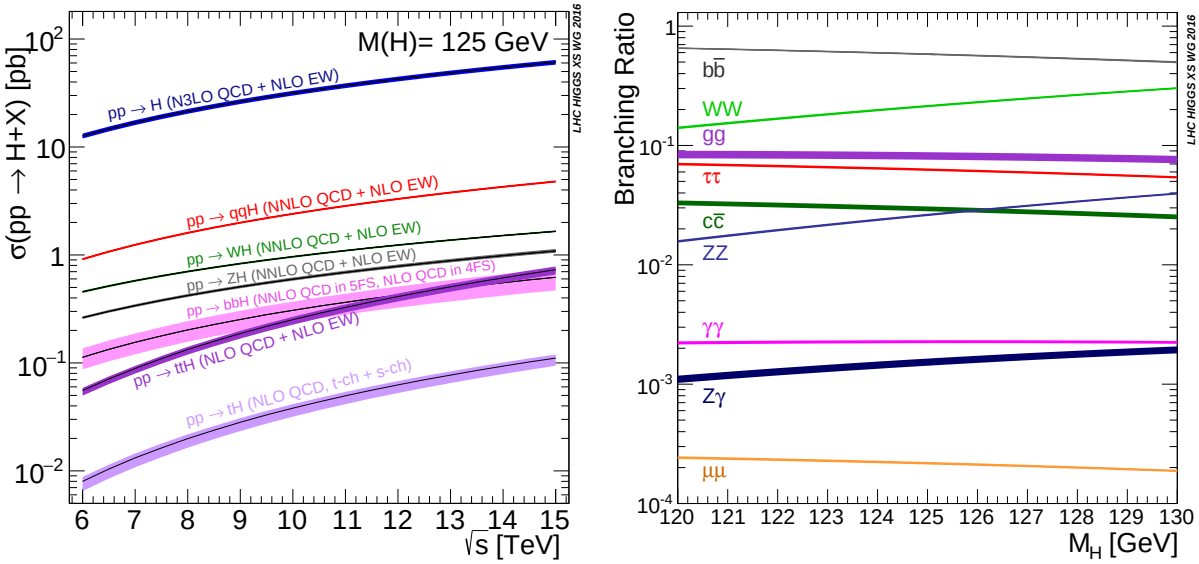

Figure 1.5: (Left) Higgs boson production cross section for several production modes, as a function of the $\sqrt{s}$ center-of-mass energy. (Right) Branching fractions of Higgs boson decays, as a function of m_H. Taken from [49].

charged W bosons. The VBF production mode, despite its small cross section, has a striking experimental signature with its two additional jets, which usually have a large di-jet invariant mass, and are produced back-to-back. VBF and VH also enable to probe the Higgs coupling to vector bosons, which is not present in ggF. The ttH and tH channels have an even smaller cross section, but can provide additional information on the top quark Yukawa coupling. Among the listed production modes, only bbH and tH have not yet been measured at the LHC. The former recently had a dedicated CMS search observing a 95% confidence level (CL) upper limit of 3.7 times the expected SM cross section value [50]. Once measured, it will serve as a further test for the Higgs coupling to bottom quarks. Facts worth noting are the absolute and relative dependencies of the various production modes with the Higgs boson transverse momentum (p_T) shown in Fig. 1.6. As the energy of the Higgs boson increases, the relationship between production modes is drastically affected. Above $p_T^H \approx 550\,\text{GeV}$ VH becomes the second most common production mode, and it even surpasses ggF at the admittedly very high $p_T^H \approx 1250\,\text{GeV}$ value. This illustrates the dynamic nature of Higgs physics.

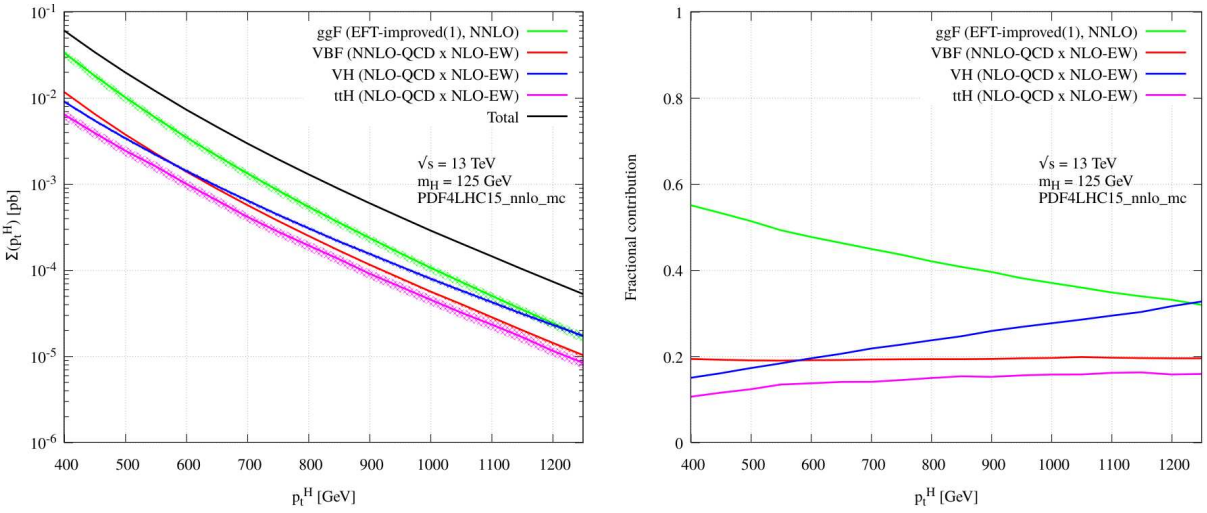

Figure 1.6: Cumulative cross section for the production of a Higgs boson as a function of the lowest Higgs boson transverse momentum. The cross section due to ggF (green), VBF (red), vector boson associated (blue) and top quark pair associated (magenta) production mode are shown in absolute values (left) and relative size (right). Taken from [51].

Decay mode	$\mathcal{B}$ [%]
$\mathrm{H} \to \mathrm{bb}$	$58.09^{+0.72}_{-0.73}$
$\mathrm{H} \to \mathrm{W}^{\pm}\mathrm{W}^{\mp *}$	21.53 ± 0.33
$\mathrm{H} \to gg$	8.18 ± 0.42
$\mathrm{H} \to \tau^{+}\tau^{-}$	6.27 ± 0.10
$\mathrm{H} \to \mathrm{cc}$	$2.88^{+0.16}_{-0.06}$
$\mathrm{H} \to \mathrm{ZZ}^{*}$	2.641 ± 0.040
$\mathrm{H} \to \gamma\gamma$	0.2270 ± 0.0047
$\mathrm{H} \to \mathrm{Z}\gamma$	0.1541 ± 0.0090
$\mathrm{H} \to \mu^{+}\mu^{-}$	$0.02171^{+0.00036}_{-0.00037}$

Table 1.2: Largest single Higgs boson branching ratios for $m_{\mathrm{H}} = 125.09\,\mathrm{GeV}$ [49]. The asterisk $*$ refers to virtual particles.

The final state decays of the Higgs boson is "maximally rich" [37], presenting a vast potential for exploration. We list the most common decay channels in Table 1.2 and show their Feynman diagrams in Fig. 1.7. The $\mathrm{H} \to \mathrm{bb}$ decay dominates, but suffers from an exceedingly large multijet background, plus significant resonant $\mathrm{Z} \to \mathrm{bb}$ background, and has never been observed in ggF. Similar issues arise in VBF, despite the additional handles provided by the jets, since the cross section is greatly reduced. Instead, associated VH production is used, since the additional boson can decay to leptons, which leave a clean signature in the detector. The multijet background is even larger for $\mathrm{H} \to \mathrm{cc}$, given the lower masses involved. Again, its production in association with vector bosons decaying leptonically is explored, imposing constraints on its cross section and on the charm Yukawa coupling [52]. It is important to note that such a decay channel would have seemed impossible to measure just a few years ago. Less surprisingly, the "golden channels" $\mathrm{H} \to \gamma\gamma$ and mostly $\mathrm{H} \to \mathrm{ZZ}^{*} \to 4\ell$ continue to provide excellent results, despite

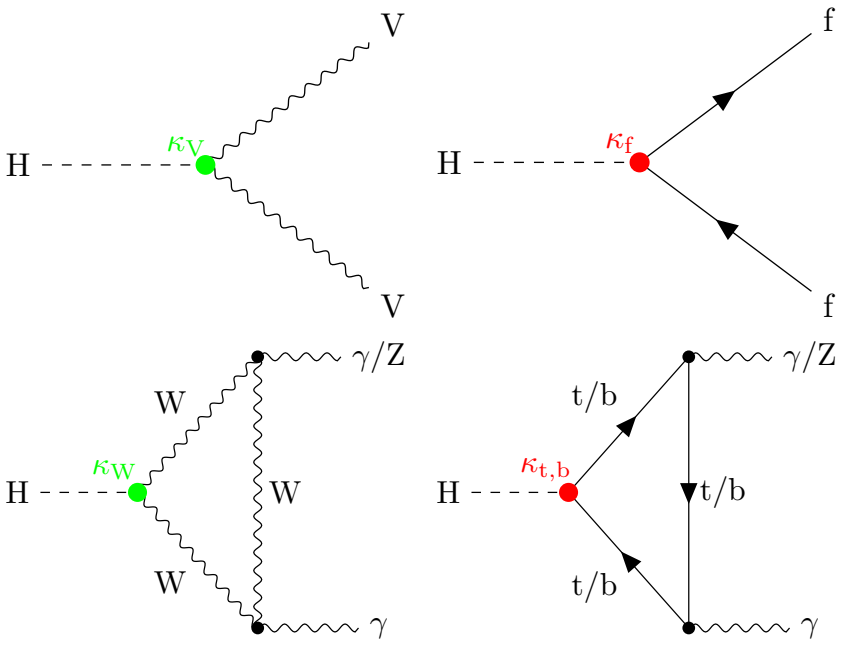

Figure 1.7: Feynman diagrams for the leading Higgs boson decay channels. Top left: decay to two vector bosons. Top right: decay to two fermions. Bottom: decay to two photons or a photon and a Z boson, via vector bosons (left) or (fermions) right. Yukawa coupling modifiers are highlighted.

their very small couplings, but relying on their experimentally clean signature. The very rare $H \to Z\gamma$ process has been recently combined between CMS and ATLAS, leading to an evidence ($3.4\,\sigma$) for the first time [53]. Notice that most decays with particles lighter than muons, such as light quarks or electrons, are extremely hard to measure in current or future detectors unless BSM effects come into play. These final states have extremely small predicted couplings to the Higgs boson. To give an example, the $H \to e^+e^-$ decay is predicted to have a branching ratio $\mathcal{B}$ of $\sim 5 \times 10^{-9}$, but the most recent CMS 95% CL upper limits excludes 3×10^{-4} only [54], with similar results by ATLAS [55]. However, the measurement of the electron Yukawa coupling at a future electron-positron collider cannot be for the moment completely excluded [56]. The two experiments have also been probing Lepton Flavour Violation (LFV) processes such as $H \to e\mu$, $H \to \tau\mu$ and $H \to \tau\mu$, reaching better and better sensitivities of up to 10^{-5}, and other rare decays, such as Higgs boson decays to J/ψ or its excited states. Of particular relevance for CMS are the observation of $H \to \tau\tau$ [57–59] and the first evidence for $H \to \mu\mu$ [60].

To conclude, Higgs physics comes with an extremely rich programme, and all production and decay channels so far measured agree with SM predictions. A summary is shown in Fig. 1.8, where the two panels report the agreement of CMS measurements with the SM using signal strength modifiers μ, defined as $\mu_i \equiv \sigma_i/(\sigma_i)_{\text{SM}}$ for the production modes i on the left and as $\mu^f \equiv \mathcal{B}^f/(\mathcal{B}^f)_{\text{SM}}$ for the decay channels f on the right. Additionally, and as already discussed in Section 1.1.3 and illustrated in Fig. 1.3, the predicted proportionality between the Higgs couplings and the fermion and gauge boson masses is remarkable, being confirmed by both CMS and ATLAS. Finally, the phenomenology of cubic and quartic interactions of the Higgs boson, as introduced by Eq. (1.39), is still left to be discussed. Given its tight connection with this Thesis' subject, it receives a dedicate treatment in the next Section.

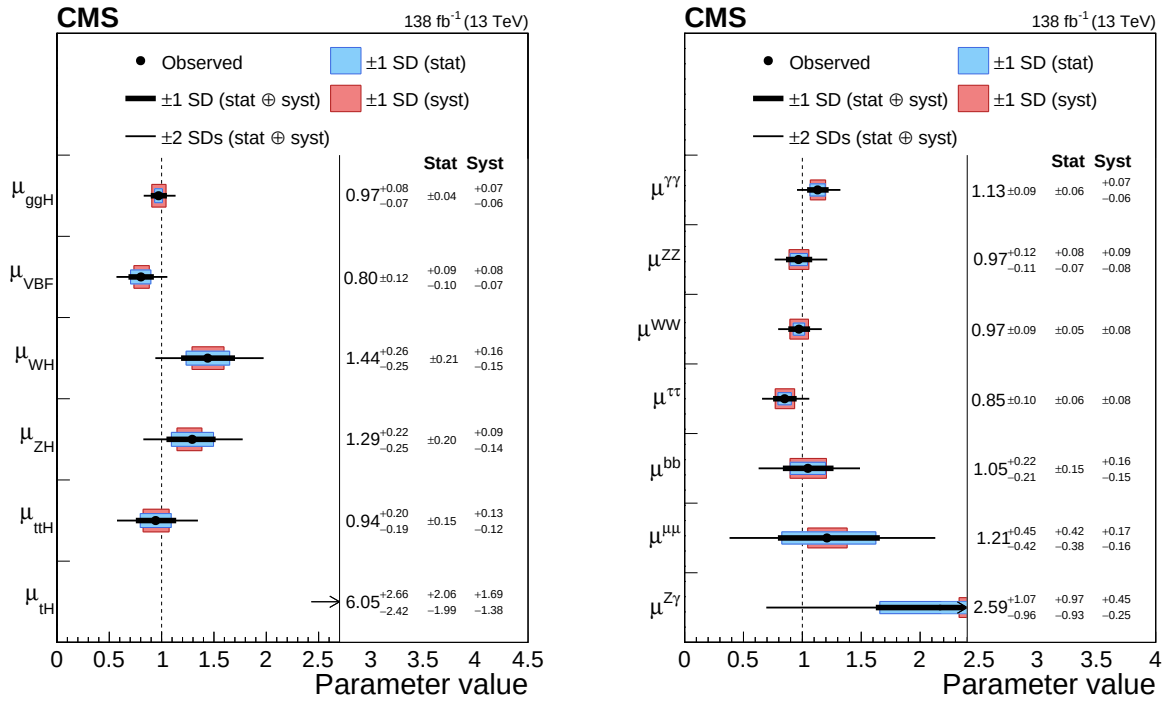

Figure 1.8: Signal strength parameters extracted for various production modes μ_i, assuming $\mathcal{B}^f = \mathcal{B}^f_{\mathrm{SM}}$ (left), and decay channels μ^f, assuming $\sigma_i = (\sigma_i)_{\mathrm{SM}}$ (right). The thick (thin) black lines indicate the 1 (2) standard deviation (s.d.) CLs, with the systematic and statistical components of the 1 s.d. interval indicated by the red and blue bands, respectively. The vertical dashed line at unity represents the values of μ_i and μ^f in the SM [38].

1.2 Higgs Boson Pair Phenomenology in the SM

The Higgs self-coupling λ_{HHH} plays a central role in the spontaneous symmetry breaking of EW theory, and can notably be probed using Higgs boson pair production. The connection to the SM is established via the Higgs potential in Eq. (1.39), where the triple and quartic Higgs couplings are predicted, given m_{H} and G_F. The self-couplings represent the last fundamental SM pieces yet to be measured, and provide a unique SM consistency test, enabling the crucial validation of the EWSB mechanism [49]. The shape of the Higgs potential itself would be affected in case of BSM deviations, leading to dramatic consequences for our understanding of the Universe.

Experimentally, one aims to verify if the relationships in Eq. (1.40) hold. However, as discussed in Section 1.2.1, the available experimental signatures often involve different types of interactions, which do not simply include the Higgs self-couplings. It is therefore necessary to disentangle λ_{HHH} contributions from other effects. Occasionally, simultaneous fits of multiple couplings are performed, when a particular process originates from more than one production mode. Furthermore, HH events are some of the rarest processes probed at the LHC, carrying a very small cross section, three orders of magnitude lower than the already rare single Higgs processes. Finally, challenges can also be found in the final states explored, being often complex, including decay products with jets and/or missing transverse energy (MET), as discussed in Section 2.4.

In this Section we discuss the phenomenology of HH physics, looking at its production in Section 1.2.1 and its final states in Section 1.2.2.

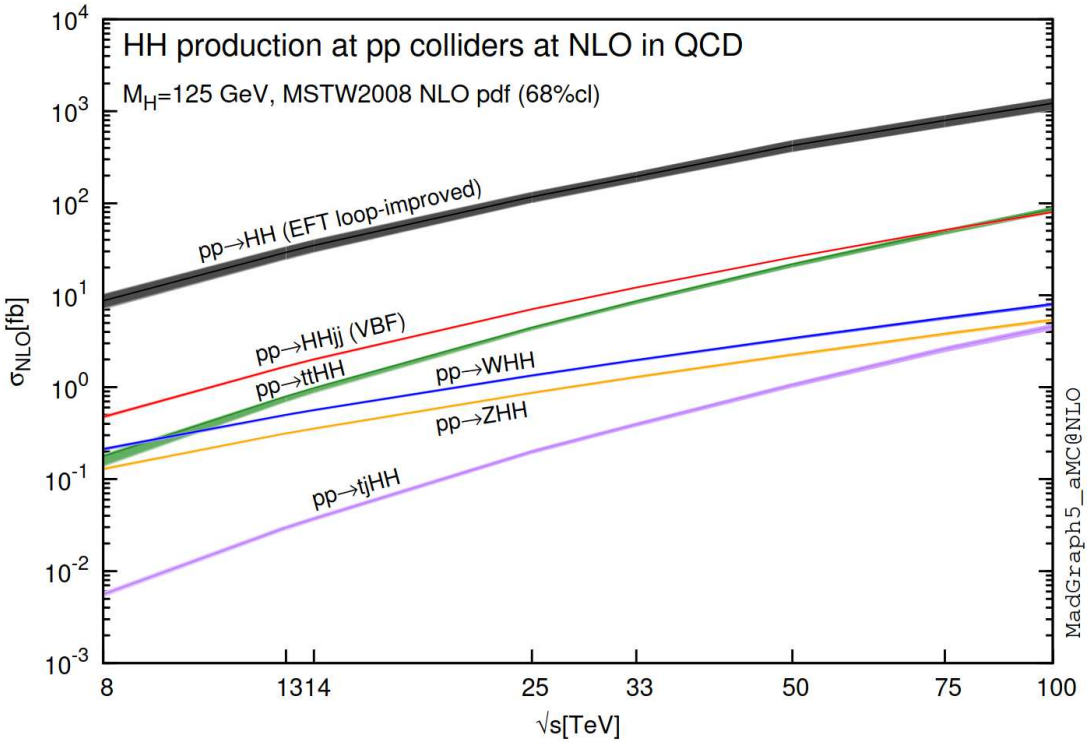

Figure 1.9: HH production cross section as a function of the center of mass energy for the six largest HH production channels at pp colliders. The thickness of the lines corresponds to the scale and PDF uncertainties added linearly. Gluon fusion dominates for the entire energy range. The figure is taken from [61].

1.2.1 Production Modes

The double Higgs production mechanisms are in many ways similar to the production modes of single Higgs physics covered in Section 1.1.4. The main differences consist of the larger number of possible Feynman diagrams for each production mode at the lowest loop order, the multiplicity increase in the final state decay products, and the dramatic cross section reduction. The five most common production modes are listed in Table 1.3, along with their cross sections for $\sqrt{s} = 13\,\text{TeV}$ and $m_\text{H} = 125\,\text{GeV}$. In Fig. 1.9, we show the evolution of those production modes with available center-of-mass energy. Modes with top quarks increase their cross section faster than processes including vector bosons, an observation that might have to be taken into account in future detectors at higher center-of-mass energies. Gluon fusion will nevertheless remain dominant in any future hadron collider, just as it is at the LHC. For the same reason, ggF is also the production mode considered in the analysis reported in Chapters 4 and 5. In the following, we provide a more detailed description of each production mode:

Gluon fusion

The gg → HH gluon fusion is the most important production mechanism of Higgs boson pairs at hadron colliders [63]. Contrary to single Higgs ggF, in HH phenomenology two diagrams participate at lowest order. The first involves a heavy quark "triangle" loop, just like for single Higgs production, radiating two Higgs bosons from another Higgs particle. It is shown in Fig. 1.10 (left), where the dependence on the self-coupling λ_HHH and on the top Yukawa coupling y_t is clear. A second "box" diagram exists, shown in Fig. 1.10

Prod. mode	Fraction [%]	Cross section [fb]
ggHH	90.22	$30.77^{+2.2\%}_{-5\%}(\text{scale})^{+4.0\%}_{-18.0\%}\%(m_t) \pm 1.5\%(\text{PDF}) \pm 1.7\%(\alpha_S)$
VBF HH	4.95	$1.687^{+0.05\%}_{-0.04\%}(\text{scale}) \pm 2.7\%(\text{PDF} + \alpha_S)$
ZHH	1.06	$0.363^{+3.4\%}_{-2.7\%}(\text{scale}) \pm 1.9\%(\text{PDF} + \alpha_S)$
W^+HH	0.96	$0.329^{+0.32\%}_{-0.41\%}(\text{scale}) \pm 2.2\%(\text{PDF} + \alpha_S)$
W^-HH	0.51	$0.173^{+1.2\%}_{-1.3\%}(\text{scale}) \pm 2.8\%(\text{PDF} + \alpha_S)$
ttHH	2.22	$0.756^{+4.3\%}_{-15.0\%}(\text{scale}) \pm 3.3\%(\text{PDF} + \alpha_S)$
tqHH	0.08	$0.0289^{+5.5\%}_{-3.6\%}(\text{scale}) \pm 4.7\%(\text{PDF} + \alpha_S)$

Table 1.3: Relative abundances and inclusive cross sections plus uncertainties for the most significant HH production modes at $\sqrt{s} = 13\,\text{TeV}$ and $m_H = 125\,\text{GeV}$ in the next-to-next-to-leading order (NNLO) Full Theory (FT) approximation for $m_H = 125\,\text{GeV}$. Taken from [62].

(right), involving only the top Yukawa coupling. The amplitude of the two diagrams is overwhelmingly dominated by the top quark contribution; the mass of the second-heaviest quark, the b, is ∼40 smaller, and its contribution to the HH ggF cross section amounts to ∼0.3 %, being usually neglected due to the current experimental sensitivity [64]. Importantly, given the identical final states, the two production diagrams interfere, and the interference is destructive. Taking also into account the additional diagram vertices compared to single Higgs processes, the HH ggF cross section becomes extremely small. On the other hand, the same effect increases the sensitivity to NP, since BSM effects, including new resonances, can affect the balance between diagrams, and boost HH rates. BSM HH studies are thus strongly motivated, as the one presented in Chapter 4.

Vector Boson fusion

The VBF HH process, or $qq' \to jjHH$, where j stands for a jet, besides including λ_{HHH}, exhibits a dependence on the cubic (κ_V) and quartic (κ_{2V}) couplings of the Higgs boson to vector gauge bosons, as shown in Fig. 1.11. The two forward jets provide a powerful discriminative handle. In addition, theoretical arguments lead to suppressed hadron production in the central regions of the detector, between the two jets [65]. Therefore, and despite its cross section more than 18 times smaller than for ggF, VBF HH provides a clear signature that suppresses the multijet background.

Associated production, or Higgs-strahlung

The $qq' \to VHH$ process involves the same couplings appearing in the VBF HH process, and an additional vector boson is present in the final states, as shown in Fig. 1.12. The extra boson is used as a powerful handle to discriminate against multijet background, so far tested in analyses where Higgs bosons decay to b quarks [66, 67]. Additionally, and especially in the $4 < \kappa_\lambda < 7$ range of the coupling modifier, the cross sections get closer to the VBF and ggF ones, creating some complementarity worth exploring. The limits become indeed comparable to other sub-leading searches for those high κ_λ values.

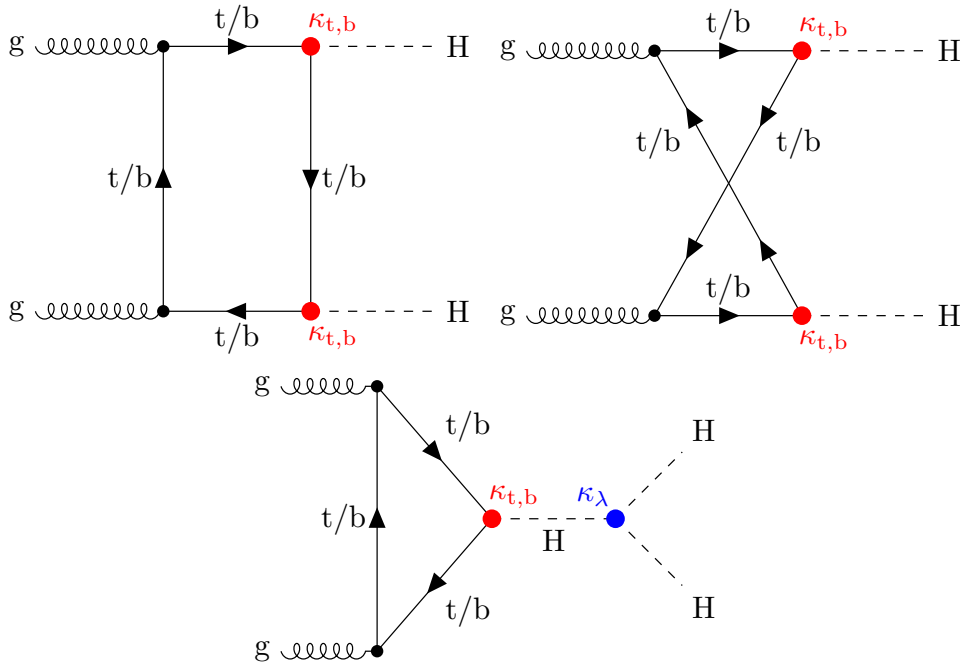

Figure 1.10: Feynman diagrams for double Higgs gluon fusion production. The box (top) and triangle (bottom) diagrams have amplitudes of similar magnitudes, but interfere destructively. The quark loops are dominated by the top quark. The triangle diagram is characterized by the Higgs self-coupling and the coupling of the Higgs to top quarks. Box diagrams include two contributions from Higgs-top couplings.

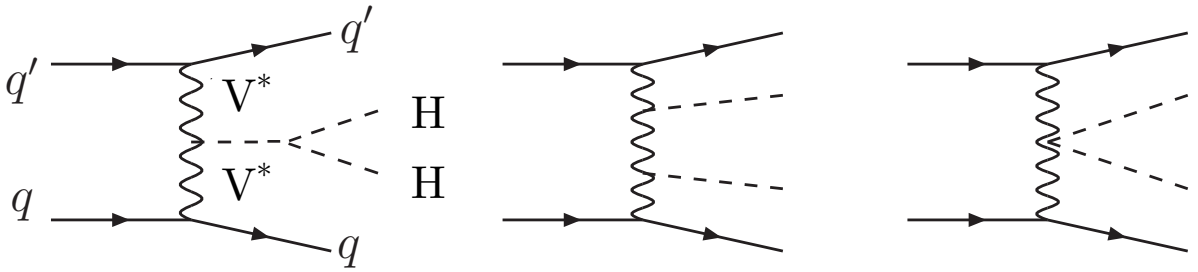

Figure 1.11: Feynman diagrams for double Higgs VBF production.

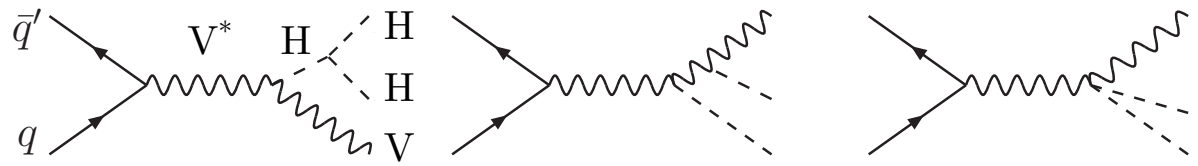

Figure 1.12: Feynman diagrams for double Higgs associated production.

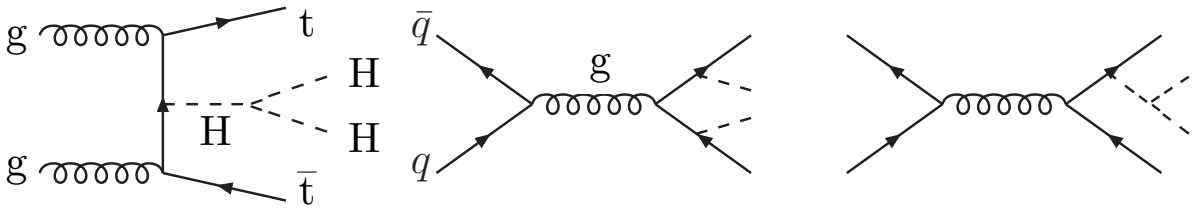

Figure 1.13: Feynman diagrams for double Higgs $t\bar{t}$ associated production.

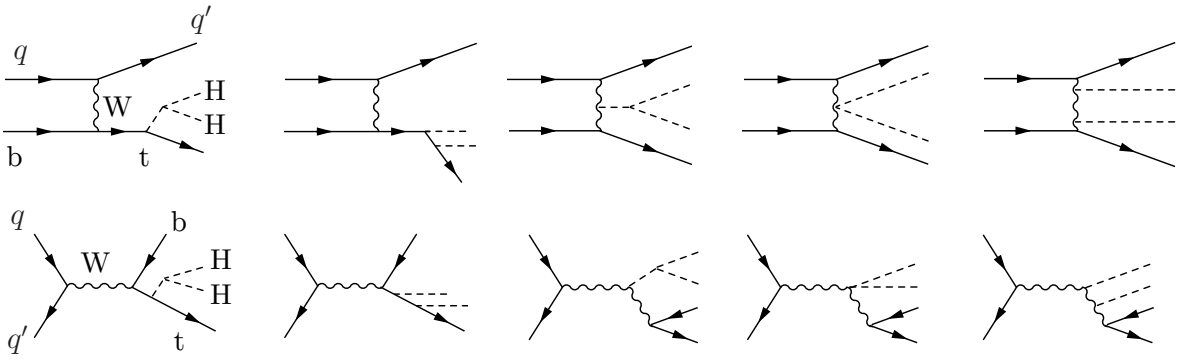

Figure 1.14: Feynman diagrams for double Higgs single top quark associated production.

Top/anti-top associated production

The $qq'/gg \to t\bar{t}HH$ process can be both gluon- or quark-initiated, and is characterized by the presence of an additional top quark pair. In this mode, either two Higgs bosons are radiated from the top quarks, or are otherwise produced from the virtual Higgs particle, as shown in Fig. 1.13. The ttHH cross section exceeds the VBF HH one starting from $\sqrt{s} \approx 70\,\text{GeV}$. This behavior is very different from what happens for single Higgs processes, where ttH remains very subdominant also for high energies (see Fig. 1.5, left).

Single top associated production

The single top $qq' \to tq''HH$ process can proceed via a multitude of t- or s-channels, as shown in Fig. 1.14, where s-channels are around one order of magnitude rarer. The cross sections are so small that it can be hardly investigated at the LHC. Future colliders can nevertheless open a window on this production mode, because the cross section increases faster with energy than the associated production, and reaches the same value at $\sqrt{s} \approx 100\,\text{GeV}$. Additionally, the two jets will provide background reduction capabilities. The interest resides from being able to probe HH couplings to vector bosons and top quarks simultaneously.

1.2.2 Final States

The sensitivity of certain HEP processes is occasionally dominated by a single final state. In single Higgs physics, for instance, the $H \to ZZ^* \to 4\ell$ is often regarded as a "golden channel", since it has a large signal-to-background ratio, with a flattish background under the Higgs boson peak, and possesses an extremely detector-friendly final state, which benefits from the good electron and muons resolution in CMS and ATLAS. Another

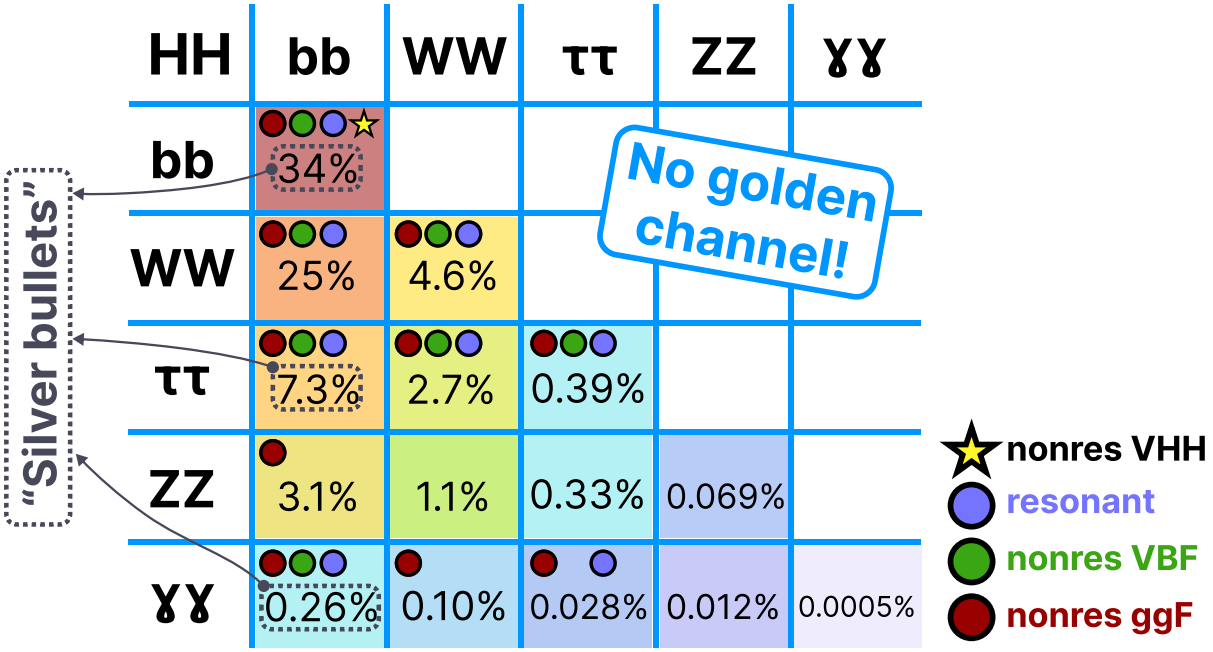

Figure 1.15: Illustration of the branching ratios (in %) for the most common HH decay final states. There is no single "golden channel" that fully dominates the sensitivity. The three most sensitive channels, called "silver bullets", are highlighted. The legend refers to the type of measurements performed within CMS at the time of writing. A first VHH measurements has been recently published [66]. ttHH topologies are starting to be explored, but no results are yet available. Taken from [5].

example can be provided in the context of B-Physics, where the $B_s^0 \to \mu\mu$ and $B^0 \to \mu\mu$ decays are equally seen as "golden", taking once again into account the exquisite muon resolutions in CMS, and are highly suppressed in the SM, making these rare channels particular sensitive to BSM effects [68]. On the contrary, double Higgs production does not include a single sensitivity-dominating channel. Instead, a vast measurement programme has to be established, considering multiple final states, and foreseeing their subsequent combination. Three channels are nevertheless more sensitive than the others: bbbb, bb$\gamma\gamma$ and bb$\tau\tau$. They are collectively referred as "silver bullets", since their combined measurement drives most of the sensitivity in HH physics:

- $HH \to bbbb$: characterized by the largest branching ratio at 34%, this decay channel is unfortunately overwhelmed by multijet background, much like $H \to bb$ and $H \to cc$ processes discussed in Section 1.1.4. Its relative importance has been increasing, given the recent and very significant improvements in the performance of jet taggers, mostly driven by the exploitation of novel machine learning techniques. It is particularly suited for the high energy boosted regime, when the jets are collimated, but it is successfully used across the entire phase-space.

- $HH \to bb\gamma\gamma$: just like for the $H \to \gamma\gamma$ process, a sub-percent branching fraction is here compensated by a pure, low background final state. The irreducible di-photon background can be mitigated by applying a selection on the photon pair invariant mass.

- $\mathrm{HH} \to \mathrm{bb}\tau\tau$: represents an ideal compromise between a sizable branching fraction at 7.3% and a background benefiting from the presence of a relatively pure $\tau\tau$ pair. It is the channel considered for the analysis later described in this Thesis.

The remaining channels have a limited impact on the final sensitivity, but are nevertheless explored. They potentiate the development of new techniques that can be later transferred to other analyses, they can be sensitive to very particular and unexpected BSM signatures, and they can provide the combined additional small gain required to observe a new process. The available final states are illustrated in Fig. 1.15.

1.3 Going Beyond the SM

The SM represents a remarkable triumph of modern physics, beautifully explaining all known EM, weak and strong interactions, and establishing a profound connection between symmetries and governing equations which precisely predict the entire set of measurements so far performed at collider experiments. Precision tests confirm the SM validity up to EW scale energies, and the observed Higgs boson resembles the theorized one up to great precision. Models predicting complex "Higgs families", such as multiple flavors of SUSY theories or Two-Higgs Doublet Model (2HDM) types, are getting less favored as more data is collected and analyzed.

It is however striking that some SM elements are empirically added to reproduce experimental data. Taking the Yukawa couplings as an example, discussed in Section 1.1.3, they are introduced to explain fermion masses, and coincidentally happen to work extremely well with the newly introduced Higgs doublet, but nobody knows *why* those terms are present in the first place. In total, assuming neutrinos of Dirac type (as opposed to Majorana fermions), the SM brings 26 free parameters to the table: 12 Yukawa couplings, three coupling constants, namely the weak and strong couplings and Fermi's constant G_F, the two parameters describing the Higgs potential in Eq. (1.33), eight mixing angles for the PMNS and CKM matrices, and a strong charge-parity (CP) phase. The presence of so many parameters suggests that, despite its undeniable achievements, the SM continues to be a model where parameters are chosen to match observations. Moreover, the entire theory rests on the $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ gauge group, but it remains unclear why this particular symmetry should be more relevant than a different one. A series of patterns exists between the different parameters, hinting at an underlying yet unknown symmetry principle. Also unclear is the existence of exactly three generations of fermions. Could it also be connected to some underlying general truth, or are theoretical developments in Particle Physics stalled until the next paradigm shift is attained?

From an experimental point of view, to successfully go beyond the SM one ideally needs a physical process with high discovery potential on multiple fronts. Double Higgs pair production fits such description. In this section, after discussing the many clear discrepancies the SM cannot convincingly explain, we discuss all the options currently explored to find NP in HH processes, in both nonresonant (Section 1.3.2) and resonant (Section 1.3.3) processes. The latter also establishes the motivational basis for the $\mathrm{bb}\tau\tau$ analysis work reported in Chapters 4 and 5. Notably, even if no deviations or excesses are found, results will still be able to provide information about the world we live in: strong limits on the phase-space of many BSM models, and the seemingly inevitable measurement of the Higgs boson self-coupling in a not-so-distant future.

1.3.1 SM shortcomings

Despite its success, intriguing observations or fully unexplained phenomena blatantly highlight the limitations of the SM. In this Section, we list the most prominent challenges a successful future theory should ideally address [20, 69]. That high-energy collider physics has the potential to look for answers in many of the following topics is a statement for the relevance HEP experiments still have in the general experimental landscape.

- **Dark Matter**: It has been known for almost a century that standard gravitational equations fail to describe the velocity distribution of stars orbiting galaxies as a function of the distance r from the galaxies' centers. Assuming that the largest fraction of a galaxy's mass is found at its center, the tangential velocity of the stars should decrease with $r^{-1/2}$. Observations instead indicate that it decreases much more slowly, suggesting that the mass distribution is proportional to r, implying a very significant non-luminous component: the "dark" matter. Many experiments have provided independent measurements that seem to support the existence of dark matter, like the Cosmic Microwave Background (CMB) measurements with WMAP [70] and Planck [71], or the observation galaxy clusters, including the famous "bullet cluster" [72, 73]. Some SM extensions predict the existence of weakly interacting massive particles (WIMPs), as SUSY, Warped ED (WED) and little Higgs theories. They should pervade our Universe, and so might be observable directly via dedicated experiments, or produced at the LHC. Alternative explanations for the observed discrepancies nevertheless exist, most notably modified newtonian dynamics (MOND) theories [74], but are generally disfavored by the scientific community.

- **Grand unification**: At the EW scale, the three coupling constants present in the SM have a somewhat similar value. Significantly, they evolve with energy in such a way that their unification could be achievable at some higher energy, assuming BSM symmetries. Grand Unified Theories (GUTs) have been proposed with all-encompassing gauge symmetry groups, where convergences are achieved at grand unification scales ($\approx 10^{16}$ GeV), often dependent on the presence of additional particles, which modify the running of the couplings. These theories usually require proton decay, which has not been observed so far[3]: current lower bounds are set at around $10^{33}/10^{34}$ at 90% CL [76, 77]. Even more dramatic is the absence of gravity from the current framework. The quantization of general relativity as been shown to be non-renormalizable, and cannot thus be integrated in the SM as it is.

- **CP violation and baryogenesis**: Sakharov identified three necessary ingredients to explain a process capable of reproducing the observed matter-antimatter asymmetry: *i*) lepton and baryon number violation, *ii*) a Universe was out of thermal equilibrium at some point in time, and *iii*) a CP-violating interaction. In the SM, the violation of CP symmetry can only be accommodated in the parameters of the CKM and PMNS matrices. CP violation has already been studied in great depth in the quark sector, but the same is not the case for neutrinos. However, even assuming the observation of CP violation in the PMNS matrix, it seems unlikely that the SM will be able to explain the full extent of the observed matter/anti-matter asymmetry in the Universe. The required CP violation could also be hidden in the

[3] I can't resist mentioning the futuristic possibility of using material taken from the Moon's depths to improve on current limits [75], which are otherwise slowly reaching practical experimental ceilings.

strong SM phase, an issue that can be tackled by introducing axions, which have the advantage of also being dark matter candidates.

- **The nature of neutrinos**: There is a striking disparity between neutrino masses and the masses of all other SM particles, including other fermions. Dirac-type neutrinos imply unnaturally small Yukawa couplings to the Higgs boson, since the latter is proportional to the fermion mass. Majorana-particles might circumvent this issue via a seesaw-like mechanism, violating the lepton number [78]. The most striking evidence for Majorana-type neutrinos would be the observation of neutrinoless double β-decay $(\mathrm{Z}, \mathrm{A}) \rightarrow (\mathrm{Z}+2, \mathrm{A}) + 2e^-$, with Z the atomic number and A the mass number, which cannot occur if the neutrinos are of Dirac-type. The level of required experimental precision is however very high, and no definitive answer was found so far. Even the mass ordering of the three know neutrino families is not known. Neutrino oscillation experiments have shed light on their non-zero masses, but they are unfortunately sensitive to squared-mass differences only. It is not currently clear which of the neutrinos is the heaviest, and which one the lightest.

- **Dark energy**: The Universe has been shown to be in a state of accelerated expansion [71]. The concept of *dark energy* is invoked to explain such phenomenon, as it introduces a repulsive force that counteracts the gravitational attraction, representing more than two thirds of the total energy in the present day observable Universe. This quantity can be found in general relativity, but is absent from the SM.

- **Unnaturalness of the Higgs mass**: The Higgs boson mass seems to be precisely and unnaturally fine-tuned to have a much lower value than what would be expected from arguments coming from Quantum Mechanics. Indeed, quantum corrections from Higgs interactions with other particles should place m_{H} at the level of the Planck scale, i.e. 17 orders of magnitude larger than its measured value.

Regardless of what may lie ahead, we are confident in saying that sometime, somewhere, the SM's resilience will crack, hinting at what lies beyond. In this Thesis we took the experimental risk of assuming the answer can be found in double Higgs production. Even if proved wrong, our choice will allow future work to get closer to the next big discovery.

1.3.2 Non-resonant BSM HH Production

BSM signatures can naturally be probed in nonresonant modes, i.e. production modes not created by a resonance, which in this context would decay to a pair of Higgs bosons. Physicists look for deviations from SM predictions. Causes for such deviations are substantiated by many proposed BSM theories, but experimentally amount to either a change in the HH production rate, and/or the modification of the shape of some kinematical distributions.

Often, instead of directly measuring the couplings, results are instead expressed as a function of *coupling modifiers* κ. The modifiers compare the measured couplings with SM expectations, and parameterize NP effects. This approach is called the κ *framework*. For the Higgs self-coupling, the definition of the modifier is $\kappa_\lambda \equiv \lambda_{\mathrm{HHH}}/\lambda^{\mathrm{SM}}_{\mathrm{HHH}}$, where λ_{HHH} is fully determined in the SM once v and m_{H} are measured. BSM theories are thus characterized by $\kappa_\lambda \neq 1$. A variable constraining power is associated to each production mode, depending on the phase-space region being studied. Other coupling modifiers are similarly defined, namely $\kappa_{\mathrm{V}} \equiv c_{\mathrm{V}}/c^{\mathrm{SM}}_{\mathrm{V}}$, for the coupling between two Higgs bosons and

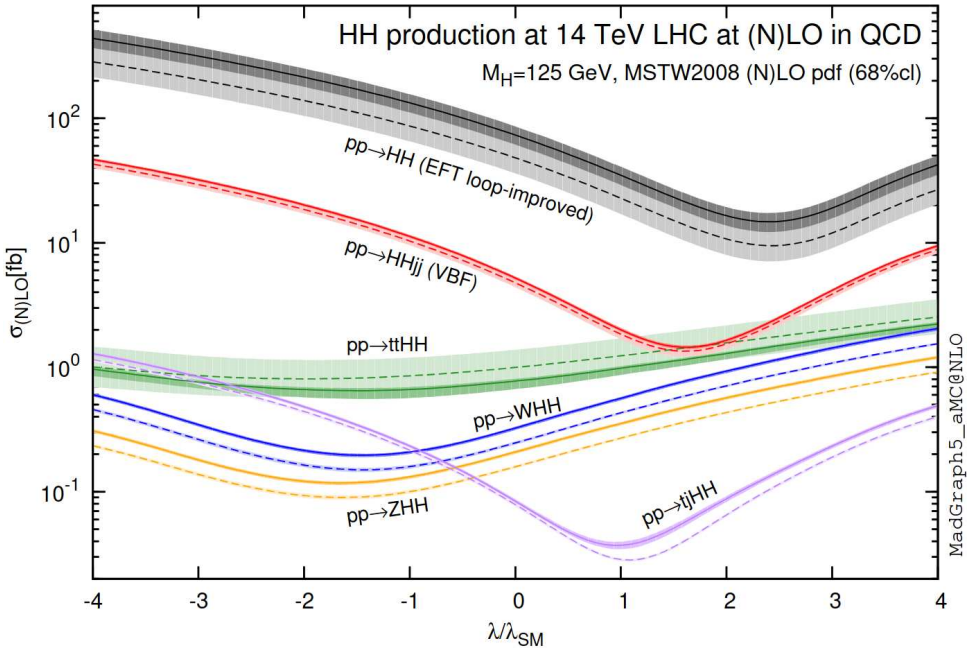

Figure 1.16: HH production cross section as a function of the coupling modifier κ_λ for several production mechanisms. The dashed and solid lines denote respectively the leading order (LO) and next-to-leading order (NLO) predictions and the bands indicate the parton distribution function (PDF) and scale uncertainties added linearly. The interference minima are not aligned for different production modes. For ggF the cross section is now known at NNLO level with finite top quark mass effects, while the figure displays the values for the NLO FTapprox calculation. The figure is taken from [61].

one vector boson, and $\kappa_{2V} \equiv c_{2V}/c_{2V}^{SM}$, a quartic coupling between two Higgs bosons and two vector bosons, which at first order can only be studied with VBF. The coupling modifiers alter the relative abundance of production modes and final states. In Fig. 1.16 we display the running of the HH cross section for several production mechanisms, as a function of κ_λ. The minima are due to diagram interferences for each production mode, as discussed in Section 1.2.1. Minima notably occur at different κ_λ values for different production modes; $\kappa_\lambda = 2.45$ for ggF. In Fig. 1.17 we show how the cross section of single Higgs production modes (left) and decays (right) are affected by a κ_λ parametrization. Then, using the recipe described in [79], we extend the computation to cover the HH branching ratio (BR) dependence on κ_λ, as displayed in Fig. 1.18 (right). The left plot in Fig. 1.18 again refers to single Higgs production and was obtained to validate the procedure; it should be compared to the right plot in Fig. 1.17. We used the assumption that the HH BR can be obtained as the product of the individual single Higgs boson BRs, which holds in the κ-framework. Relative changes δ in the BR are the same for all fermions, which are identified by the letter f. We can observe that final states of Higgs boson pairs can have a sizable dependence on κ_λ, at times than 5%. However, as we will see in Section 1.4, current κ_λ constraints are often significantly tighter than the range displayed in the plots; the relevant variations thus are of at most $\sim$2 % given current constraints. Gluon fusion HH cross section deviations can be parameterized as

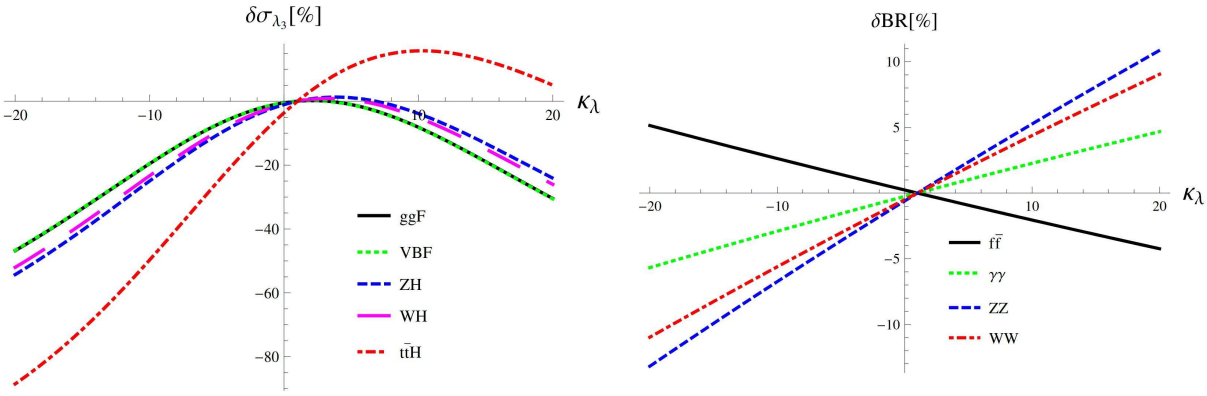

Figure 1.17: (Left) Single Higgs cross section dependence on κ_λ, for the five most common production modes, namely ggF, VBF, associated production with split contributions from the W and Z boson, and $t\bar{t}H$. (Right) Single Higgs BR dependence on κ_λ, showing the Higgs couplings to fermions ($f\bar{f}$), photons ($\gamma\gamma$) and vector gauge bosons (W and Z). Taken from [79].

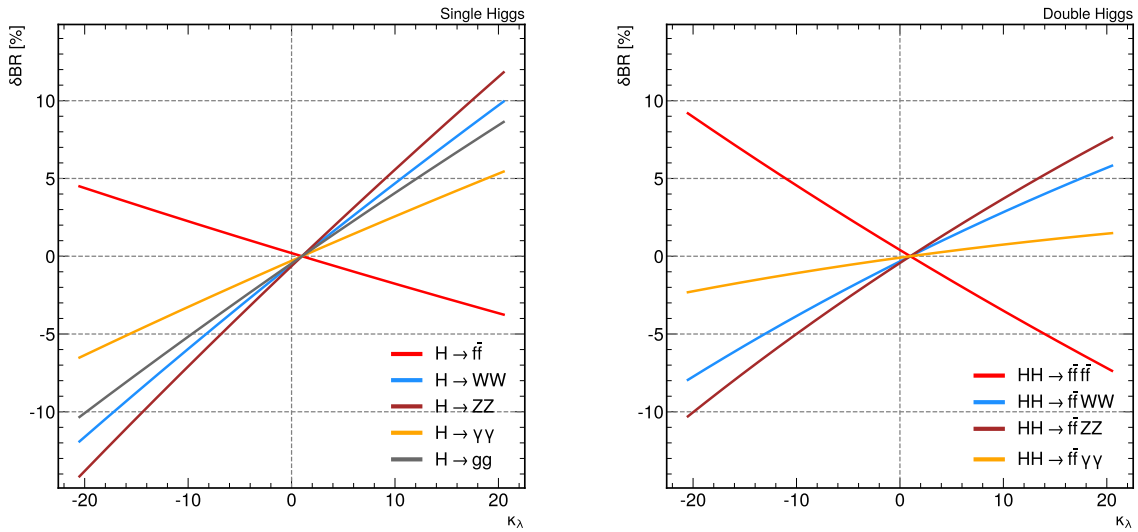

Figure 1.18: Branching fraction deviations, in percentage, of single (left) and double (right) Higgs processes as a function of the HH coupling modifier κ_λ. Different decays are included, where "g" stands for gluon and "f" for fermion. The single Higgs process is included for validation only, and matches the right plot of Fig. 1.17. The deviations are calculated following the procedure detailed in [79].

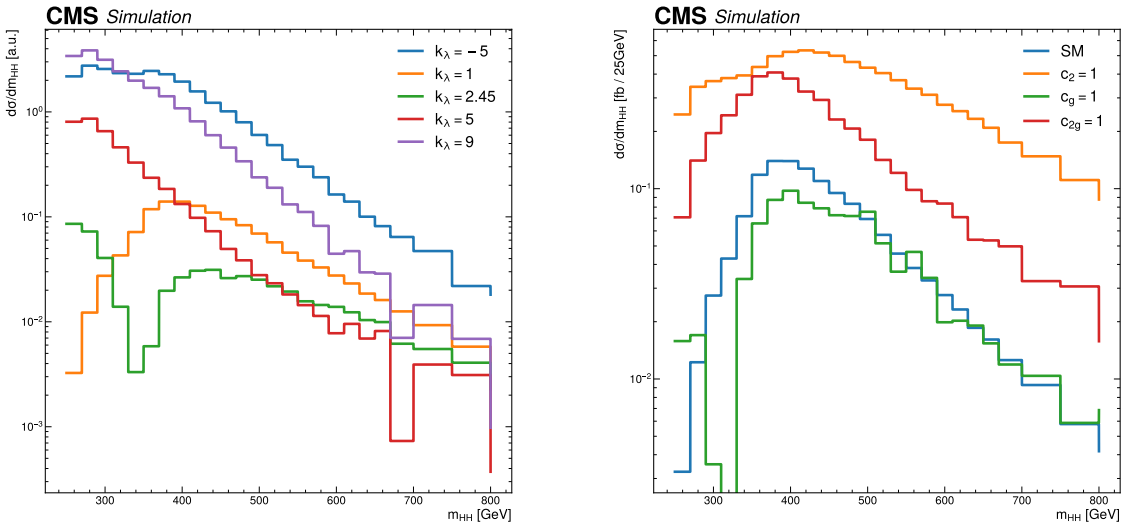

Figure 1.19: Differential HH cross section as a function of the HH system mass, for different values of coupling modifiers. (Left) m_{HH} shape dependence on κ_λ, including SM ($\kappa_\lambda = 1$) and maximum interference ($\kappa_\lambda = 2.45$), highlighting the strong impact of a deviation from the expected SM values. Larger $|\kappa_\lambda|$ values correspond to scenarios where the HH "triangle" diagram dominates. (Right) Dim-6 ggF HH Effective Field Theory (EFT) couplings set to one, compared to the SM scenario. Both the overall rate and shapes are modified.

a function of κ_λ and $\kappa_\mathrm{t} \equiv y_t/y_t^{\mathrm{SM}}$ as follows:

$$\sigma_{\mathrm{ggF}}/\sigma_{\mathrm{ggF}}^{\mathrm{SM}} \sim |\mathcal{A}_\triangle|^2\kappa_\lambda^2\kappa_\mathrm{t}^2 + |\mathcal{A}_\square|^2\kappa_\mathrm{t}^4 + \mathcal{I}_{\triangle\square}\kappa_\lambda\kappa_\mathrm{t}^3\,,$$

$$\sigma_{\mathrm{ggF}}/\sigma_{\mathrm{ggF}}^{\mathrm{SM}}\Big|_{\sqrt{s}=13\,\mathrm{GeV}} \sim 0.28\ \ \kappa_\lambda^2\kappa_\mathrm{t}^2 + 2.09\ \ \kappa_\mathrm{t}^4 - 1.37\kappa_\lambda\kappa_\mathrm{t}^3\,, \tag{1.50}$$

where $\mathcal{A}_\triangle$ and $\mathcal{A}_\square$ are the amplitudes of the triangle and box ggF diagrams, respectively, and $\mathcal{I}_{\triangle\square}$ represents the interference term. We consider gluon fusion due to its dominant cross section, but other production modes can be parameterized in similar ways using their corresponding coupling modifiers.

Variations of the coupling modifiers can have a enormous impact on the shape of kinematic distributions. Taking the mass of the HH system in Fig. 1.19 (left) as an example, we can see that the $\kappa_\lambda = 1$ configuration is characterized by a broad peak at ≈390 GeV, with a large high-energy tail. Importantly, the kinematics are completely altered by smaller or larger κ_λ values. Given the couplings present in the box and triangle diagrams, low $|\kappa_\lambda| \to 0$ is box-dominated, while large $|\kappa_\lambda|$ values are dictated by the triangle diagram. The most eye-catching variations occur for $\kappa_\lambda = 2.45$, where the interference is maximal, and the distribution goes to zero at ≈350 GeV. Despite the focus on ggF, it is worth noting that other production modes also include interferences between different diagrams, as one could already infer from Fig. 1.16. The consequence is clear: kinematic deviations represent a smoking gun for the presence of anomalous couplings.

While the κ-framework has been so far effective in quantifying differences with respect to the SM, it also bears some drawbacks:

- The involved theoretical calculations are rather complex, and become impractical for high loop orders. At same time, uncertainties below ∼10 % require NLO calculations.
- The λ_{HHH} is fully determined in the SM, and the modification of the SM with a ratio introduces an inconsistency in the theory.
- The framework cannot be used to interface results from the LHC and the Large Electron–Positron Collider (LEP), given the different energies involved.
- It cannot be straightforwardly extended to the more consistent and general approach we present in the following Section.

Effective field theories

No resonance has so far been found in the energy range covered by the LHC, i.e. up to a few TeVs. Even assuming that all possible phase-space corners have been exhausted, which is certainly not the case, it would still be possible to imagine the presence of NP at higher energies, capable of indirectly influencing the shape of distribution tails which detectors have access to. A general and systematic method has been developed to look at high-energy effects from an unknown high-energy theory in a model independent way. The idea consists in extending the dimension-4 (dim-4) SM Lagrangian with higher order operators. The extra terms are appropriately suppressed by powers of an energy scale Λ, oftentimes chosen at 1 TeV, constructing an EFT Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda}\mathcal{O}_i^5 + \sum_i \frac{c_i'}{\Lambda^2}\mathcal{O}_i^6 + \sum_i \frac{c_i''}{\Lambda^3}\mathcal{O}_i^7 + \dots , \tag{1.51}$$

where all BSM physics are parameterized by the Wilson coefficients c_i, c_i', c_i'', etc. The scale parameter Λ can be interpreted as the energy up to which only SM fields propagate, but also as the BSM scale itself. The additional terms make the theory non-renormalizable, but that is not an issue, as the EFT expansion does not represent a complete theory, but rather a low-energy approximation of an unknown, complete and renormalizable theory at a higher energy scale.

From an experimental point of view, one has to define the scale up to which to extend the effective theory, knowing that with each additional term a very significant number of coefficients might be added. A BSM can then be matched to the truncated expansion, so that an expression for each coefficient is found, depending on the parameters of the corresponding model. This approach implies that different models can be realized and compared in terms of common Wilson coefficients, as for instance 2HDM or composite Higgs models. In the most general EFT formulation, there is a single dim-5 operator, which introduces Majorana-like masses for the neutrinos, and is neglected in the HH context. Next, approximately 2800 dim-6 operators exist, which reduce to less than 100 by imposing certain restrictions. For instance, in Standard Model EFT (SMEFT)[4], one of the EFT approaches often explored, one demands additional flavour symmetries and CP-conservation to reduce the number of operators to 59. Operators with higher dimensions

[4] As part of my PhD, I gave a talk on SMEFT fitting techniques within the ATLAS, CMS and LHCb experiments, at the "QCD@LHC2022" conference, which took place in Orsay, France, from November 28$^{\mathrm{th}}$ to December 2$^{\mathrm{nd}}$ 2022 [6].

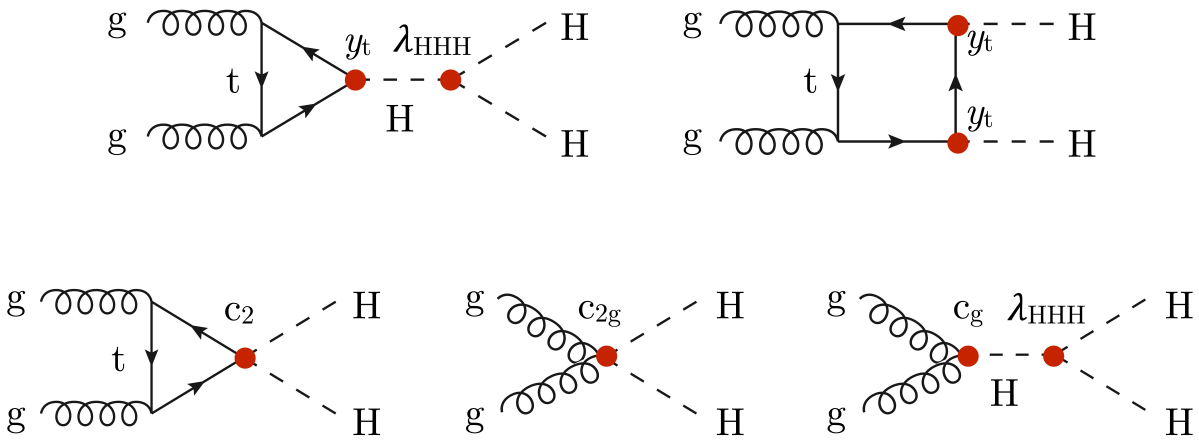

Figure 1.20: Leading order Feynman diagrams in the ggF HH EFT description, at dimension 6, for ggF production mechanism [83].

are very suppressed and usually neglected. However, some dedicated studies for dim-8 operators have recently been pursued [80], since EW quartic gauge couplings can be studied via VBF processes, where dim-8 operators are used to parameterize deviations from the SM prediction [81]. Departures from SM triple and quartic couplings lead to the growth of scattering amplitudes, signaling NP [82].

For gluon fusion HH production, a possible EFT Lagrangian can be constructed out of all dim-6 gauge-invariant operators [84], with t representing the top quark and $G_{\mu\nu}$ the eight gluons:

$$\begin{aligned}\mathcal{L}_{\text{EFT}}^{\text{HH}} = &\frac{1}{2}\partial^{\mu}\text{H}\partial_{\mu}\text{H} - \frac{1}{2}m_{\text{H}}^2\text{H}^2 + \kappa_{\lambda}\lambda_{\text{HHH}}v\text{H}^3 \\ &- \frac{m_{\text{t}}}{v}\left(v + \kappa_{\text{t}}\text{H} + \frac{c_2}{v}H^2\right)(\bar{\text{t}}_{\text{L}}\text{t}_{\text{R}} + \text{t}_{\text{R}}\bar{\text{t}}_{\text{L}}) \\ &+ \frac{\alpha_{\text{S}}}{12\pi v}\left(c_{\text{g}}\text{H} - \frac{c_{2\text{g}}}{2v}\text{H}^2\right)G_{\mu\nu}^{a}G_{a}^{\mu\nu}\,,\end{aligned} \tag{1.52}$$

where new c_{g}, $c_{2\text{g}}$ and c_2 BSM couplings appear, as illustrated in Fig. 1.20. To the SM box and triangle diagrams a set of three new diagrams is added, displaying couplings between the Higgs and quarks or gluons. Just like in Eq. (1.50), the cross section can be parameterized from the square of the amplitude, which at LO takes the form:

$$\begin{aligned}R_{\text{HH}} = \frac{\sigma_{\text{HH}}}{\sigma_{\text{HH}}^{\text{SM}}} = &\text{A}_1\kappa_{\text{t}}^4 + \text{A}_2c_2^2 + \text{A}_3\kappa_{\text{t}}^2\kappa_{\lambda}^2 + \text{A}_4c_{\text{g}}^2\kappa_{\lambda}^2 + \text{A}_5c_{2\text{g}}^2 \\ &+ \text{A}_6c_2\kappa_{\text{t}}^2 + \text{A}_7\kappa_{\lambda}\kappa_{\text{t}}^3 + \text{A}_8\kappa_{\text{t}}\kappa_{\lambda}c_2 + \text{A}_9^i c_{\text{g}}\kappa_{\lambda}c_2 \\ &+ \text{A}_{10}c_2c_{2\text{g}} + \text{A}_{11}c_{\text{g}}\kappa_{\lambda}\kappa_{\text{t}}^2 + \text{A}_{12}c_2\kappa_{\text{t}}^2 \\ &+ \text{A}_{13}\kappa_{\lambda}^2c_{\text{g}}\kappa_{\text{t}} + \text{A}_{14}c_{2\text{g}}\kappa_{\text{t}}\kappa_{\lambda} + \text{A}_{15}c_{\text{g}}c_{2\text{g}}\kappa_{\lambda}\,,\end{aligned} \tag{1.53}$$

where the SM parameterisation is recovered for $c_2 = c_{\text{g}} = c_{2\text{g}} = 0$. Further precision can be obtained by introducing quantum corrections at NLO or higher. In Fig. 1.19 (right), we show the impact some of the new couplings have on HH kinematic distributions. The $[A_1, ..., A_{15}]$ coefficients are determined from a simultaneous fit of a simulated HH LO cross section [85]. Even limiting ourselves to a simplified EFT parameterisation with only three additional couplings leads to very significant complications from an experimental,

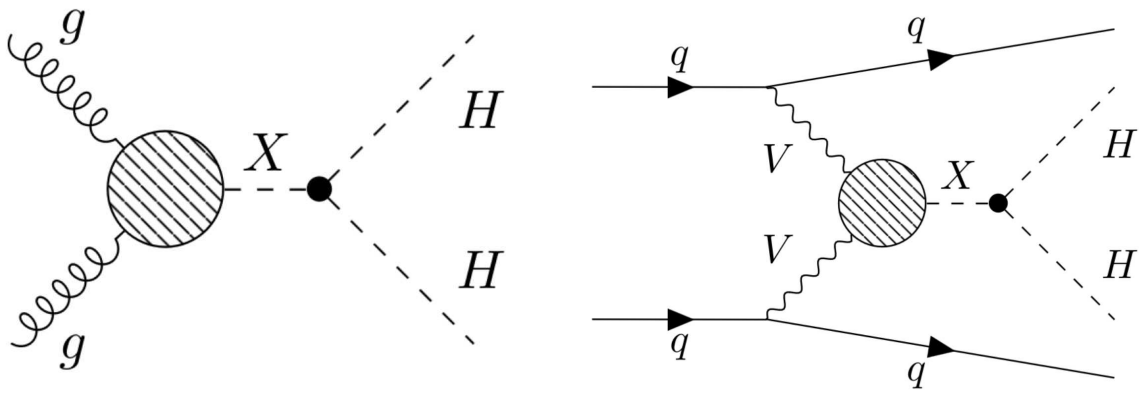

Figure 1.21: Feynman diagrams for hypothetical resonant processes contributing to double Higgs boson pair production via ggF (left) and VBF (right).

and especially computational point of view. To study the entire phase-space, samples would have to be generated to cover the five HH couplings in a 5D space, considering all possible combinations. Fortunately, kinematics vary smoothly enough between points that the definition of specific benchmarks is enough to adequately sample the entire phase-space. These *shape benchmarks* are calculated by scanning a large sample of 5D points, clustering them into regions of kinematic similarities, defined from a binned likelihood ratio test [86]. The clustering procedure also guarantees a meaningful extrapolation between different benchmarks.

1.3.3 Resonant BSM HH Production

The simplest imaginable signature of NP consists in the production of an s-channel resonance decaying to a pair of SM particles. In the context of HH searches, and in addition to the nonresonant approaches discussed in Section 1.3.2, BSM physics can be probed under the assumption that there indeed exists some new particle X with a mass within what the LHC can produce, decaying to a pair of Higgs bosons, as shown in Fig. 1.21 for ggF and VBF processes. In that case, given enough experimental sensitivity, we would naively expect to see a "bump" in double Higgs boson mass distributions. In practice, admitting such hypothetical particle exists, the excess might only be discernible through variables with a higher discriminative power, such as discriminators built out of Machine Learning (ML) procedures. The experimental challenges of such kind of analysis is highlighted in Chapter 4. Here we instead focus on the very large number of theoretical models which can accommodate $X \to HH$ or $X \to YH$ processes, with Y being yet another BSM scalar particle. It is experimentally unfeasible to address all such models separately. If that were the case, each model would require its own set of simulated data with varying parameters, such as the mass or the shapes of every resonance, adding also complications from the storage point of view. Furthermore, analysis tools would have to be optimized for each sample type, requiring independent computing-intensive trainings, which take time to process and validate. Instead, the approach usually followed, including in this Thesis, is to define a set of (as much as possible) model-independent benchmarks. Results can be subsequently reinterpreted in phenomenological studies, in the context of more specific models. This enables a single Physics analysis to explore a broad class of models at once. The simplest model possible is the singlet model, which is covered in Section 6.2, since it is directly connected with the work discussed in Chapter 6. In this

Section we instead discuss other options, which serve as a strong motivation to perform the HH searches described in further Chapters.

Warped Extra Dimensions

Generations of physicists have tried to incorporate quantum gravity in the SM without much success so far. Despite its negligible effect at the energy scales probed in our experiments, the gravitational force is bound to become significant at close range, or alternatively at very high energy, near the 10^{19} GeV Planck scale. One of the theories that tries to unify the four known forces is string theory, where the basic units of matter are multi-dimensional *branes* instead of the zero-dimensional particles present in the SM. The idea of additional dimensions was first introduced in the Kaluza-Klein model, which sought to extend general relativity to a five (4+1) dimensional space (four space dimensions and one time dimension). There, the existence of additional microscopic space dimensions was theorized. Those dimensions would be *compactified*, or equivalently "curled up" and hidden from the macroscopic scale. Multiple theories use the concept of extra dimensions (EDs) to address the shortcomings of the SM. Whenever those theories are based on distorted space-time geometries, they are referred to as WED.

Among WED models, one benchmark often used for LHC experiments is the Randall-Sundrum model [87]. In this model, two branes are defined in a five-dimensional space, with one being the SM "TeV" brane where our (3+1)-dimensional Universe is located, and the other corresponding to the Planck scale. The region between the two branes is denoted the "bulk". The Universe thus described includes an extra compactified dimension, corresponding to a line segment between the two branes. The most general solution of Einstein's equations leads to an exponential metric on the fifth dimension, bounded by the two branes. The model helps explaining the surprising difference in scales between gravity and other forces. Instead of an extreme and unnatural fine-tuning, gravity could be concentrated in specific regions of the warped space-time, leading to a very weak version of gravity in our Universe, but stronger elsewhere. Furthermore, the Higgs mechanism can be added in the same way as in the SM, confining the added doublet in Eq. (1.31) to the TeV brane.

Importantly, the model predicts two new particles [88]. A spin-2 resonance appears as the mediator of the gravitational force, appropriately named "graviton", and a new scalar is required by the size of the additional dimension, being called "radion". Assuming the Narrow Width Approximation (NWA), where the width of the new resonance is assumed to be negligible compared to the resonance's mass, the phenomenology of a ggF-produced radion is identical to the phenomenology of a heavy Higgs scenario [88], which lessens model-related dependencies. For certain regions of the model's phase-space, both particles can decay to a pair of Higgs bosons or pairs of W and Z bosons. Searches at the LHC are thus sensitive to these two hypothetical high mass particles. The Thesis here reported explores the resonant $HH \to bb\tau\tau$ decay, where the signal samples in Section 4.1 indeed refer to radion- and graviton-like particles. The radion and the Higgs boson can even be allowed to mix, although such scenario is not directly explored at analysis level. Related work in the context of a singlet model is covered in Chapter 6, where the NWA is covered in greater detail.

Two Higgs Doublet Model

The 2HDM, as its name indicates, adds a second doublet field to the one already introduced by the Higgs mechanism [89]. This *ad-hoc* choice is strongly motivated by multiple reasons. The simplest version of SUSY, the Minimal Supersymmetric Standard Model (MSSM), cannot give mass to all quarks without the introduction of a second doublet [90]. One can thus say that the phenomenology of SUSY is described by the 2HDM. Additionally, some axion models, especially some not yet ruled out by experiments, require two doublets for their effective low-energy theory to function [91]. The 2HDM can also be made to accommodate the required baryon asymmetry to explain baryogenesis, with explicit or spontaneous CP violation [92].

Following a similar recipe to what was done in the SM in Eq. (1.31), we add two complex scalar SU(2) doublets with eight fields in total:

$$\phi_a = \begin{pmatrix} \phi_a^+ \\ \phi_a^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_a^1 + i\phi_a^2 \\ \phi_a^3 + i\phi_a^4 \end{pmatrix}, \quad \phi_b = \begin{pmatrix} \phi_b^+ \\ \phi_b^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_b^1 + i\phi_b^2 \\ \phi_b^3 + i\phi_b^4 \end{pmatrix}, \tag{1.54}$$

At this point, a series of simplifications are usually done. CP is assumed to be conserved in the Higgs sector and not spontaneously broken, and some discrete symmetries are imposed to eliminate quartic terms odd in one of the two doublets, like $\phi_a^\dagger \phi_b \phi_b^\dagger \phi_b$. The most general potential one can then write is the following:

$$\begin{aligned} V(\phi_a, \phi_b) = {} & m_{aa}^2 \phi_a^\dagger \phi_a + m_{bb}^2 \phi_b^\dagger \phi_b - m_{ab}^2 \left(\phi_a^\dagger \phi_b + \phi_b^\dagger \phi_a \right) + \frac{\lambda_a}{2} (\phi_a^\dagger \phi_a)^2 + \frac{\lambda_b}{2} (\phi_b^\dagger \phi_b)^2 \\ & + \lambda_c \phi_a^\dagger \phi_a \phi_b^\dagger \phi_b + \lambda_d \phi_a^\dagger \phi_b \phi_b^\dagger \phi_a + \frac{\lambda_e}{2} \left[(\phi_a^\dagger \phi_b)^2 + (\phi_b^\dagger \phi_a)^2 \right], \end{aligned} \tag{1.55}$$

where all parameters m and λ are real. Three out of the eight scalar fields are the Goldstone bosons that give mass to the W and Z bosons, and the remaining five fields correspond to five physical Higgs bosons, one of them being identified with the scalar boson observed at the LHC. Two of the fields are CP-even and neutral scalars ρ, two are charged scalar particles $\phi^\pm$, one being the antiparticle of the other, and the last is a CP-odd neutral scalar η [69]. Given the known mass of the Higgs and its VEV, 2HDM models have six free parameters. Using the minima v_a and v_b of the potential in Eq. (1.55):

$$\langle \phi_a \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_a \end{pmatrix}, \quad \langle \phi_b \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_b \end{pmatrix}, \tag{1.56}$$

one can write the mass-squared terms of the Lagrangian, which for the charged scalars take the following form:

$$\mathcal{L}_{\text{mass}}^{\phi^\pm} = \left[m_{ab}^2 - (\lambda_d + \lambda_e) v_a v_b \right] \begin{pmatrix} \phi_a^- & \phi_b^- \end{pmatrix} \begin{pmatrix} v_b/v_a & -1 \\ -1 & v_a/v_b \end{pmatrix} \begin{pmatrix} \phi_a^+ \\ \phi_b^+ \end{pmatrix}, \tag{1.57}$$

where the mass-squared of the additional "charged Higgs" bosons is given by $m_{\text{H}^\pm}^2 = [m_{ab}^2/(v_a v_b) - \lambda_4 - \lambda_5](v_a^2 + v_b^2)$, and the other eigenvalue coming from the diagonalization is zero and corresponds to a charged Goldstone boson, which is eaten by the $\text{W}^\pm$. The term for the pseudo-scalars is given by:

$$\mathcal{L}_{\text{mass}}^{\eta} = \frac{m_{\text{A}}^2}{v_a^2 + v_b^2} \begin{pmatrix} \eta_a & \eta_b \end{pmatrix} \begin{pmatrix} v_b^2 & -v_a v_b \\ -v_a v_b & v_a^2 \end{pmatrix} \begin{pmatrix} \eta_a \\ \eta_b \end{pmatrix}, \tag{1.58}$$

with $m_{\mathrm{A}}^2 = [m_{ab}^2/(v_a v_b) - 2\lambda_e](v_a^2 + v_b^2)$. The diagonalization also provides a massless pseudo-scalar, which is eaten by the Z boson. Finally, the mass-squared terms for the scalars are given by:

$$\mathcal{L}_{\mathrm{mass}}^{\rho} = -\begin{pmatrix}\rho_a & \rho_b\end{pmatrix}\begin{pmatrix} m_{ab}^2(v_b/v_a) + \lambda_a v_a^2 & -m_{ab}^2 + \lambda_{cde} v_a v_b \\ -m_{ab}^2 + \lambda_{cde} v_a v_b & m_{ab}^2(v_a/v_b) + \lambda_b v_b^2 \end{pmatrix}\begin{pmatrix}\rho_a \\ \rho_b\end{pmatrix}, \tag{1.59}$$

where $\lambda_{cde} = \lambda_c + \lambda_d + \lambda_e$. The mass-squared matrix can also be diagonalized, with an additional parameter α representing the corresponding rotation angle. Finally, a second angle β can be defined, by diagonalizing together the mass-squared matrices of the charged scalars and pseudo-scalars. An important relationship with the two VEVs is defined, often used for parameterizations:

$$\tan\beta = \frac{v_b}{v_a}\,. \tag{1.60}$$

The two angles determine all interactions between the five "Higgs" fields and the vector gauge bosons and fermions, the latter by also adding Yukawa-like terms. In the end, the model predicts mass-diagonalized H and X CP-even neutral bosons (where H corresponds to the SM Higgs boson), a CP-odd pseudo-scalar A, and two charged Higgs bosons $\mathrm{H}^{\pm}$.

One should not forget that the potential in Eq. (1.55) made some assumptions which, if relaxed, reveal even more complexity. For instance, Flavour Changing Neutral Currents (FCNCs) become possible at tree-level, despite being strongly constrained by experiments. Scenarios with natural flavour conservation are currently preferred. One can also require each Higgs doublet ϕ to only couple to charged fermions of one type, either up-type quarks, down-type quarks, or charged leptons. Adding on top the requirement of no FCNCs, as in the mass terms above, four types of 2HDMs emerge, based on which doublets couple to which fermion types:

- Type I: all charged fermions couple to ϕ_b;
- Type II: only up-type quarks couple to ϕ_b;
- Type X (lepton specific): only up-type and down-type quarks couple to ϕ_b;
- Type Y (flipped): only up-type quarks and charged leptons couple to ϕ_b.

Further types can be defined based on FCNCs-related requirements.

Current constraints from EW precision data severely limit the allowed phase-space of the six free parameters. For one, scenarios with mass degeneracy are preferred, where the masses of at least two of the additional Higgs bosons should be very similar [93]. Flavour observables also lead to strong constraints, in particular to lower bounds on $m_{\mathrm{H}^{\pm}} \sim 600\,\mathrm{GeV}$ in Type II and Type Y models. Measurements of Higgs boson couplings has so far constrained $\cos(\beta-\alpha)$ to small values, where $\cos(\beta-\alpha) \to 0$ is the *alignment limit* in which the boson H becomes SM-like. The couplings of the heavy Higgs bosons also depends on $\cos(\beta-\alpha)$, and for 2HDMs models without mass degeneracy, decays of the type $\mathrm{A} \to \mathrm{ZX}$ and $\mathrm{X} \to \mathrm{ZA}$ exist, with large BRs that are not suppressed in the alignment limit. Searches performed at CMS exclude high $\tan\beta$ values [94]. In Fig. 1.22 we show $\mathrm{X} \to \mathrm{HH}$ BR examples for Type I and II 2HDMs.

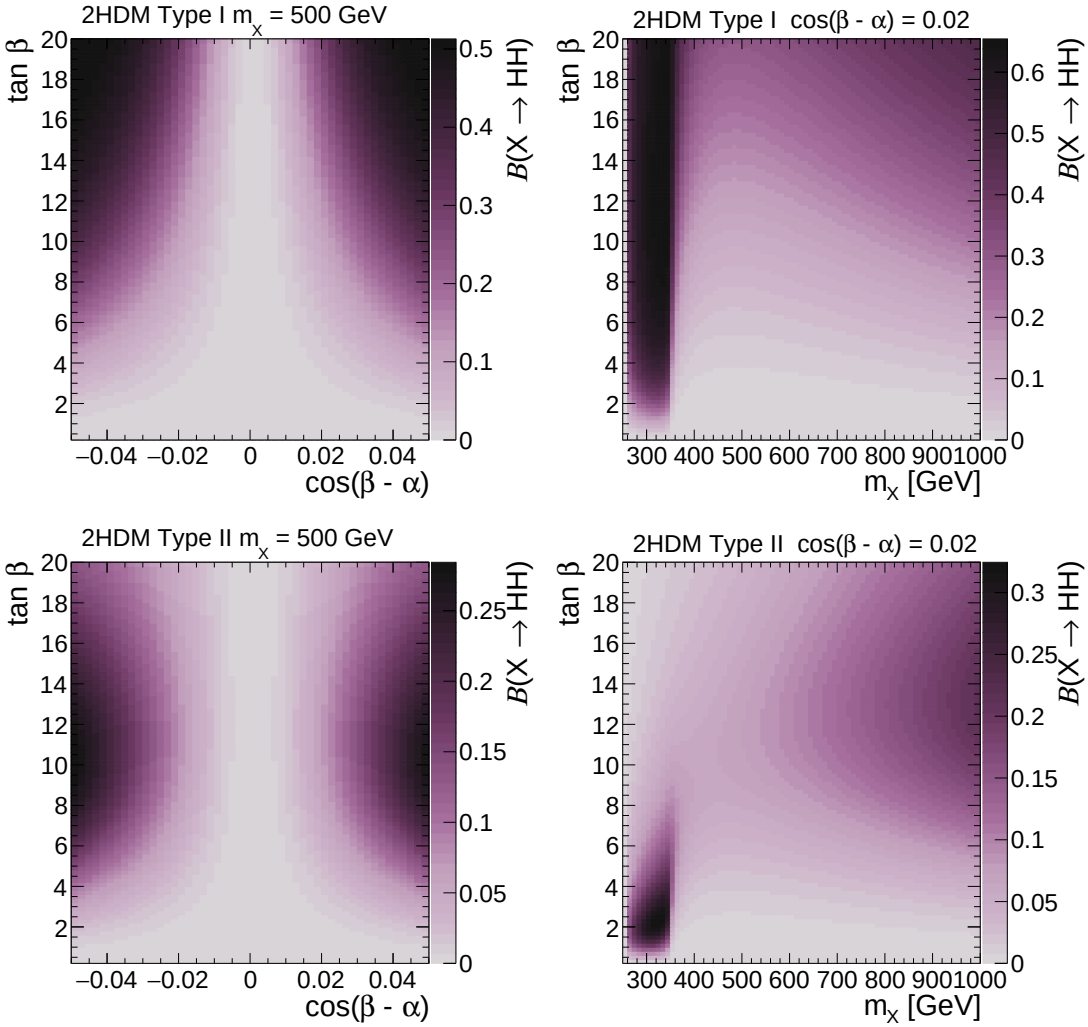

Figure 1.22: Branching fractions of $X \to HH$ decays in 2HDMs of Type I (top) and Type II (bottom) in the $\cos(\beta - \alpha)$ vs. $\tan\beta$ plane for $m_X = 500\,\text{GeV}$ (left) and in the m_X vs. $\tan\beta$ plane for $\cos(\beta - \alpha) = 0.02$ (right). The branching fractions have been calculated with `2HDMC v1.8.0` [95, 96]. Taken from [7].

Additional singlets and doublets

Many extensions of the 2HDM have been explored. One option is to simply add an extra real or complex singlet, creating the next-to-minimal 2HDM (N2HDM) or 2HDM+S models [97–99]. Requiring CP conservation, the N2HDM leads to six new Higgs bosons, adding four free parameters to the 2HDM [100]. The same types as in 2HDM can be defined based on the added Yukawa couplings. The phenomenology of N2HDMs is also very similar to the 2HDM one, but the presence of an additional scalar Y enables new decay topologies on top of $X \to HH$, namely $X \to YY$ and $X \to YH$, the latter being often searched by ATLAS and CMS (see, for instance, [101]). These new topologies are not suppressed in the alignment limit, and can thus dominate as along as they are kinematically permitted. Moreover, the 2HDM+S model is also experimentally consistent with the $X \to YH$ decay. Just like the MSSM is linked to the 2HDM, the next-to-minimal MSSM (NMSSM) corresponds to a 2HDM+S model of Type II. Importantly, however, SUSY models tend to be more strongly constrained by the measurements.

One may continue adding doublets and singlets, for ever richer phenomenologies. A good example is the 3HDM and its variations (for instance, [102]). These models go however well beyond the scope of this Thesis.

1.4 Experimental Status of HH Physics

As I write these lines, data-taking at the LHC is on-going, and many results will be complemented by updated analysis iterations at 13.6 TeV, exploiting more data and more performant experimental techniques. This will also be the case for HH resonant and nonresonant processes. Nevertheless, one cannot forget the plethora of results Run 1 and Run 2 provided, consistently beating experimental expectations. Many searches for Higgs boson pair production have been performed by the CMS and the ATLAS experiments. Direct measurements enable to probe λ_{HHH} using HH production, and the current experimental status is covered in Section 1.4.1. There is also the option to investigate the Higgs self-coupling indirectly, namely via higher-order loop. This idea is described in Section 1.4.2. Finally, the direct and indirect approaches can be combined, as detailed in Section 1.4.3.

We briefly mention that some analyses are starting to probe λ_{HHHH}, which can be done via the production of three Higgs bosons. There is no experimental guarantee that $\lambda_{HHH} = \lambda_{HHHH}$, with vital consequences for the shape of the Higgs potential in Eq. (1.39), and ultimately for our Universe, as covered in Section 6.1. The shape of the Higgs potential can really be only fully determined from a combined measurement of both HH and HHH processes. Given the correlations between λ_{HHH} and λ_{HHHH}, any hypothetical κ_λ deviation will provide information on the quartic coupling [103]. Despite the vanishingly small cross section at $\sim$0.1 pb and smaller branching ratios ($\mathcal{B}(HH \to bbbb) \approx 0.33$ while $\mathcal{B}(HHHH \to bbbbbb) \approx 0.19$, for instance), triple production can add some constraining power to κ_λ without requiring the quartic Higgs coupling modifier to be set to one in HH + HHH combinations. CMS is currently initiating studies with the bbbbbb and bbbb$\gamma\gamma$ final states. However, given current sensitivity constraints, HHH analyses can become relevant in future colliders only.

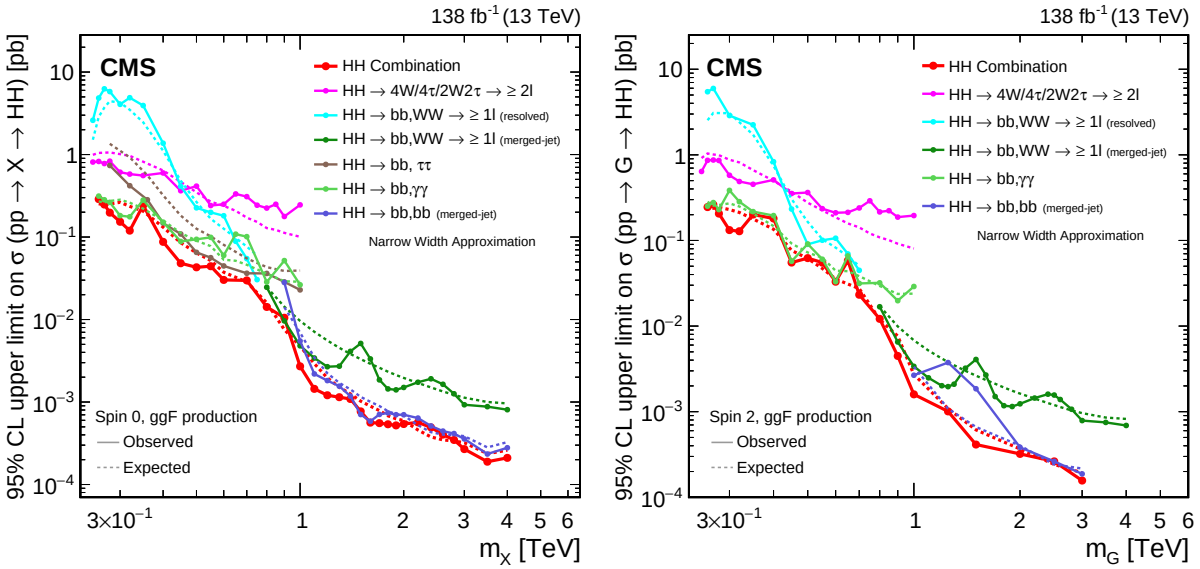

Figure 1.23: Searches for Higgs boson pairs resonant production. Observed (solid lines) and expected (dashed lines) 95% CL upper limits on $\sigma\mathcal{B}$ for the ggF production of a spin-0 resonance X (left) and a spin-2 resonance G (right). Individual results and statistical combinations are shown, where observed limits are indicated with solid lines and expected limits with dashed lines. Taken from [7].

1.4.1 Direct Searches

The κ_λ modifier is currently one of the least constrained properties of the SM. While mentioning some results from the ATLAS Collaboration in Appendix A.1, in the following we provide an overview of the most up-to-date resonant and nonresonant CMS results involving the direct production of two Higgs bosons. Refer to Fig. 1.15 for a glimpse of all results so far produced, with the many explored final states. All results are compatible with the SM.

Resonant searches

The CMS Collaboration has very recently performed a combination of multiple Run 2 resonant analyses [7][5], considering both the $X \to HH$ and the $X \to YH$ topologies, with X and Y being BSM scalars which could reflect the theories described in Section 1.3.3. In the same work, interpretations are done for different theories, establishing exclusion limits on the model's parameters. Individual analyses extract 95% CL upper limits on the resonant HH production as a function of the resonance mass m_X. When YH limits are drawn, an additional dependence on m_Y is included. In total, six searches are considered, targeting a variety of final states. On the one hand, $X \to HH$ searches include the bbWW decay with resolved [104] and merged [105] topologies, and also a multilepton search [106]. On the other hand, the searches considered for the $X \to YH$ process are $X \to Y(bb)H(\tau\tau)$ [101], $X \to Y(bb)H(\gamma\gamma)$ [107] and $X \to HY \to bbbb$[108]. YH searches are included in the HH results by evaluating the Y scalar at the mass of the SM Higgs boson, when such a phase-space point is defined, or at a nearby point otherwise. All results use the NWA and SM Higgs branching fractions.

[5]This is the same study which includes the material presented in Chapter 6, and which I helped reviewing during my Thesis.

We show the $X \to HH$ combination in Fig. 1.23, where 95% CL $\sigma\mathcal{B}$ upper limits are drawn as a function of m_X, in a 0.28 to 4.5 TeV mass range. The left (right) plots show the limits considering spin-0 (spin-2) signals. The exclusion in terms of $\sigma\mathcal{B}$ goes down to ∼0.1 fb for both probed spin scenarios. The best sensitivity at low masses is obtained by the $X \to HH \to bb\gamma\gamma$ search, while at high masses the two searches with b-tagged merged jets dominate, mostly $X \to HH \to bbbb$. The $X \to HH \to bb\tau\tau$ result becomes more important at intermediate masses, demonstrating the complementarity between the three silver bullets. The results of the statistical combination are shown as solid red lines. No deviation larger than $2\,\sigma$ from the expected limits is observed. The combination brings good limit improvements in the intermediate m_X range, with respect to individual channels. This happens in phase-space regions where no single channel dominates the sensitivity. A recent combination of HH searches performed by ATLAS can be found in [109], where the spin-0 result uses the three most sensitive channels and obtains comparable results to CMS.

We show the $X \to YH$ combinations in Figs. 1.24 and 1.25, where 95% CL $\sigma\mathcal{B}$ upper limits are drawn as a function of m_Y, and different m_X limits are rescaled to make the curves visually distinct. The two plots differ on the m_X range, with the split happening at 1 TeV. No correction was made for the unknown $Y \to bb$ BR, which is the same for all analyses. At low m_X, $X \to Y(bb)H(\tau\tau)$ and $X \to Y(bb)H(\gamma\gamma)$ dominate, while for $m_X = 1$ TeV and higher, the merged $X \to HY \to bbbb$ provides the best sensitivity. This ceases to be true for high m_Y, since the larger mass implies a smaller Y boost, making the boosted analysis less performant; the $X \to Y(bb)H(\tau\tau)$ dominates once more. The exclusion upper limits tend to be of around 50, 5, and 0.3 fb for $m_X = 0.5$, 1, and 3 TeV, respectively.

The $X \to HH \to bb\tau\tau$ result obtained in this Thesis represents an enormous improvement over the first of such studies in CMS, which used only 2016 data [110]. This work also improves on the techniques of its nonresonant counterpart [111]. As reported in Section 5.4.1, the expected results are even better than the most recent CMS resonant combination at an intermediate mass range. A $X \to Y(bb)H(\tau\tau)$ has also been published by CMS [110], not obtaining limits for a 125 GeV Higgs boson. An interpolation was nevertheless performed and is shown in Fig. 1.23. Comparisons are discussed in Section 5.4.1.

Nonresonant searches

For nonresonant analyses, we start by mentioning the most sensitive decay channels, namely the "silver bullets", and focus on the most recent results for each channel. The ggF and VBF $HH \to bbbb$ processes have been studied in their boosted [112] and resolved [113] topologies, relative to their b-jets, the former managing to exclude $\kappa_{2V} = 0$ at 6.3 standard deviations. For the first time, associated production with a vector gauge boson was also considered [66]. Jointly studying ggF and VBF, $HH \to bb\tau\tau$ [111] and $HH \to bb\gamma\gamma$ [114] results were also produced with good performances.

The above channels drive the sensitivity of combinations, but are complemented by others, now mentioned. In the ggF production process only, CMS obtained results on $HH \to bbZZ$ [115] and on the extremely rare $HH \to \tau\tau\gamma\gamma$ decay ($\mathcal{B} \approx 0.028\%$) [116]. The latter is not used in the combinations reported below. Taking into account both ggF and VBF, CMS also has multilepton (HH decaying to WWWW, WW$\tau\tau$ and $\tau\tau\tau\tau$) [106], $HH \to bbWW$ [104] and $HH \to WW\gamma\gamma$ [117] analyses. The most up-to-date Run 2 HH

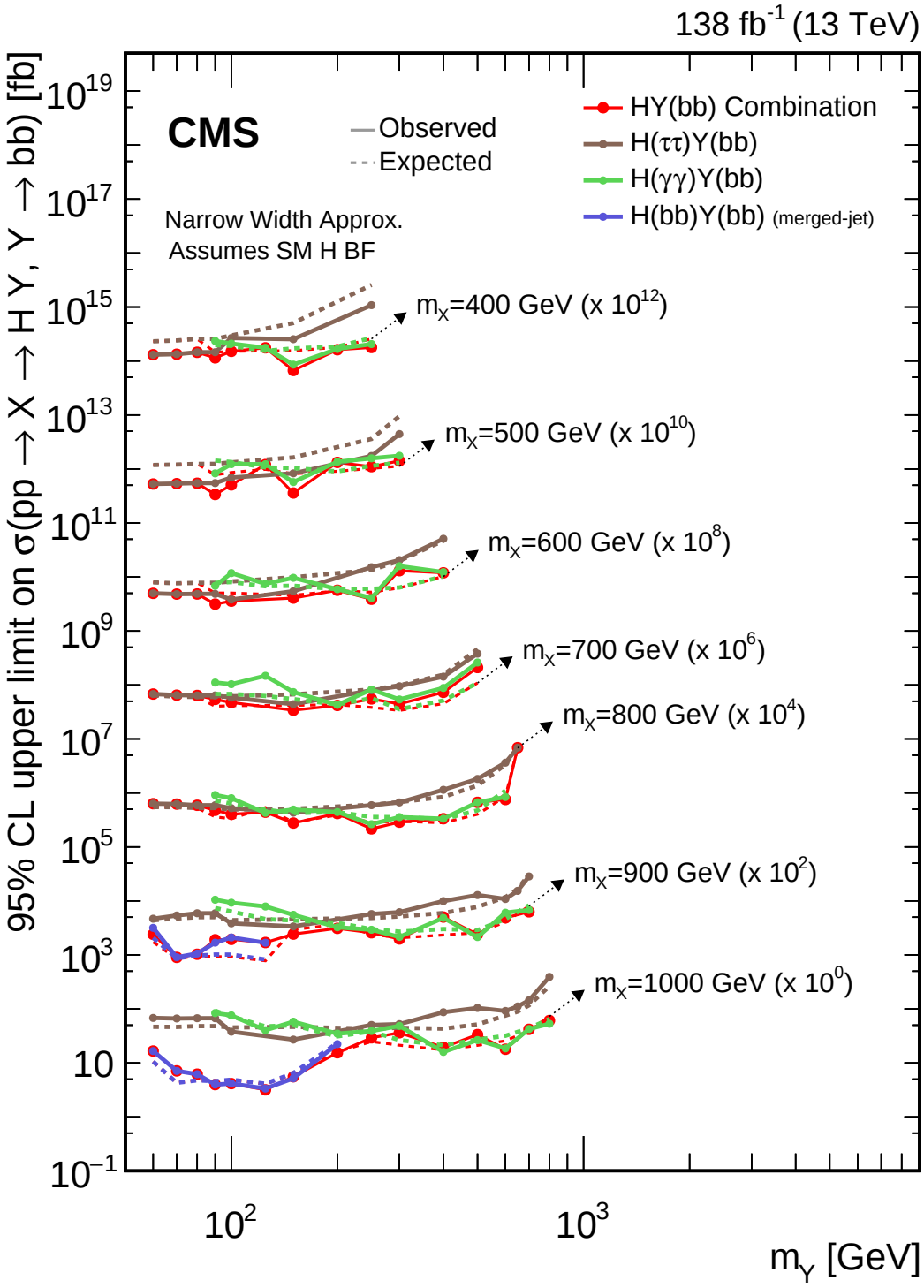

Figure 1.24: Observed and expected $X \to YH$ upper limits, at 95% CL, on the product of the cross section σ for the production of a resonance X via ggF and the BR $\mathcal{B}$ for the $X \to Y$ (bb)H decay. The SM BRs of the $H \to \tau\tau$, $H \to \gamma\gamma$ and $H \to bb$ decays are assumed. The results from the individual analyses and their combination are shown as functions of m_Y and m_X for $m_X \leq 1\,\text{TeV}$. Observed (expected) limits are indicated by markers connected with solid (dashed) lines. For an easier visualization, the limits have been scaled in successive steps by two orders of magnitude each. For each set of graphs, a black arrow points to the corresponding m_X value.

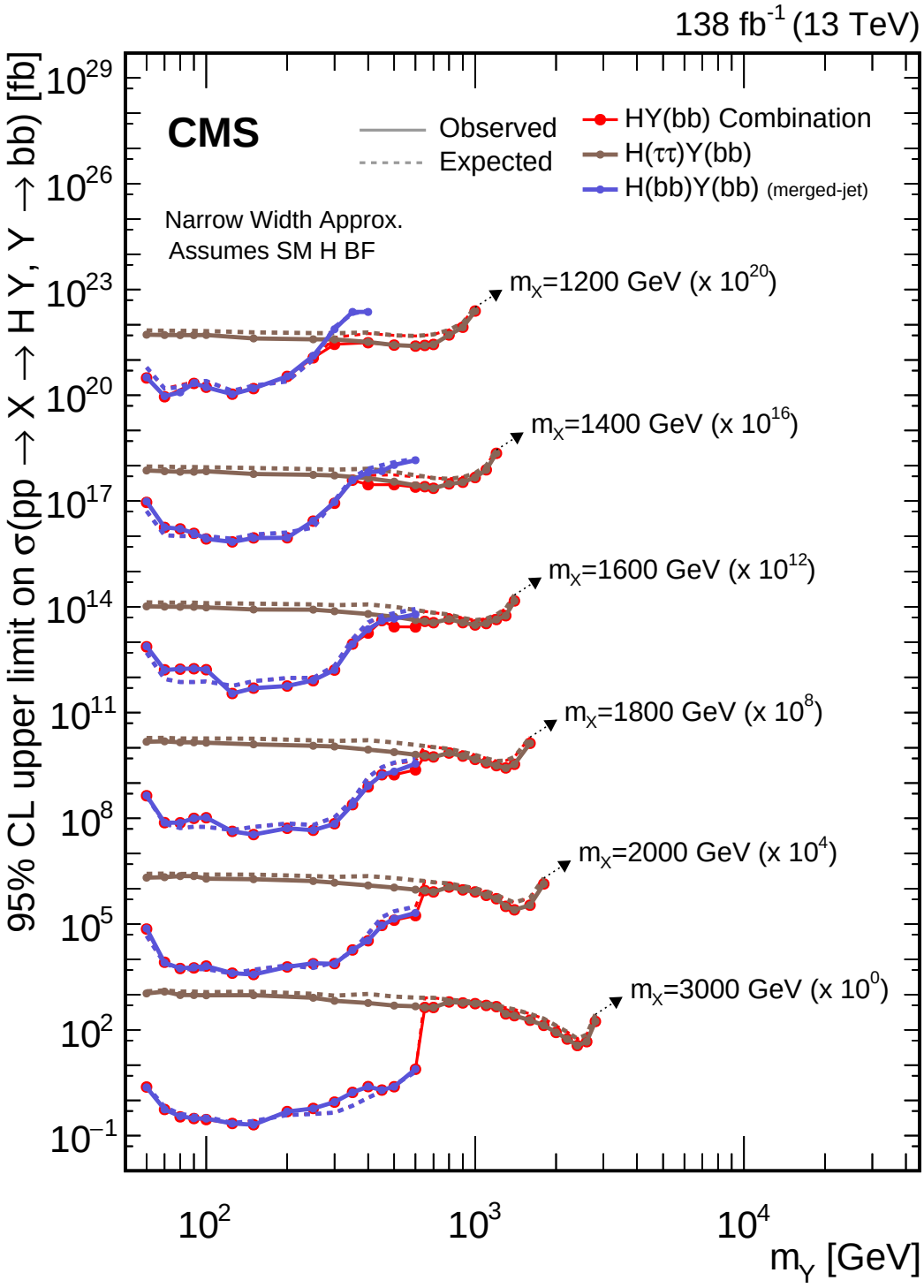

Figure 1.25: Observed and expected $X \to YH$ upper limits, at 95% CL, on the product of the cross section σ for the production of a resonance X via ggF and the BR $\mathcal{B}$ for the $X \to Y$ (bb)H decay. The SM BRs of the $H \to \tau\tau$, $H \to \gamma\gamma$ and $H \to$ bb decays are assumed. The results from the individual analyses and their combination are shown as functions of m_Y and m_X for $m_X \geq 1.2\,\text{TeV}$. Observed (expected) limits are indicated by markers connected with solid (dashed) lines. For an easier visualization, the limits have been scaled in successive steps by two orders of magnitude each. For each set of graphs, a black arrow points to the corresponding m_X value.

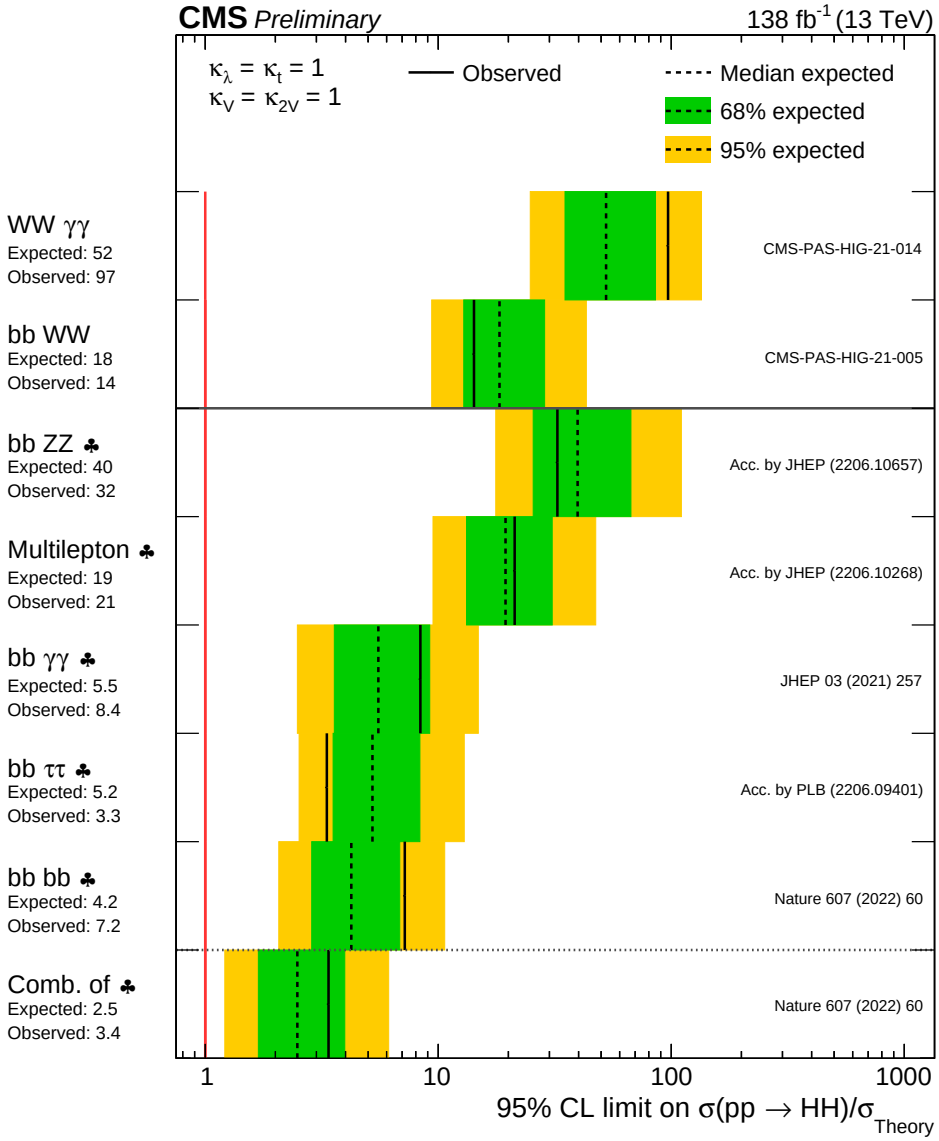

Figure 1.26: Upper limits at 95% CL on the SM signal strength $\mu = \sigma_{\mathrm{HH}}/\sigma_{\mathrm{HH}}^{\mathrm{SM}}$. The inner (green) and outer (yellow) bands indicate the regions containing 68% and 95%, respectively, of the limits on μ expected under the background-only hypothesis. The quoted expected upper limits are evaluated with the post-fit values of the uncertainties. Figure taken from [118].

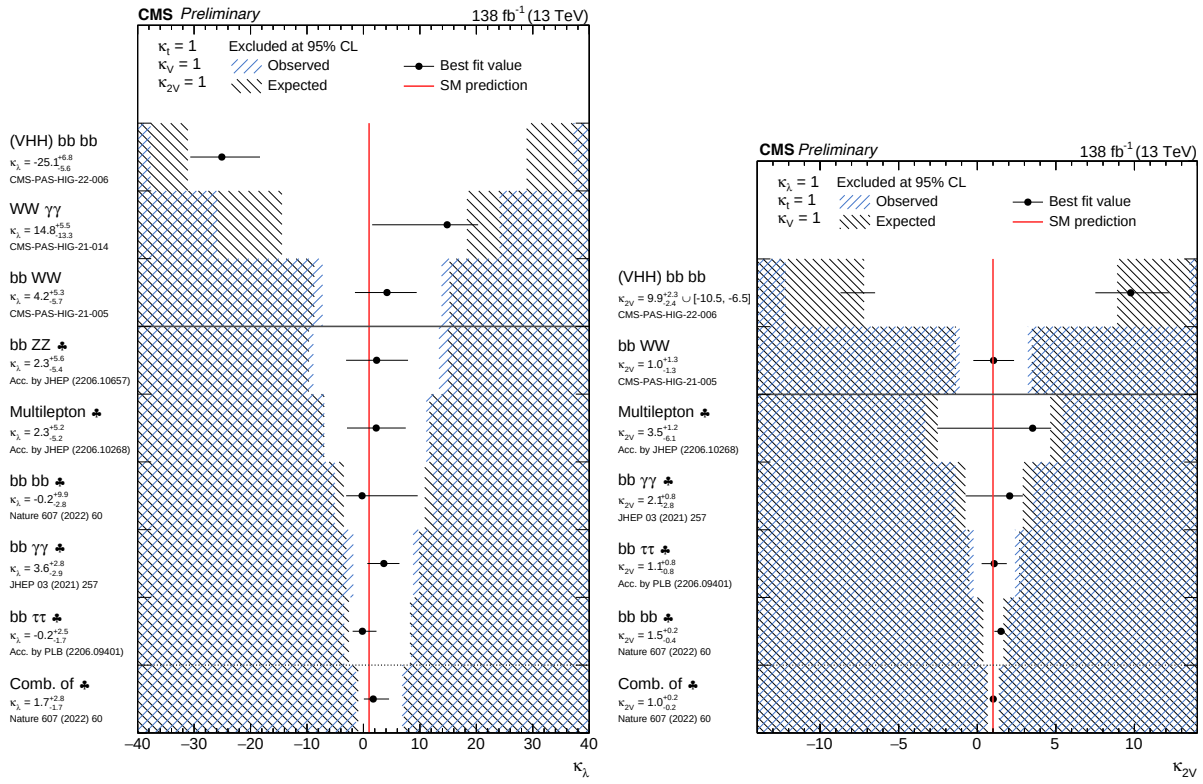

Figure 1.27: 95% confidence intervals on κ_λ (left) and κ_{2V} (right), superimposed by the best fit value on this parameter, assuming SM values for the κ_t, κ_V and κ_{2V} coupling modifiers. The blue (black) hashed band indicates the observed (expected) excluded regions, respectively. The band around the best fit value corresponds to a $1\,\sigma$ interval. The quoted expected upper limits are evaluated with the post-fit values of the uncertainties. Results are taken from the references marked next to each individual measurement.

combination constrains the observed (expected) HH cross section to $\sigma_{HH} < 3.4$ (2.5) σ_{HH}^{SM}, as shown in Fig. 1.26. The limit is also interpreted as a function of the κ_λ and κ_{2V} coupling modifiers, observing $-1.24 < \kappa_\lambda < 6.49$ and $0.67 < \kappa_{2V} < 1.38$, respectively, at 95% CL. The coupling modifier results are shown for individual channels in Fig. 1.27, and compared to the theory prediction in Fig. 1.28. Importantly, the κ_{2V} result allows the exclusion of $\kappa_{2V} = 0$ at 6.6 standard deviations. The above combination fixes all coupling modifiers except one to have their SM value. One can additionally perform simultaneous fits in ggF and VBF HH processes of $(\kappa_\lambda, \kappa_{2V})$ and $(\kappa_\lambda, \kappa_t)$, as shown in Refs.[38] and [114], respectively.

We conclude by mentioning two very recent results, produced after the combinations above. The first refers to the HH → bbVV analysis [119], which covers so far unexplored all-hadronic final states, with a 13% BR. Despite the absence of a relevant κ_λ constrain, the analysis focus on the boosted regime to better constrain κ_{2V}, almost excluding $\kappa_{2V} = 0$ by itself. The analysis represents the first use of the novel Particle Transformer algorithm [120], briefly covered in Section 5.5, to tag a merged H → VV → 4q jet. The second result consists of using vector-boson scattering (VBS) production of $W^\pm W^\pm H(\to b\bar{b})$ to constrain HHWW couplings, considering only W leptonic channels [121, 122]. The diagrams also include λ_{HHH}, but given the experimental sensitivity the analysis focuses instead on κ_{2V}, being mostly sensitive to V = W due to the lepton selection. The final constraints are competitive with other sub-leading channels. These new ideas show the

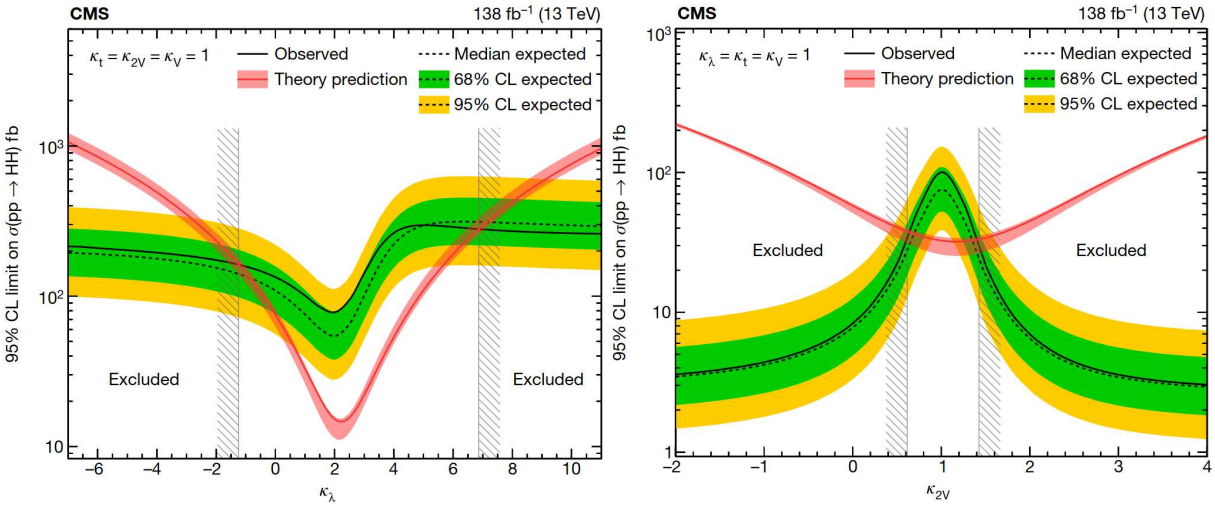

Figure 1.28: Combined expected and observed 95% CL upper limits on the HH production cross section as a function of κ_λ (left) and κ_{2V} (right), assuming SM values for the κ_t, κ_V and κ_{2V} coupling modifiers. The green and yellow bands represent the $1\,\sigma$ and $2\,\sigma$ extensions beyond the expected limit, respectively; the red solid line (band) shows the theoretical prediction for the HH production cross section (its $1\,\sigma$ uncertainty). The areas to the left and to the right of the hatched regions are excluded at the 95% CL. Taken from [38].

potential for a continued deliver of high-quality results in the years to come, always hoping for larger improvements than anticipated.

1.4.2 Indirect Searches

On top of the already mentioned HH direct searches, κ_λ can be alternatively constrained by the exploitation of single Higgs processes, such as $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$. Indeed, despite not depending on λ_{HHH} at LO or at higher orders in QCD, single Higgs production does depend on κ_λ in EW interactions at NLO. We display some of those dependencies in Fig. 1.29, for production, and Fig. 1.30, for decays [79]. Searches probing the Higgs self-coupling with higher-order loops are denoted *indirect*. Single Higgs cross sections can be quite sensitive to κ_λ variations, especially in the VH and $t\bar{t}H$ production modes, where up to 10% differences are expected [123]. The fact that a higher loop order is needed is in part counterbalanced by the much higher single Higgs cross section, leading to competitive limits with respect to HH production. Having the same goal in mind, there has been an additional proposal, consisting on using processes with higher-order EW diagrams, such as the ones used in m_W and $\sin\theta_W$ measurements [124].

Modifications of the Higgs self-coupling are parameterized with the κ-framework presented in Section 1.3.2, which is known to be equivalent to an EFT approach with the introduction of higher-dimensional operators only affecting λ_{HHH}. Additional sensitivity can be obtained by measuring differential cross sections, which do not have a flat dependence on κ_λ. The transverse momentum p_T of the Higgs boson is found to be the most sensitive observable to detect BSM effects on the trilinear coupling [123, 125]. The bounds obtained from indirect searches are competitive with the ones from double Higgs boson production [126], as shown in Fig. 1.31, and provide additional constraints for single Higgs couplings. We describe in Section 1.4.3 the ample complementarity between single and double Higgs boson measurements.

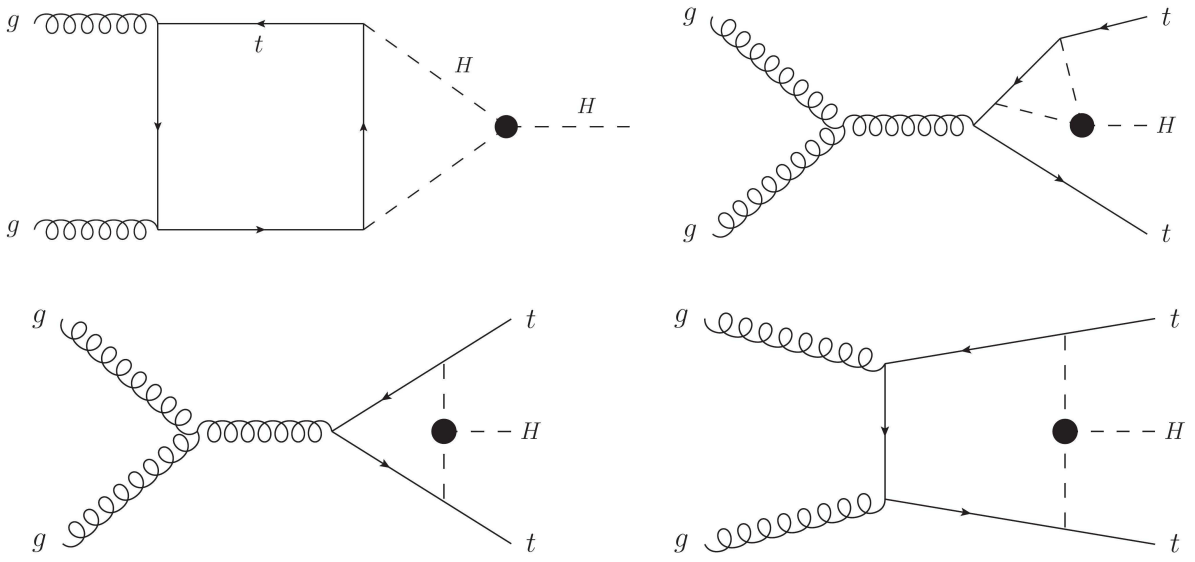

Figure 1.29: Examples of single Higgs boson production processes at NLO contributing to the Higgs boson self-coupling. The one in the top left is a ggF process, while the others refer to $t\bar{t}H$. Taken from [79].

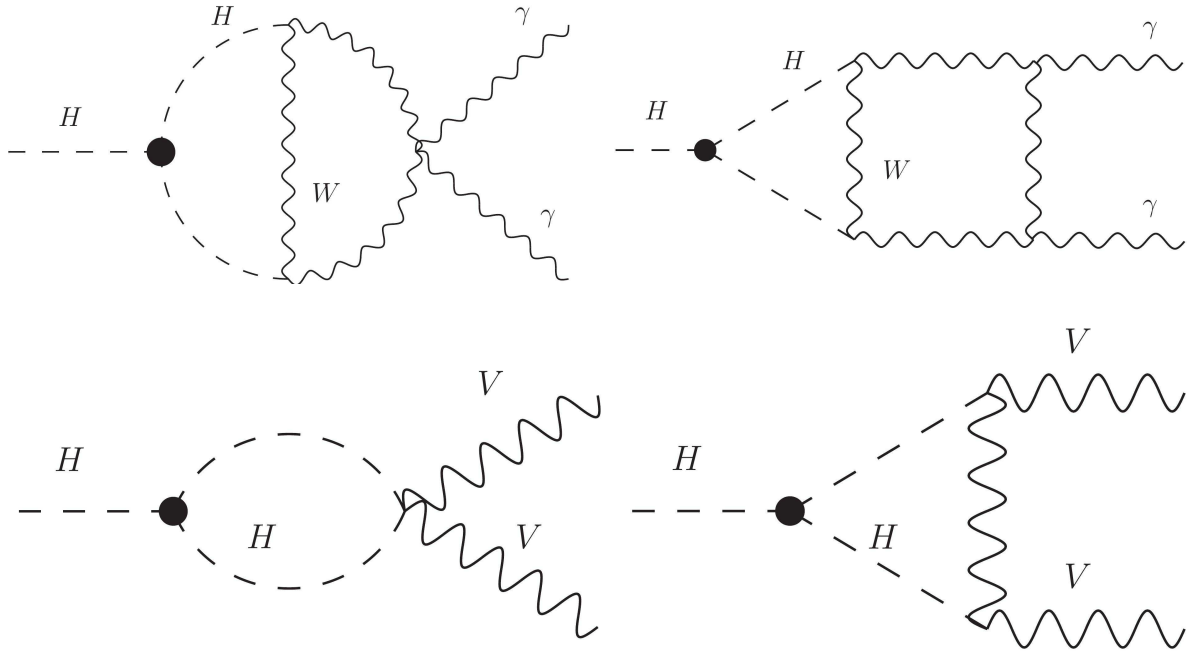

Figure 1.30: Examples of single Higgs boson decay processes at NLO contributing to the Higgs boson self-coupling. The diagrams in the top (bottom) row refer to $\gamma\gamma$ (VV) decays. Taken from [79].

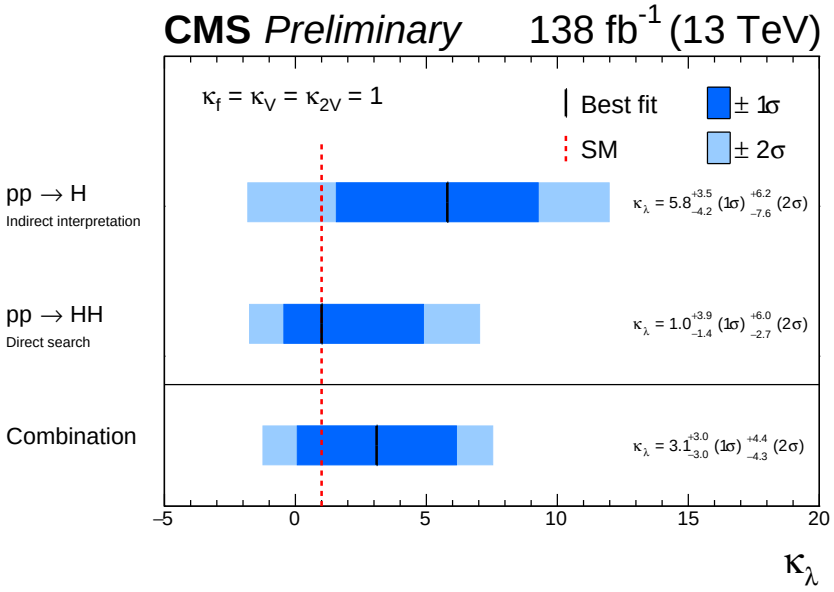

Figure 1.31: Constraints on the Higgs boson self-coupling modifier κ_λ from single and pair production of Higgs bosons. Single Higgs boson processes lead to competitive κ_λ measurements relative to HH processes. Taken from [38].

1.4.3 Combining Direct with Indirect Searches

There is a strong complementarity between direct and indirect κ_λ studies [123, 127]. Given their lower cross section, double Higgs boson searches generally provide weaker constraints on the Higgs boson couplings to fermions and vector bosons with respect to single Higgs analyses. On the other hand, HH processes provide stronger constraints on κ_λ. Combinations drawn from the two approaches maximize sensitivity, often enabling the relaxation of additional constraints during the fit, for example when setting some coupling modifiers to their SM value. Global fits also further validate individual studies, mostly at the level of the statistical methods, as the final results should be compatible with individual analyses.

CMS has recently performed a combination of single and double Higgs production analyses [128]. The analyses with the most sensitive production modes are included. The main challenge consists in estimating and efficiently removing overlaps between signal regions of different analysis. Whenever overlaps exist, one of two approaches is taken: either additional selections are applied, or the least sensitive category or analysis is removed. As an example, the bbZZ analysis is removed in favor of $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ due to the former's relatively low κ_λ sensitivity. Concerning systematics, their modeling in HH processes is generally simpler when compared to single H, due to the limited number of HH events.

CMS observed exclusion intervals at 95% CL of $-0.4 < \kappa_\lambda < 6.3$, assuming other Higgs couplings to follow the SM, or $-1.4 < \kappa_\lambda < 6.1$, otherwise. For comparison, ATLAS observed $-0.4 < \kappa_\lambda < 6.3$ assuming SM couplings, and $-1.4 < \kappa_\lambda < 6.1$ otherwise [129]. We show (κ_λ, κ_t) and (κ_V, κ_2V) scans in Fig. 1.32 (top), where the complementarity between the two types of processes is clearly highlighted. CMS was also able to once again [38] exclude $\kappa_\mathrm{2V} = 0$ at more than $5\,\sigma$, this time without fixing $\kappa_\mathrm{V} = 1$. Finally, the bottom bar of Fig. 1.31 shows that even κ_λ benefits, even if only slightly, from the inclusion of indirect processes in the overall result.

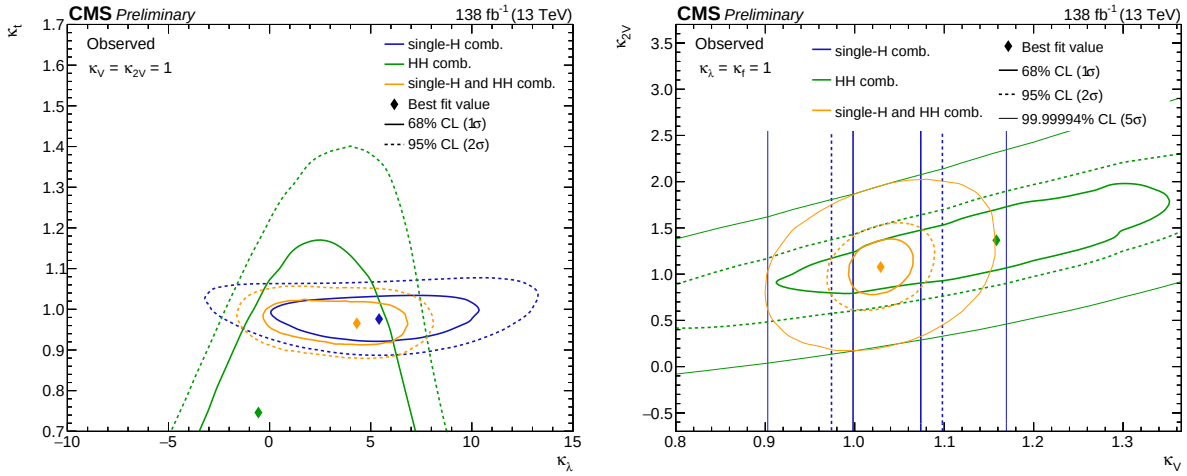

Figure 1.32: (Top) Observed two-dimensional likelihood scans of (κ_λ, κ_t) (left) and (κ_V, κ_{2V}) (right). The strong complementarity between the single and double Higgs processes is well illustrated. The combination slightly improves on HH results. The remaining coupling modifiers are set to their SM value. Taken from [128].

2 The CMS Detector at the LHC

Following the end of the Second World War, there was the need for Europe to stop the brain exodus to North America, and to establish a world-class research facility which could foster unity. The European Organization for Nuclear Research (CERN) was thus born in 1954, engaging in purely scientific endeavors, and excluding itself, as per its Convention [130], from any connection to military actions, in a spirit of peace and open scientific exchange.

Home of many atomic and particle physics experiments, CERN is nowadays widely known for its massive particle accelerator, the LHC, designed to achieve a center-of-mass energy of 14 TeV. The LHC sits at the border between France and Switzerland, close to Geneva, and spans an impressive 27 km ring of vacuum pipes, where beams of protons are accelerated in opposite directions almost at the speed of light. The tunnel is located at a depth ranging between 45 and 170 m, and was formerly used by LHC's predecessor, LEP, which provided the high-energy electron and positron beams leading to the precise measurements of the W and Z boson masses, among other SM parameters [131]. The beams at the LHC collide in four Interaction Points (IPs), producing a myriad of decay products which are analyzed by large and complex experiments. One of such machines is known as CMS, and represents a general particle physics experiment producing the data which serves as basis for the Thesis you are now reading.

This Chapter offers a condensed summary of the elements and ideas behind the data production flow taking place at the LHC and CMS, from circulating proton bunches to the definition of measured physics objects, expanding the current knowledge on subatomic physics. We describe the LHC in Section 2.1 and CMS in Section 2.2. Given the scope of the Thesis, namely the trigger work described in later Chapters, we cover the CMS trigger system in detail in Section 2.3, and provide a concise description of how offline object reconstruction works in Section 2.4, giving special emphasis to the Particle Flow approach. We finalize the Chapter by discussing Monte Carlo generation within CMS, also in light of the Monte Carlo activities developed within the time-frame of the Thesis here reported.

2.1 The Large Hadron Collider

The LHC is one of the largest scientific instruments ever built, being the current most powerful particle accelerator in the world. Designed to provide a maximum center-of-mass energy of 14 TeV, it has gathered since its 2010 debut a user community of ∼9000 scientists working in fundamental particle physics [132]. One of the main scientific driving factors to built such a powerful and expensive machine was to search for the conjectured Higgs boson and understand the mechanism of EWSB, which we covered in Section 1.1.3. The LHC is also intended to explore the TeV scale, either via the existence of new resonances, or by the indirect manifestation of higher-energy and yet unknown particles, in order to answer some of the most profound questions in modern physics. Beyond its

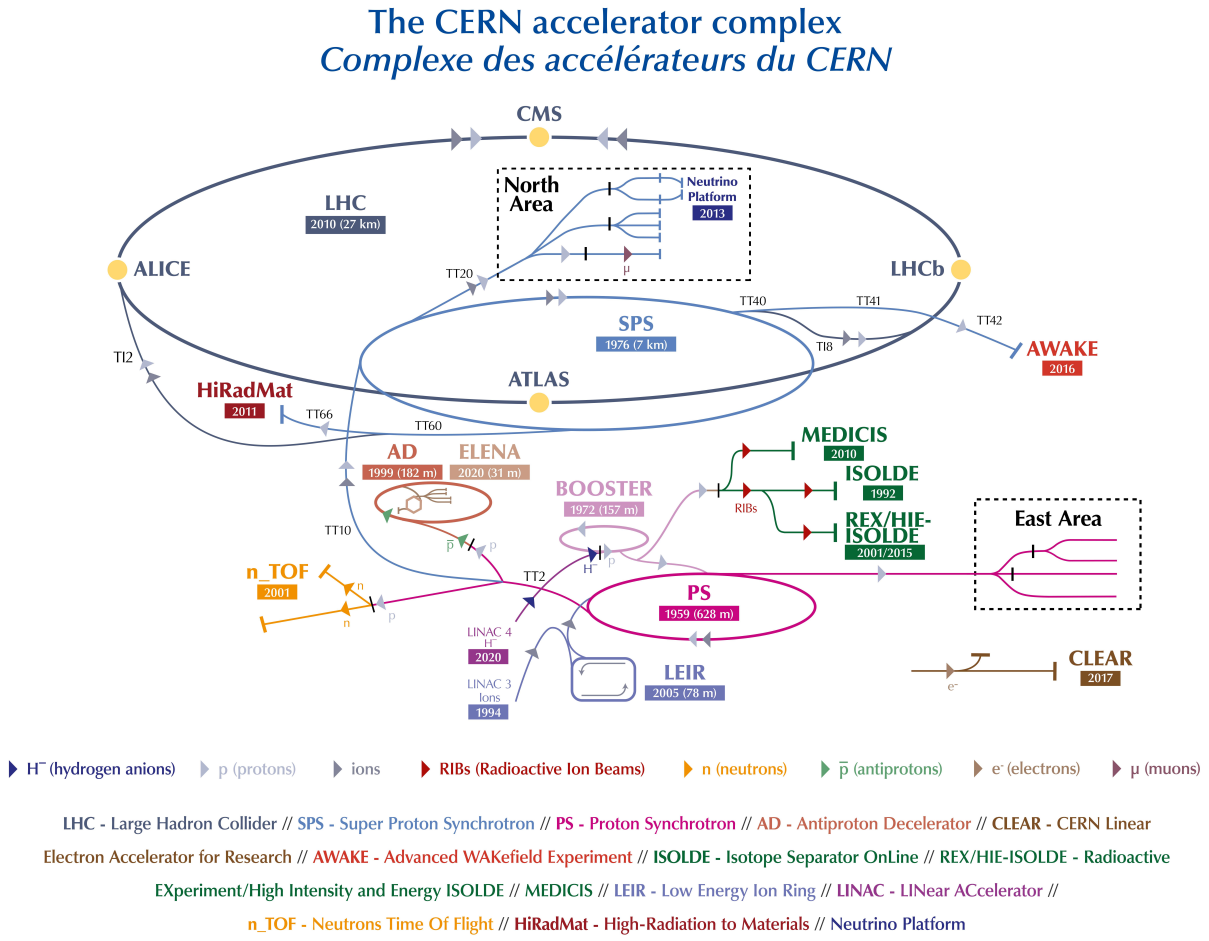

Figure 2.1: The CERN accelerator complex. The LHC is the 27 km ring shown at the top in dark blue. Its the last of a complex chain of particle accelerators with progressively higher beam energies. The smaller accelerators are used in sequence to accelerate the proton beams that collide at the four IPs, corresponding to the center of the ALICE, ATLAS, CMS and LHCb experiments, located at CERN's North Area. The East Area hosts instead multiple experiments connected to the Proton Synchrotron. Taken from [133].

proton-proton (pp) collisions, the LHC also collides heavy ions, like Pb or Xe, enabling the study of the *quark-gluon plasma*, an extremely dense and hot state of matter where quarks and gluons do not experience the strong force. In the following, we cover LHC's design and operations, and briefly describe some of its most iconic active experiments.

2.1.1 Design

The record-breaking 13.6 TeV proton center-of-mass energies at the LHC cannot be achieved without a complex chain of injection. The chain starts with the extraction of protons from a simple hydrogen bottle[1], within Linear Accelerator (Linac4), and continues in accelerator rings of progressively higher energies, going first to the Booster, then to the Proton Synchrotron (PS), and finally into the Super Proton Synchrotron (SPS), before being injected into the LHC. The complete accelerator complex is illustrated in Fig. 2.1.

The Linac4 [134] is designed to boost negative hydrogen ions [135] to high energies via radio-frequency cavities, and it has been the proton source of the LHC since 2020.

[1] Around three 1.5 m-tall bottles are used per year of operation.

It is 86 m long and is located 12 m underground, and accelerates ions in a series of four stages. First, the magnetic field of a radio-frequency quadrupole reaches 3 MeV, then drift tubes arrive at 50 MeV, later other linear accelerators attain 100 MeV, and finally so-called Pi-mode structures provide the final 160 MeV energy. During the injection in the Booster, the ions are stripped from their two electrons, leaving only the individual protons. This enables larger particle multiplicities in the synchrotron, reduces beam losses and simplifies the injection procedure. The beam continues its path in the 150 m-long Booster, being accelerated to up to 2 GeV. From there the beams are directed to the PS, which is also connected to physics experiments. The acceleration procedures continues with the PS, CERN's first synchrotron, and the SPS, reaching energies of 25 and 450 GeV, respectively. Specially built transfer lines connect the SPS and LHC ring tunnels, and fast kicker magnets are used to split the proton beam into two beamlines, which are then accelerated to their target energies using additional high frequency cavities. The SPS, with its 7 km ring length and 1317 electromagnets, also directly provides proton beams for some experiments, such as NA61/SHINE, NA62 and COMPASS. It has also notably been involved in the discoveries of the W and Z boson, since in 1983 it was running as a proton-antiproton collider for the UA1 and UA2 experiments.

The LHC consists of a ring of ∼1600 Niobium-Titanium (NbTi) superconducting magnets cooled with He at 1.85 K with a number of accelerating structures to boost the energy of the particles along the way. The two separate beam pipes are kept at an ultrahigh vacuum reaching of $\sim 10 \times 10^{-10}$ mbar, a value almost as low as the pressure found at the surface of the Moon. The protons are packed into bunches separated by a time spacing Δt_{b}, forming *trains*. The more bunches per train, the higher the integrated luminosity becomes, and thus the beams are packed in long trains as close to each other as possible. However, there is a limit to it, since the kicker magnets have specific rise and fall times that must be taken into account, leading to complex *bunch schemes*. Once the beams achieve the desired energy, they are focused and put into collision in four specific IPs along the LHC's circumference, where each IP is surrounded by a large physics experiment: ATLAS, ALICE, CMS, and LHCb. After some time, in which the luminosity steadily decreases due to losses in the collisions, the beam is dumped, and the injection procedure restarts. The ATLAS and CMS experiments are two multi-purpose experiments with a broad and similar physics programme. LHCb and ALICE target instead more specific branches, namely flavor physics and heavy-ion physics, respectively. These and other detectors are succinctly described in Section 2.1.3. CMS in particular is described at length in Section 2.2, due to its immediate connection to this Thesis.

A fundamental parameter of collider experiments is its *instantaneous luminosity* $\mathcal{L}$, a quantity measuring the number of collisions taking place per unit of time and unit of area. If $\mathcal{L}$ is integrated over time, one obtains the *integrated luminosity*, a useful parameter to know how many collisions occurred during a specified time period. The total number of times a physics process takes place in a collider experiment, often referred to as the number of *events*, is given by the product between the process' cross section and its integrated luminosity:

$$N_{\mathrm{events}} = \sigma \int \mathcal{L}\,, \tag{2.1}$$

where the time integral refers to some unspecified time period, such as a full week, month or data-taking year. The equation shows that the observation of rare processes requires the production of more data, which can amount to longer data-taking periods and/or larger instantaneous luminosities. Assuming that the two LHC beams are identical, that

their transverse section is perfectly round, and that they are highly collimated, $\mathcal{L}$ is given by:

$$\mathcal{L} = \mathcal{F}\,\frac{N_{\mathrm{b}}^2\, n_{\mathrm{b}}\, f_{\mathrm{rev}}\,\gamma}{4\pi\,\epsilon_n\,\beta^*} \quad , \qquad \mathcal{F} = \left(1 + \frac{\theta_{\mathrm{c}}\sigma_z}{2\sigma_{xy}}\right)^{-1/2} , \tag{2.2}$$

where N_{b} represents the number of particles in each of the n_{b} bunches in the beam, f_{rev} is the revolution frequency of the beam, $\gamma \equiv (1 - v^2/c^2)^{-1/2}$ is the relativistic factor with v being the beam linear velocity, $\mathcal{F}$ stands for a geometric factor accounting for the luminosity reduction caused by the crossing angle θ_{c} at the interaction point between the two beams with transverse and longitudinal root mean square (rms) bunch sizes σ_z and σ_{xy}, respectively, and $\epsilon_n\,\beta^*$ is related to beam optics, where ϵ_n stands for the average rms normalized emittance of the two beams, being the emittance a measure of the volume covered by the beam, and β^* is the beta-function at the collision point [136, 137]. Some of these quantities are listed in Table 2.1, specifically for the design values of the LHC.

Description	Parameter	Value
Center-of-mass energy	$\sqrt{s}$	14 TeV
Number of particles per bunch	N_{b}	1.15×10^{11}
Number of bunches per beam	n_{b}	2808
Revolution frequency	f_{rev}	11.2 kHz
Transverse beam emittance	ϵ_n	3.75 µm
Beta function	β^*	0.55 m
Bunch time spacing	Δt_{b}	25 ns
Collision angle	θ_{c}	285 µrad
Transverse bunch rms at IP	σ_z	7.55 cm
Longitudinal bunch rms at IP	σ_{xy}	16.7 µm

Table 2.1: Nominal design parameters of the LHC machine in its pp collisions configuration.

The desired luminosity increase inevitably leads to a larger number of multiple pp interactions per bunch crossing, a phenomenon known as *pile-up (PU)*. The average of such quantity is given by:

$$\langle\mu\rangle = \frac{\mathcal{L}\,\sigma_{pp}^{\mathrm{inelastic}}}{n_b f_{\mathrm{rev}}} , \tag{2.3}$$

where $\sigma_{\mathrm{pp}}^{\mathrm{inelastic}}$ is the inelastic pp cross section, measured to be $68.6 \pm 0.5\,(\mathrm{syst}) \pm 1.6\,(\mathrm{lumi})$ mb at 13 TeV [138]. The PU represents a limitation on the physics sensitivity in any collider experiment. Firstly, the larger it is, the larger will the radiation levels be for the detectors. When the increase in luminosity is very significant, new or upgraded detectors might be required, as extensively discussed in Chapter 3. Secondly, more sophisticated data analysis techniques are also needed, to disentangle the interesting processes produced at the primary vertex (PV) from all generally low-energy products coming from PU vertices. Therefore, unless considerable efforts are undertaken, the increase in luminosity can negatively impact the sensitivity to the observation of NP.

2.1.2 Operations

The LHC inaugural run took place in September 2008, when the first protons successfully circulated around the machine. Unfortunately, nine days later, an unexpected and violent leakage of a He coolant due to a faulty electrical connection caused the severe damage of multiple equipment, including superconducting magnets. The morose and expensive repairs required a one-year technical stop, after which 450 GeV proton beams were initially injected for testing, and the energy was steadily increased without further issues.

The LHC physics programme started on March 30$^{\text{th}}$ 2010, with collisions at 3.5 TeV per beam. The first data period used for Physics analyses, or Run 1, considered the years of 2011 and 2012, with a total of 29.4 fb^{-1} at 20 MHz, where the first 6.1 fb^{-1} were collected in 2011 at 7 TeV, and the following 23.3 fb^{-1} were obtained at 8 TeV in 2012 [139]. These datasets were responsible for, among other measurements, the observation of the anticipated Higgs boson at $m_{\text{H}} \sim 125$ GeV [1–3], announced on July 4$^{\text{th}}$ 2012. It followed a two-years-long Long Shutdown (LS), where the magnet interconnects responsible for the 2008 incident were consolidated. Experiments also took advantage of this period to implement detector upgrades.

For Run 2, the center-of-mass energy was increased to 13 TeV at a 40 MHz bunch crossing frequency (or, equivalently, a 25 ns spacing, roughly doubling the number of bunches in a train), totaling 137.2 fb^{-1} between 2015 and 2018. Per year, the LHC delivered 4.2 fb^{-1} in 2015, 41.0 fb^{-1} in 2016, 49.8 fb^{-1} in 2017 and 67.9 fb^{-1} in 2018, a large part of which was recorded by the experiments. The Run 2 dataset is used for the physics analysis discussed in Chapters 4 and 5, except for the dataset taken in 2015, which was employed solely for commissioning of the updated LHC configuration. In Figs. 2.2 and 2.3, we show the luminosities delivered by the LHC and recorded by the CMS experiment, and also show the measured CMS PU values. After the end of Run 2, it was time for further detector refurbishments. For CMS, this notably included new electronics for the hadronic calorimeter.

The last LHC run, naturally called Run 3, is currently ongoing, having started in 2022 and with an end planned for next year, in 2025. It features a slightly higher center-of-mass energy, sitting at 13.6 TeV, approaching LHC's design value. Unfortunately, another He leak was responsible for a complete halt of operations on September 2023. The culprit was eventually identified: a tree fell on Swiss power-lines, originating a power glitch that lead to the quench of an inner triplet magnet and the mentioned leak. Once the issue was addressed, the rest of the year was devoted to Pb-Pb collisions at 5.36 TeV. Despite the leak, operations in 2024 have been extremely efficient, with the LHC's 10000$^{\text{th}}$ fill reached on August 21$^{\text{st}}$. The 250 fb^{-1} delivered luminosity target that had been initially set for Run 3 will likely be achieved, given the current value above 150 fb^{-1}. 2024 is the most productive year in LHC's history so far. We refer to the collection of the first three runs as "Phase 1", while the period in which the next version of the LHC will be operational is called "Phase 2". The LHC upgrade is discussed in detail in Section 3.1,

It is also interesting to note some of the challenges tackled by the LHC over the years. We briefly mention a few, all currently under control:

- **Unidentified Falling Objects**: 10 µm-sized particles occasionally fall from the top of the vacuum chambers and interact with the positively charged beam, causing significant beam losses and potentially leading to quenches. The study and simulation of these microscopic particles has strongly mitigate their negative effects.

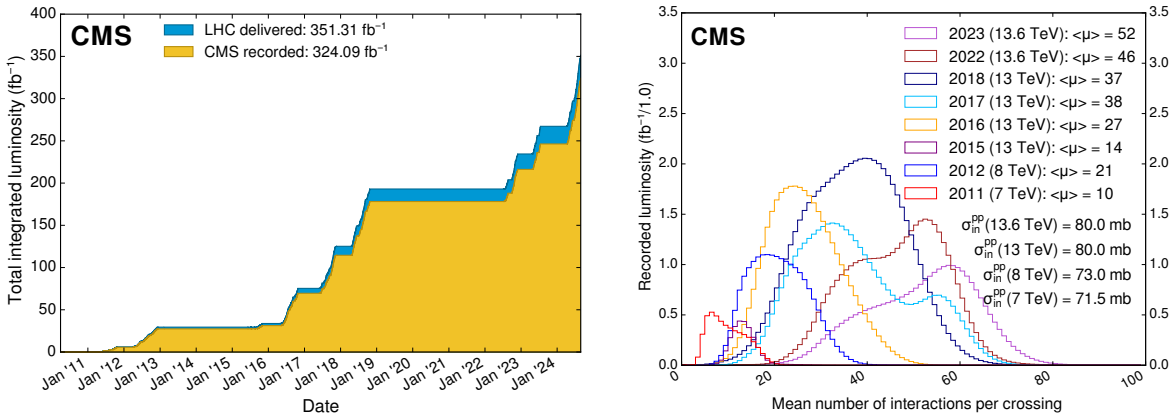

Figure 2.2: Luminosity and PU values. All plots include the still ongoing Run 3, and are thus subject to future changes. (Left) Luminosity delivered by the LHC and recorded by CMS as a function of time, from the start of Run 1 to the present day. (Right) Average number of pp interactions per bunch crossing for all data-taking years. The mean values per data-taking year are shown, together with the inelastic pp cross sections for all different center-of-mass energies considered at the LHC. Taken from [140].

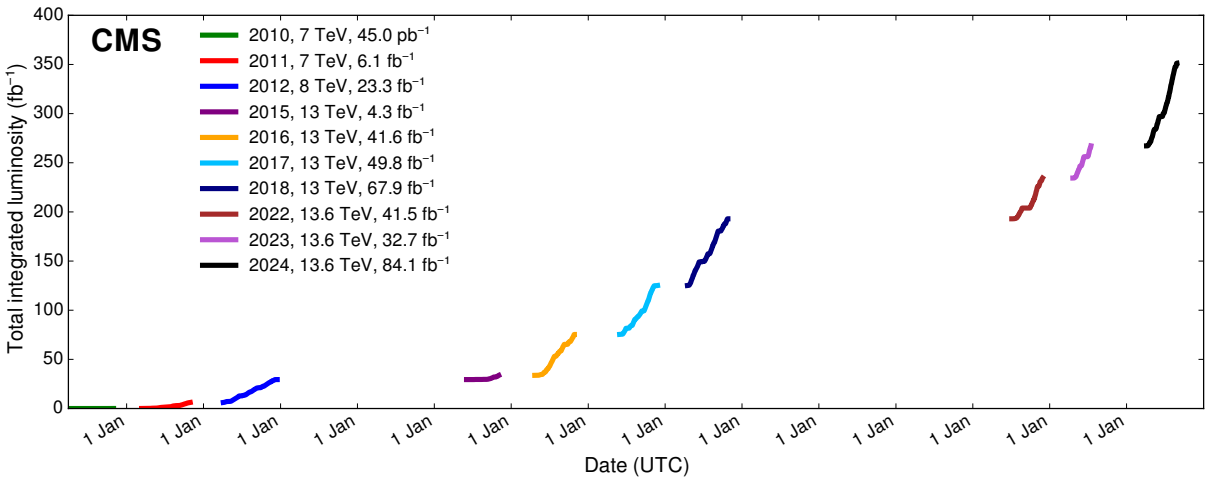

Figure 2.3: Cumulative luminosity delivered to CMS during stable beams at nominal center-of-mass energy, as a function of time, for all data-taking years. Gaps in time correspond to regular end-of-year shutdowns or LSs. The plot includes the still ongoing Run 3, and is thus subject to future changes. Taken from [140].

- **Electron clouds**: Electrons can be produced in ionization or photo-emission processes and interact with the walls of the beam chamber, producing more electrons, and leading to beam instabilities. This can be fixed by a process known as *scrubbing*, where an intentionally created electron cloud is impinged on the problematic surfaces. It has been shown that bombarding a material with electrons dramatically reduces its secondary electron yield.

- **Beam induced heating**: The increase in luminosity also increases the temperature and pressure felt by the equipment located near the LHC beams. This happens due to the exposure to all sorts of beam-related radiation: synchrotron radiation, lost protons, photoelectrons and even radio signals [139].

2.1.3 Experiments

At the LHC, proton beams circulating in counter-wise directions along two separate beam pipes converge in four IPs. Such locations, numbered IP 1, 2, 5 and 8, are located in underground experimental caverns, each attached to one or more experiments. They are part of CERN's North Area, where SPS provides proton beams. Other IPs were used by the LEP experiments: ALEPH[2], at IP4, and OPAL[3], at IP6. The IPs for DELPHI[4], at IP8, and L3[5], at IP2, were reused by the LHC beauty (LHCb) and the A Large Ion Collider Experiment (ALICE) detectors, respectively [141]. In this Section, we provide a summary for some of the physics experiments exploiting LHC beams, before describing the CMS detector at length in Section 2.2.

ATLAS

The ATLAS detector [142] is, along with CMS, one of the two general-purpose detectors at the LHC. Located at IP1, it comprises a total length of 46 m, a diameter of 25 m, and a weight of 7000 t. It includes a series of concentric and hermetic detection layers to measure the different kinds of particles produced at the PV. The Inner Detector (ID) is the first of such subdetectors, devised to measure the charge, direction and momentum of charged particles. It includes a Pixel Detector, a Semiconductor Tracker (SCT), and a Transition Radiation Tracker (TRT). The $\sim 92 \times 10^6$ silicon (Si) pixels are located just 3.3 cm away from the beamline, spread across four detection layers, reaching a 10 µm precision. The SCT follows, containing $\sim 6 \times 10^6$ strips, for a total of 60 m^2 of silicon material. It attains a precision of 25 µm. The last component of the ID is the TRT, which uses a drift-tube gaseous technology. The TRT exploits the production of transition radiation to extract information on the particle type. The more relativistic a particle is, the more photons are released at the interface between two materials with different diffraction indexes. Since lighter particles travel faster, they will also produce more radiation. This principle enables the distinction of particles with equal charges and energies but different masses. For the specific case of ATLAS, the TRT provides a significant discrimination between electrons and charged pions in the 1 to 100 GeV energy range [143]. A central NbTi solenoid magnet surrounds the ID, and provides the 2 T needed to bend the trajectories of charged tracks. Encircling the solenoid, a homogeneous liquid

[2] Apparatus for LEP PHysics
[3] Omni-Purpose Apparatus for LEP
[4] DEtector with Lepton Photon and Hadron Identification
[5] Third LEP Experiment

argon calorimeter serves as EM calorimeter [144]. The usage of noble liquids is a common choice in calorimeters due to their radiation resistant properties, ease of replacement, and good response uniformity. Argon in particular, despite the not-so-short radiation length, is cheap, and thus easily replaceable in the advent of contamination. Argon also provides an intrinsically stable and linear energy response, and is relatively easy to segment in the longitudinal and transversal directions. The ATLAS calorimeter notably features an "accordion" geometry with a honeycomb pattern, and works by ionization. This particular geometry is motivated by the desire to eliminate projective azimuthal cracks [145]. It is interesting to note that around half the signal comes in the form of scintillation light due to the recombination of electron-ion pairs. However, scintillation readouts are not included due to technical and geometrical challenges when extracting both charge and light signals. An excellent energy resolution can nevertheless be achieved [146]. The calorimeter features three segments in depth, the first having an extremely fine segmentation of 0.003 along pseudorapidity (η)[6], to clearly separate prompt photons from π^0 photons, up to $p_{\mathrm{T}} \sim 70\,\mathrm{GeV}$. The segmentation was optimized for the $\mathrm{H} \to \gamma\gamma$ analysis. The calorimeter is preceded by a presampler which corrects the electron and positron energy losses due to the presence of upstream material, including the solenoid. To keep the argon liquid, the calorimeters are kept at $-184\,^{\circ}\mathrm{C}$. A tile hadronic calorimeter follows, made of alternating steel absorber and active scintillator (Sci) plastic layers, read out by photomultiplier tubes (PMTs). Around 420 000 tiles are present, weighting almost 2900 t. Finally, the outer regions of ATLAS are occupied by a large muon spectrometer. Five different subdetectors provide precise ($\sim$100 µm) and fast ($\sim$2.5 µs) muon momentum measurements: Monitored Drift Tubes, Resistive Plate Chambers (RPCs), Thin Gap Chambers, Small-Strip Thin Gap Chambers, and Micromegas. The muon trajectories are bent thanks to additional magnets, including two endcap magnets and the largest toroidal magnet ever built, located in the center of the experiment, which took almost a year to install. The magnets provide up to 3.5 T, and are located in between the muon subdetectors.

We conclude by mentioning the two-tiered trigger system [147]. The first level, or L1, is based on two independent systems using information from the calorimeters or muon detectors. The decision of whether to keep the event at L1 is made in combination with multiple subsystems. Selected events are then sent to the ATLAS online software-based High-Level Trigger, where algorithms perform event reconstruction at a 3 kHz output rate. ATLAS has implemented "trigger-level analysis" and "delayed stream" modes [148], which are respectively similar to the CMS scouting and parking strategies discussed in Section 2.3.3.

LHCb

Instead of surrounding the entire collision point with an enclosed detector as ATLAS and CMS do, the LHCb experiment [149], located at IP 8, employs a single-arm forward spectrometer. The experiment includes large aperture subdetectors, from 10 to 300 mrad, placed perpendicularly to the beam axis. Given that LHCb's main purpose is the study of b-flavored baryons as probes for NP, its distinctive design can be readily explained: at current energies, the decay particles of b hadrons tend to travel close to the beam pipe. Phenomena studied by LHCb include rare b meson decays, the possible existence of CP violating asymmetries in b and c hadron decays, the precise measurement of the

[6] The pseudorapidity η is defined in Section 2.2.1.

three interior angles of the CKM matrix and of B_s^0 mixing, and tests for lepton flavour universality, among many others [150].

Given the asymmetric geometry of LHCb, with all detectors located on the same side, the LHC optics are modified so that the collision point is displaced 11.25 m from the center, which maximally exploits the volume of the underground cavern for detector components [151, 152]. Starting from the collision point, and moving outwardly through 21 m and 5.6 t of subdetectors, LHCb presents an array of semi-circular silicon-based detectors composing the Vertex Locator (VELO), followed by the first Ring-imaging Cherenkov detector (RICH), focused on low-momentum tracks. We note that particle identification is essential to distinguish pion, kaon and proton tracks, notably in flavour physics [150]. Several layers of the tracker systems follow, separated by a warm dipole magnet [153], and a second RICH detector lies immediately behind the tracker, to measure high-momentum tracks. One then finds a Shashlik electromagnetic calorimeter and a hadronic calorimeter composed of iron and scintillator tiles. The muon system finalizes the design, enabling for instance the study of the rare $B_s^0 \to \mu\mu$ process.

Concerning its trigger system, large beauty and charm cross sections required the reduction of the input maximum rate from ∼30 MHz down to ∼12.5 kHz, in Run 2. LHCb interestingly uses a trigger deferral system, where the time when the LHC is idle is repurposed for further processing. This time is part of a normal data-taking year, with technical stops, maintenance work or time spent preparing for each new fill. To maximally utilize the deferral, the trigger is split into two layers, where the buffering only takes place after the first level, receiving the entirety of the selected data. Additional trigger improvements enabled the inclusion of offline-quality information inside the trigger, raising the question of whether an additional offline reconstruction step was still required. A new "turbo" stream was therefore made available [154], where the full physics analysis is performed, producing an output format one order of magnitude smaller than what is possible with the "full", standard stream. Many analyses have started using datasets produced with the new trigger stream, benefiting from the additional unprescaled events. We note that the "turbo" approach is extremely similar to the scouting methodology in CMS, which is described in Section 2.3.3. For Run 3, the turbo mode share has been extended, and the hardware trigger was removed, meaning that LHCb now supports a purely software trigger at 30 MHz!

LHCb is also the sole LHC experiment capable to run both in collider and fixed-target mode [155]. The System for Measuring the Overlap with Gas (SMOG) provides the means to inject noble gases (He, Ar, Ne) into VELO. A fixed-target experiment brings many advantages. One can explore collisions with targets of mass number between the ones of protons and Pb ions, and with an 30 to 115 GeV energy range, which corresponds to an uncharted territory, between the energies provided by SPS and the higher energies reached by the LHC or the Relativistic Heavy Ion Collider, located at Brookhaven National Laboratory, in New York. Finally, the new approach can improve the understanding of nuclear and charm PDFs, and thus test different aspects of QCD. Understanding charm production also has important consequences for neutrino astronomy, since charmed hadrons from cosmic rays produce neutrinos that represent a large background to extra-galactic neutrino signals [156]. Some other results made possible by SMOG include fixed-target J/ψ and D^0 production, and direct measurements of antiproton production [157], which is relevant for dark matter (DM) searches.

ALICE

The ALICE experiment [158] is dedicated to heavy-ion physics, built to study strongly interacting QCD matter at extreme energy densities, namely the Quark Gluon Plasma (QGP). It mostly considers Pb ions, but also features lighter ions, proton-nucleus runs, and runs at lower energy. Located 56 m underground at IP2, weighting 10 kt, and measuring 26 m long, 16 m high, and 16 m wide, its design sets it apart from the other three main LHC experiments. Clearly, it must cope with the extreme particle multiplicity anticipated in central Pb-Pb collisions, which could be up to three orders of magnitude larger than in typical pp interactions at the same energy. The tracker was planned to be especially robust, featuring radiation-resistant technologies with a silicon vertex inner tracking system and gas electron multipliers in a time-projection chamber, where the latter can also help for particle identification (PID). ALICE incorporates most known PID technologies, such as RICH and transition radiation detectors, complemented by ionization and time-of-flight (TOF) detectors, with additional contributions from calorimeters and muon chambers. The transition radiation detector also helps with the tracking, where a charged pion rejection factor of ∼100 is achieved to provide good electron identification capabilities. A small single-arm high-granularity photon spectrometer, made of lead tungstate crystals, measures photons and neutral mesons, while a larger, lower-granularity Pb-scintillator sampling calorimeter is responsible for measuring jets. A forward muon arm in turn measures the decay products of heavy resonances on one side of the experiment, having a large enough mass resolution to distinguish between all heavy quarkonia resonances decaying to muon pairs. The muon detector is located at high rapidities to enhance the sensitivity to low p_{T} muons, and consists of an arrangement of absorbers, tracking stations, RPCs and a large 3 T dipole magnet. Other small and specialized detectors are installed in the forward region, for instance enabling calibration and alignment with cosmic rays, or providing information on particle multiplicity. Triggering is ensured by a series of subsystems, and notably allows the full readout of the data at the required interaction rates without the need for hardware trigger selections, i.e. ALICE can run in a triggerless, continuous mode for some subdetectors [159, 160]. The experiment is surrounded by a giant octagonal solenoid magnet, reused from the L3 experiment. Interestingly, the magnet is not perfectly centered around the LHC beam, since the electron-positron beam at LEP had a slightly lower alignment. Steel elements are therefore inserted to partially compensate the off-axis effect.

FASER

The ForwArd Search ExpeRiment (FASER) [161] is designed to study the interactions of TeV neutrinos via charged currents, and to look for long-lived BSM particles, such as dark photons or axions. It studies an extremely forward $\eta > 8.5$ region, but is located at 480 m from the ATLAS IP, a distance at which most SM model particle rates are negligible, with the exceptions of neutrinos and muons. The detector benefited from the reuse of spare LHCb calorimeter modules and ATLAS SCTs. The experiment also includes trigger and timing detectors, and scintillator veto systems. Downstream, we find FASERν [162], a subdetector consisting of multiple emulsion chambers interleaved with tungsten passive material, totaling approximately 220 X_0, for a spatial resolution of ∼400 nm per hit and a angular resolution of 0.06 mrad for 1 cm long tracks. The emulsion films are replaced every ∼3 months to mitigate PU effects. The detector achieved the

first ever direct observation of neutrino interactions at a collider [163], and has recently released the first neutrino cross section measurements [164], consistent with the SM.

SND@LHC

The Scattering and Neutrino Detector at the LHC (SND@LHC) [165] is a compact and dedicated experiment designed to study neutrinos from all flavors produced in pp collisions. It specifically focuses on neutrino cross sections between 350 GeV and 10 TeV energies, a range currently unexplored. The detector also has the potential for measuring hypothetical Feebly Interacting Particles (FIPs). Located about 480 m downstream from the ATLAS IP, in a previously unused tunnel linking the LHC to the SPS, SND@LHC is strategically positioned to capture particles in the $7.2 < \eta < 8.4$ region, which remains out of reach for the larger experiments, and is complementary to the coverage of the FASERν experiment. At this η range, electron neutrinos and anti-neutrinos are overwhelmingly produced by the decay of charmed hadrons, which can therefore be studied by SND@LHC. The detector consists of a hybrid system that combines upstream emulsion cloud chambers for tracking, each followed by a scintillating fiber plane for timing, with a downstream hadronic calorimeter and muon identification system. Tungsten is used as passive material to increase the neutrino containment. A veto system for muons is included in front of the emulsion chambers. This setup allows for the precise tracking and identification of different types of neutrino interactions. After its approval in 2021, SND@LHC recently detected neutrinos for the first time [166].

LHCf

The LHC forward (LHCf) detector [167] is designed to study particles emerging in the very forward region of collisions. Its primary goal is to improve our understanding of cosmic rays, by simulating ultra-high-energy cosmic ray interactions using particles thrown forward by collisions at IP1, the same IP used by ATLAS. These studies can help in the calibration and interpretation of larger cosmic rays experiments, such as the Pierre Auger Observatory [168], or the Telescope Array Project[7] [169]. The setup of LHCf consists of two small detectors located 140 m from the ATLAS IP, weighting 40 kg only, and measuring $30 \times 80 \times 10\,\text{cm}^3$. Each detector includes a sampling calorimeter tower to measure the energy and transverse momentum of neutral particles, particularly photons and neutral pions, produced at very small angles relative to the proton beam direction. The detectors also include a tracker system, with silicon layers or scintillating fibers, depending on the location.

TOTEM

Being the fifth approved LHC experiment, the TOTal Elastic and diffractive cross section Measurement (TOTEM) experiment [170], as its name implies, is dedicated to the precise measurement of pp cross sections in the forward region of the LHC, and to understand the dependence of the cross section with center-of-mass energy. The detector can also be exploited to study the structure of the proton, by measuring its elastic scattering over a wide range of momenta, hopefully shedding some light on low-energy QCD mechanics.

[7] The TAP measured the "Amaterasu" particle, the third most energetic particle ever observed, but the pole position still belongs to the 1991 "Oh-My-God" particle, boasting an impressive $(3.2 \pm 0.9) \times 10^{20}$ eV energy.

All of its subdetectors, Roman Pots (RPs) and two particle telescopes, detect charged particles emitted at CMS's IP and include trigger capabilities. The silicon-made RPs localize the trajectory of scattered protons within a 20 µm precision, being located on both sides of the IP at distances of 147 m and 220 m. The telescopes, instead, measure the rate of inelastic scattering, and are located in the CMS experimental cavern, at 10.5 m and 13.5 m from either side of IP5. Their goal is to measure the tracks of the charged particles produced in the PV. TOTEM measurements can strongly impact the theoretical models feeded into the generators used by the larger LHC experiments, such as CMS.

MoEDAL

The Monopole & Exotics Detector At the LHC (MoEDAL) experiment [171], together with its Run 3 update MoEDAL Apparatus for Penetrating Particles (MAPP), is located at IP8, sharing the experimental cavern with LHCb. It directly searches for the theorized magnetic monopole, together with other exotic particles that could indicate BSM, such as "Q-balls", "dyons" (particles containing both electric and magnetic charges), particles with multiple charges, and even black hole remnants. The detector is composed of an array of 10 sheets of plastic nuclear track detectors: an hypothetical new particle would break the molecular structure of the plastic during its traversal and reveal its trajectory thanks to a series of holes. MoEDAL is positioned around the LHCb VELO subdetector, for a maximum surface area of 25 m^2.

MilliQan

The recent milliQan detector [172] has been installed 33 m above the CMS experimental cavern, in a tunnel at $\eta \sim 0.1$, and started taking data at the beginning of Run 3. The detector looks for hypothetical milli-charged particles, which due to their tiny electric charge are hard to detect by conventional apparatuses. The detection mechanism is based on plastic scintillator arrays coupled to photomultiplier tubes, optimized for the light output of milli-charged particles. The tiles are arranged in two geometries, the "bar" and the "slab", covering different ranges of electric charge. Rock shielding between the IP and the detector ensure the suppression of beam backgrounds, and cosmic rays are discriminated based on the placement of the scintillators. If discovered, these particles would represent strong DM candidates. Exclusion limits provided by a milliQan demonstrator [173] suggest good future performances.

2.2 The CMS Detector

CMS can be found in the french village of Cessy, located 100 m underground, and is undoubtedly one of the most complex scientific instruments ever built. It derives its perceived "compactness" from a comparison with ATLAS, since it is twice as small but twice as heavy. Its acronym is also relate to two fundamental components of CMS's structure, namely an imposing 4 T solenoid magnet, and a large system of muon detectors which provide an exquisite muon resolution. A community composed of almost 6300 people works for the success of CMS, where approximately 2100 physicists, 1400 undergraduate students, 1200 PhD students and 1100 engineers are included. The experiment spans five continents, with the involvement of 57 countries and regions, totaling 247 institutes at the time of writing.

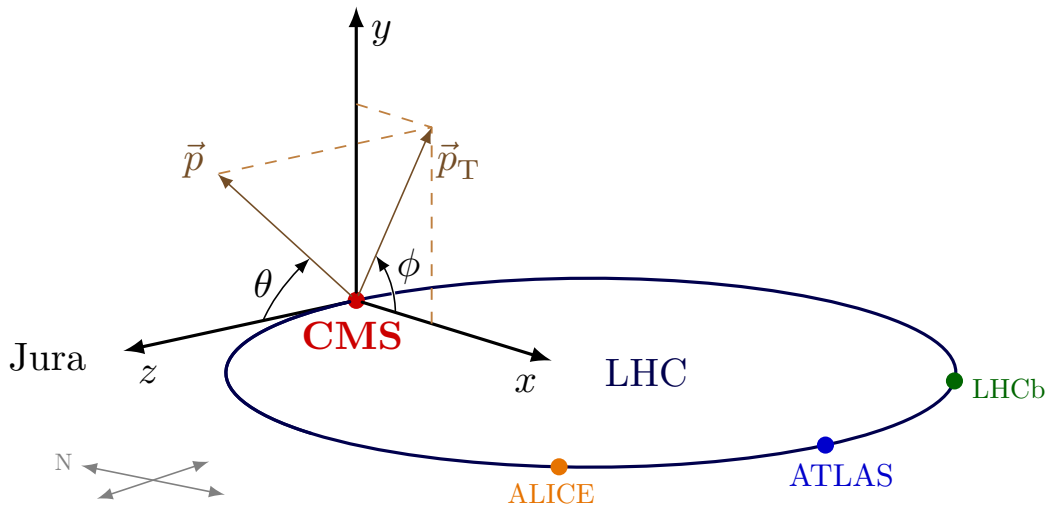

Figure 2.4: The coordinate system of the CMS detector, with the IP at its origin. The geographical location of CMS with respect to the other large LHC experiments is also provided. Adapted from [17].

The CMS experiment was officially proposed in 1992, with the primary goals of discovering the Higgs boson and exploring BSM physics at the TeV scale. Key to this objective were the $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ and $\mathrm{H} \to \gamma\gamma$ physics channels, which required an excellent EM calorimeter for effective electron and photon detection. Nevertheless, all subdetectors in CMS were designed to be versatile, and capable of measuring a wide range of physics processes across various energies, making CMS truly a multi-purpose detector, as the more than 1000 diverse papers so far published can attest.

In this Section we start by defining the CMS coordinate system which will be used across the remaining Chapters of this Thesis. Next, a detailed description of individual components of CMS is provided, since they represent the basis for understanding the trigger, reconstruction and simulation strategies later discussed.

2.2.1 Coordinate system

CMS follows a right-handed coordinate system, where the origin is placed at the IP. Following the diagram in Fig. 2.4, the x-axis points towards the center of the LHC ring, the y-axis points vertically upward, and the z-axis points towards the Jura mountains. The cylindrical structure of CMS naturally favors the use of a cylindrical coordinate system, given by:

$$\begin{cases} x = \mathrm{R} \sin\theta \cos\phi \\ y = \mathrm{R} \sin\theta \sin\phi \\ z = \mathrm{R} \cos\theta \end{cases} \tag{2.4}$$

where x, y and z are the standard 3D Cartesian coordinates, the radial coordinate is given by $\mathrm{R} = \sqrt{x^2 + y^2}$, the azimuthal angle ϕ is measured in the x-y plane from the x-axis, and the polar angle θ is measured from the z-axis in the y-z plane.

To characterize particle interactions at the IP, another quantity is often considered, namely the *rapidity* of a particle:

$$\text{rapidity} = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right), \tag{2.5}$$

where E is the particle's energy and p_z its longitudinal momentum. A Lorentz-boost along the z-axis shifts the rapidity by a constant term which depends on the boost itself. This means that rapidity differences between pairs of particles are invariant with respect to Lorentz boosts along z. In most HEP pp collisions, the distribution of final state particles is approximately uniform in rapidity [174]. This explains why detector elements are often uniformly spaced in rapidity.

However, we note that, contrary to what happens in accelerators colliding fundamental particles, rapidity uniformity does not hold true for individual pp collisions, given the partonic distributions within each colliding proton. Different collisions can have wildly different rest frames. Additionally, due to the absence of detector components along the longitudinal direction close to the beamline, which would interfere with the pp beams, the momentum along the z-axis remains unknown. Most importantly, the momentum fraction carried along the z-axis by each colliding parton is also unknown, as well as the longitudinal boost of the rest frame of the collision. It is thus beneficial to use variables that do not depend on p_z, being also Lorentz-invariant for those longitudinal boosts. The full energy and 3D momentum of a particle is thus replaced by the transverse momentum p_T and the transverse mass m_T:

$$\begin{cases} p_\mathrm{T}^2 = p_x^2 + p_y^2 \\ m_\mathrm{T}^2 = m^2 + p_x^2 + p_y^2 = E^2 - p_z^2 \\ E_\mathrm{T}^2 = m^2 + p_\mathrm{T}^2 \end{cases}, \tag{2.6}$$

where the transverse energy E_T is defined based on the mass m and transverse momentum. For the same reasons, the rapidity of a particle is not easy to measure. A new quantity is thus defined, dependent solely on the angle θ and illustrated in Fig. 2.5 (left), called *pseudorapidity*, or η:

$$\text{rapidity} \approx \frac{1}{2}\ln\left[\frac{E(1+\cos\theta)}{E(1-\cos\theta)}\right] = -\frac{1}{2}\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \equiv \eta \tag{2.7}$$

In the limit of large momenta, the rapidity and pseudorapidity become equal, but $y < \eta$ is always verified. Differences of η are thus also Lorentz invariant for $p \gg m$. Note that, despite the rapidity being finite, η can be infinite for particles produced at $\theta = 0°$ or $\theta = 180°$. One can also use the above to define a Lorentz-invariant spatial separation between two particles, which will be often used in later Chapters:

$$\Delta\mathrm{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2.8}$$

Due to mechanical limitations, subdetectors in HEP are frequently split according to their η coordinates, where the *barrel* section stands for a central region with low η values, and a *forward* region or an *endcap* refers to a high $|\eta|$ region, corresponding to the bases

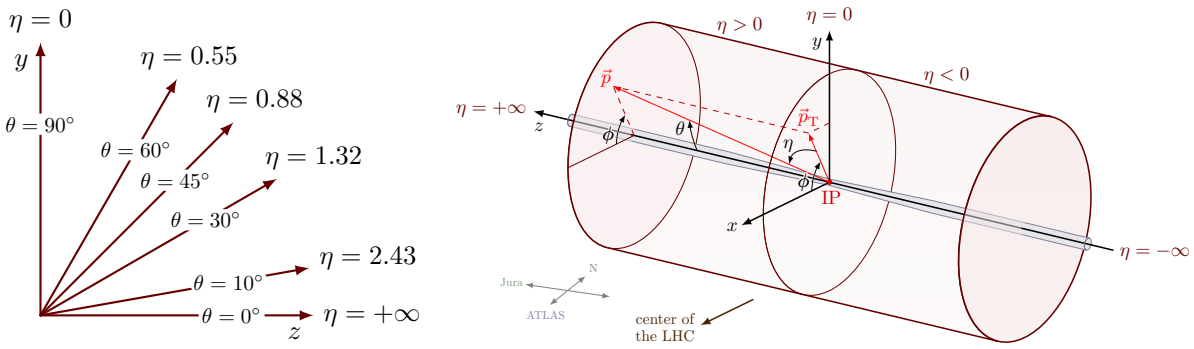

Figure 2.5: (Left) Schematic of different η values and their correspondence with the polar angle θ. (Right) Standard coordinate system at CMS, represented in relation to the LHC and its experiments, and including the pseudorapidity. Courtesy of Izaak Neutelings [17].

of the cylinder roughly defined by CMS. We show a visual conversion between θ and η values in Fig. 2.5 (right), together with its inclusion in the full CMS coordinate system.

2.2.2 Detector Structure

The architecture of the CMS detector is depicted in Fig. 2.6. It consists of a cylindrically-shaped apparatus with a length of 21.6 m and a diameter of 14.6 m, weighting 14 kt, and comprising a central section known as the barrel, and two forward sections called endcaps. To fulfill its goals, CMS includes multiple concentric detection layers, each tailored to identify specific decay products of high-energy collision events. Enveloping the calorimeters, the CMS superconducting solenoid provides a powerful magnetic field exploited by all subdetectors. Occasional redundancy in the data from the various subdetectors is intentional, to maximize the reconstruction efficiency. This Section details the individual components of the CMS detector, ordered by distance to the PV, paying particular attention to each subdetector's role in the detection of a specific particle type.

Tracker

The CMS tracker [176, 177] is the first layer of detection a charged particle sees after being produced at a pp collision. It is an all-silicon detector, conceived to withstand very large fluences of up to $3.6 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ after $500\,\mathrm{fb}^{-1}$ of integrated luminosity. Its working principle is well known: charges moving in the semiconductor electric field will generate a detectable electric current in an external circuit [178]. The tracker's highly granular semiconductor technology is organized in several layers, and allows the precise reconstruction of the trajectories of charged objects which, in conjunction with the strong magnetic field later discussed, provide information on charge, direction, PV assignment and PU. Furthermore, the tracker enables the reconstruction of secondary vertices originating from b or τ decays, and facilitates studies on displaced tracks [179]. The tracker is composed of two detection technologies, pixels and strips, located in separate regions, as shown in Fig. 2.7. In total, it measures 5.6 m in length and 2.4 m in diameter. Pixels usually provide higher resolutions and lower noise, but are also more expensive than strips due to interconnect technologies [178]. In CMS, pixels occupy the inner region and are immediately surrounded by the strips.

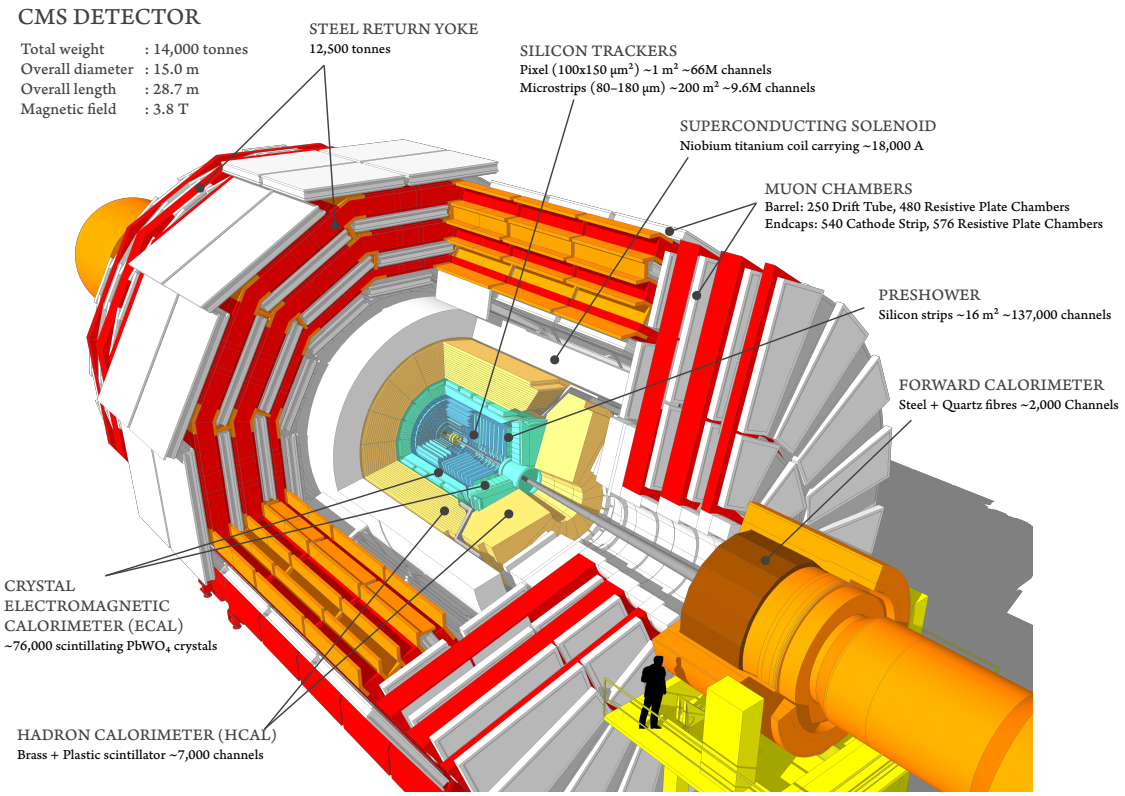

Figure 2.6: Cutaway 3D model of the CMS detector. All subdetectors are visible and labeled, and are described in detail in the text. The black figure gives a sense of the sheer scale of the detector. Taken from [175].

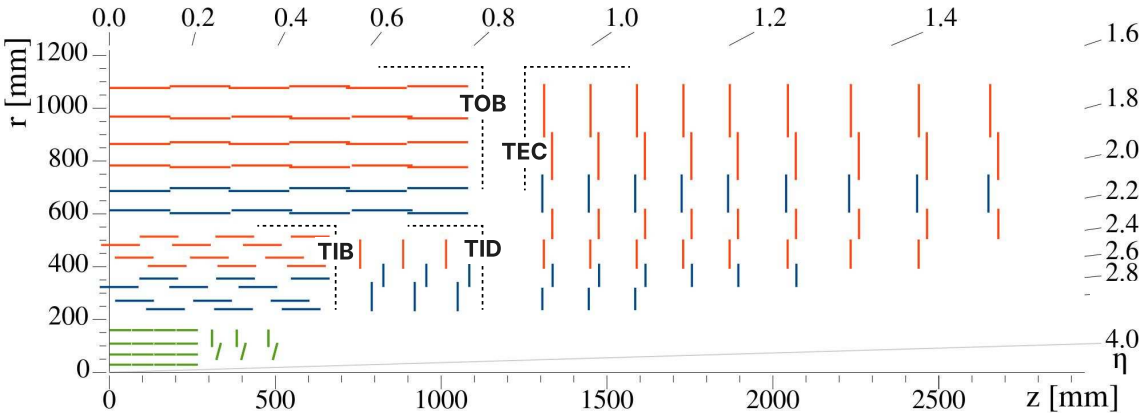

Figure 2.7: 2D R vs. z projection of one quarter of the CMS tracker. The pixel detector is depicted in green, while single-sided and double-sided strip modules are shown as red and blue segments, respectively. The four components of the strip detector are indicated with dashed lines: Tracker Inner Barrel (TIB), Tracker Inner Disk (TID), Tracker Outer Barrel (TOB) and Tracker EndCap (TEC). Adapted from [180].

- **Pixels**: The pixel detectors are organized in modules, for a total of 128×10^6 readout channels, each one measuring 100×150µm. In total, pixels cover a Si area of 1.1 m^2. The pixel detector is optimized to have four-hit coverage up to $|\eta| \sim 2.5$, improved pattern recognition and track reconstruction, and added redundancy to cope with hit losses. The innermost layer in the barrel pixel (BPIX) sits at only 29 mm from the IP. This represents an upgrade over the original Phase 1 pixels, installed in the 2016/2017 technical stop, and which also required the installation of a new and tighter beam pipe. Three other layers follow at 68 mm, 109 mm and 160 mm radii. The forward pixel (FPIX) endcaps are covered by three disks on each side, placed at 296 mm, 396 mm and 516 mm from the IP. The detectors are surrounded by "service cylinders", which provide all needed readout and control circuits, as well as cooling tubes and power lines (Fig. 2.8, left). Despite the presence of one extra layer, the material budget has been substantially reduced to a minimum in the endcaps and more dramatically in the barrel, by a factor of ∼60 %: electronic boards on the cylinders were placed outside the tracker acceptance, and special materials were employed for the mechanical structure and CO_2 cooling system (Fig. 2.8, right). This is intended to decrease phenomena such as unintended multiple scattering, pair production, and nuclear interactions, which can create *preshowers*, i.e. particle showers developing before reaching the calorimeters. The cooling itself runs at −10 °C and 30% relative humidity. The cooling, besides dissipating the generated power, also helps to mitigate the effects of damage to the Si pixels. Each pixel provides a spatial resolution of 10 µm and 20 µm along the transverse and longitudinal directions, respectively.

- **Strips**: The strip tracking detector is located at larger distances from the IP, starting at 20 cm, and is made of $\sim 10^6$ microstrip sensors, also organized in modules. Its η coverage resembles the one from the pixels. The strips are subdivided into four sections, which are clearly delimited in Fig. 2.7. The TIB and TIDs are made of 4 concentric layers and 3 disks, respectively. The TOB (TEC) are located a bit further, and comprise 6 layers (9 disks). Roughly half of the modules are double-sided, made of two independent single-sided modules glued back-to-back with a relative rotation of 100 mrad. This allows the determination of the third spatial component, while single-sided modules access 2D positions only [181]. All components are cooled at −15 °C. The barrel extends up to 116 cm in radius, and the endcaps reaches 282 cm in depth. The strips achieve a resolution between 20 and 50 µm in the transverse plane, and between 200 and 500 µm in the longitudinal direction. The CMS strips are the world largest tracking system ever, with 93 000 strips covering 198 m^2.

Electromagnetic Calorimeter

The scintillating lead tungstate ($PbWO_4$) electromagnetic calorimeter (ECAL) [183] is located inside the 3.8 T solenoid magnet, consisting of 75 848 channels in total, split across 61 200 crystals in the barrel and 7324 crystals in each of the two endcaps. The medium was chosen based on its compact $X_0 = 0.89$ cm radiation length, providing a total containment of 25 X_0, and on its small 2.19 cm Molière radius, providing full coverage of all EM showers. Its hermetic design targets the measurement of the properties of photons and electrons, via their scintillation and absorption in the crystals. These particles are

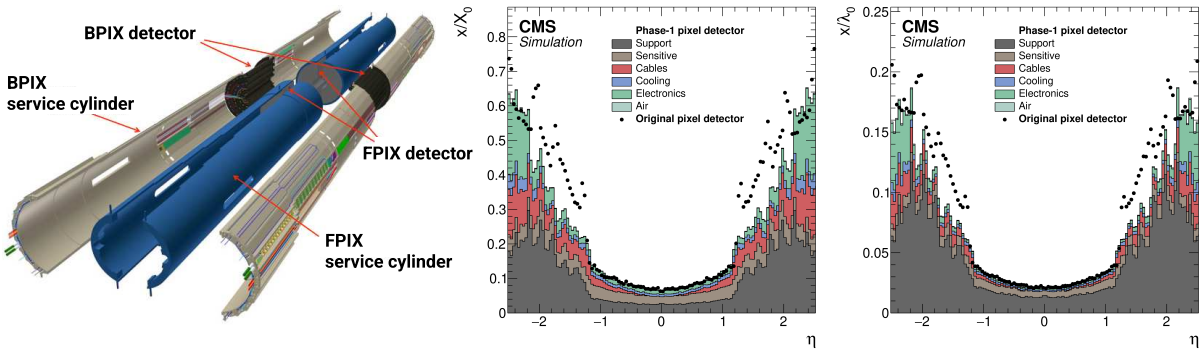

Figure 2.8: (Left) 3D layout of the Phase 1 BPIX and FPIX detectors with their respective service half-cylinders. (Middle, Right) Material budget in units of radiation length and hadronic interaction length, as a function of η, as obtained from simulation. The material budget of the CMS Phase 1 pixel detector is split into the contributions of the different categories, and the black dots display the original material budget before the technical stop update in 2016/2017. The disk structure of the endcaps leads to the observed peaks. Adapted from [182].

reconstructed over a wide range of energies, from ∼5 GeV electrons to hypothetical multi-TeV resonances. The ECAL Barrel (EB) covers a $|\eta| < 1.479$ region. EB crystals are organized into *modules*, in turn arranged into *supermodules*, each spanning half the length of the EB and covering 20° in the azimuthal angle (ϕ). The supermodules also contain Avalanche Photodiodes (APDs) and readout electronics. Each module houses 425 crystals, which measure $\eta \times \phi = 0.0174 \times 0.0174$ and 23 cm in depth. In each ECAL Endcap (EE), crystals are instead arranged in two semi-circular support called *dees*, and form 5×5 groups named *supercrystals*, extending up to $\eta = 3.0$. The crystals are read out by two kinds of photo-detectors: APDs in the barrel and Vacuum Phototriodes (VPT) in the endcaps. All crystals are placed perpendicularly to the IP in a quasi-radial geometry. This layout mitigates the negative impacts of structural gaps, which are nevertheless present, mostly at $\eta = 0$ and $\eta = 1.479$. A EM preshower detector is present right in front of the EE disks to increase the spatial resolution in the $1.65 < |\eta| < 2.6$ region. Between ∼6 % and ∼8 % of the energy of EM showers deposit their energy here, on average. The preshower is composed of two layers of radiation-hard Si strips alternated with Pb absorbers, in the guise of a short, ∼1 X_0 sampling calorimeter. Its main purpose is to distinguish single photons from $\pi^0 \to \gamma\gamma$ double photons. Despite its name, the ECAL also measures hadronic energy deposits and the EM component of hadronic showers, since it is placed in front of the hadronic calorimeter (HCAL). The ECAL also provides timing capabilities for EM showers, which are used for noise rejection and in the removal of debris from time-adjacent bunch-crossings (BXs). Additionally, timing is useful in long-lived particle (LLP) searches, especially the ones with a lifetime larger than 1 ns [184]. The structure of the ECAL is illustrated in Fig. 2.9, and the origin of the gaps is shown in Fig. 2.10, where a photo of a supermodule is included.

The energy response of calorimeters is usually parameterized via three separate terms [146], which have been empirically observed to precisely describe resolution curves:

- **Stochastic term**: Arises due to intrinsic Poissonian fluctuations in the shower development, which for ECAL corresponds to the number of produced scintillation photons. Since the number scales with energy, the Poissonian uncertainty of the mean scales with the energy's square root.

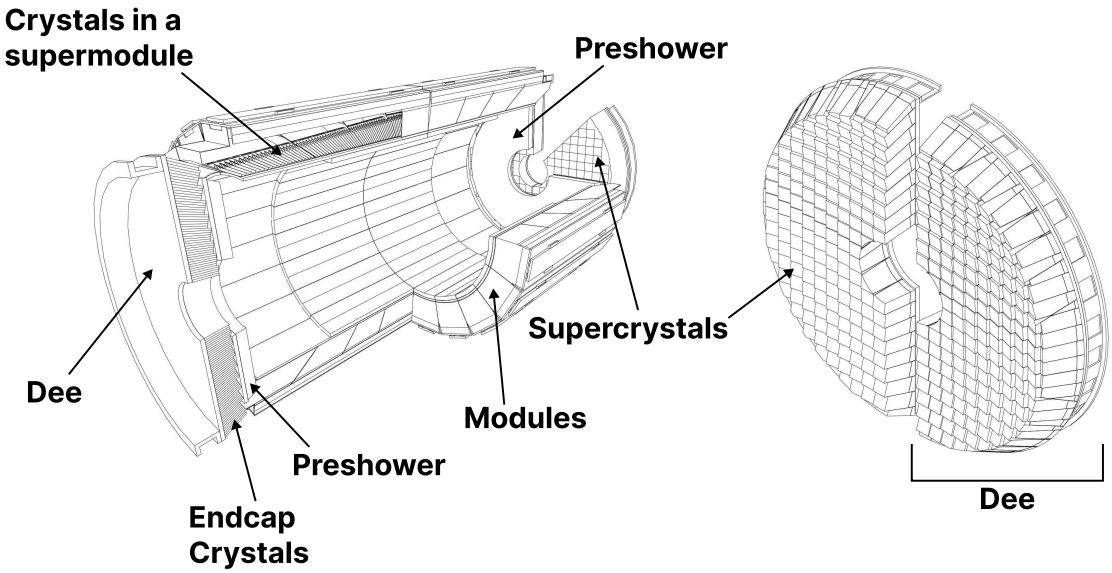

Figure 2.9: (Left) Conceptual representation of the ECAL mechanical structure. The lead-tungstate crystals are housed in the modules and supermodules of the barrel, while in the endcap they are arranged between the preshower and the support dees, grouped in supercrystals. (Right) A single endcap with Dees apart, showing its supercrystals. Adapted from [183].

- **Noise term**: Comes from the electronic noise of the readout chain, and scales linearly with the energy. It depends on features such as the detector capacitance and cabling. The noise tends to be larger for charge-collecting devices, due to the presence of preamplifiers. Scintillation-based detectors such as ECAL have photosensitive devices at the start of their circuit, lowering the noise.
- **Constant term**: Includes all sort of contributions which do not depend on the energy of the particle, like instrumental effects, nonuniformities and structural imperfections. As the center-of-mass energy of particle accelerators increases, this term will become more and more significant.

In the case of ECAL, the energy resolution of 3×3 or 5×5 crystal arrays was measured by a dedicated electron test beam [185], with energies varying from 20 to 250 GeV, and was parametrized as:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2\,, \tag{2.9}$$

where the first, second and third terms are, respectively, the stochastic, noise and constant term. The actual ECAL resolution was also measured, done in-situ with Run 1 7 TeV data [186], and was found to be excellent, ranging from 1.1 to 2.6% in the barrel and 2.2 to 5% in the endcaps for photons coming from a Higgs boson decay.

Hadronic Calorimeter

Immediately at the back of ECAL appears HCAL [187], a sampling calorimeter which measures final-state hadrons and hadronic showers depositing only a part of their energy in ECAL. Showers initiated by hadrons are governed in part by the strong interaction, which significantly complexifies the involved dynamics [188]. In a calorimeter, a hadron

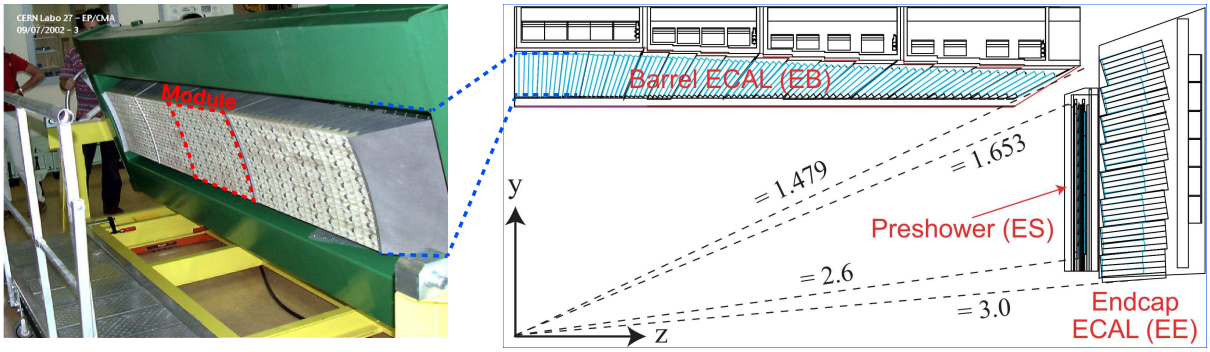

Figure 2.10: (Left) Photograph of one supermodule with its modules clearly visible. (Right) Structure of a quarter of ECAL, highlighting individual modules, supermodules and supercrystals. The spacings between supermodules and supercrystals explain the η gaps at 0 and 1.479. Adapted from [183].

can lose its energy either via continuous ionization (if it is charged) or abruptly via nuclear interactions. To make things more complicated, neutral hadrons can only interact strongly, and so their interactions with matter enormously differ from what happens with EM showers. It can thus happen that an hadronic shower loses its energy mostly via EM interactions, but a single hadron can also penetrate the full ECAL length without releasing any energy whatsoever. Another option is for a shower to lose its energy by releasing nucleons from the medium. The binding energy required to release those nucleons is called *invisible energy*, since it is inherently undetectable and thus lost for calorimetric measurements. All of these reasons lead to a hadronic resolution that is generally much worse than the EM resolution, and can only be improved by correlating the information from multiple subdetectors with software techniques described in Section 2.4.1. On average, only 1/3 of the energy of a hadron shower is deposited in ECAL. The scale of the shower development is largely dominated by the nuclear interaction length, λ_0. Importantly, the absorption of hadronic showers tends to require much more material than the absorption of EM showers at the same energy. This fact explains the location of HCAL behind ECAL, an approach used to achieve full containment, with its $\sim 7\,\lambda_0$ at $\eta = 0$ up to $\sim 12\,\lambda_0$ at $\eta \sim 1.2$, then stabilizing at $\sim 10\,\lambda_0$ for the endcaps.

The lower radiation levels that HCAL experiences when compared to ECAL or even the tracker make possible the usage of inexpensive plastic scintillator tiles as active medium, themselves interleaved with brass absorber layers. The calorimeter is composed of five separate sections: the HB, two HCAL Endcaps (HEs) on either side of CMS, the HO and the HF. The HB is mechanically segmented into 36 wedges, each spanning half the length of the HB, covering an azimuthal sector of 20°, and weighting about 25.7 t. Inside each wedge, plastic scintillators are organized into 16 η regions. In the endcaps, the HEs are instead arranged such that their absorber plates are bolted together to form a single 18-sided polyhedral structure (see Fig. 2.12, middle) with insertion gaps for the scintillator trays. Each of the 18 sides weighs $\sim$300 t, being fastened to CMS to avoid gaps at the HB-HE interface, in order to avoid geometric inefficiencies, just like the ones in ECAL. The HB is complemented by the HO for $|\eta| < 1.4$, located just outside the solenoid magnet, but using the return yoke as absorber material, while the active medium remains Sci plastic. The HO captures the tails of hadronic showers, adding $1\,\lambda_0$ to the $10\,\lambda_0$ from the HB, which was judged not high enough to fully contain hadronic showers. The light produced in the plastic scintillators is wavelength-shifted and captured in fibers. The HCAL is completed by the HF, which is placed at $z = \pm 11.2$ m from the IP and

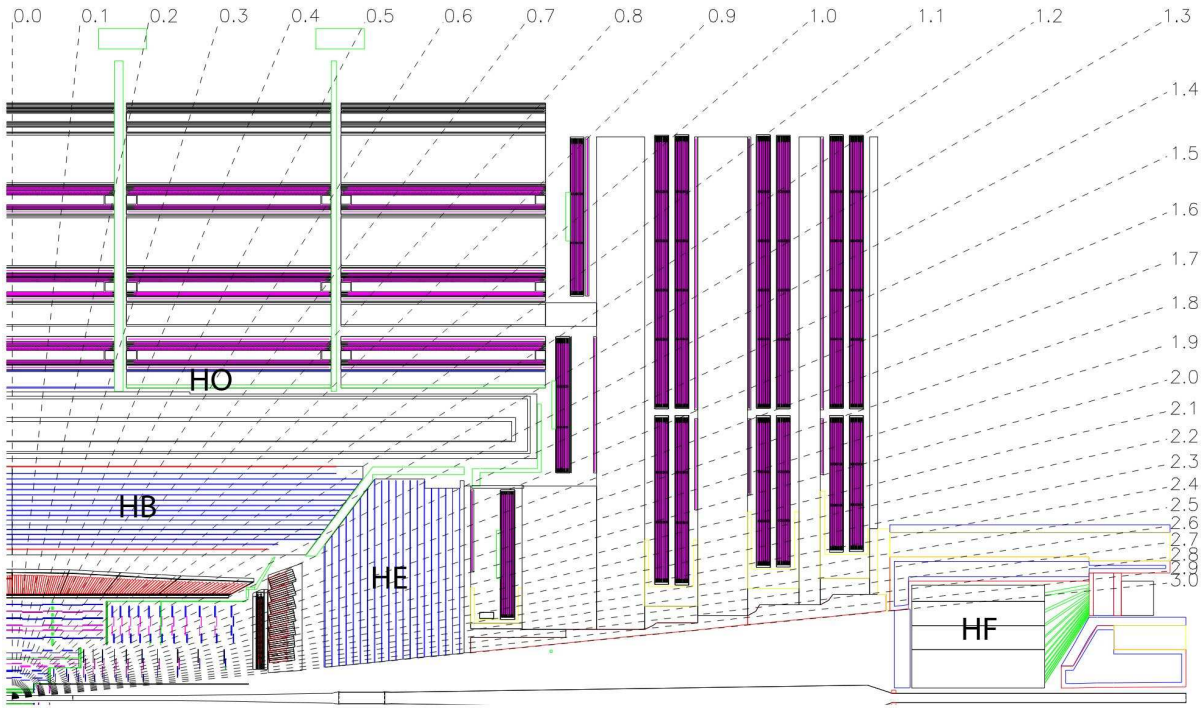

Figure 2.11: Schematic view of a quarter of the hadronic calorimeter, along the longitudinal direction. The four section are shown: HCAL Barrel (HB), HCAL Outer barrel (HO), Hadron Forward calorimeter (HF) and HF. The dashed lines provide visual guidance for the η coordinate. Taken from [189].

covers up to $|\eta| = 5.2$. The HF is also a sampling calorimeter, with alternating layers of steel absorber and quartz fibers as active material. The fibers produce scintillating light via the Cherenkov effect and send it to PMTs. The relative orientation of some of ECAL's and HCAL's components facilitates the calorimetric trigger processing chain. Indeed, both the HB and HE are segmented into projective towers, aligning, respectively, with the EB and EE. In the HB, $\eta \times \phi = 0.087 \times 0.087$ towers match the 5×5 ECAL crystal arrays, while the HE is segmented into coarser $\eta \times \phi \sim 0.17 \times 0.17$ towers. A schematic representation of HCAL can be seen in Fig. 2.11 and individual components are shown in Fig. 2.12.

All components in HCAL used to included Hybrid Photo-Diodes (HPDs) for $|\eta| < 3$. Initially seen as beneficial due to their high magnetic field tolerance and large gains, issues on high voltage electric discharges lead to significant increases in the overall HCAL noise. As a consequence, it was decided to progressively replace HPDs with silicon photomultipliers (SiPMs), and the procedure was completed during the Long Shutdown 2. The upgrade also introduced more performant electronics and data linking, which increased the segmentation in HB and HE, and also improved timing measurements. Hadronic shower development is thus measured more precisely, boosting the performance of analyses targeting signatures containing delayed or displaced jets.

Magnet

The large, 220 t Nb-Ti superconducting solenoid magnet is the defining feature of the CMS design, delivering an axial and uniform magnetic field of 3.8 T over a 12.5 m length and a 3.15 m radius [190]. The radius is large enough to accommodate both EM and

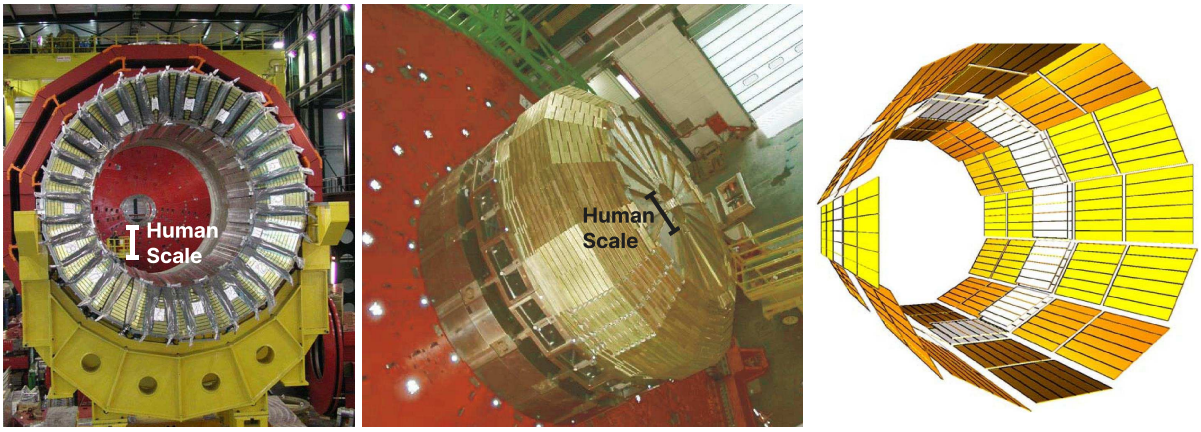

Figure 2.12: (Left) Assembled HCAL half-barrel. (Middle) Partially assembled HEm without the absorber, where Sci trays can be seen inserted in some of the outer sectors. (Right) Layout of all the HO trays in the CMS detector. Adapted from [189].

hadronic (HAD) calorimeters, reducing the material budget in front the calorimeters. This eliminates charged particle preshowers in the coil material, facilitating the matching between energy deposits and tracks. At normal incidence, the bending power of 4.9 T m a provides a strong separation between energy deposits of charged and neutral particles. As an example, a 20 GeV p_{T} charged particle deviates ∼5 cm in the transverse plane at the surface of ECAL (at 1.29 m from the PV), which is enough to distinguish it from a photon coming from the same direction. The precise bending is estimating using a 3D magnetic field map, with an accuracy of less than 0.1 % [191]. We can roughly get the right numbers by applying the R = p/q B formula, where R is the radius of the trajectory in the transverse plane, q the particle's charge and p it's momentum, and B the value of a constant magnetic field. The magnet is cooled by liquid He, and must thus operate at −269 °C. It is for this reason enclosed in a vacuum vessel made of two stainless steel cylinders. In order to contain the magnetic flux, the solenoid is surrounded by a return yoke, which is conveniently interleaved with the muon chambers to additionally provide structural support and increase muons momentum resolution.

Muon Chambers

CMS is specifically optimized for muon measurements, which are performed by Drift Tubes (DTs) in the barrel region and Cathode Strip Chambers (CSCs) in the forward region. RPCs are also available for triggering and redundancy. The entire system is based on gaseous detectors, and is located outside the solenoid, where the distance to the PV is large enough so that only muons are expected. Indeed, muons produced at the LHC, with energies ranging from a few MeV to several GeV, are the closest a particle becomes from being a MIP, as shown in Fig. 2.13, and thus traverse large quantities of matter remaining mostly undisturbed. In particular, they are not stopped by the calorimeters. We note that muons have a mass ∼200 times larger than the electrons, rendering bremsstrahlung effect comparatively minor. The barrel section of the muon chambers is composed of four muon stations interleaved with the steel return yoke, which provides mechanical support. The ∼1.8 T magnetic return flux can thus be used to measure muon momenta, independently from the tracker. A dedicated muon based trigger is thus possible, and combining muon position and p_{T} measurements with the tracker becomes a powerful tool. However, the extreme proximity to the return yoke also creates negative effects, namely

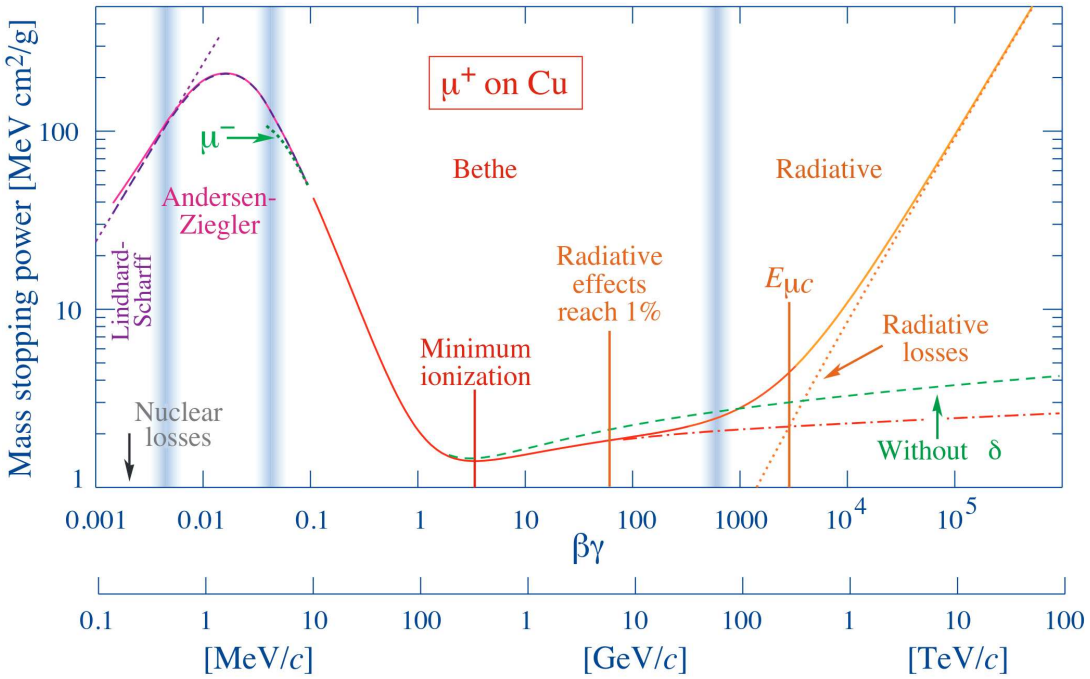

Figure 2.13: Mass stopping power, in MeV $\mathrm{cm}^2\,\mathrm{g}^{-1}$, for positive muons in Cu as a function of $\beta\gamma \equiv \mathrm{p/M}$ and energy, with p being the momentum and M the energy, over 12 orders of magnitude in energy. Muons produced at the LHC behave similarly to minimum ionizing particles (MIPs). Solid curves indicate the total stopping power. Vertical bands indicate boundaries between different approximations. The mass stopping power in the radiative region is not simply a function of $\beta\gamma$. Further discussion available in [13, Chpt. 34], where the figure was taken.

the presence of EM showers induced by muon bremsstrahlung, which degrades momentum resolution. A highly redundant muon system is therefore found ideal to preserve physics performance. We can indeed find RPCs present both in the barrel, together with DTs, and in the endcaps, with CSCs. The redundancy also plays a role in reducing the impact from acceptance blind spots introduced by the support mechanisms and cabling of such large detectors [192]. The structure of the muon chambers, including future upgrades, is shown in Fig. 2.14.

- **Drift Tubes:** Present in the barrel section, they consist of drift chambers aiming at providing position resolutions of the order of 100 µm and time resolutions lower than 6 ns. A single 4 cm-wide tube contains a stretched wire within a gas volume. When a charged particle passes through the gas, it knocks electrons off the gas atoms. The electrons drift along the electric field's direction, reaching the anode and producing a signal. The DTs ensure a constant drift velocity along the entire drift path, which enables the identification of the two-dimensional point in space where the charged particle, a muon in this context, crossed. Each DT module ranges from $2\times2.5\,\mathrm{m}^2$ to $4\times2.5\,\mathrm{m}^2$ in size, and is composed of two or three superlayers (SLs). Each SL contains in turn four aluminum layers of staggered DTs. A SL thus provides four two-dimensional points to measure the muon's position. SLs within a module are aligned in two perpendicular directions, which allows a three-dimensional measurement of the position of the muon track. The modules are ultimately laterally arranged in five sections, or *wheels*, with a depth of four stations.
- **Cathode Strip Chambers**: Stationed in the endcaps, CSCs consist of arrays of positively-charged wires perpendicular to negatively-charged Cu strips, all within a

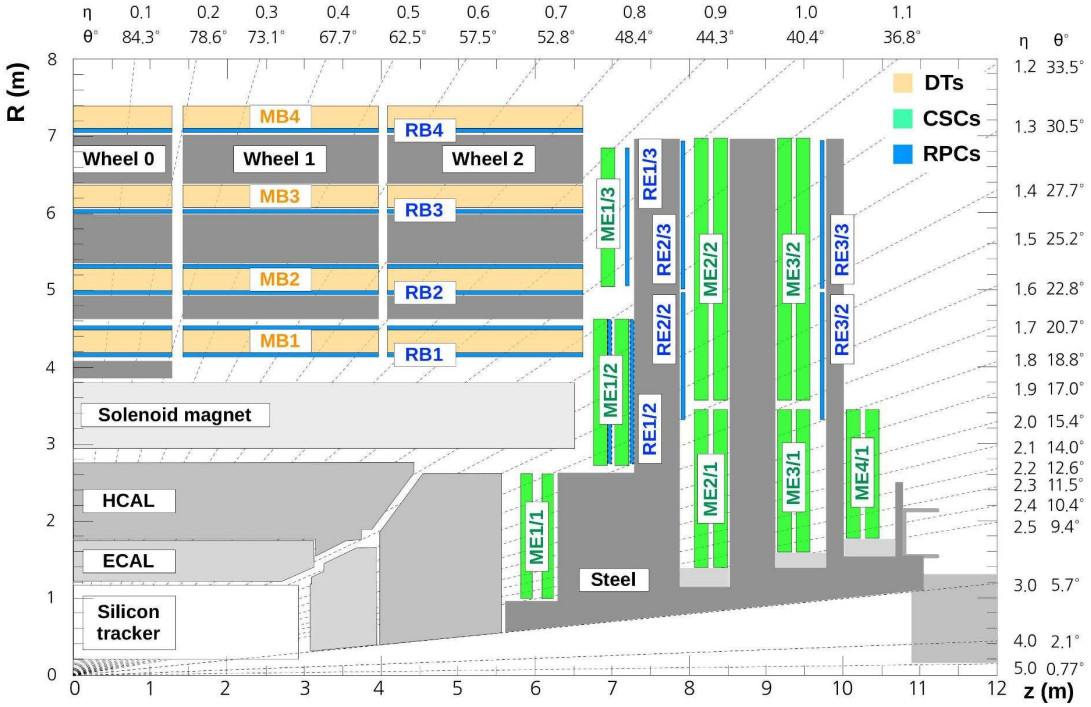

Figure 2.14: Schematic longitudinal view of a quadrant of the R-z cross section of the CMS detector during Run 2, when the Gas Electron Multiplier (GEM) detector was not yet present. All muon subdetector are shown: DTs (yellow), CSCs (green) and RPCs pseudorapidity values are given with dashed lines. Taken from [193].

gas volume. When muons pass through, electrons get knocked off the gas atoms. Both displaced electrons and ions follow the electric field, inducing signals in the wires and strips, respectively. The relative positioning of wires and strips enables a 2D position measurement for each passing muon. The existence of six layers per CSC module significantly increases the precision of the measurement. The resolution for one layer is in the 80 to 450 µm range, and approaches 50 µm when combined. The timing resolution is similar to the one in the DTs.

- **Resistive Plate Chambers**: They are present in both the barrel and in the endcap, and provide trigger redundancy with respect to DTs and CSCs. RPCs are made of two parallel plates defining an electric field separated by a thin gas volume. Like for the other muon detectors, when muons pass through an RPC, they knock out some of the gas electrons, creating electron avalanches. Those electrons traverse the plates without interacting and, after a precisely known time delay, are picked up by external metallic strips. Despite providing a limited spatial resolution, the time resolution goes as low as 1 ns. All muon stations are equipped with at least one RPC, but two are present in the inner barrel to compensate for the lower resolution of low p_{T} muons. The additional resolution extends the CMS trigger low-p_{T} reach to ∼4 GeV in the barrel and ∼2 GeV in the endcaps.

- **Gas Electron Multipliers**: A first batch of 144 GEM chambers, called GE1/1, was introduced in the CMS muon system during the Long Shutdown 2. They are located very close to the beampipe, subject to the highest radiation doses among all muon detectors. The GEMs improve the measurement of the muon polar bending angle, extending previous trigger capabilities. The provided η coverage of the muon

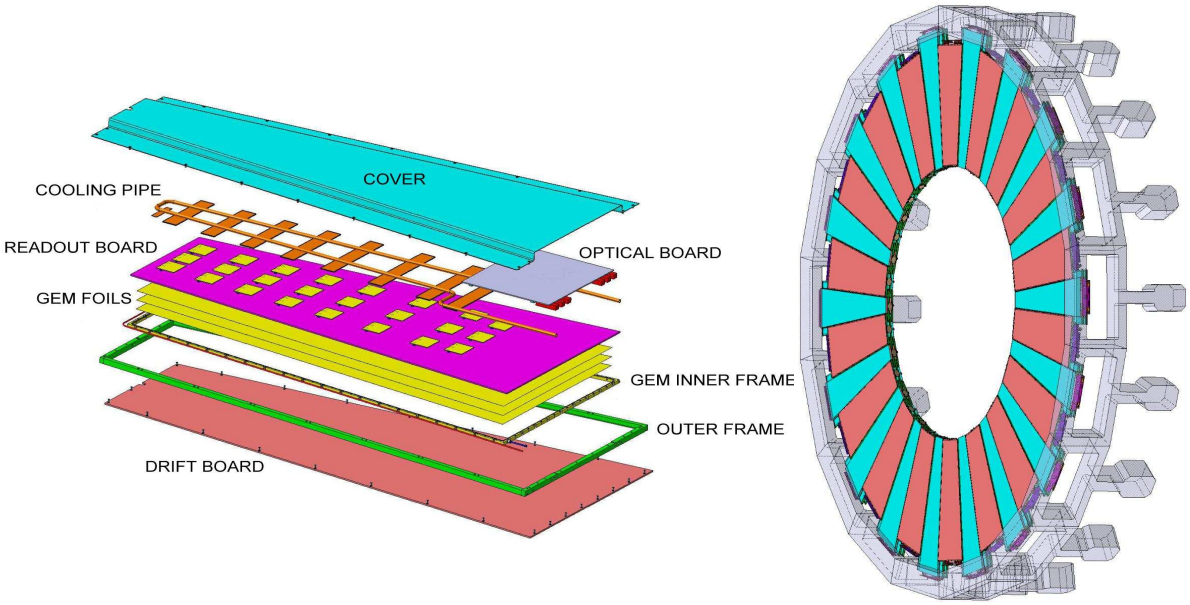

Figure 2.15: (Left) Mechanical design blowup of the a triple-GEM chamber, following the description in the text. (Right) Positioning of short and long chambers in the CMS endcap. During the next phase of the LHC, the HL-LHC, GEM detectors will be placed right at the back of HGCAL, the novel calorimeter discussed in Section 3.2.3. Adapted from [194].

detectors will also be extended, up to the forward $1.55 < \eta < 2.18$ region. The chambers come in two alternating sizes, in order to maximize the η coverage while fitting in the available volume, which is constrained by the support structure. Each GEM chamber includes a stack of three GEM foils, which consist of a 50 µm-thick insulating polymer covered on both sides by thin copper conductive layers. A strong electric field is applied between the two conductors. In total 36 superchambers have been installed, where each superchamber is made of two chambers and covers 10 degrees in ϕ. The chambers are filled with a 70/30 Ar/CO_2 mixture, which is ionized by incident muons, and are segmented in strips along ϕ. The electrons created during the ionization process drift towards the foils creating avalanches. The resulting electron avalanche induces a readout signal on the finely spaced strips. The structure of the GE1/1 chambers can be seen in Fig. 2.15. Its location is shown in red in Fig. 3.8, right behind the future endcap calorimeter, presented in Section 3.2.3. The CMS GEMs are the largest GEM system ever installed, with an area of $\sim 0.5\,\mathrm{m}^2$ per chamber. They bring a combined spatial resolution of $\sim 100\,\mu\mathrm{m}$ and a timing resolution of $\lesssim 10\,\mathrm{ns}$ [194]. The greatest benefit of the early installation of part of this system is a Level-1 (L1) muon trigger improvement before the upgrades planned for the tracker and its trigger [195].

2.3 The CMS Trigger System

With proton beams crossing every 25 ns, and considering the average PU measured during Phase 1 (Fig. 2.2, right), around 10^9 pp interactions are expected every second at the CMS IP. Given $\sim 1\,\mathrm{Mb}$ zero-suppressed events, a $\sim 40\,\mathrm{Tb\,s^{-1}}$ throughput follows. Meanwhile, the maximum rate which can be realistically archived in real time by the CMS online

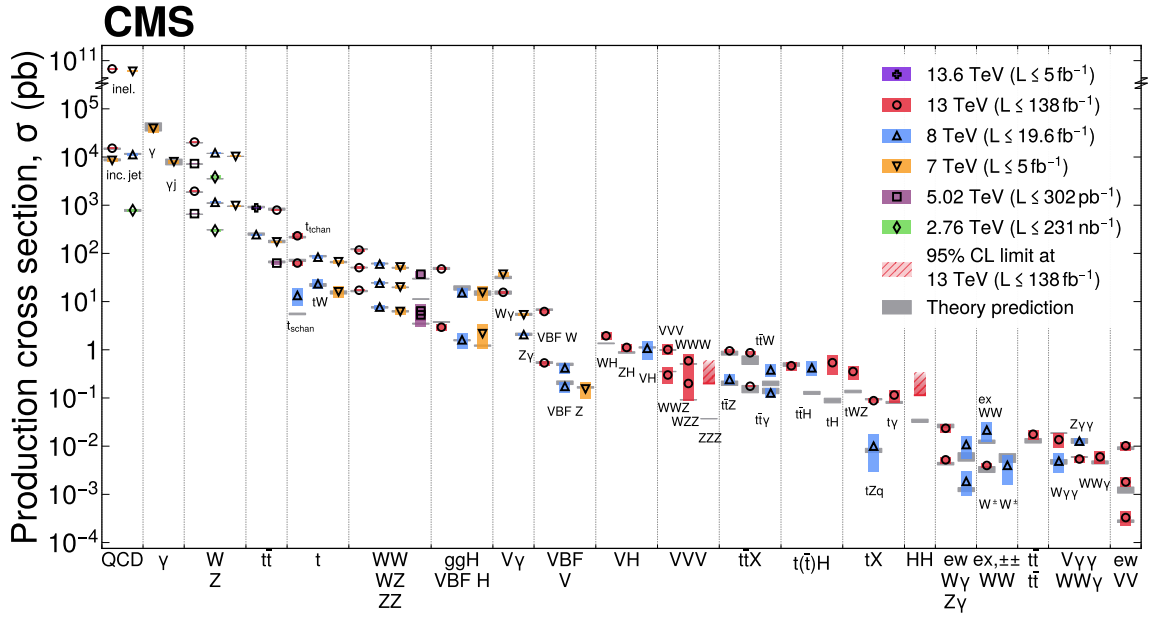

Figure 2.16: Summary of selected CMS cross section measurements of high-energy processes. Measurements performed at different LHC pp collision energies are marked by unique symbols and the coloured bands indicate the combined statistical and systematic uncertainty of the measurement. Grey bands indicate the uncertainty of the corresponding SM theory predictions. Shaded hashed bars indicate the excluded cross section region for a production process the measured 95% CL upper limit on the process indicated by the solid line of the same colour. Values span around 14 orders of magnitude. Taken from [196].

computer farm sits at ∼5 kHz, which already represents a 50-fold improvement with respect to Run 1. Collision and data rates are thus much higher than the rate at which data can be written to mass storage. This implies a throughput reduction by a factor of 10^4.

However, among all events produced in CMS, only a small fraction carries physical significance. Most events are related to well known physical interactions which are not meant to be covered by the LHC programme. As perfectly illustrated in Fig. 2.16, the processes studied by CMS span at least 14 orders of magnitude. The dominant inclusive proton-proton interactions tower at 10^{14} fb, followed closely by their inelastic counterparts at $\sim 6 \times 10^{13}$ fb [138, 197]. On the other end, extremely rare triboson, VBF, VBS, top and Higgs processes, including the HH process studied in Chapters 4 and 5, sit at the edge of what is currently achievable, with cross sections of the order of some fb, or even less. These huge differences showcase the richness of the LHC programme. They also show that a big fraction of the interesting processes have small cross sections. The large rate reduction provided by the trigger is therefore justifiable on physics grounds, besides being absolutely crucial given current technological limitations.

The needed data reduction is the reason for existence of the CMS trigger system: only events deemed interesting are selected for further processing. The CMS trigger adopts a two-tiered approach. In the first level, or L1, the trigger constrains the LHC rate from 40 MHz to 100 kHz, with a 3.8 µs latency. In the second level, the High-Level Trigger (HLT), a further reduction down to 1 kHz with a ∼300 ms latency is performed. Latency here refers to the time taken for the data to be processed in a particular trigger system. It includes the time taken to read the data from the detectors, the algorithmic

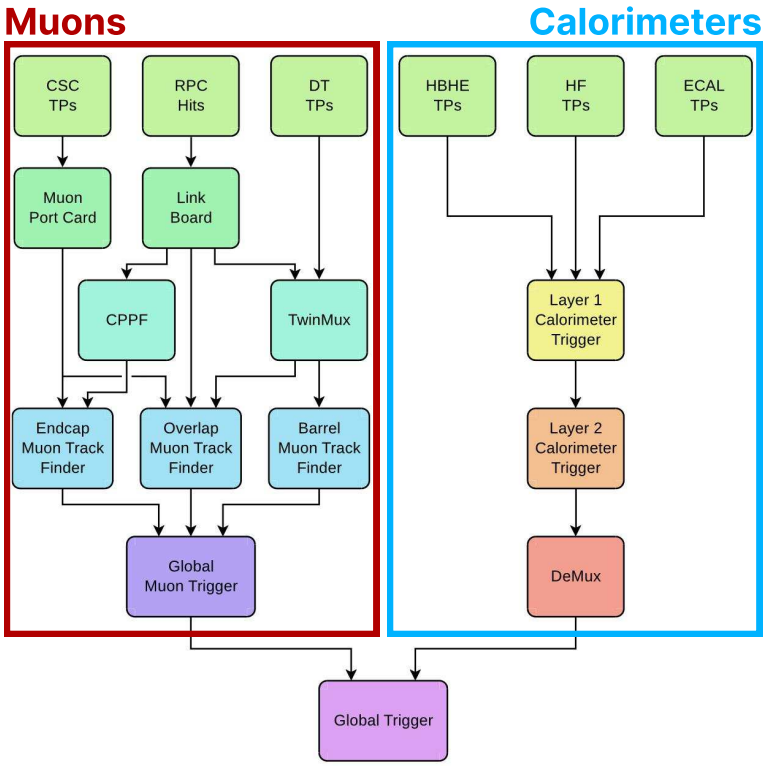

Figure 2.17: Diagram of the Phase 1 CMS L1 during Run 2. No L1 tracking is present. TPs are generated from the DT, RPC and CSC muon systems and from the HF, ECAL and hadronic barrel and endcap calorimetric subdetectors (where the latter refers to the HCAL). The two separate paths are merged into the GT, which make a L1A decision on whether each particular event should be kept. TPs from GEM are currently being validated, but are not yet used. Adapted from [200].

time, the L1-Accept (L1A) signal distribution time, and further delays introduced by potential limitations at the level of buffer capacity [192, 198–202].

2.3.1 The Level-1 Trigger

The first trigger level is implemented in custom electronics and Field Progammable Gate Arrays (FPGAs), and uses coarsely segmented data from calorimeter and muon detectors to ensure a low latency. All high-resolution data are nevertheless held in pipeline memories in the front-end (FE) electronics, waiting for a L1A signal to be issued. The data is stored for 3.8 µs, equivalent to 152 BXs, including ∼1 µs dedicated to trigger algorithms. Each event forwarded to the HLT takes ∼2 MB which implies, at a rate of 100 kHz, a bandwidth usage of 200 GB s^{-1} supported by the data acquisition (DAQ) system. The L1 trigger receives as input the raw data from the FE readout electronics. The first processing stages are handled literally on the detector, to allow for direct access and thus faster processing. The remaining stages are taken care of by racks of FPGAs located in the experimental cavern. The small time budget implies a compromise on the quality of the reconstruction, which is a recurrent theme in all trigger-related matters. The muon and calorimeter trigger paths produce data called trigger primitives (TPs), which are merged into the Global Trigger (GT), and used to issue a L1A decision. TPs from the GEM subdetector are currently being validated, but are not yet used. The tracker does not participate in the Phase 1 L1 trigger decisions. The L1 rate has been increased to

∼110 kHz during Run 3 to favor the introduction of the new parking techniques discussed in Section 2.3.3.

Calorimeter Trigger

The calorimeter trigger is split in two layers. The first one receives, calibrates, and sorts the TPs provided by the HF, ECAL and HCAL subdetectors. The second layer reconstructs and calibrates physics objects such as electrons and photons, tau leptons, jets, and energy sums. It should be noted that the calorimeter trigger cannot distinguish photons from electrons; they all correspond to energy deposits named *candidates*, which are formed based on information collected in groups of 5×5 crystals in the ECAL and one projective readout unit in the HCAL, jointly known as Trigger Towers (TTs). L1 reconstruction algorithms work with a granularity dictated by the TTs, using discriminative quantities such as the ratio of HCAL to ECAL energy, or the shape of energy deposits as seen by data-aggregating sliding windows. The calorimeter trigger takes advantage of a spatially-pipelined mechanism, the Time Multiplexed Trigger (TMT), which enables the treatment of all event data in a single processing board, in contrast with the parallelism that is usually required [203]. It works by converting the parallel processing of spatially decoupled detector regions in a ΔT time period into the sequential processing of the full detector in N ΔT time. N corresponds to the "depth" of the TMT, which is equivalent to the number of available boards with enough memory to process one full event. The benefit of using the TMT is the removal of region separations in the event. It was found that the trade-off between the delay from organizing the data in series and the increased throughput and reduced buffering time brought by the TMT is largely favorable [202]. A demultiplexer board then reorders, re-serializes, and formats the events for the GT processing. The running algorithms are fully pipelined and start processing the data as soon as it is received.

Muon Trigger

The redundancy of the muon system in terms of DTs, CSCs and RPCs is used to robustly define the Global Muon Trigger (GMT). The subdetectors identify track segments from the information produced by the hits in the gas chambers. The segments are collected and transmitted via optical fibers to three muon track finders: the Endcap Muon Track Finder (EMTF), the Barrel Muon Track Finder (BMTF) and the Overlap Muon Track Finder (OMTF). These run pattern recognition algorithms to identify candidates as muons and extract their momenta. The track finders play the role of regional triggers, and send the best muon candidates to the GMT. The GMT receives the information and tries to correlate the muon candidates from the three regions in order to make a decision. It also merges compatible muon candidates found by more than one single system, and can reject candidates based on their identification quality or the lack of confirmation from other muon systems [204].

Global Trigger

The GT collects all L1 muon and calorimeter candidates, and executes multiple selection and identification algorithms, in parallel, for the final trigger decision. Kinematical quantities of the candidates are exploited, such as the invariant mass of e/γ or μ objects, or the angular distance between two objects.

2.3.2 The High Level Trigger

The HLT is provided by a subset of the thousands of commercial Central Processing Units (CPUs) and Graphical Processing Units (GPUs) which constitute the online farm located at the IP5, running the full reconstruction software framework of CMS, CMS Software (CMSSW). The goal of the HLT is to reduce the rate from 100 to ∼1 kHz, while keeping the most interesting events. More than 400 trigger paths are available, targeting a broad range of physics signatures. Given the tight event time budget of ∼300 ms, the reconstruction is performed starting from previously selected L1 seeds, which correspond to events that issued a L1A. The HLT farm processes the data starting with Builder Units (BUs), which gather data from multiple subdetectors, and later with Filter Units (FUs), which decompress, reconstruct and filter the events. The full granularity information is available, together with tracker information, which is absent from the L1. The selected events are acquired by the DAQ and streamed to the Tier-0 at CERN, where they are prepared for offline reconstruction and organized into Primary Datasets (PDs). The latter are defined based on collections of HLT trigger paths, as for instance muon or e/γ triggers. The reconstructed data is eventually sent for permanent tape storage in Tier 0 and Tier 1 sites, managed by the Worldwide LHC Computing Grid (WLCG) [205].

In Run 3, the addition of GPU processing at the HLT improved the performance of some triggers, which lead to an increase of data throughput from 1 to 5 kHz. This was due to several advancements in CMSSW, which can now leverage more parallelism while exploiting a new heterogeneous architecture. CMSSW can now transparently switch between CPU and GPU implementations [206]. The reconstruction of several subdetectors has been offloaded to GPUs, including the ECAL [207], HCAL [208], and some parts of the reconstruction of pixels and vertices, including an improved track seeding. As a consequence, the HLT timing and throughput improved by 40% and 80%, respectively, and the computing power consumption was reduced by 30%. The CMS HLT system is constantly subject to updates and optimizations, in order to improve algorithmic efficiency and increase allowed rates, while maintaining an excellent physics performance, despite the increasingly more challenging running conditions [184, 209].

2.3.3 Parking and Scouting

The quest for ever higher event rates given current technology limitations and experimental benefits has lead CMS to explore a non-standard use of triggers. There are various constraints imposed on the trigger system and data processing framework which limit the number of events that can be selected, recorded and analyzed. Some examples include *i*) the ∼100 kHz L1 acquisition rate, which is limited by the need to avoid losing recording time if the readout system is not ready for a new event, *ii*) the HLT latency, which is constrained by the available number and speed of processing cores, or *iii*) the available permanent storage space, which is distributed across disks and tape, the former providing faster access but reduced storage size. The absolute and relative costs of all hardware trigger components has a strong impact on the overall capacity and structure of the computing farm. In the following, the parking and scouting strategies are presented [210–212]. They are also summarized in Fig. 2.18.

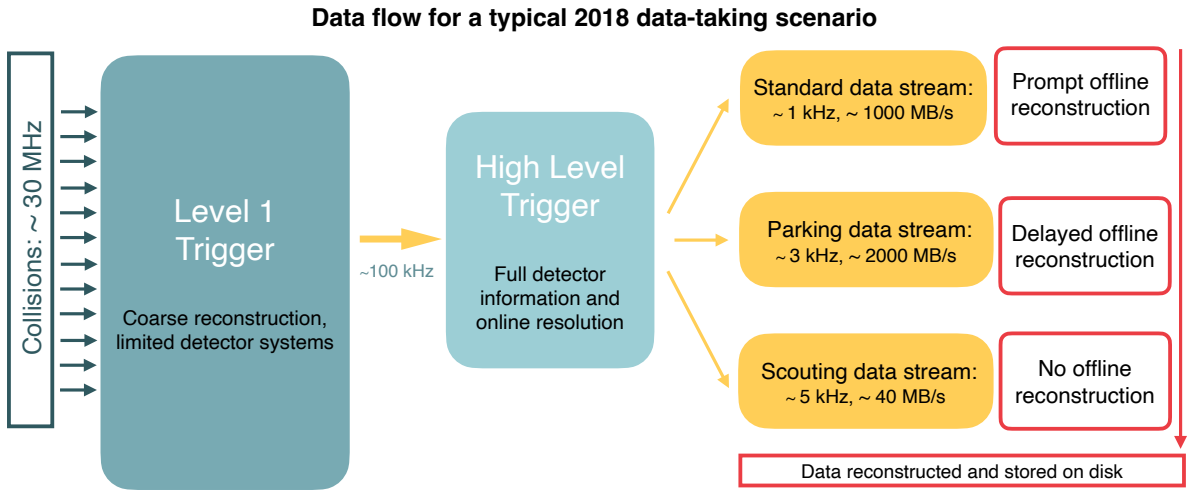

Figure 2.18: A schematic view of the typical Run 2 data flow during 2018, showing the data acquisition strategy with scouting and parking data streams, together with the standard data stream. A value of $\mathcal{L}_{\text{inst}} = 1.2 \times 10^{34}\,\text{cm}^2\,\text{s}^{-1}$ over a typical 2018 fill, corresponding to an average pileup of 38, is considered. The average collision rate lies below the 40 MHz frequency due to occasional but required gaps between consecutive bunch trains. The parking and scouting data streams have been significantly extended during Run 3. Taken from [210].

Parking

The essential idea of data parking is to record as much data as possible, and process it later, as soon as resources used for the ∼48 hours long prompt reconstruction become available, which often happens between data-taking periods. One of the positive effects of this strategy is to lower the kinematical thresholds used by algorithms, and thus increase the acceptance to low-mass physics signals. To sustain the high rates which necessarily arise from lower thresholds, the data is written directly to tape, usually during the latest stages of an LHC fill, when the rate has substantially decreased due to lower instantaneous luminosities.

CMS has exploited parking strategies since its inception to explore NP in VBF, Higgs, B-Physics and SUSY scenarios. During Run 3, an enhanced *b parking* program has been envisaged, with the goal of collecting di-muon and di-electron final states from b hadron decays, in the wake of what had been done in Run 2. The physics goal is strongly centered around searches for LFV. Additionally, three new parking strategies have been introduced:

- **VBF**: The VBF production mode of the Higgs boson drives the sensitivity of some of its decay modes, like $\text{H} \to \mu\mu$ or $\text{H} \to \tau\tau$. It also brings an important constraining power to EFT coupling measurements, and provides unique access to $\kappa_{2\text{V}}$ in double Higgs processes. VBF triggers offer an alternative to raise thresholds in order to keep rates under control: the selection of the two forward jets.

- **LLPs**: Particles with long lifetimes are central to many current BSM scenarios, but imply large displacements relative to the PV, which renders its final states often inadequate to be measured with standard trigger techniques. A group of dedicated HLT paths has thus been introduced, targeting displaced jets and di-jets, taking into account their time delay with respect to prompt decays.

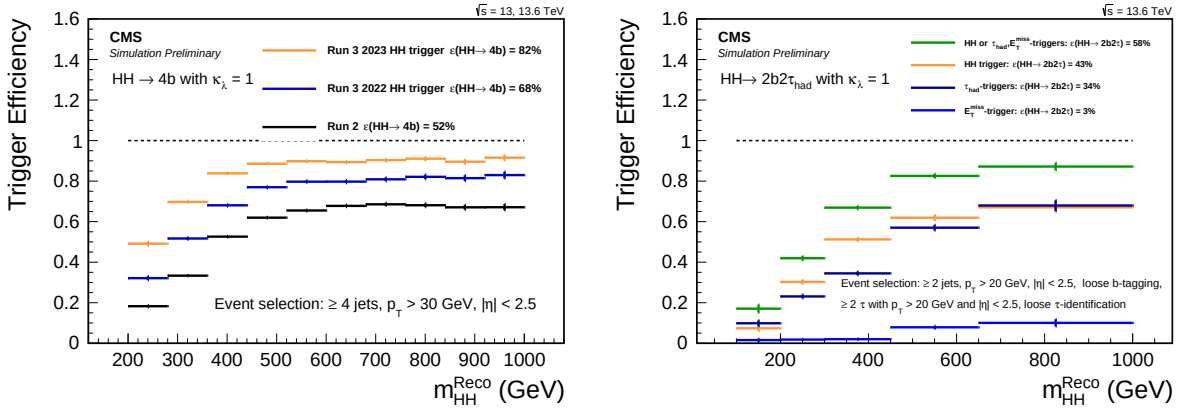

Figure 2.19: Trigger efficiencies as a function of the invariant mass of Higgs boson pairs for simulated samples. $\kappa_\lambda = 1$ is assumed in both plots. Very significant efficiency improvements are observed in the HHH $\to$ bbbbbb and HHH $\to$ bbbb$\tau\tau$ analyses, driven by the inclusion of the HH parking stream. (Left) HH $\to$ bbbb, considering the full Run 2, 2022 and 2023 data-taking periods. (Right) HH $\to$ bb$\tau\tau$ in 2023 only. Taken from [213]; also available in [212].

- **HH**: As discussed in Section 1.2, the study of the Higgs boson self-coupling is one of the most relevant, if not the most important measurement to be pursued at the LHC in the foreseeable future. CMS has successfully deployed the Particle Net (PNet) b-tagging discriminant at HLT level, which made possible to loosen jet p_T thresholds, H_T selections and b-tagging requirements. These updates were combined with a reduction of the H_T requirement at L1. As shown in Fig. 2.19, double and triple Higgs analyses benefit enormously from these developments, with much increased signal efficiencies.

Scouting

The offline reconstruction in CMS notably increases the quality of the collected dataset. Unfortunately but unsurprisingly, it also brings a significant growth in bandwidth and storage needs. It is not possible to perform the offline reconstruction on all events processed and accepted by the HLT. The trigger scouting strategy proposes to save some events using HLT reconstruction only, which provides a smaller event size at the cost of data resolution. This effectively allows to keep events that would otherwise be lost, or to enhance the sensitivity to low-energy processes by lowering HLT thresholds. The strategy depends on the performance of the HLT algorithm which must, as much as possible, approach the quality of the offline reconstruction. In the future, there will always be the possibility to extend the current standard triggers to some specific scouting phase-space regions if some unexpected and promising trend is observed in the data.

The scouting stream was premiered in Run 1 to search for di-jet resonances with jets reconstructed only from calorimetric energy deposits. This was considerably extended in Run 2, with the addition of jet, muon and electron candidates to the scouting event record, closing the gap with respect to the standard CMS data streams. As an example, studies of BSM low-mass di-muon resonances were able to reach a threshold close to twice the muon mass [214]. The scouting stream in the on-going Run 3 has strongly benefited from the inclusion of GPU processors and related software infrastructure at HLT level [215]. The scouting bandwidth is currently $\sim$30 kHz, ten times higher than the standard

data stream. This implies the exploration of lower kinematic thresholds, with the increase in physics sensitivity as a consequence. There are plans to extend the scouting strategy to L1, as will be discussed in Section 3.2.2.

2.4 Offline Reconstruction

Reconstruction is the operation of building physical quantities from the raw data collected in the experiment. Similarly to other HEP detectors, CMS consists of consecutive barrel and endcap detector layers surrounding the pp IP. As depicted in Fig. 2.20, the design includes a tracker, calorimeters, the solenoid and muon chambers, with different detector elements establishing a synergy to measure distinct particle types. Charged particles, such as electrons[8], muons or charged hadrons, leave energy deposits, or *hits*, in the tracker. The relative bending of their trajectories in the solenoid's magnetic field conveys information on their charge and momentum. Most particles, either charged or neutral, deposit all their remaining energy in the calorimeters. Calorimeter deposits separated from the extrapolated position of tracker hits constitute a clear signature of neutral particles, such as photons. Instead, particles with energy deposits matching the hits' trajectories hint at the passage of charged particles, such as electrons. Besides neutrinos, which in all practical terms do not interact with the detector material, muons are the only particle type which manages to traverse the solenoid and leave hits along the muon chambers. Combining all information above, it should be possible to distinguish many particle types.

Historically, the first steps in reconstruction at CMS foresaw a somewhat rigid structure, where the focus went to localized information in each subdetector [216]. One would, for instance, reconstruct jets using solely their calorimetric energy deposits, while the identification of individual particles within jets was not pursued. Ideas on how to perform reconstruction, based on the layered-structure present in all HEP experiments, targeted single physics objects, identified by specific subdetectors:

- muons were identified mostly based on information from the muon chambers;
- electrons and isolated photons were measured by the ECAL;
- jets, MET and their properties were all measured in the ECAL and HCAL;
- τ and b jet tagging exploited information from the tracker.

It is now clear that a reconstruction relying only on a few subdetectors per particle necessarily provides a lower performance than a holistic approach. As an example, the energy resolution of the CMS HCAL is $\sim 100\%/\sqrt{E}$, which leads to a poor offline resolution if unaccompanied by information from other subdetectors.

2.4.1 Particle-flow

Inspired by what had been done at LEP, specifically ALEPH [217], the CMS Collaboration decided to adopt a novel reconstruction paradigm, named Particle Flow (PF), for the first time in hadron colliders. The PF approach is characterized by an enhanced energy,

[8] In this Section, whenever we mention electrons we also mean their anti-particles, positrons, for brevity. The same logic is used for other particle types, when applicable.

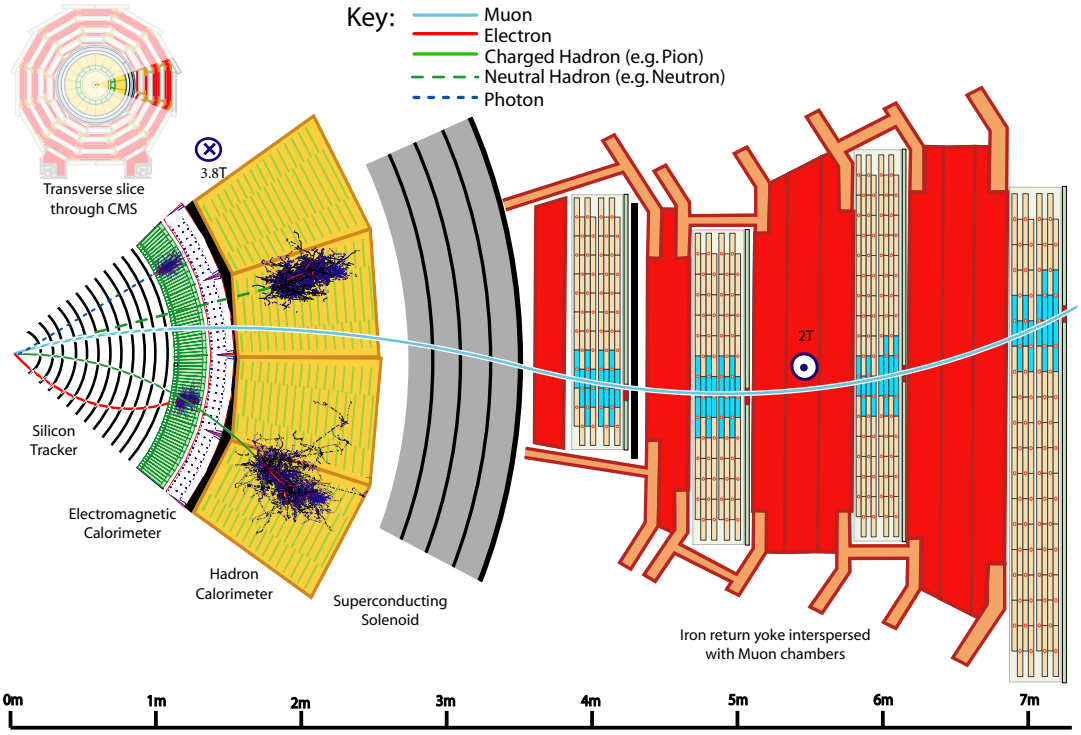

Figure 2.20: Schematic of a transverse slice of the CMS detector, from the pp IP, on the left, to the muon detectors, on the right. The muon and the charged pion are positively charged, and the electron is negatively charged. Particles interact in different subdetectors according to their type, producing different signatures. The detector's structure is described in Section 2.2. Taken from [191].

position and time resolution from the optimized combination of all CMS subdetectors [191, 218, 219]. Despite not being designed with PF in mind, CMS seriously benefited from it, with very substantial improvements in calorimetric resolution and response, in PU mitigation, and in particle type identification. Among the features that enable a successful PF reconstruction at CMS, we find:

- a highly-segmented tracker, producing trajectories of charged particles up to $p_T \sim 1\,\text{TeV}$, and with an efficiency of $\sim$90 %, down to $\sim$500 MeV;
- a granular EM calorimeter, providing a clear separation between neighboring particles, the capability of matching calorimetric deposits with tracks, and the discrimination between different particles in jets, namely charged hadrons, neutral hadrons and photons, up to $p_T \sim 1\,\text{TeV}$;
- a strong magnetic field, disentangling contributions from charged and neutral particles; the 4.9 T m bending power can be favorably compared to ATLAS, ALEPH, or to Tevatron experiments, all with less than 3 T m.
- very precise muon detectors.

Some limitations are nevertheless present, and historically caused concern on the successful implementation of PF, especially given the complex pp and heavy ion environments. The existence of a significant amount of tracker material in front of the calorimeters represents a limitation, inasmuch as it increases the probability for a charged particle to preshower, making the matching between tracks and energy deposits in the calorimeter harder to accomplish. This is the case in spite of the placement of the calorimeters inside the solenoid, which is meant to improve the matching. Additionally, the HCAL is known to have a relatively poor energy and position resolutions. However, the holistic PF approach was ultimately able to surmount all challenges, proving itself robust, and sitting at the basis of all CMS reconstruction techniques.

The PF reconstruction follows the dataflow schematically shown in Fig. 2.21. Its building blocks are calorimeter hits, and trajectories in the tracker and muon chambers, also known as *tracks*. A series of CMSSW "producers" process the data, sequentially converting the inputs into higher-complexity objects, starting from PF *tracks* and PF *clusters*. Eventually, the information between all subdetectors is linked, producing PF *blocks* that are used to create the final PF *candidates*. The latter correspond to the building blocks of all physics objects used in CMS analyses.

PF Tracks

Reconstructing tracks consists in recreating their trajectories from individual hits, and extracting physical properties from those trajectories. At CMS, an iterative tracking algorithm was devised, where a combinatorial track finder procedure based on Kalman Filters (KFs) [220] is sequentially applied $\sim$10 times. The algorithm is particularly suited to provide a high tracking efficiency without affecting the misreconstructed track rate, which is generally challenging. Indeed, the increase in misreconstructed tracks reaches $\sim$80 % when the track p_T threshold is reduced from 900 to 300 MeV and the total number of hits to build a track is decreased from 8 to 5. The resolution degradation comes from the additional hits, which can be randomly assigned in such a way as to produce a high momentum track.

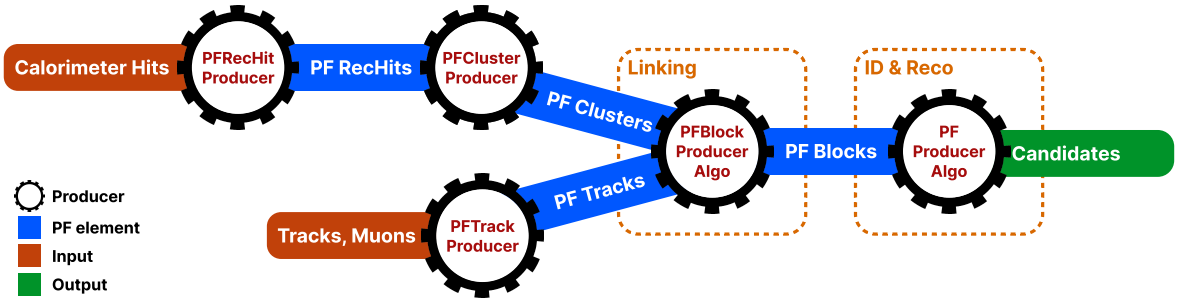

Figure 2.21: Illustration of the processing steps of the PF reconstruction. Its building blocks are energy deposits in the calorimeter and tracks. Calorimetric and track information is only merged at a later stage into blocks, from which candidates are created. The term "producer" refers to a CMSSW processing code element which produces output collections from a set of input collections. Refer to the text for details.

At every iteration of the KF-based algorithm, a set of hits is selected based on quality criteria imposed on tracks and hits. Selected hits are masked and not considered in subsequent iterations. The progressive decrease of hit density in the event enables to loosen the applied quality criteria to lower and lower values, thus improving the overall efficiency, without affecting track purity[9]. Tracks with p_{T} as small as 200 MeV can be measured. Given the progressively lower energy of hits after each iteration, the purity is kept stable by running more and more refined algorithms. The iterations separately target tracks with low hit multiplicity, displaced tracks, nuclear interactions and their secondaries, or the core of high-p_{T} jets. The last iterations specifically focus on the reconstruction of muon tracks, by exploiting hits in the muon detectors.

PF Clusters

The PF clustering algorithm runs separately in most subdetectors of the calorimeter: barrel and endcaps for the ECAL and HCAL, and the two preshower layers. The task is particularly challenging given constant overlaps between photons, neutral and charged hadrons, and electrons with their bremsstrahlung energy deposits. Clustering also plays an essential role in cases where the tracker underperforms, which happens for low quality and high p_{T} tracks. The algorithm starts by defining *seeds*, which correspond to detector elements with an energy larger than its neighbors and larger than a predefined threshold. Topological clusters are then built, centered on the seeds, based on the physical connection of neighboring cells with energies larger than a given S/N[10] threshold. An iterative algorithm based on a Gaussian mixing model is used to reconstruct clusters within the topological clusters. The algorithm postulates a fixed number of seeds, and associates a Gaussian function to each, allowing some room for energy sharing across clusters. The positions and energies of the clusters are obtained via a maximum likelihood fit.

Once computed, clusters need to be calibrated to obtain the correct energy scale. A precise calibration of the calorimetric response facilitates PID by removing calorimetric overlaps, such as the ones between photons and hadrons. A first and generic calibration step exploits cosmic rays, radioactive decays and testbeam data to improve the energy

[9] By *efficiency* we mean the fraction of selected items from all available items, and by *purity* we instead mean the fraction of "true" items, as defined by some ground truth, with respect to all selected items.
[10] Signal over noise.

scale in the calorimeters. The raw energy measured by the detector is expected to be lower than the real energy due to inefficiencies and acceptance "holes", and given the numerous threshold cuts applied in the clustering steps. A more detailed calibration is first performed at ECAL, also taking into account the two preshower layers. The calibration estimates the ECAL response to electrons and hadrons with a χ^2 minimization fit. For low energies, the corrections can be as large as $\sim$20 %. A second, similar calibration step is run at HCAL, on top of the ECAL calibration. Separate calibrations are needed given the different responses of calorimeters to hadrons, as explained in the next paragraph. Both calibrations are done as a function of cluster energy and η. The HAD calibration is also done separately for HAD showers leaving energy in both calorimeters or in the HAD section alone.

HAD showers are in general much more complex than EM showers, due to the involvement of the strong interaction. Neutral hadrons, for instance neutrons, can only lose energy via nuclear reactions, while charged hadrons can also ionize the atoms of the medium being traversed. An immediate consequence of the introduction of the strong force is the existence of a fraction of dissipated energy which is fundamentally undetectable, the so-called *invisible energy*. Indeed, high energy hadrons can break apart the calorimeter's nuclei via nuclear interactions, and the binding energy connecting those nuclei is lost for calorimetric purposes. The effect is not negligible: 30 to 40 % of the non-EM energy of hadronic showers is lost via these undetected processes. Additionally, hadronic showers include both EM and non-EM components. It is known that the EM component represents around one third of the shower energy, at low energies. However, the EM fraction increases significantly with energy, mainly due to additional π^0 mesons decaying to $\gamma\gamma$. The observed dependence, coupled with the invisible energy phenomenon, implies that all homogeneous and most sampling calorimeters are non-linear in what concerns their energy response. Whenever the response to the EM component is different than the response to the non-EM component, the calorimeter is said to be *non-compensating*. The CMS calorimeter, in particular, *undercompensates*, since its non-EM response is always lower than the EM one. Would the invisible energy fraction be the same for every event, and it would be possible to fix the resolution degradation. Unfortunately, HAD showers are prone to very large event-to-event fluctuations, caused by the large variety of strong interactions that can occur during the shower's development. Because of these fluctuations, the energy resolution of hadron calorimeters is usually significantly worse than the EM energy resolution [188]. All this justifies the existence of separate calibration steps. Additionally, barrel and endcap regions are calibrated separately to deal with different cell sizes and thresholds.

PF Linking

When a particle crosses CMS, it usually produces many PF elements in various subdetectors, namely PF tracks and PF clusters. As an example, at least one PF track in the inner tracker and one PF cluster in the ECAL are expected to be formed by an electron. A linking algorithm proceeds to connect PF elements coming from different subdetectors into *PF blocks*, using only its (η, ϕ) nearest neighbors to reduce time complexity. Once a link is found, depending on selection criteria associated to the particles being linked, a distance or quality metric is associated to it. Links are established in a detector- and particle-dependent way. In total, five link types exist, connecting the individual PF tracks and PF clusters, based on proximity conditions:

- link tracks to clusters;
- link all bremsstrahlung photons emitted by the same electron;
- link clusters to other clusters, specifically ECAL to preshower clusters, HCAL to ECAL clusters, and ECAL clusters into superclusters;
- link tracks to other tracks sharing a common secondary vertex;
- link tracks to muon tracks, forming global muons and tracker muons.

Identification and reconstruction

Once PF blocks are formed, the identification and reconstruction sequence follows a fixed order:

1. muon candidates are identified and reconstructed, and their PF tracks and PF clusters are removed from the block;
2. electrons, including the collection of all bremsstrahlung photons, plus energetic and isolated photons, are also identified, and all corresponding PF elements are masked;
3. remaining tracks with large uncertainties are masked, decreasing the track misreconstruction rate, but increasing the inefficiency for some high-p_{T} charged hadrons, which are anyways more precisely measured in the calorimeters;
4. the PF elements still left in the block are reconstructed as photons and as charged and neutral hadrons, including hadrons interacting strongly in the tracker.

When the above has been run for all PF blocks, a final post-processing, or cleaning step corrects residual identification and reconstruction inefficiencies. The particles produced by PF can be directly used in physics analyses. They are assembled into offline *physics objects*, which we describe in the following sections.

2.4.2 Muons

A distinctive feature of CMS is the presence of muon chambers for additional tracking, enabling a clear separation between muons and other charged particles. This happens in light of the low probability for a particle, other than a muon, to reach the muon detectors without being absorbed in the calorimeters. The interplay between tracker and muon chambers leads to three different muon signatures:

- **Standalone Muons:** DT and CSC hits are clustered into tracks, which serve as seeds for pattern recognition algorithms that also exploit the RPCs; GEMs are not used, since the benefit outside the tracker acceptance is minor, except for calibrations.
- **Global Muons:** If geometrically compatible, standalone muons are matched to tracks in the inner tracker, increasing the momentum resolution for tracks with $p_{\mathrm{T}} \gtrsim 200\,\mathrm{GeV}$.
- **Tracker Muons:** Tracks satisfying $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and $p > 2.5\,\mathrm{GeV}$ in the inner tracker, where a geometrical match exists with at least one muon segment in the muon chambers.

The tracker muon reconstruction is more efficient than the global one when muon segments are present in a single muon detector plane. This happens more often for muons with $p_T \lesssim 10\,\text{GeV}$, due to scattering on the steel return yoke. Only $\sim$1 % of muons within the acceptance of the muon detectors is reconstructed as a standalone muon, and they consistently have the worse resolution. This once again highlights the advantages brought forward by the PF approach. Occasionally, and despite the CMS calorimetric density, some energetic charged hadrons reach the muon systems and are reconstructed as muons. A balance must thus be established between muon identification efficiency and purity. The obtained muon p_T resolution ranges between 1% and 6% for $p_T < 100\,\text{GeV}$ muons and is of around 10% for central muons of $p_T \sim 1\,\text{TeV}$.

The identification of muons is based on a set of selections driven by the properties of global and tracker muons, such as the track fit χ^2, the number of hits per track in the inner tracker and/or in the muon system, the degree of compatibility between tracker tracks and muon tracks, or the compatibility with the PV [221]. Firstly, isolated global muons are identified using criteria based on the inner tracks and calorimetric deposits within a certain distance from the muon direction. Concerning muons inside jets, tighter criteria are required, since PF tends to create spurious neutral particles whenever charged hadrons are identified as muons. On the reverse side, when muons are identified as charged hadrons, neutral particles will have their energy "eaten out" by the algorithm. Finally, non-isolated global muons are selected using criteria which aim at suppressing muons from in-flight decays and hadronic longitudinal leaks. The full selection is rather complex and detailed, and is fully explained in Ref. [221]. At the level of the analyses, the criteria above are encoded in so-called working points (WPs), to be discussed in Section 4.4. We note that the muon reconstruction and identification steps can be revisited if significant mismatches are found for charged hadrons between their measured tracker momenta and calorimetric energies.

2.4.3 Electrons

Given the significant material budget in the tracker, most electrons lose a sizable fraction of their energy via bremsstrahlung emissions. The latter are extremely dependent on η, due to the increasing amount of material budget a particle traverses for higher angles. A series of calorimeter energy clusters is thus created in the ECAL, originated by all emitted photon clusters, plus the one from the electron. All clusters put together form an ECAL *supercluster*, via merging windows which are narrow along η but extended in ϕ, in order to consider the electron bending under the magnetic field. The success of the PF reconstruction resides on how complete the measurement of the full electron shower energy is, while avoiding the inclusion of unrelated energy deposits coming from other showers or PU. However, position and energy resolutions are hindered by isolation thresholds, required mostly due to overlaps of superclusters with energy deposits from hadronic activity. The energy radiated by low p_T electrons is also hard to supercluster, given the position spread of the produced bremsstrahlung clusters. Additionally, track combinatorics complicate matters when trying to unambiguously assign superclusters to specific tracker hits. It is for all the above reasons that PF electrons take an enormous advantage from the inclusion of tracker information in the reconstruction algorithms, especially at low p_T.

A tracker-based electron seeding method was developed, starting from the iterative tracking algorithm already described. The method uses Gaussian-sum filters (GSFs)

[222] rather than a KF, since the former provides better trajectory fits when the particle radiates, while the latter cannot describe the sudden and significant energy losses in the electron's trajectory. The GSF track fitting algorithm is CPU intensive, thus requiring a seeding technique to avoid running on all tracker hits. The seeding highlights the hits more likely to be associated to the particle's trajectory. For the tracks to form an electron seed, matching criteria are imposed between the track and ECAL clusters. In PF blocks, GSF tracks seed electron candidates, as long as the respective ECAL cluster is not linked to three or more extra tracks. In that case, the clusters are instead used as seeds. Electron candidates are further required to pass HCAL energy limits and momentum compatibility checks, together with Boosted Decision Tree (BDT) cuts containing information on HCAL and ECAL energy ratios, GSF and KF fit χ^2, hit multiplicity, and more.

Given their similarities, isolated photon and electron reconstruction steps are done together [223]. ECAL superclusters with missing transverse energy above 10 GeV can be used to seed photon candidates, but only when no GSF track is linked. Photon candidates must also satisfy a series of quality criteria, similar to the ones mentioned for electron candidates. The energy of superclusters must in general be corrected, leading to up to 25% more energy, which is assigned to photons. The full history of electron and photon reconstruction is propagated to the analyses, given that particle interpretations can differ between the general PF approach and more dedicated studies. Since photons are not exploited in the $X \to HH \to bb\tau\tau$ analysis, we do not cover photon identification and reconstruction in detail.

Exploiting the available ECAL granularity, electrons with $p_T \sim 45$ GeV reach an energy resolution between 2 and 5%, depending on electron η and energy loss in the detector. The energy scale uncertainty is smaller than 0.1% and 0.3% in the barrel and endcaps, respectively [223]. The benefits arising from the all-encompassing PF approach can be appreciated in Fig. 2.22 (left), where very significant efficiency increases come from the tracker-based electron seeding, both for electrons and pions within b-jets. The improvement in the association of converted bremsstrahlung photons to their parent electron also minimizes double counting in later PF steps.

2.4.4 Hadrons

The particles left to be identified after muons, electrons and isolated photons have been removed are the neutral and charged hadrons, non-isolated photons from hadronic decays (most notably π^0s), and occasional muons from early decays of charged hadrons [191]. Photons and neutral hadrons are built from calorimetric clusters not linked to any track. Precedence is given to photons, given their significant energy fraction in hadronic jets and hadronically decaying τs. Outside the tracker acceptance ($|\eta| > 2.5$), neutral and charged hadrons are undistinguishable, and the above precedence is not longer justifiable. In those regions, the presence of HCAL clusters serves as a discriminative factor between photons and all kinds of hadrons, both being associated to ECAL deposits. Remaining HCAL clusters are linked to remaining tracks, which are in turn linked to remaining ECAL clusters, forming single charged hadrons. After the calibration, whenever the track momenta does not match the calorimetric energies, the difference is interpreted as the presence of an additional photon and eventually a neutral hadron, based on energy and resolution considerations. If no mismatch is found, no additional neutral particle is identified, and the estimate for the charged hadron momenta is improved by performing a χ^2 fit to the associated tracks and clusters. The fit is particularly helpful for situations with

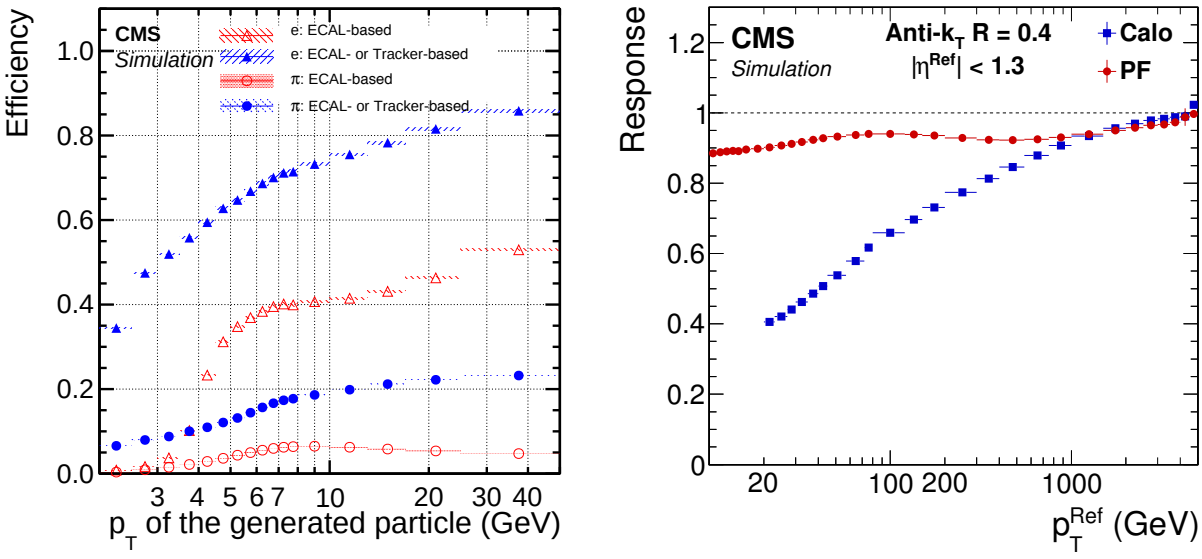

Figure 2.22: PF performance. Taken from [191]. (Left) Electron seeding efficiency for electrons (triangles) and pions (circles) as a function of p_{T}, from a simulated event sample enriched in b quark jets with $80 < p_{\mathrm{T}} < 170\,\mathrm{GeV}$, and with at least one semi-leptonic b hadron decay. One can compare the efficiencies between the ECAL based seeding with (solid symbols) and without (hollow symbols) the tracker-based seeding. (Right) Jet energy response of Calo and PF jets, as a function of the momentum of the reference jet, $p_{\mathrm{T}}^{\mathrm{Ref}}$, using the anti-$k_{\mathrm{T}}$ algorithm with $\mathrm{R} = 0.4$ (see Eq. (2.10)). The reference jet is defined as the result of the jet algorithm applied to all stable particles produced by the event generator, excluding neutrinos.

low resolution tracks, usually at high energies or large η values, and ensures a gradual shift between tracker-dominated low energy and cluster-dominated high energy measurements. Additional muons are searched for if the calibrated cluster energy is much smaller than the sum of the track momenta, but this rarely happens. The description in this Section serves as the basis for the complex algorithms that reconstruct hadronic tau decays, and which are briefly described in Section 4.3, proving essential for the $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis.

2.4.5 Tau Leptons

The tau lepton, despite having similar quantum numbers to the two other leptons, behaves in a fundamentally different manner. For one, and contrary to the electron, which is stable, and to the muon, which can travel large distances before decaying, the very short $(290.3 \pm 0.5) \times 10^{-15}\,\mathrm{s}$ [13] lifetime of the tau lepton implies that it cannot currently be observed directly, as a τ lepton will travel on average just a few millimeters before decaying. Secondly, its mass makes it the only hadronically-decaying lepton. In fact, it does so approximately 2/3 of the time, while the remaining decays proceed leptonically, to electrons or muons. The decay modes of the τ are listed in Table 2.2, and can be categorized in terms of the number of charged particles, or *prongs*, the final state includes. Due to charge conservation, τs can only decay to an odd number of prongs, and usually decays with five or more prongs are neglected due to their extremely low BRs. The decay itself can happen via a meson resonance, or directly to the hadrons, which consist of pions on ∼98 % of occurrences, with the rest being kaons [13].

Decay mode	Meson resonance	$\mathcal{B}$ [%]
$e\nu_e\nu_\tau$		17.8
$\mu\nu_\mu\nu_\tau$		17.4
all leptonic decays		35.2
$h^\pm\nu_\tau$		11.5
$h^\pm\pi^0\nu_\tau$	$\rho(770)$	26.0
$h^\pm\pi^0\pi^0\nu_\tau$	$a_1(1260)$	9.5
$h^\pm h^\mp h^\pm\nu_\tau$	$a_1(1260)$	9.8
$h^\pm h^\mp h^\pm\pi^0\nu_\tau$		4.8
other hadronic decays		3.2
all hadronic decays		64.8

Table 2.2: Branching fractions of the τ lepton, where $h^{\pm/\mp}$ symbolizes any charged hadron [13].

In general, the hadronic decays can be differentiated from quark or gluon jets by the isolation of the decay products, their collimation and their multiplicity [191]. Tau leptons can also be faked as electrons and muons by decays with one charged track, and by decays with one prong plus a few photons, which can be interpreted as an electron plus its bremsstrahlung radiation. The individual particles reconstructed by PF are fed into the CMS-specific Hadron Plus Strips (HPS) algorithm [224–226], which is responsible for hadronic τ reconstruction. The task is rather complex, since a plethora of decay particles exists. Neutral hadrons, on one hand, decay to photons ($\pi^0 \to \gamma\gamma$) which quickly convert to electron-positron pairs in the material of the tracker. Charged hadrons, instead, amount to kaons and pions, and can be produced with different multiplicities. The algorithm starts by reconstructing neutral hadrons by collecting energy deposits in *strips*, i.e. dynamically-sized regions along (η, ϕ) which are created by the magnetic bending of electrons and positrons. The momenta of the strips amounts to the vectorial sum of all its components. Next, charged hadrons satisfying $p_\mathrm{T} > 0.5\,\mathrm{GeV}$ and coming from the PV are reconstructed. They can fit into multiple topologies, as shown in Table 2.2, and when appropriate, the combination of charged hadrons with strips is required to be compatible with the masses of the ρ or a_1 resonances. Candidates with particles outside the so-called *signal cone* are also rejected, where the cone is defined as $3\,\mathrm{GeV}/p_\mathrm{T}(\tau_\mathrm{h})$, capturing the dependence with the system's boost. The cone size is bounded between 0.1 at low p_T and 0.05 at high p_T. An *isolation cone* with a process-dependent radius of 0.3 or 0.5 is also defined, in order to reduce the misidentification probability of τ_hs as jets. A series of BDT classifiers is also exploited for further discrimination. Finally, in case multiple τ_h candidates satisfy the requirements, only the one with the highest p_T is kept, such that only one candidate is defined per jet.

The performances of HPS during Run 2 are very dependent on the process and on the kinematics, and are presented in detail in Ref. [224]. One is generally concerned with the misidentification probabilities of jets, electrons and muons, and with the identification efficiencies of hadronically-decaying τs. In a nutshell, the misidentification probabilities for leptons are extremely small, often at sub-percent level given appropriate WP choices. Jet misidentification tends to be a bit higher, but still usually below 1%, and never above 2%. Concerning efficiencies, they are WP-dependent too, lying always above 90% for

Drell-Yan (DY) events and above 86% for $t\bar{t}$ events. Finally, the HPS algorithm can recover events with two prongs, where one extra track was "lost". This happens for 19% of the 3-prong decays and for 13% of the 3-prong decays with an additional π^0. We finalize by noting that 2-prong recovered events are not considered in the analysis, since for most analyses the background increase brought by the extra events outweighs the increase in efficiency. We thus follow the general CMS recommendations, and consider only 1-prong and 3-prong events, with or without an additional neutral pion.

2.4.6 Jets

Since quarks and gluons carry color charge, they must obey QCD *confinement*, according to which only colorless states are allowed. As a consequence, once a hard collision takes place, ejected quarks and gluons are immediately associated to other colored objects, forming colorless hadrons in a process called *hadronisation*. The group of all newly formed colorless hadrons is called a *jet*, since all its constituents travel in approximately the same direction, outlining a narrow cone shape. In CMS, jets are reconstructed from charged and neutral PF candidates, 85% of which are photons and charged hadrons, via the anti-k_{T} algorithm [227]. This algorithm, designed to be insensitive to the effect of soft radiation, defines the following two distance parameters, themselves inspired by previous clustering algorithms [228–230]:

$$\begin{cases} d_{ij} = \min(1/k_{\mathrm{T},i}^2, 1/k_{\mathrm{T},j}^2)\,\Delta_{ij}^2/\mathrm{R}^2 \\ d_i^{\mathrm{B}} = 1/k_{\mathrm{T},i}^2 \end{cases}, \tag{2.10}$$

where d_{ij} is the distance between two PF candidates or pseudo-jets i and j, d_i^{B} is the distance between i and the beam B, and $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, with y being the rapidity, ϕ the azimuthal angle and k the transverse momentum. $\Delta_{ij}^2/\mathrm{R}^2$ encodes the jet angular extension, where R determines the radius of the clustering. The fact that the distances are inversely proportional to the transverse momentum of the particles or pseudo-jets ensures that the soft particles will first cluster around the hard particles in the event. This is meant to avoid the modification of the jet shape by soft radiation. In the limit where there is only a single hard particle, the jet will be perfectly conical, accumulating all soft particles within R. If two hard particles exist, the algorithm will instead produce two clipped cones, with their overlap divided by a straight line. The algorithm proceeds iteratively as follows:

- identify the smallest of the two distances between all available particles or pseudo-jets i and j: *a*) if $d_{ij} < d_i^{\mathrm{B}}$, then i and j are combined to form a new pseudo-jet; *b*) if $d_{ij} > d_i^{\mathrm{B}}$, the iteration ends and i is defined as a new jet and removed from the list of particles and pseudo-jets;
- the procedure is repeated until no particles or pseudo-jets are left and a list of jets has been produced.

In CMS, PF objects are feeded to the `FASTJET` package [231], which runs the anti-k_{T} with R = 0.4, 0.8 or 1.5 depending on the intended jet cone size. In the resonant HH analysis detailed in this Thesis, the first two values are employed, for resolved and boosted topologies, respectively. The jet four-momentum is computed as the vector sum of all the four-momenta of clustered PF candidates. As shown in Fig. 2.22 (right),

the anti-k_T together with the PF approach allows the reconstruction of ∼90 % of the jet constituents, representing a dramatic performance improvement especially at lower energies. Jet resolutions in CMS are quite energy-dependent, at around 15-20% for $p_T \sim 30\,\text{GeV}$, 10% for $p_T \sim 100\,\text{GeV}$, and 5% for $p_T \sim 1\,\text{TeV}$ [232]. Additional jet identification criteria can be introduced to target specific jet signatures, such as c and b jets, or hadronically decaying τs. A detailed description of such criteria is discussed in Section 4.3, in the context of the $X \to HH \to bb\tau\tau$ analysis. Required corrections such as jet energy scales and resolutions are also covered there.

2.4.7 Missing Transverse Energy

Due to the extremely low cross-sections involved, neutrinos are not detected by the CMS experiment or any of other four large LHC experiments. In a collision event, the presence of a transverse momentum imbalance, which could be naively seen as the violation of momentum conservation, can therefore be attributed to neutrinos, neglecting other effects arising to due imperfect efficiencies and acceptances for other objects. Another source for the imbalance might instead come from hypothetical BSM particles interacting very weakly with the detector. These effects lead to *missing transverse momentum*, or MET, which is defined as the negative vector sum of all N reconstructed PF particles in an event:

$$\vec{p}_T^{\,\text{miss}} = -\sum_i^N \vec{p}_{T,i}\,, \tag{2.11}$$

which are the visible particles in the final state of the collisions. We also mention a closely-related quantity called MHT, but applied only to the event's PF jets only:

$$\overrightarrow{\text{MHT}} = -\sum_i^{N_{\text{jets}}^{\text{PF}}} \vec{E}_{T,i}\,. \tag{2.12}$$

which represents the vectorial sum of the jet transverse energies.

Despite the optimized combination of PF tracks and clusters previously described, the probability of particle misreconstruction and misidentification is not zero, albeit small. In some rare cases, an artificially large $\vec{p}_T^{\,\text{miss}}$ is reconstructed in the event, most often caused by a misidentified or misreconstructed high-p_T muon. These usually correspond to genuine cosmic muons that traverse CMS in coincidence with the LHC BX, and are identified by comparing their trajectories with the beam axis. Other reasons include a completely wrong muon momentum estimate due to a wrong inner track association, a decay in flight, significant synchrotron radiation, or interactions in the steel yoke. For all the above, a dedicated event post-processing is employed in the PF approach, as mentioned at the end of Section 2.4.1.

We finalize this Section by noting that the measured MET is systematically different from the true MET. Applied corrections are described in the Chapters where the $X \to HH \to bb\tau\tau$ is discussed, namely in Section 4.3.

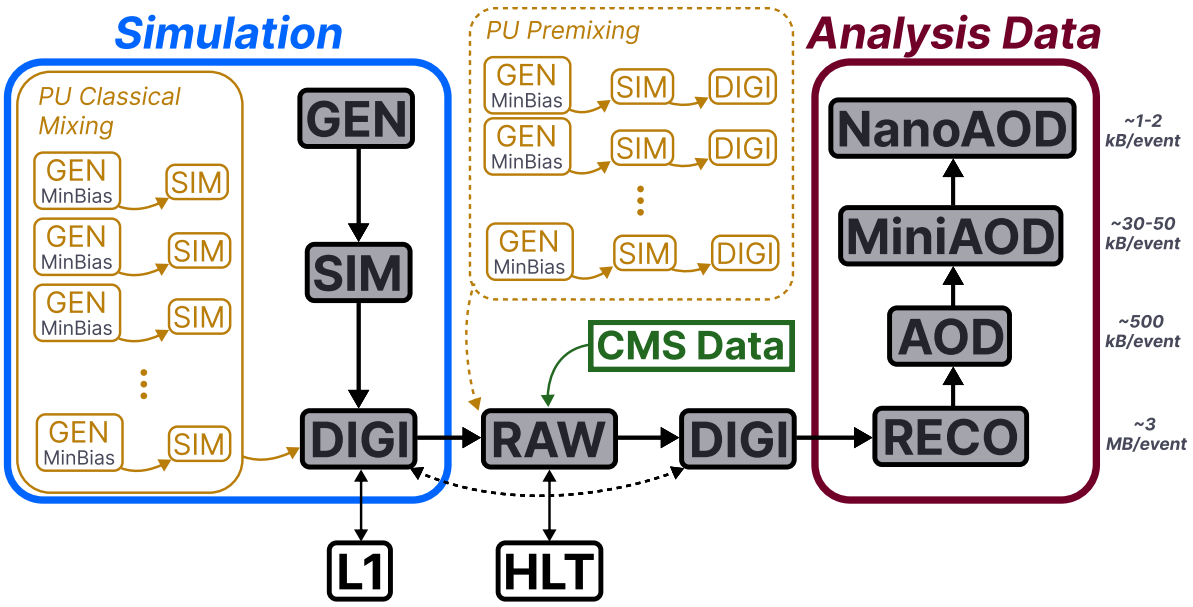

Figure 2.23: Illustration of the generation workflow of MC and data samples as defined within the CMS Collaboration. After the simulated RAW format is produced and the hits and clusters are measured, the chain is identical for data and simulated samples. The PU premixing technique is drawn using a dashed square to remember the reader that the classical and premixing techniques are not used simultaneously. The dashed double-headed arrow between the two DIGI formats represents the communication required to validate the MC chain, which in addition emulates the L1 and HLT triggers. The vast majority of CMS analyses uses either the MiniAOD or NanoAOD formats. The latter is becoming more dominant, being more lightweight and faster to process.

2.5 MC Generation and Data Processing

The usage of simulated Monte Carlo (MC) samples is indispensable in HEP, from modeling signal processes to assessing the impact of new analysis methods and studying the behavior of particles in a detector. The development of future experiments would be impossible without dedicated simulations, which estimate fundamental quantities such as responses and resolutions before the hardware is put in place, and provide training data and optimization potential to a series of algorithms. MC samples are also essential to compare measured data to what is expected, giving handles for all sorts of calibration procedures and correction techniques. It goes without saying that MC samples are a fundamental ingredient in this Thesis, for the work related to the CMS detector in Chapter 3 as well as for the $X \to HH \to bb\tau\tau$ analysis in Chapters 4 and 5.

The so-called full simulation, or `FullSim`, is used throughout this Thesis for creating samples with a single particle in dedicated detector studies, or to perform instead the complete simulation of SM and BSM $HH \to bb\tau\tau$ events. The `FullSim` corresponds to a detailed simulation of the entire experimental setup, including the propagation of particles through all detector components, the interaction of particles with the detector material, and the response of the various detectors components. The simulation of MC events follows a well-defined processing chain in CMS, being structured such that simulation and measured experimental data are processed in similar ways. Indeed, after the simulation phase, the two chains are identical, and produce identical data formats to be used by analysts. Output objects are eventually stored in different files, where files with the same

event content share a common data tier, such as RAW, RECO or AOD. The general procedure is illustrated in Fig. 2.23, and described in what follows.

2.5.1 GEN

The first step involves the *generation* (GEN) of the physics processes of interest from theoretical principles [233]. Generated processes, handled by an event generator, exhibit an enormous variety. CMS relies on a rich set of event generators such as `Herwig` [234–236], `Sherpa` [237], `Pythia` [238, 239], `PowHeg` [240] or `MadGraph` [241]. The analysis discussed in this Thesis most often uses the last three. Event generators try to accurately describe the intrinsically random nature of HEP collisions or particle decays, using a set of probability distributions to model the subatomic particles produced, along with many of their properties, storing intermediate stages of the event modeling along the way [239]. Generators aim at describing what an event looks like before interacting with detector elements. A generic pp collision is usually generated with separate components for practical purposes, namely the hard scattering of two partons, radiative corrections, parton showering (initial and final state radiation), multiple-parton interactions (MPIs), Beam-Beam Remnants (BBRs) and hadronisation. All of the above components, with the exception of the hard collision, are collectively referred as the Underlying Event (UE) which, contrary to PU, is characterized by having the same vertex as the hard scattering event. The BBRs are what remains after a parton is scattered out of each of the two initial beam hadrons, while the MPIs are additional soft or semi-hard parton-parton scatterings which occur within the same hadron-hadron collision [242]. The UE is the unavoidable accompanying activity to hard pp scattering processes, representing a background for most measurements and searches. To further complicate matters, the UE activity is not constant on an event-by-event basis, and so its contribution cannot be subtracted. The MC modeling can however be tuned by leveraging measurements sensitive to UE activity. In fact, the full properties of an event cannot be calculated from first principles only. A set of QCD parameters is thus derived to precisely describe the UE. These parameters are known as *tunes*, resulting from a complex fitting procedure which, at least in the case of `Pythia`, uses data from the Collider Detector at Femilab (CDF) and CMS at different energies. Fig. 2.24 illustrates the many pieces in the generation of an event in `Pythia`, for the case of $t\bar{t}$ production.

2.5.2 SIM

As soon as the process under study is generated, it is the task of the *simulation* (SIM) step to use the GEN output to model the interaction of particles with the detector material, producing energy deposits in different detector elements. In CMS, this is done using one of two frameworks. The first one, known as `FastSim`, relies on a simplified version of the CMS geometry with infinitely thin material layers, and includes simple parameterized interaction models. On the other hand, a fully detailed but time consuming simulation is carried out by `Geant4` [244, 245] within the `FullSim`, which is a widely used software toolkit to simulate the passage of particles through matter, and which has access to the full CMS geometry. Comparing `FastSim` with `Geant4`, the former is ∼20 times faster, processing and reproducing the same output within a 10% uncertainty [246]. It is used whenever the additional precision brought by `Geant4` is not required, as is for instance the case for certain systematic uncertainty studies, and also when the speed-up

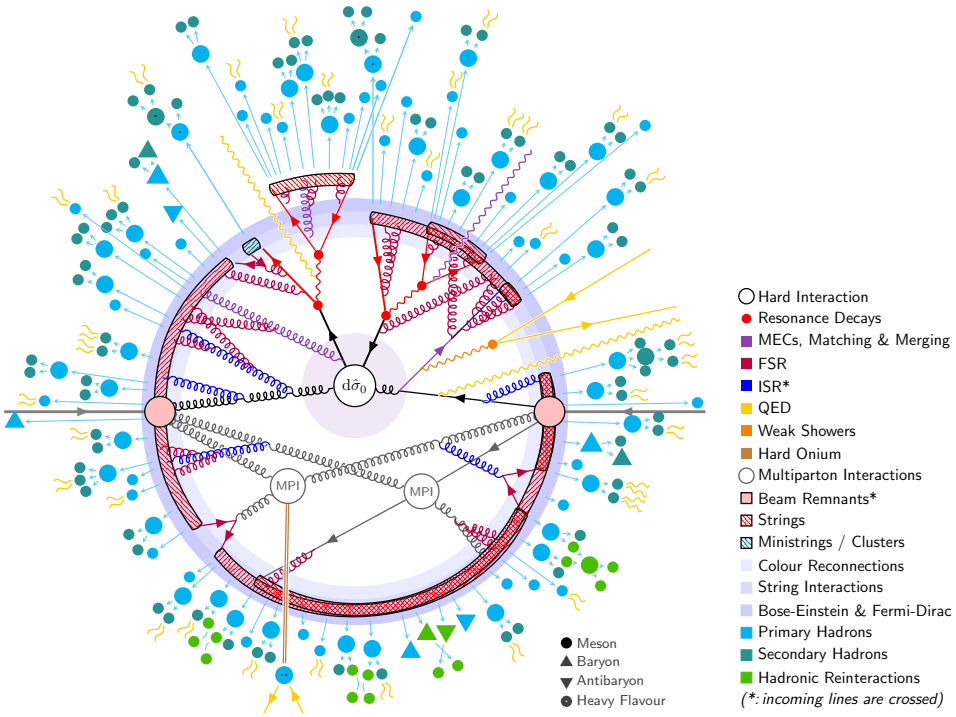

Figure 2.24: Simplified illustration of the structure of a $t\bar{t}$ event, as modeled by `Pythia`. Incoming momenta are depicted as crossed ($p \to -p$) in order to avoid BBRs and outgoing initial state radiation (ISR) emissions to criss-cross the central part of the diagram. Taken from [243].

is essential, which can happen for large parameter scans, as in some SUSY signal scans. In this Thesis, all MC samples were simulated using `Geant4`. The geometry embedded in `Geant4` is continuously updated, reflecting the changes that are currently taking place for the High Luminosity LHC (HL-LHC).

2.5.3 DIGI

The analog output of the SIM step is converted into digital signals, in a step called *digitization* (DIGI). The goal is to mimic the digital output format produced by the detector electronics, including noise, readout logic, pulse shaping and digitization, and L1 emulation. The `FastSim` digitization is very similar to the `FullSim` one [246]. The L1 trigger is emulated at this stage, ensuring consistency with the actual data processing. Importantly, the PU is also simulated at this step, and depends on the concept of Minimum Bias (MB) events. MB events model inelastic pp collisions, named after the very loose and thus "unbiased" triggers used to select them. In CMS, the MB trigger is based on the HF calorimeter ($3 < |\eta| < 5$), specifically on the energy registered by its TTs.

There are two ways of producing samples with simulated PU: *premix* and *classical mixing* [247]. Classical mixing requires the prior production of multiple MB events with the GEN and SIM steps. The primary simulated event is digitized together with the MB events, where information regarding the PU distribution is needed. For the premix technique, instead, the PU sample is digitized separately, before the primary event generation takes place, and can be reused multiple times. As a consequence, the digitization steps needs to be run only for the primary event, and not for the PU, making the premix method much faster and less CPU consuming than the classical mixing. The

premix technique is an essential requirement in the CMS offline processing chain during the HL-LHC, given the prohibitive I/O levels required to sustain the classical approach.

2.5.4 RAW and Reconstruction

The output of the DIGI step is reformatted and packed into the RAW data format, which is the format required to emulate the HLT, and which also exactly matches the format the detector outputs. From this moment on, the MC and data processing chains follow identical steps. Next, the RAW data is unpacked back to the DIGI format. This is required in order to compare the DIGI output with the simulated one for validation purposes. From the DIGI format the offline *reconstruction* begins, being explained in detail in Section 2.4. The output of the reconstruction is called RECO, containing detailed information on the physics objects that were reconstructed, and is reprocessed a few times per data collection period. When using `FastSim`, the same reconstruction is used except for the tracker, where a simplified version aims at reducing CPU time [246]. The first processing iteration is called `PromptReco` and occurs within ∼48 hours of data collection, and a second named `ReReco` follows at the end of the yearly data-taking period. During the LSs, additional reprocessing iterations can be requested, called first Legacy and then Ultra Legacy (UL). Every new iteration improves on the detector calibration with respect to previous iterations. However, analyses usually do not need all the detail provided by the RECO format, which is large (∼ 3 MB/event) and thus inefficient. Instead, smaller and less detailed formats are available with progressively less information and reduced precision, focusing on the quantities most analyses use. The data formats are called AOD, MiniAOD [248] and NanoAOD [249]. The $X \to HH \to bb\tau\tau$ analysis reported in this Thesis used the UL MiniAOD data format. Future iterations of the same analysis in Run 3 will upgrade to NanoAOD, which should provide faster processing times. Central processing tools are naturally compatible with the MiniAOD and NanoAOD data formats.

From now up to the end of the HL-LHC program, we can anticipate a virtuous feedback loop between accumulated data and the theoretical work improving the quality of MC samples. Updates will be also driven by the need to generate larger and larger samples for the most common processes, and to efficiently manage parameter scans for uncertainty studies. We should expect some developments along the directions of a precision increase for inclusive observables, technical improvements for fast and efficient generation of events, and improvements in the modeling of the hadronization and UEs [63].

2.5.5 Activities as CMS Monte Carlo Contact

I started my currently on-going activities as CMS ttH+HH MC contact on June 2023, as part of the Higgs MC group. During this time I generated more than 1000 samples, covering mostly nonresonant HH Run 3 MC requests, for 2022 and 2023, and contributed to the common Higgs MC software tools. The requests cover the vast majority of HH samples that will be used for Run 3 nonresonant HH analyses, including ggF, VBF and BSM samples, for ∼15 final states. Besides sample generation, the contact position also requires following requests closely, request submission priority updates, and serve as an intermediary between users, who request the samples, and overall MC contacts and CMS Higgs conveners, who approve requests and priority changes.

3 Reconstruction of Trigger Primitives for the HGCAL Level-1 Trigger

There is no doubt the crux of a measurement lies in its detector, and in how it converts a physical quantity into a quantifiable signal. Likewise, as most scientists can readily appreciate, the decisions made by a triggering step have dramatic consequences for the data which is ultimately recorded. Yet, in between these two stages, a simple, relatively unknown, but meticulous series of extremely constrained operations takes place, creating the conditions which define whether a specific measured signal is interesting from a Physics perspective. We are naturally referring to Trigger Primitives (TPs), the essential ingredient of any L1 Trigger decision. The reconstruction of TPs is severely limited by bandwidth and latency budgets, by the cost of hardware and the space it occupies, and even by cable connections.

With the upcoming HL-LHC and its intense radiation levels, the computation of TPs will gather even more relevance, as it will enable the control of otherwise unmanageable data rates. This Chapter covers reconstruction methods applied to the groundbreaking High Granularity Calorimeter (HGCAL), a new subdetector to be placed in the CMS endcap regions. Its extreme longitudinal and lateral segmentations, despite potentiating fantastic resolutions, also brings enormous data rates, which can only be contained with the definition of discriminative quantities, the development of clever algorithms and the implementation of reliable software and firmware.

In this Chapter, my contributions to the reconstruction of TPs for the HGCAL are described, together with how they impact future developments. I have started by designing and implementing mitigation techniques for the splitting of energy clusters, an event topology where the TP chain mistakenly interprets a single particle as many. These developments have in turn lead to the introduction of promising new reconstruction approaches, for instance considering different coordinate systems. Future improvements are now straightforward, given the software infrastructure put in place.

We start by introducing the HL-LHC in Section 3.1, giving a strong focus on the physics reasons motivating its installation, ultimately providing a meaning for all the efforts reported in this Chapter. We proceed with a detailed description of HGCAL in Section 3.2.1, a fantastic detector bound to revolutionize calorimetry. We cannot stress enough how HGCAL breaks the frontiers of what can be done with a calorimeter, paving the way for exciting physics-inspired software and firmware developments. In Section 3.3, we detail how the reconstruction is currently defined. Such knowledge is useful to conceptualize the framework described in Section 3.4, which is then exploited in Section 3.5 to produce a series of results related to cluster splitting and beyond. We conclude the Chapter by discussing possible future steps.

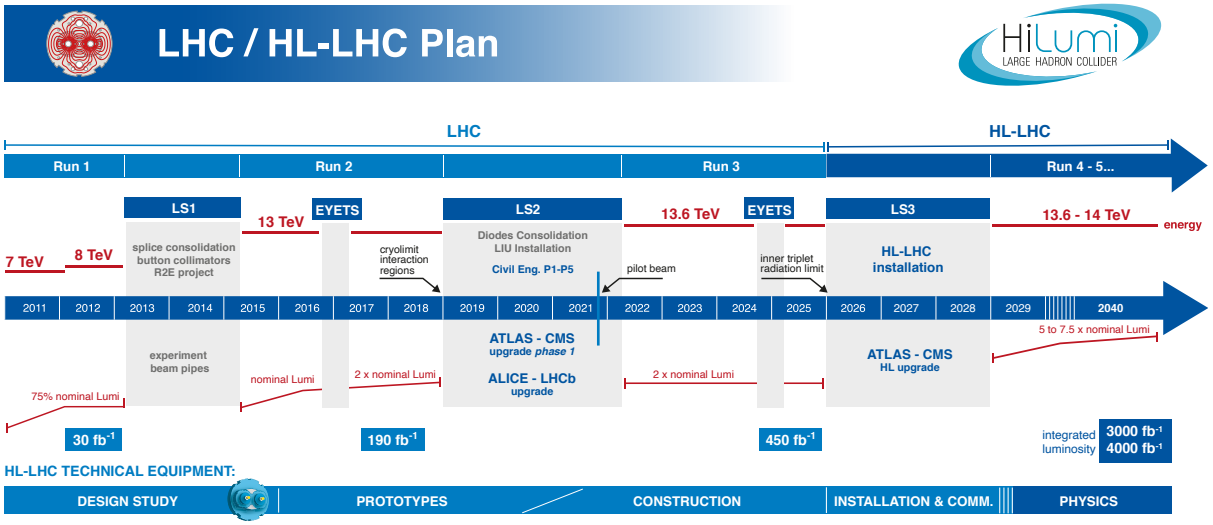

Figure 3.1: The HL-LHC project timeline [253]. Run 3 is currently on-going, and the HL-LHC is planned to start collecting data in 2029, following three years of LHC shutdown for detector upgrades. Data collection should end in the early 2040's. Current plans are still subject to change.

3.1 The High Luminosity LHC

The first phase of the LHC has been running since 2010, spanning three data runs, the last of which, Run 3, is currently approaching its end of life, foreseen in the last quarter of 2025. Phase-2 will soon follow with the new HL-LHC [132], which is planned to start taking data in 2029, extending HEP studies well into the future (see Fig. 3.1) [250, 251]. An integrated luminosity of more than $3\,\mathrm{ab}^{-1}$ is envisaged to be collected over a period of 10 years, corresponding to 90 % of the total data collected by the LHC. The HL-LHC is designed to operate at a center-of-mass energy of 14 TeV, achieving unprecedented instantaneous luminosities of $7.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This is more than twice the LHC's current value. The conditions correspond, in the ultimate HL-LHC configuration, to up to 200 PU, 2 MGy of total ionizing dose[1] and a fluence of up to $10^{16}\,\mathrm{n_{eq}/cm^2}$. The neutron equivalent fluence, in this context, is defined as the number of 1 MeV neutrons per cm^{-2} traversing the detector material, and serves as a measure for potential radiation damage. The normalization at 1 MeV is chosen both due to historical reasons and the fact that expected neutron spectra at the HL-LHC typically have a probability density peaking in that region. For comparison with the HL-LHC, the LHC currently brings an average of around 50 PU interactions [140], a dose of the order of 0.1 MGy and a fluence of about $10^{15}\,\mathrm{n_{eq}/cm^2}$ (see Section 2.1.1) [252]. To give an example of how dramatic the changes are expected to be, current CMS endcap calorimeters are designed to sustain up to $500\,\mathrm{fb}^{-1}$, 8 times less than what the HL-LHC may bring. Extensive detector upgrades are therefore required for many LHC experiments, which will be detailed in Sections 3.2.1 and 3.2.4. The operational scenario of the HL-LHC is continuously evolving, with some uncertainties still present. Given past delays and current unknowns, it is still soon to definitely confirm current operational plans.

The reason for a more powerful LHC is the expected enlarged physics program. The accuracy of current measurements will necessarily increase, and the sensitivity to rare processes through both direct and indirect searches will be enhanced, with more and

[1] One Gray is defined as the absorption of one Joule of radiation energy per kilogram.

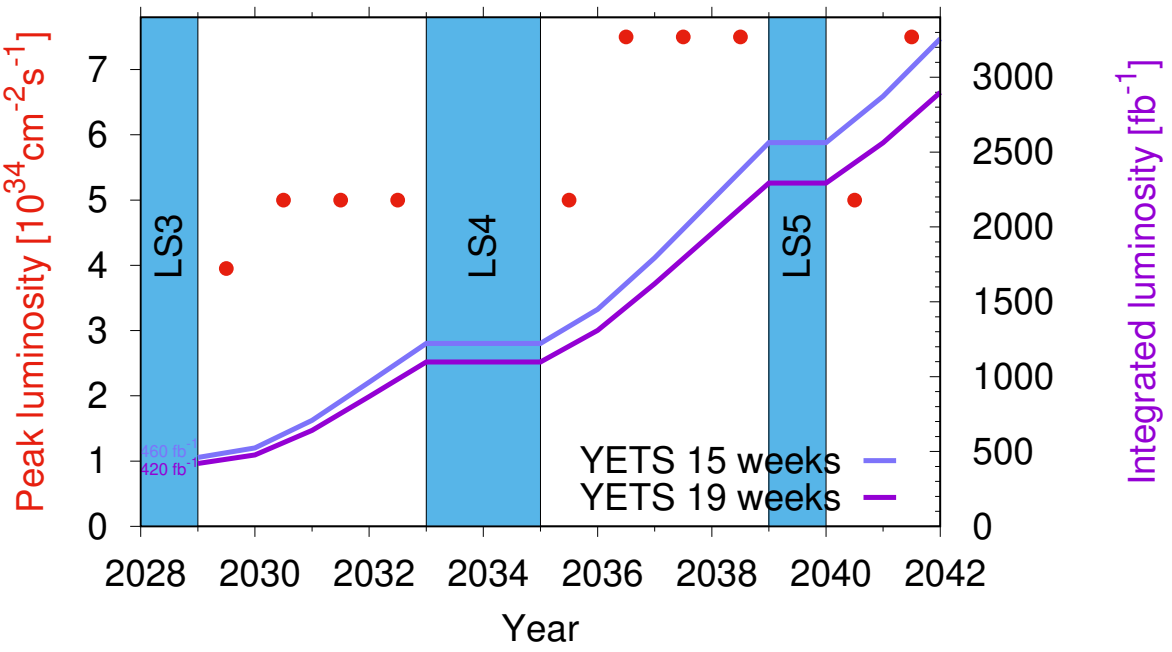

Figure 3.2: Planned peak and integrated luminosities during the HL-LHC. Three data-taking runs are foreseen, interspersed by three LSs. Two scenarios with different Year-End Technical Stop (YETS) durations are shown, where shorter stops have the potential for significant integrated luminosity increases. The schedule was updated on January 2022. Taken from [251].

more challenging signatures occurring below the current sensitivity level. A virtuous cycle often takes place between data collection and improvements in simulation and theory calculations. In the following we provide a brief and necessarily incomplete overview of some of the physics motivations to build the HL-LHC. A detailed description is out of the scope of this work.

3.1.1 EW Processes

The increase in luminosity provided by the HL-LHC will provide a significant boost in sensitivity to measurements of (small) electroweak couplings, especially when dealing with low cross section processes. This notably applies to W mass m_{W} and weak mixing angle θ_{W} measurements, but also to VBF, VBS, and tri-boson production measurements. Circa 2×10^6 W boson events will be collected per week, leading to an increased statistical sensitivity for the W mass, and more precise cross section measurements. The added luminosity might also help solving the tension between old θ_W measurements performed by LEP and Tevatron, adding to the recent CMS result [254], which agrees with the SM, just as previous LHC-hosted measurements. Concerning VBF and VBS processes, despite the overall lower cross section, of about one order of magnitude with respect to the ggF production, the presence of additional hard and forward jets enables a very significant reduction of the multi-jet background. This is particularly true for large rapidity separations $\Delta y_{\mathrm{jj}} \gtrsim 4.5$ and di-jet invariant masses $m_{\mathrm{jj}} \gtrsim 700\,\mathrm{GeV}$, where VBF production dominates over ggF. In VBS, measurements of longitudinal scattering of two Vs (V = W, Z) are particularly enticing, as it is still unknown whether the Higgs alone can avoid a divergent behavior of the scattering amplitude at high energies, violating unitarity [255]. The longitudinal component of the V bosons is in particular tightly con-

nected to the Higgs mechanism and EWSB, since massless particles only have transversal components. However, differential measurements of gauge boson polarization is harder given the reduced cross section of the longitudinal-only components. Would unitarity be violated, particle probability would no longer be conserved: the calculated number of V pairs produced in the interaction would exceed the incident flux [69]. There is also the possibility for the longitudinal components of the V bosons to be strongly coupled at around 1 TeV center-of-mass energies [256], allowing for instance the formation of resonance states in technicolor models. VBS also forms the ideal field for anomalous gauge couplings measurements, especially quartic couplings. Finally, even rarer tri-boson processes ($\sigma \sim 0.1\,\text{pb}$) brings an extra handle on SM testing.

3.1.2 Strong Interactions

The HL-LHC will enable an increase in the kinematic reach for jet and photon production. This will in turn reduce the associated experimental uncertainties, improving PDF modeling. High-p_T b-jets are sensitive to higher-order corrections, parton shower modeling and PDFs. Therefore, investigating the b-jet contribution to the total jet cross section enables to test the available theoretical predictions. The shape of the distributions of jets of various flavors will also be known with a higher precision, especially in the tails. On the photon side, di-photons with high invariant mass are ideal probes to test the SM and search for BSM. In particular, prompt photons do not interact with other final state particles, enabling a precise study of QCD [257]. The HL-LHC will enable a more precise knowledge of photon related cross sections. For the first time, the HL-LHC will make possible the measurement of quantum correlations between partons in the proton by precisely studying differential distributions. This brings an improvement in Double Parton Scattering (DPS) modeling, which has been so far very limited. DPS can be as relevant as the single proton scattering for same-sign WW final states involving $\text{b}\bar{\text{b}}$ or $\text{c}\bar{\text{c}}$ quark pairs.

3.1.3 Top Physics

The heaviest particle in the SM plays an important role in EWSB and in BSM searches. With the HL-LHC, it will be possible to improve the precision of the m_t measurement, which plays a role in understanding the stability of the EW vacuum, as discussed in Chapter 6. The $t\bar{t}$ differential distributions will also be studied in more detail, reaching $m_{t\bar{t}} \sim 7\,\text{TeV}$ and $p_\text{T} \sim 2.5\,\text{TeV}$. These will benefit PDF measurements. One will also be able to explore BSM for $m_{t\bar{t}} > 7\,\text{TeV}$ due to the low SM background. An interesting link might be drawn between the results of LHCb and ATLAS/CMS, given the increase in η reach up to about 4. Further studies will be possible on the rare $t\bar{t}t\bar{t}$ production, despite its $\mathcal{O}(10)\,\text{fb}^{-1}$ cross section, following its recent observation [258, 259]. $t\bar{t}t\bar{t}$ can be used to constrain some of the Wilson coefficients associated to EFT dimension-6 operators, to further study the top Yukawa coupling constrain, including its magnitude and CP properties, since it can occur via the mediation of a Higgs boson [260], and even to assess the presence of the non-SM top quark dipole moment.

3.1.4 Forward Physics

Assuming concepts similar to the current CMS TOTEM Precision Proton Spectrometer (CTPPS) [261] and the ATLAS Forward Proton Project (AFP) [262] are extended in

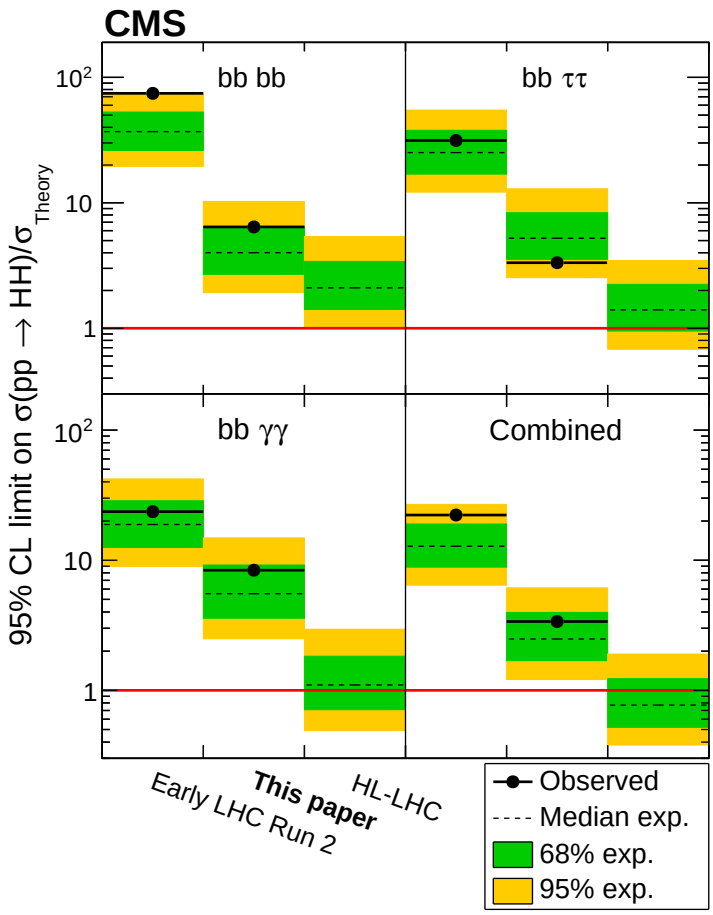

Figure 3.3: Evolution of the expected and observed 95% CL upper limits on the HH production cross section, with the CMS detector. The figure compares results from early LHC Run 2 data ($35.9\,\text{fb}^{-1}$) with full LHC Run 2 data ($138\,\text{fb}^{-1}$), and HL-LHC projections ($3000\,\text{fb}^{-1}$). At the end of the HL-LHC it should be possible to challenge the SM prediction (red line) with the result of the analysis of multiple final states, possibly in combination with ATLAS. Taken from [38].

the HL-LHC phase, Central Exclusive Production (CEP) $\text{pp} \to \text{p}\,(\gamma\gamma \to \text{X})\,\text{p}$ phenomena, among which light-by-light scattering ($\text{X} = \gamma\gamma$), will be further explored. Other processes include, for instance, $\text{X} = \mu\mu$, $\tau\tau$, Z, H, WW, ZZ, and enable the study of anomalous gauge couplings and the magnetic moment of the τ, among other studies [263, 264]. CEP processes carry particular interest since they bring production of charged particles initiated only by photons, into what amounts to using the LHC as a $\gamma\gamma$ collider. In parallel, a whole plethora of QCD-related measurements can be performed in CEPs. The HL-LHC will push CEP processes to higher masses and lower cross sections, increasing their discovery potential.

3.1.5 Higgs Physics

The Higgs boson programme will remain at the forefront of CMS research, with the 150 million Higgs bosons and 120 thousand Higgs boson pairs expected to be produced at the end of the HL-LHC phase. The precision of the Higgs couplings will improve by a factor of about 4, with most couplings measured to a 2 % precision. After the expected obser-

vation of $H \to \mu\mu$ already at Run 3, the $H \to Z\gamma$ channel will share the same fate during Phase 2 [38], despite its uncertainty remaining statistically driven. However, the long-awaited flagship measurement of HL-LHC concerns di-Higgs production. Indeed, current projections (see Fig. 3.3) suggest that the combined analysis of multiple final states and experiments might challenge the SM prediction of the Higgs boson self-coupling. Comparing Run 2 results with current full Run 2 upper cross section limits, we see that an improvement of a factor of around 7 was obtained, much above what a naive luminosity scaling would provide, given the four-fold increase in collected data. The improvement over the luminosity baseline is due to multiple areas, such as trigger, identification, and reconstruction algorithms, but also to the increase of explored finals state channels [265]. Past results have consistently surpassed initial estimates, and thus provide an optimistic view on future triple and quartic Higgs coupling measurements, which will surely be the most precise ever. A series of new techniques should provide significant improvements, starting from Run 3 and extending into the HL-LHC, including new machine learning approaches or better estimates of the QCD multi-jet background. The usage of PNet [266] for τ-initiated jets and the application of transformer technology to jet tagging [120] are expected to boost HH sensitivity. Additionally, an improved trigger strategy has been implemented, considering both data scouting and parking [212], and including PNet b-tagging and τ-tagging at trigger level, as discussed in Section 3.2.2. We also expect some HH analyses to benefit from the inclusion of synthetic datasets [267]. As discussed in Section 1.4.2, indirect searches can also add a contribution in the quest for better HH sensitivity. The precision of EFT couplings will also benefit from the increased number of events. Finally, yet unexplored HH production modes and decay channels are expected to be added, given the phase-space extension brought by the HL-LHC [38]. In conclusion, the next decade looks extremely promising for Higgs Physics.

3.2 The High Granularity Calorimeter and Other Detector Upgrades

3.2.1 Detector Upgrades in CMS

The extreme radiation levels and high PUs mandate extensive updates of most LHC detector systems, and CMS is no exception. For the scope of this Thesis, the drastic upgrades put forward in the trigger system and endcap calorimeters take precedence, and will be described in Sections 3.2.2 and 3.2.3, respectively. However, given the intricacies of the online trigger chain, and the ideas behind PF reconstruction, considering subdetectors as independent elements is naive and ultimately wrong. We therefore provide a brief summary of all other subdetector upgrades, which can put into perspective all recent and substantial efforts of the CMS Collaboration [268].

Tracker

The pixel and strip tracking detectors are subject to the highest radiations and doses among all CMS components, and their performance will abruptly degrade after $500\,\mathrm{fb}^{-1}$. The radiation levels are expected to increase by one order of magnitude during the HL-LHC. The pixel detector originally installed has already been replaced during the technical stop between the 2016 and 2017 data-taking periods [182, 269]. Its structure

is designed with accessibility in mind, such that an update of the pixels does not imply moving the strips. The channel occupancy is required to stay within 0.1% for the pixels and 1% for the strips, which is quite demanding at 200 PU. The track separation is also expected to be improved relative to the one currently observed for high-energy jets, where hits are merged at the level of the pixels. Following the recent upgrade, the material budget of the tracker should also be decreased, since it negatively affects the downstream calorimeters and overall PF reconstruction. The tracker will also be integrated in the L1 trigger chain for the first time, enabling the reconstruction of trigger tracks up to $|\eta| = 2.4$ in a PF-like approach (see Section 3.2.2), and requiring updated reconstruction algorithms, to avoid losing selection efficiency.

The tracker will thus be entirely replaced, with considerably increased granularity and η coverage extended up to $|\eta| \sim 4.0$, both for the pixels and strips [180]. The number of layers will be kept unchanged, ensuring particles traverse on average at least 6 layers, even for the highest η values. The updates will bring better resolutions and lower fake rates.

- **Inner Tracker**: Planar pixel modules will still be present, but will be significantly thinner and lighter: 100- to 150 µm-thick sensors compared to the current 300 µm ones, developed in partnership with ATLAS [270]. The increase in granularity, which comes with a reduction in area by a factor of 6, improves the impact parameter resolution and limits the occupancy. The pixels are organized in 336×432 arrays in a dedicated radiation-hard chip, which is designed to also mitigate single event effects (SEEs) [271]. The flexible design of the Inner Tracker (IT) will continue to allow replacements during technical stops.

- **Outer Tracker**: For the Outer Tracker (OT), the update is mainly driven by the requirement to introduce the information from the tracker at L1. The strips will be grouped into over 13 000 modules with two different configurations: *2-strip* and *pixel-strip*. The former is limited since it lacks segmentation in the z direction [272], and is therefore included in the outer OT layers only. For the three inner layers, *pixel-strips* are used, where each module has ∼1000 strips and ∼30 000 pixels. The barrel pixel-strips measure a third spatial coordinate, constraining the z coordinate of the primary vertex to a 1 mm region, helping in vertex discrimination. Contrary to the current tracker, the future barrel OT will have its first three layers tilted, perpendicularly oriented relative to the IP. This mechanically challenging layout is meant to recover from inefficiencies arising due to particles traversing near-adjacent double-sided modules, in consecutive layers, which cannot communicate with each other (see Fig. 3.4, and compare it with Fig. 2.7).

MTD

The MIP Timing Detector (MTD) [273, 274] has no analog in the current Phase 1 CMS detector. Its addition is driven by the 4-fold PU increase brought by the HL-LHC. Indeed, the primary vertex has to be disentangled from all PU vertices, and timing measurements have been shown to provide the required extra handle, as expected from adding a new uncorrelated dimension to the analysis. The MTD will have the ability to measure the production time of MIPs for vertex location at a 30 ps level, although a deterioration of a factor of 2 is expected by the end of the HL-LHC. This is to be compared with the average

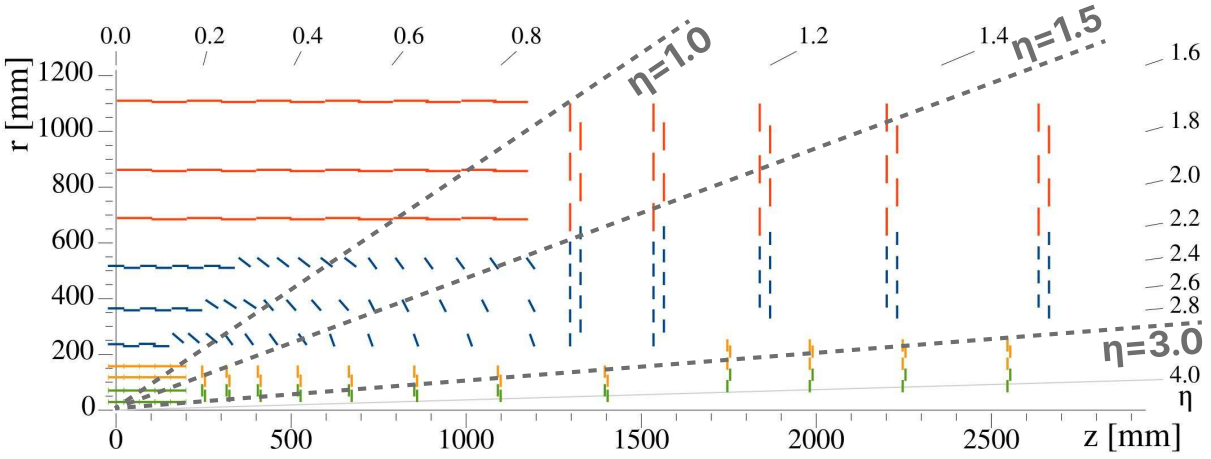

Figure 3.4: Diagram of one quarter of the Phase 2 tracker layout in R-z view. The IT green (yellow) lines correspond to pixel modules made of two (four) readout chips. In the OT, the blue and red lines represent the two types of modules described in the text. The dashed lines provide visual guidance for the η coordinate. Adapted from [180].

$\sim$200 ps root mean square of individual interactions in a BX. The synergy between the tracker and the MTD will enable the association of tracks to the right vertices more often, even for scenarios with closely packed vertices, helping in LLP reconstruction, charged hadron identification and b-jet identification. The latter is tightly connected with the di-Higgs measurement performed in this Thesis, as detailed in Chapter 4, where precision timing can bring an additional $\sim$20 % signal yield. Another interesting cooperation strategy can be envisaged with the calorimeters, especially HGCAL (introduced in Section 3.2.3), which will also provide timing measurements with a similar precision, and can thus be used to ensure compatibility with MTD measurements. To be noted that, contrary to HGCAL, the MTD will not impose energy thresholds, and so should be able to give meaningful timing measurements for lower momentum charged particles. Other benefits are foreseen, like better measurements of τ lepton decays due to the improved MET reconstruction quality, or improved VBF tagging.

The MTD will be located in the gap between the tracker and the calorimeters, in both the barrel as the Barrel Timing Layer (BTL), and in endcap regions, as the Endcap Timing Layers (ETLs), having a 28 m^2 surface area. In the BTL, crystal bars read out by a pair of SiPMs have a granularity implying a 7% occupation per crystal at 200 PU. The low occupation limits the probability of double hits during a BX. Bars are grouped in modules, which are organized in Readout Units (RUs), which in turn are structured in trays, supported by the tracker support tube (see Fig. 3.5). In the ETLs, located at a distance of 3 m from the IP, a two-disk structure of Si sensors will achieve the desired time resolution with two hits per track. The Si technology was chosen due to the higher radiation at the endcaps. The same radiation also requires the addition of a neutron moderator, which shields the tracker from back-scattered particles. The MTD will consist of more than 300 000 SiPMs, which will cover an $\eta < 3$ area on both sides. It will remain accessible to allow for repairs or future updates, again related to the high radiation levels. All in all, the MTD will help CMS maintaining its current performances in a harsher environment.

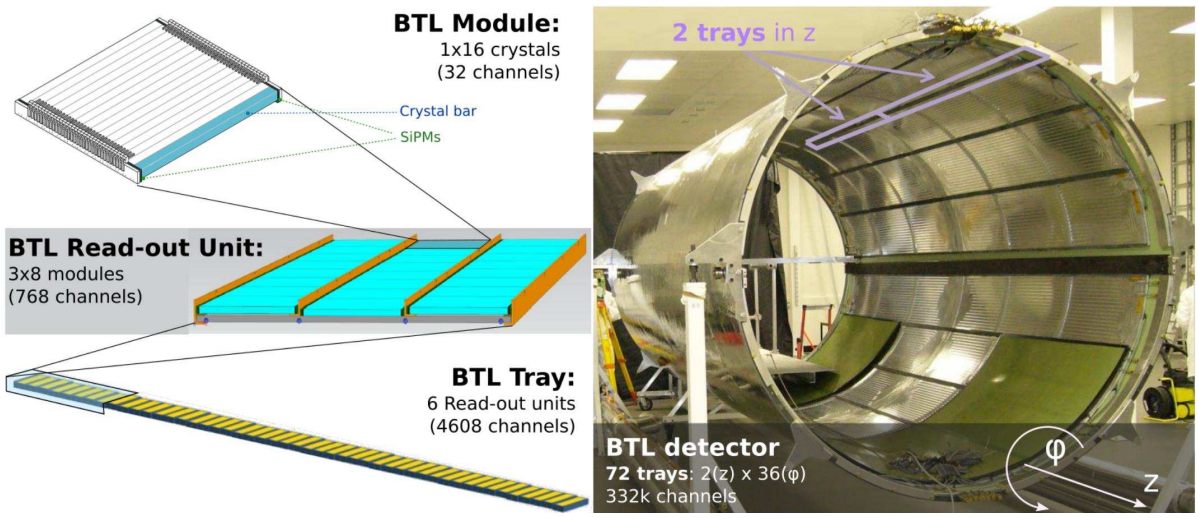

Figure 3.5: Overview of the barrel side of the MTD. (Left) The hierarchical arrangement of the various components, bars, modules, and readout units. (Right) The tracker support tube with highlighted MTD trays (in purple). Taken from [273].

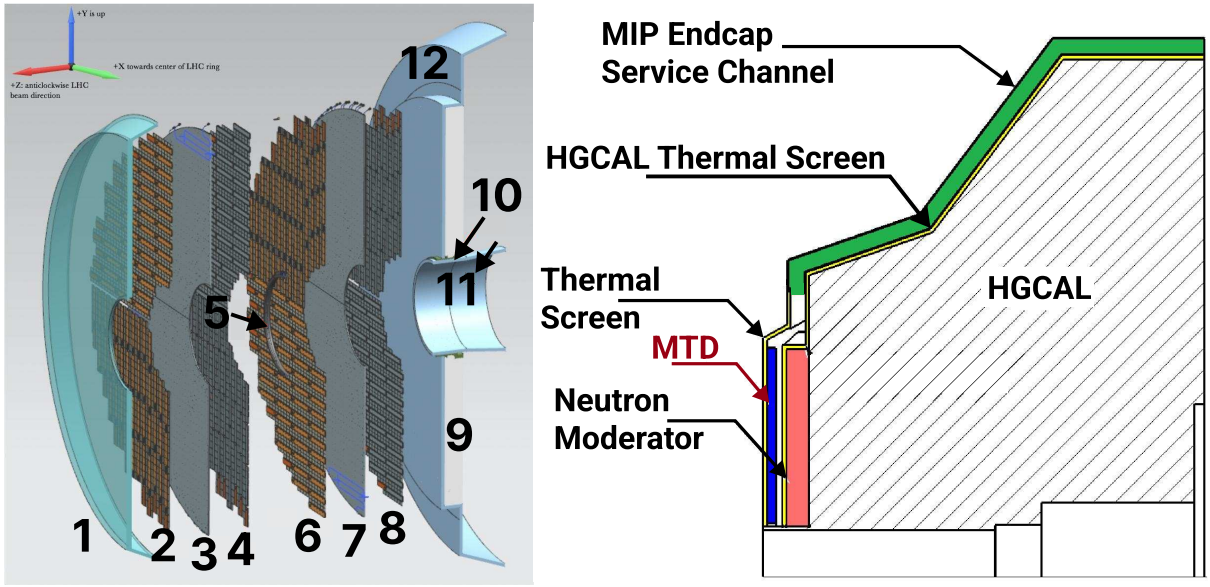

Figure 3.6: (Left) Cross-sectional view of the endcap timing layer (ETL) along the beam axis. The IP is to the left of the image. Shown are two endcap disks populated with modules on both faces, along with the support structure. 1: thermal screen; 2, 4, 6, 8: disk faces; 3, 7: support plates; 9: HGCAL neutron moderator; 10: support cone; 11: insulation of the support cone; 12: HGCAL thermal screen. (Right) Location of the BTL, relative to the future HGCAL. The ETL is placed in front of the neutron moderator of HGCAL, but is kept in a separate volume, enabling its independent access for repairs or upgrades. Adapted from [273].

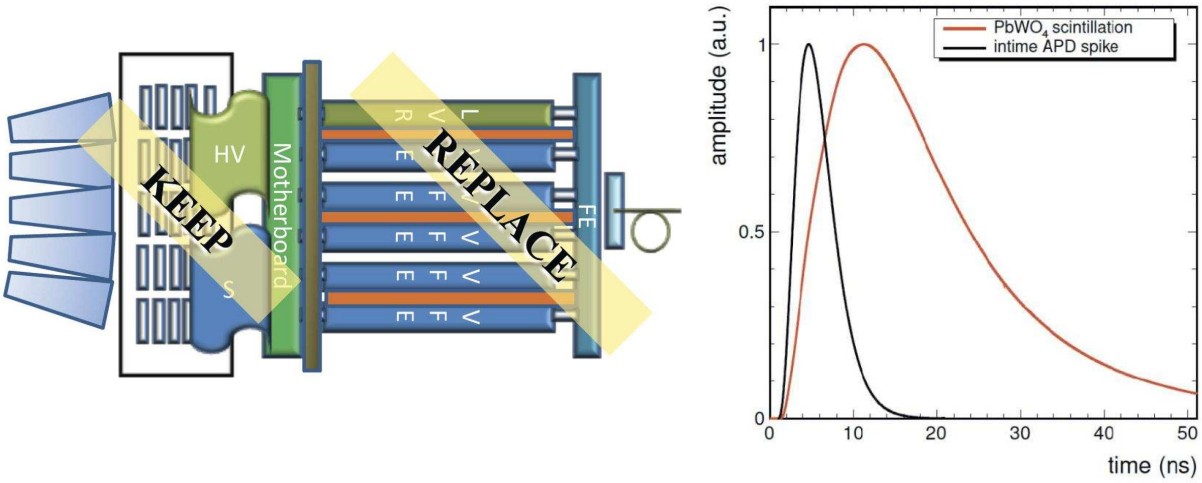

Figure 3.7: (Left) Illustration of the upgrade strategy for the EB: the mechanical structure will remain, while the full electronics will be replaced. (Right) Illustration for the shape difference between a spike pulse shape (black) and a scintillation pulse shape (red), enabling a clear separation. Adapted from [277].

Calorimeters

By far the most important upgrade in the CMS calorimeters will happen in the endcaps, namely the HGCAL. This revolutionary calorimeter is described in full in Section 3.2.3, and we here focus on the remaining updates, most of them happening in the barrel section [275]. ECAL control and safety systems will also be completely replaced [276].

Unlike for most CMS subdetectors discussed in this Section, the changes are not directly driven by the increase in radiation levels. Contrary to what happens in the endcaps, the crystals in the barrel region are capable of handling the environment of Phase 2, especially when cooled from a nominal Phase 1 temperature of 18 °C down to 9 °C, as it is currently planned. Indirectly, however, the higher luminosities and PUs, lead to modifications at the L1 trigger level, namely the latency and rate increase to 12.5 µs and 750 kHz, respectively. Both the FE and back-end (BE) of the EB will be fully replaced, as illustrated in Fig. 3.7 (left). The upgrades are meant to provide precise timing measurements, limit power dissipation, and increase the granularity of barrel TPs to a single crystal level, versus the current 5×5 arrays. This last point reduced backgrounds significantly, enabling the calorimetric trigger thresholds to remain at levels which allow performant physics analyses, namely on precision Higgs physics. The improved timing will be interestingly also responsible for a better noise filtering and, in conjunction with the increased trigger granularity, is expected to completely remove occasional spurious energy "spikes" in the APDs, since they have a different time development with respect to standard scintillation pulses, as shown in Fig. 3.7 (right). Spikes are formed from direct ionization in the APD which results in a faster signal rise and decay time [277]. The timing resolution targets ∼30 ps, enabling constraints of ∼1 cm on the longitudinal position of the collision vertices.

In what concerns the barrel HCAL, detailed longevity studies allowed to conclude that the replacement of scintillators or fibers will not be necessary, given the detector location further away from the high radiation zone. The high η regions will still suffer a degradation of a factor of 2 or 3, but this will be largely offset by the recent FE SiPM upgrade, which addressed all concerns on signal degradation. The SiPM have a

higher efficiency than HPD, and will continue to be used during the full duration of the HL-LHC. However, the BE cannot support the rate increase planned for L1, and so it will be updated accordingly, following the EB design.

The HF will not undergo any Phase 2 upgrade. In order to cope with the larger rates, some boards and crates that had been allocated to the old HB or HE will be reused for the HF, ensuring that the number of channels per board is low enough for the system to be sustainable even at full HL-LHC conditions.

We also mention an entirely new addition with respect to the Phase 1 detector. A much simpler version of HGCAL is being considered for the very forward endcap regions, covering $3 < |\eta| < 4.2$, called Forward HGCAL (HFnose). It would be located in front of the HF, being subject to a maximum absorbed dose of 4.6 MGy and to a fluence of $2 \times 10^{16}\,\mathrm{n_{eq}/cm^2}$, both twice as large as the levels expected in HGCAL. Its motivation stems from improved VBS scattering measurements of longitudinally polarized channels and from enhanced single- and di-Higgs VBF production, both of which include two very forwards jets.

Muon detectors

Following the Run 3 upgrade discussed in Section 2.2.2, with the introduction of the GE1/1 GEM superchambers, two additional GEM detectors will be installed before the start of the HL-LHC. The first, named GE2/1, consists of superchambers similar to the already installed but larger, covering 20 degrees in ϕ. The η coverage will also be slightly extended, from 2.15 to 2.4. The second type, called Muon Endcap 0 (ME0), refers to superchambers located closer to the beamline, at $2.0 < |\eta| < 2.8$, which also span 20 degrees in ϕ [194, 195]. It will be installed on the back of HGCAL, improving the muon tagging at higher η values.

The remaining muon subdetectors, the DTs, CSCs and RPC will have upgraded electronics to cope with the 10 times higher rates and larger chamber occupancies, also improving their performance. For the case of RPCs and DTs, the new electronics have already been at least partially modernized during the Long Shutdown 2, and are being tested with real detector conditions in Run 3. To give an example, the current RPC boards and FE are certified radiation-hard for the HL-LHC, but their link system is composed of outdated components, with no spare boards available, most of which are also no longer available in the market [193]. The link boards will thus be replaced with faster FPGAs. Finally, the addition of the GEMs will be complemented by the installation of additional improved RPC, to be placed in the outer region, at higher η values [278]. These new RPCs have been specifically designed to handle high hit rates.

BRIL

The existing Beam Radiation, Instrumentation, and Luminosity (BRIL) project is responsible for 14 technical systems. They monitor beam losses and radiation in the CMS cavern, provide beam timing, enable the abortion of the beam to protect individual detector components, and measure the beam-induced background and luminosity values. For the HL-LHC, BRIL aims at measuring the luminosity within a 1 % uncertainty, against the 1.4 % obtained in 2022. The improvement is for instance required by multiple single Higgs cross section measurements [278]. The precision increase will be achieved thanks to the new bunch-by-bunch Fast Beam Condition Monitor luminometer, which will be

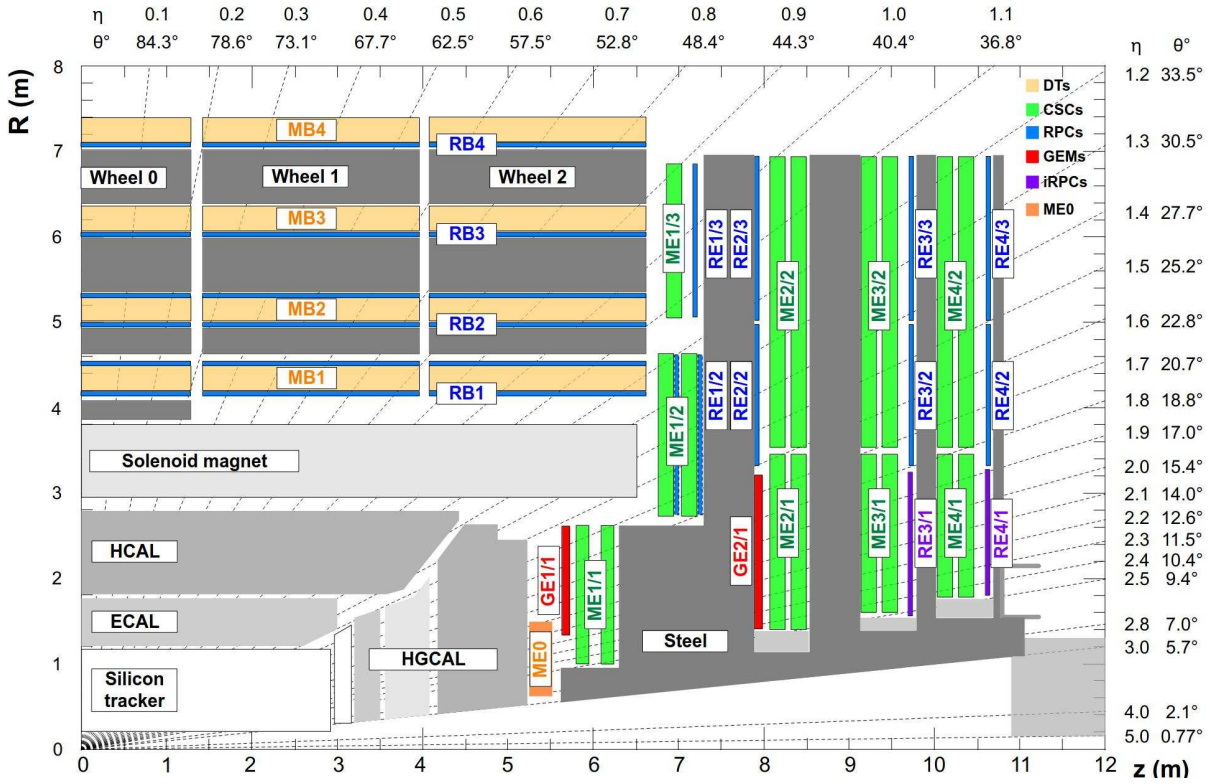

Figure 3.8: Schematic longitudinal view of a quadrant of the R-z cross section of the CMS detector during the HL-LHC, to be compared with Fig. 2.14. All muon subdetector are shown, including future additions: DTs (yellow), CSCs (green), RPCs and GEMs. Additions on the muon side feature the GE2/1 and ME0 superchambers, which are part of GEMs, and the improved Resistive Plate Chamberss (iRPCs). ME0 will be installed on the back of HGCAL. Pseudorapidity values are given with dashed lines. Taken from [194].

independent from central CMS services and will be able to provide measurements also when LHC beams will not be stable. Providing CMS with a fast feedback on LHC beam conditions is crucial, since a continuous beam assessment can protect the sensitive CMS subdetectors.

3.2.2 The CMS Phase-2 Trigger System

Following the current trigger design, already described in Section 2.3, the CMS trigger and DAQ will continue to feature an online synchronous hardware-based L1 trigger, consisting of custom electronic boards, followed by the HLT, running software algorithms offline and asynchronously [279]. Hand in hand with critical detector upgrades, a complete redesign of the L1 CMS trigger system is foreseen [198]. For the first time, the tracker will be taken into account for L1 decisions, and the granularity of the endcaps will increase dramatically. The all-hardware data processing chain focus on latency reduction by carefully optimizing available resources, from FPGA memory and throughput allocation, to high-speed links for data transfer. The new trigger algorithms and architecture must deal with the higher luminosity and increased pile-up, while accessing all the discovery potential the new or updated detectors can provide. To cope with the above, the total latency of the system is increased from 3.8 µs in Phase 1 to 12.5 µs, and the output trigger rate from the current ∼100 kHz to 750 kHz, maintaining the efficiency of the signal selection to the levels exhibited in Phase 1, while keeping or enhancing the sensitivity to NP. The Phase 2 trigger architecture will become more stratified and interdependent, consisting on four initially separate data processing paths. The implementation of global triggers (Global Calorimeter Trigger (GCT), GMT and Global Track Trigger (GTT), plus the final GT to where all paths converge) allows for resource reduction when implementing the PF-like algorithms in the newly introduced correlator layers. The correlators combine the information from multiple subdetectors to achieve near-HLT sensitivity, exploiting a novel suite of intricate algorithms. This structure is complemented by the introduction of an innovative data scouting system, which further attests to the almost real-time analysis L1 will provide. The full architecture of the Phase 2 L1 trigger is illustrated in Fig. 3.9, and Fig. 3.10 displays the complex interplay between TPs from multiple subdetectors, trigger objects, L1 algorithms and the physics being targeted.

Calorimeter Trigger

The upgrade of the barrel and endcap calorimeters will be exploited to process high-granularity data to produce cluster-related variables, leading to calorimetric TPs. Additionally, using the HFs will extend the η coverage up to $\eta = 5$. The GCT assembles all TPs and builds calorimeter-only objects, such as hadronically decaying taus, e/γ candidates, jets and energy sums. The data is then sent to the correlators with a 5 µs latency, which also corresponds to the HGCAL allowed latency, of which 1 µs is used only for transferring the data from the BE to the GCT. Some global calorimetric quantities are sent directly to the GT, within 9 µs maximum. A detailed description of the reconstruction of trigger primitives in the future HGCAL will be provided in Section 3.3, as it is closely related to part of the work developed in the context of this Thesis.

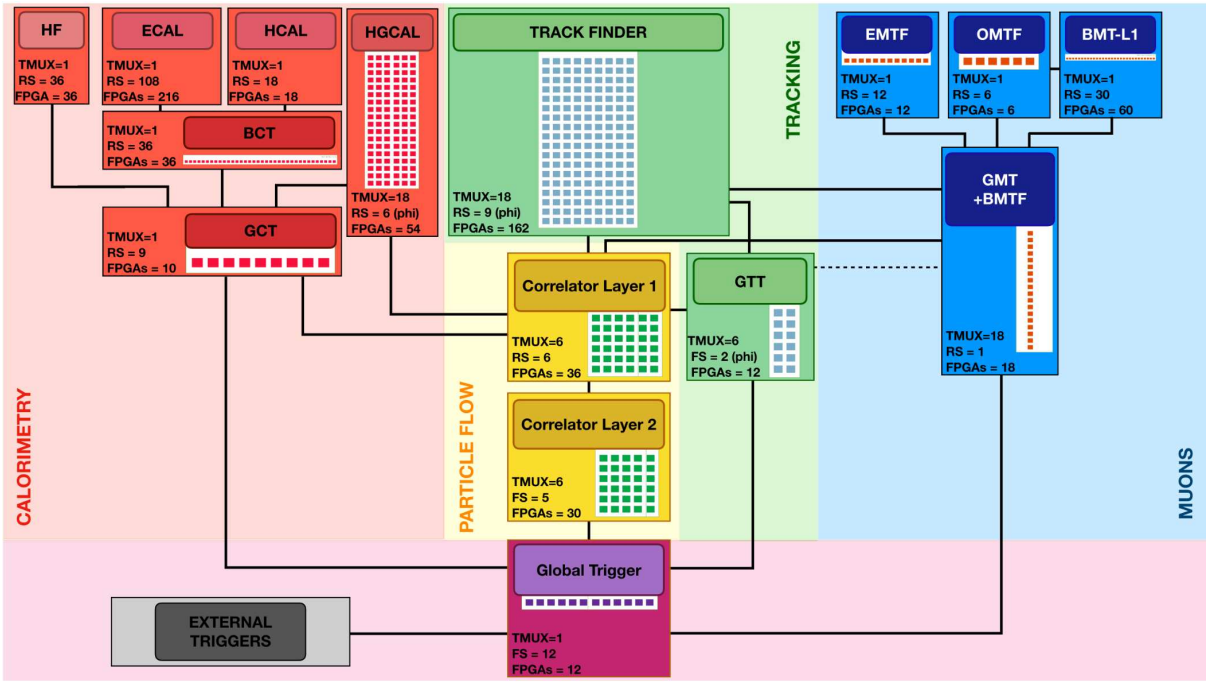

Figure 3.9: Diagram of the CMS L1 Phase 2 trigger design, to be compared with Fig. 2.17 (left). The calorimeter trigger is represented on the left. The track finder in the center sends tracking information to the correlator, the GTT, and the GMT. The muon trigger architecture is represented on the right and composed of three muon track finders. The correlator in the center is composed of 2 layers for PF processing. The GT receives all trigger information for the final decision. For each architecture component, the information about the time multiplexing period (TMUX), the regional segmentation (RS) in η or ϕ, the functional segmentation (FS), and the number of FPGAs are specified. Taken from [198].

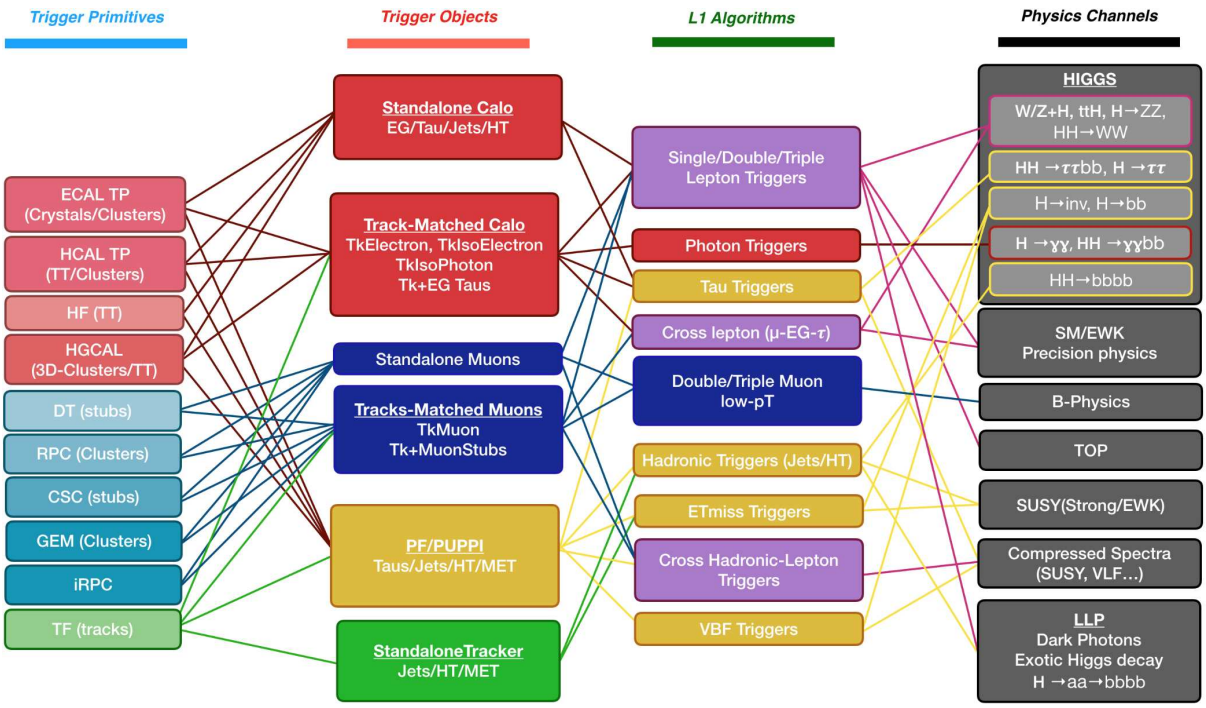

Figure 3.10: Summary diagram showcasing the interdependence of TPs, among which the ones coming from HGCAL, and physics, including HH processes. The links between TPs, trigger objects, L1 algorithms and physics channels are depicted. TPs include crystals, towers and clusters from calorimeters (ECAL, HCAL, HF and HGCAL), stubs and clusters from the muon detectors (DT, RPC, CSC, GEM and iRPC), as well as L1 tracks from the track finder. The trigger objects types produced by the Phase 2 L1 trigger system are represented: standalone, track-matched, tracker-based and PF/Pileup Per Particle Identification (PUPPI)-based. Taken from [198].

Track Trigger

For the first time in CMS, it will be possible to reconstruct charged particle tracks in the tracker at 40 MHz. TPs are produced in the track finder processors of the OT BE, and are sent to the GTT. There, the reconstruction of the PV and tracker-only objects like jets and missing transverse momentum takes place. If needed, the GTT can also propagate extra copies of tracks to other parts of the trigger system. The new track trigger will be able to mitigate the negative effects of PU in PF via a precise measurement of the position of the PV at L1.

Muon Trigger

The muon detectors' redundancy will be kept, and additional GEM stations will offer a coverage of up to $\eta = 2.8$. The processing of TPs by the muon track finders will follow the same region-based structure as in Phase 1. The GMT thus receives data from the EMTF, OMTF and BMTF, but also directly from the tracker to build muon candidates. The ability to reconstruct tracks at L1 will bring some muon reconstruction capabilities previously available for offline reconstruction only. Even without considering the tracker, new possibilities will open-up: triggering on muon-like LLP decays or slow phenomena, such as NP with heavy stable charged particless (HSCPs). Once all processing is finalized, the GMT propagates the information to the GT.

Particle Flow Trigger

Consisting on two flexible correlator layers, a PF-like approach called Correlator Trigger (CT) is introduced for the first time in CMS at L1. It optimally combines data from the calorimetry, tracking and muon trigger paths. The first layer builds the high-level PF candidates, and the second implements identification and isolation algorithms, building objects such as electrons, muons, jets, hadronic taus and energy sums. The correlator might also host a simplified version of PUPPI, with the goal of reducing the energy resolution degradation from PU. The CT finally sends a list of sorted objects to the GT.

Global Trigger

The outputs from the GCT, GMT, GTT, and CT are merged, and correlations between variables from all processed objects are exploited to further enhance the sensitivity of the scalable GT. Examples of such correlations are $\Delta\eta$, ΔR, and invariant masses between trigger objects. An event accept or reject decision is made based on a menu of $\sim$1000 algorithms, expected to continuously evolve based on lessons learned during data-taking. Improved hardware capabilities, together with striking advancements in available software libraries for firmware implementation and testing, like `hls4ml` [280], bring new Neural Network (NN)-based algorithms to the table. Some examples include auto-encoders to keep potential NP events from being rejected. Despite NN's high resource consumption, current hardware capabilities suggest that their inclusion is feasible, also given that the GT's logic is expected to run at frequencies at least 6 times larger than the proton collision rate. Other functionalities of the GT include the setting of prescales and the monitoring of trigger rates. The GT also serves as an interface to external triggers, such as the ones related to the CTPPS experiment, or to beam and luminosity monitoring. External triggers can be used in conjunction with other triggers in the same algorithms.

Each of the GT boards sends a readout record to the DAQ, containing information like GT inputs and algorithm-specific bits. This is in turn useful for the HLT, which can use the information to seed its own reconstruction with L1 objects, and reject seemingly interesting events at HLT level which were found by the L1. In total, the GT is expected to have a latency of $\sim$1 µs. Future upgrades including MTD timing information are also being envisaged. The addition would bring a powerful handle to observe displaced signals.

Scouting system

A serious advancement of the future trigger capabilities is the inclusion of the data scouting strategy already at L1. Spare optical outputs from various FPGAs will perform zero suppression and preprocessing at the beam collision rate of 40 MHz. This is the first time that CMS is able to process L1 data in a triggerless fashion. It must be noted that the data stored by the scouting stream will have the same limitations of L1 data in what concerns purity and resolution. The added scouting data will enable the study of processes lacking a clear signature for data reduction at L1, and also phenomena where the definition of TPs is not essential for competitive measurements. Moreover, the scouting will enable a detailed monitoring of the entire trigger system, and bring anomaly detection in quasi-real-time. Finally, the scouting will introduce some complementary luminosity measurements based on observed physics processes.

HLT

The Phase 2 HLT will analyse the full 750 kHz L1 output, which translates to a $\sim$50 Tb s^{-1} throughput, with the event size increasing to approximately 8.5 MB. The 30-fold throughput increase with respect to Phase 1 IS mostly driven by the upgraded tracker and the new HGCAL, which bring a tremendous increase in granularity. The reconstruction of these high granularity objects will be based on iterative procedures, namely Kalman filter algorithms with deterministic annealing for the tracker, and The Iterative CLustering algorithm (TICL) for HGCAL [281]. The goal of the HLT online selection will be to balance the following three key elements [279]:

- preserve and possibly improve the CMS physics reach for the most important processes, without strongly modifying current thresholds and efficiencies;
- reduce the event rate by a factor of 100, just like it is being currently done in Phase 1, since a 7.5 kHz HLT output rate is considered the maximum supported for permanent storage and offline processing;
- implement new algorithms to achieve the above within available resources.

The decision by CMS to adopt a heterogeneous HLT farm already in Run 3 inaugurates a novel trigger approach which will be further explored and extended during Phase 2. The offline reconstruction workflow is expected to be offloaded to GPUs by 50% and 80% by the end of Run 4 and Run 5, respectively [282]. This estimate includes the detectors not yet installed, such as HGCAL. The adoption of a heterogeneous architecture also potentially reduces the computing cost necessary to satisfy the CMS physics programme, since computation on GPUs might be cheaper than on CPUs [283]. Other potential improvements are currently being explored, such as the migration from traditional CMS

data formats to Structures of Arrays (SoAs), for better CPU and GPU utilization. Additionally, the rewriting of many algorithms to parallelization-friendly versions will boost their performance. Furthermore, some approaches to write truly heterogeneous code via abstract interfaces is being considered, in order to improve code reuse (and this avoid code duplication), and avoid the excessive dependence on particular vendor of CPU and/or GPU processors. The `alpaka` [284–286] and `Kokkos` [287, 288] portability libraries represent some of the most promising solutions. to reduce code duplication and dependency on the architecture of a particular vendor [289, 290].

3.2.3 The High Granularity Calorimeter

Running conditions during the HL-LHC data collection will be much harsher and crowded than what the LHC entails. This is particularly true for the CMS endcap regions, given the large radiation levels expected closer to the beam axis. The forward ECAL and HCAL were however designed to sustain an integrated luminosity of up to 500 fb^{-1}. This is because scintillation-based solutions are very sensitive to ionizing radiation [291]. The formation of color centers during irradiation has already lead to a stark η-dependent reduction of the transparency of existing $PbWO_4$ endcap crystals, often by values above 90%. Despite the possible recover through spontaneous or thermal annealing which can mitigate the opaqueness brought by photons [292], irreversible damage is caused by hadrons, mostly charged pions of $\sim$1 GeV [188, 293]. Given the 6 to 8 times larger luminosities expected for the full HL-LHC data-taking period, with respect to the LHC, a drastic degradation of the physics performance seems unavoidable (see Fig. 3.2). Negative effects would be evident almost from the start [294].

The CMS experiment thus foresees the complete replacement of its endcap calorimeters, introducing the ambitious HGCAL project [295]. The HGCAL will be a sampling calorimeter, with fine transverse and longitudinal granularity, capable of fully exploiting the physics events produced under the expected extreme radiation conditions. It will prominently feature Si as active material in the regions closer to the pp interaction vertex, and thus more impacted by incoming radiation. This approach departs from other more established technologies, for instance the use of liquefied noble gases, as done in ATLAS [146]. Si has the potential to bring a high signal-to-noise ratio, since only 3.6 eV are needed to produce an electron-hole pair [146], but it was mainly chosen for its ability to cope with fluences 50% higher than the ones expected by the end of Phase-2. Alternatives to HGCAL were originally considered, such as an optical Shashlik calorimeter [296] and a Dual Readout calorimeter, but were in the end judged less performant.

The proposed novel HGCAL design includes a silicon electromagnetic calorimeter (CE-E) with Si as active material in the first 26 layers, as shown in Fig. 3.11. Each layer is composed of $\mathcal{O}(300)$ 8″ hexagonal Si modules, which shape is chosen to cover the available circular transversal area more efficiently than rectangles, keeping the ability to tesselate the plane perpendicular to the beamline without introducing unmanageable hardware and mechanical challenges. The choice is also related to the circular cross section of grown Si cylindrical ingots, leading to less wasted material. Each Si module is composed of a succession of hexagonal layers featuring, in order, a printed circuit board (PCB), or hexaboard, a Si sensor, a 105 µm Kapton foil, and a CuW baseplate. The Kapton foil provides insulation between the sensor back-plane and the baseplate, which is held at ground [295]. Full modules are mounted on either side of a 6.2 mm thick Cu cooling plate which forms, together with the module baseplates, the absorber. Each module groups

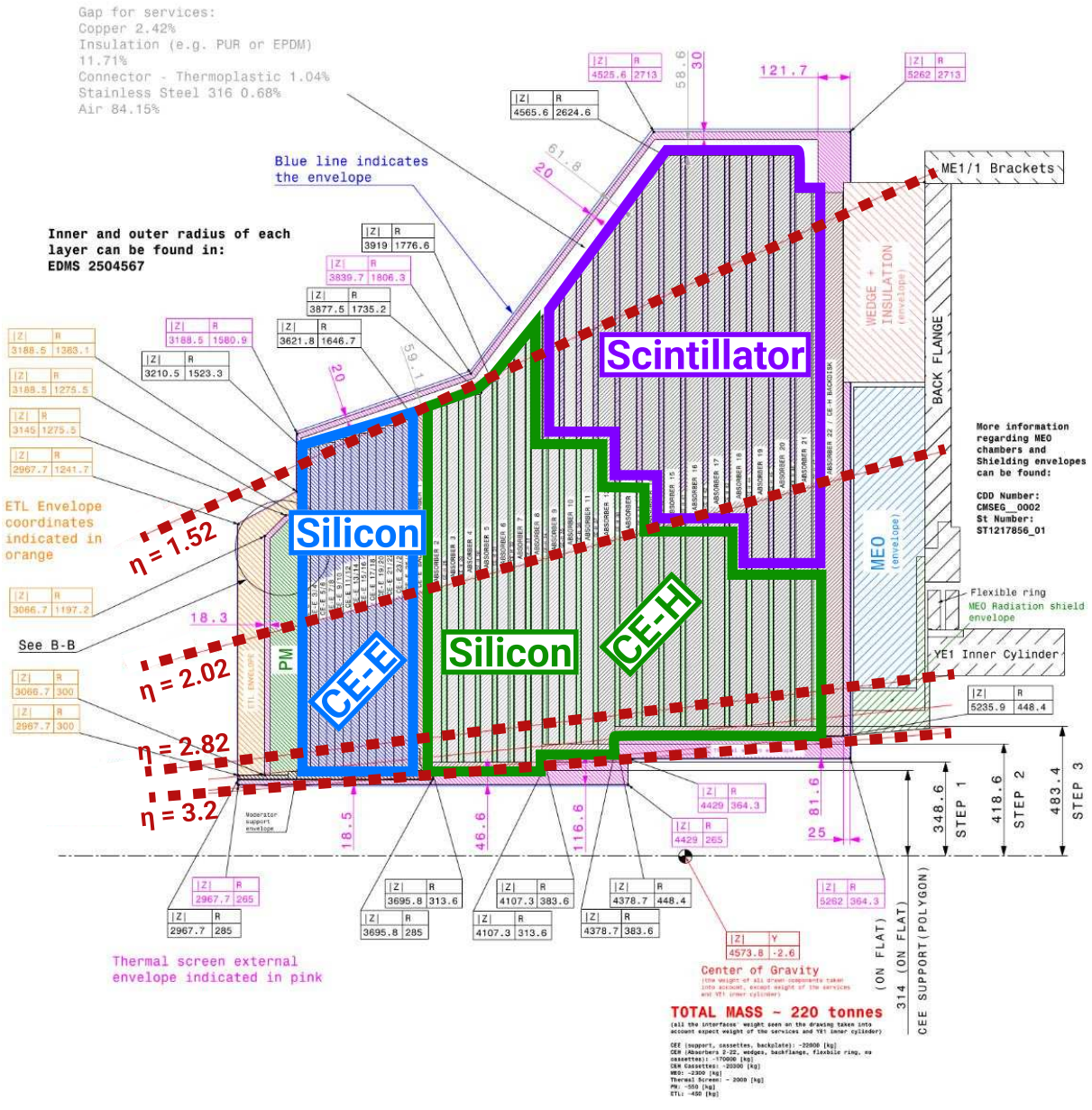

Figure 3.11: The longitudinal profile of the positive endcap of HGCAL in its latest design version. The first 26 layers, in blue, are part of the CE-E. The CE-H follows, in green, and some mixed layers lie deeper in the calorimeter, where purple refers to the region with plastic Sci tiles. The active material alternates with absorber material, varying according to the detector location, as described in the text. Adapted from [297], which is partially based on [295].

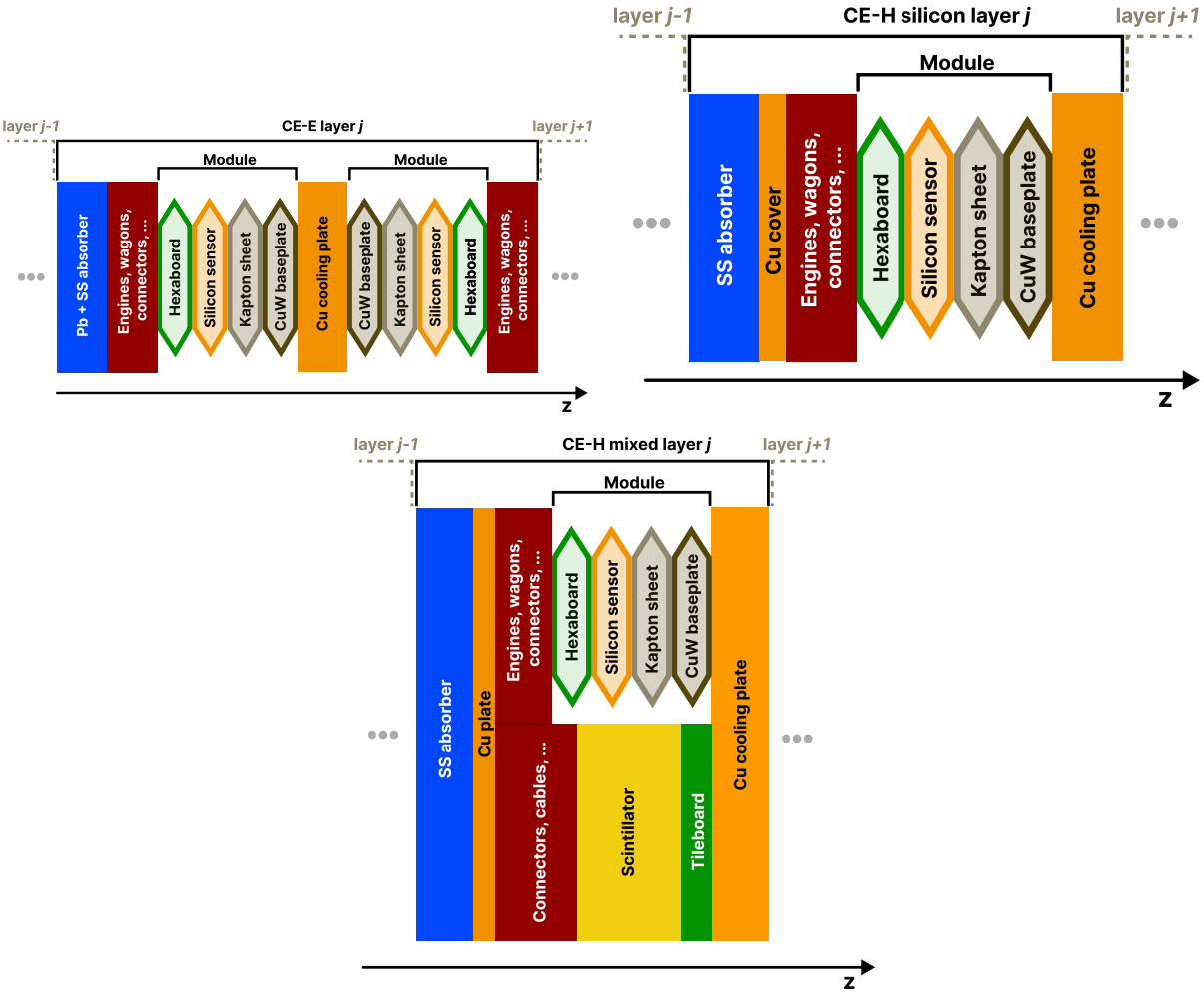

Figure 3.12: Illustration of one layer in the longitudinal structure of the CE-E (top, left), Si-only CE-H (top, right), and mixed CE-H (bottom) sections. Each CE-E cassette comprises two sampling layers. The CuW baseplate provides rigidity and cooling to the Si module. The Pb+stainless steel (SS) absorber in the last CE-E cassette is replaced by a 1 mm Cu cover. Si layers in the CE-H are mounted to a single side of the Cu cooling plate. The Sci tiles nearer the beam line are significantly smaller than those at the outer edge. The tileboards include the SiPMs. The components are not drawn to scale.

individual Si cells, or channels, which constitute the atomic detection unit of HGCAL. A series of electronic components groups the modules into larger physical and logical units. The *engines* convert electrical signals from the hexaboards into optical data, while the *wagons* are PCBs which connect the engines to the modules. In total, the 6 million channels are organized in ∼30 000 modules, which are in turn assembled into 60° triangular-shaped units called *cassettes*. Each layer thus includes 6 such cassettes, for a total of 78 in the CE-E, for one endcap. The CE-E alone extends for ∼27 X_0. We note that the total absorber material is around 5 orders of magnitude thicker than the active material. Its layout can be seen in Fig. 3.12.

Sensors are split in two different density regions, depending on their active cell size. This is required due to the η-dependent fluence. The *high density* region comprises 0.56 cm^2 cells, which are part of 120 and 200 µm thick sensors (thinner than the current CMS tracker!), and is located closer to the beam axis. The *low density* region, located at $|\eta| \lesssim 2.15$ ($R \lesssim 70$ cm), is instead made of 1.26 cm^2 cells, in modules with thicknesses of

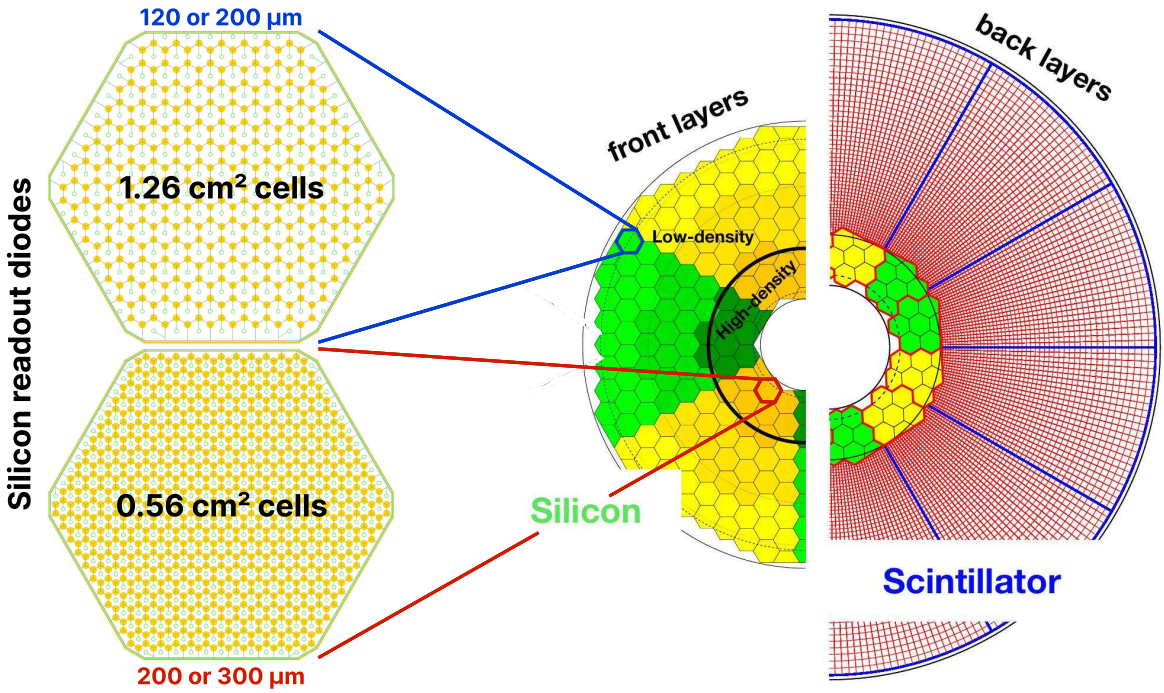

Figure 3.13: Representation of the Si sensor layout in CE-E and CE-H layers. While CE-E layers are made only of Si as active material, CE-H layers are hybrid, containing both Si and Sci. Two possible cell sizes are defined, constituting the low- and high-density regions in each layer. The radial changes in color transparency indicate different silicon thicknesses: 300, 200, and 120 µm. The solid black line marks the boundary between the high-density and low-density regions. The succession of green and yellow colors delimit the 60° cassettes. For the hybrid layer, the blue lines in the Sci section and the red lines in the silicon section delimit the 30° cassettes. Taken from [298].

200 or 300 µm, depending on their location, as shown in Fig. 3.13. Despite the benefits brought by an increase in transversal granularity, the cell area is ultimately bounded by the need to dissipate power, with its limits on cooling requirements [295]. The exact choice of the area is also connected to the size of trigger cells (TCs), as discussed in Section 3.3. The sensors are operated at −30 °C to reduce their leakage current, and are placed under a 800 V voltage to reduce signal loss. It is expected that an increase of bias voltage throughout the lifetime of the detector will be necessary to maintain a good S/N. The S/N also benefits from using Si cells with low capacitance, since the latter increases electronic noise.

Different thicknesses are considered (300, 200, and 120 µm) given the compromise between signal and noise in a semiconductor. The number of electron-hole pairs created by a passing charged particle increases with sensor thickness. However, charge collection is optimized in thinner sensors, given the high radiation which creates traps in the Si. At the same time, noise depends on the presence of leakage current, which is linearly proportional to the sensor thickness and to the integrated fluence, which in turn also increases power dissipation within the sensors [295, 299]. Thinner sensors are thus preferred for regions exposed to higher fluences.

The HAD section includes 7 silicon-only layers followed by 14 hybrid layers composed of both Si sensors and Sci plastic tiles, where the scintillation light is read out by SiPMs. We refer to the Si layers in the HAD section as silicon hadronic calorimeter (CE-H). The Si sensors are placed closer to the beamline, and the cheaper plastic tiles at lower η values

Thickness [µm]	**120**	**200**	**300**
Cell size [cm^2]	0.56	0.56 or 1.26	1.26
Maximum expected fluences [$\times 10^{15}$ n_{eq}/cm^2]	7	7	1.5
Total area [m^2]	72	181	245
Capacitance [pF]	48	29 or 67	45
Initial S/N for a MIP	4.5	6	11
Smallest S/N for a MIP after 3000 fb^{-1}	2.2	2.3	4.7

Table 3.1: Features of the Si sensors. The Si cell size defines two regions, namely the high-density and low-density regions [295, 300].

instead (see Fig. 3.11). Concerning the Si sensors, their baseplate is made of carbon fiber, with a negligible contribution to the CE-H stainless steel absorber material. The scintillators cover an area of ∼400 m^2 for a total of ∼240 thousand channels. Individual tiles have varying sizes due to the triangular shape of the cassettes, running from ∼4 cm^2 closer to the beam axis, to ∼32 cm^2 near the outer edges. Cassettes placed in the HAD section have a similar architecture to the ones in the CE-E, but they are much larger and thus span a 30° angle instead of 60° to facilitate the final assembly. CE-H cassettes also have modules mounted on a single side only. They also include an additional Cu cover meant to provide protection during the horizontal insertion. Stainless steel is used as absorber material. This configuration amounts to a total of 10 λ_0, 1.3 λ_0 for the CE-E and 8.5 λ_0 for the CE-H. The rest comes from the Polyethylene Moderator (PM). The layout of the CE-H section can be seen in Fig. 3.12 (bottom).

Despite silicon's radiation hardness, photons and hadrons still cause damage in the sensors. Charge traps can be created between the valence and conduction bands, negatively impacting the resolution of the semiconductor device, and bit flips can corrupt the stored data. These effects can be reversed by software corrections and/or redundancy, or by applying temperature annealing. However, more permanent damage is created when particles, often hadrons, knock Si atoms from their lattice site, creating long-lasting defects. To avoid the deterioration of the crucial (and expensive) Si sensors, a PM neutron moderator is added in front of the CE-E, to reduce the number of neutrons coming from the tracker [144, 301, 302]. Cooling manifolds, pipes, gas supply and cables are placed along the absorber's outer surface. Finally, a thermal screen covers each endcap, to insulate the cold detector volume from the experimental cavern. These and other components can be seen in the 3D schematic of HGCAL in Fig. 3.14.

The extremely high granularity facilitates particle identification and high resolution measurements of the position, energy and time of high-energy collision products, exploiting the CMS PF reconstruction. The high transversal granularity enables the separation of close-by showers and the observation of narrow jets. It also minimizes the PU contribution to energy measurements. In parallel, the fine longitudinal granularity is beneficial for the EM energy resolution, pattern recognition, discrimination against PU and increased pointing capabilities, including new dedicated triggers at L1 for displaced objects, crucial for the correct identification of LLPs. The longitudinal segmentation was originally chosen to ensure a $H \to \gamma\gamma$ mass resolution as high as the one obtained at the time of the Higgs boson discovery [295]. The overall granularity is such that HGCAL will be the first calorimeter ever able to perform tracking.

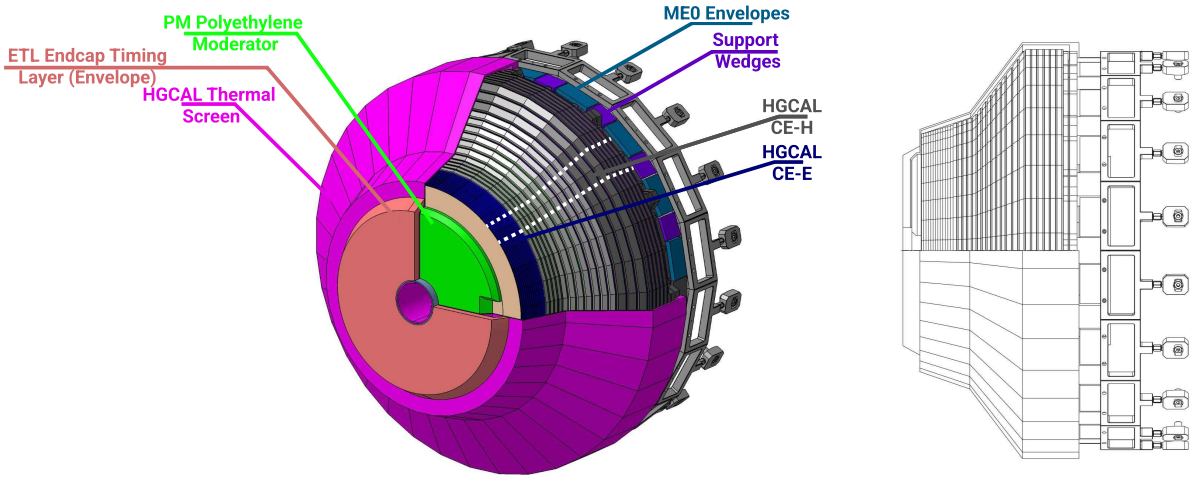

Figure 3.14: Schematic 3D view of one endcap of the HGCAL. Different subdetectors can be seen, such as the CE-E, the CE-H and the ETL. Other sections are required for structural reasons, as for instance the brackets, on the right-most region of the right plot, which are meant to attach the HGCAL to the muon chambers. The PM, or neutron moderator, placed just in front of the CE-E, reduces the number of neutrons coming from the tracker. The two dashed lines give a rough idea on the location of one pair of cooling supply and return tubes, which are connected to the layers, and are placed every 30°. The picture on the right provides a side view of the same endcap. Adapted from [295].

For the first time, high-precision timing will be used in calorimetry. Timing information is considered one of the very few options to mitigate performance degradation arising from PU [293]. In the HL-LHC environment, a robust vertex identification becomes possible with a time resolution of ∼30 ps, with an accuracy of a few cm.

In summary, while sustaining the doses and fluences[2] shown in Fig. 3.15, the HGCAL fulfills the following requirements for its upgrade [295]:

- ensure the lateral compactness of showers thanks to sufficient material density;
- ensure high transverse and longitudinal granularities;
- preserve the energy resolution until the end of Phase 2, with an effective inter-cell calibration within ∼3 %;
- enable precise timing measurements;
- include L1 trigger capabilities.

The features above turn HGCAL into a truly five-dimensional sampling calorimeter: energy, three-dimensional position and timing measurements. A summary of some of the parameters of HGCAL is provided in Table 3.2.

The most recent 2018 and 2021 testbeams hint at the future performance of HGCAL. The used prototype resembles the final calorimeter in the CE-E and silicon CE-H sections, including 300 µm-thick modules with 1.1 cm^2 Si cells, for a total of 28 CE-E layers. Beam energies ranging from 20 to 300 GeV were tested. The Analogue Hadron Calorimeter (AHCAL) prototype [306] is used as a proxy for the future HGCAL scintillator layers, and is placed downstream the silicon sections. For positrons, the stochastic term of

[2]The fluence is relevant for silicon detectors, while the dose is an important parameter for electronics and scintillators.

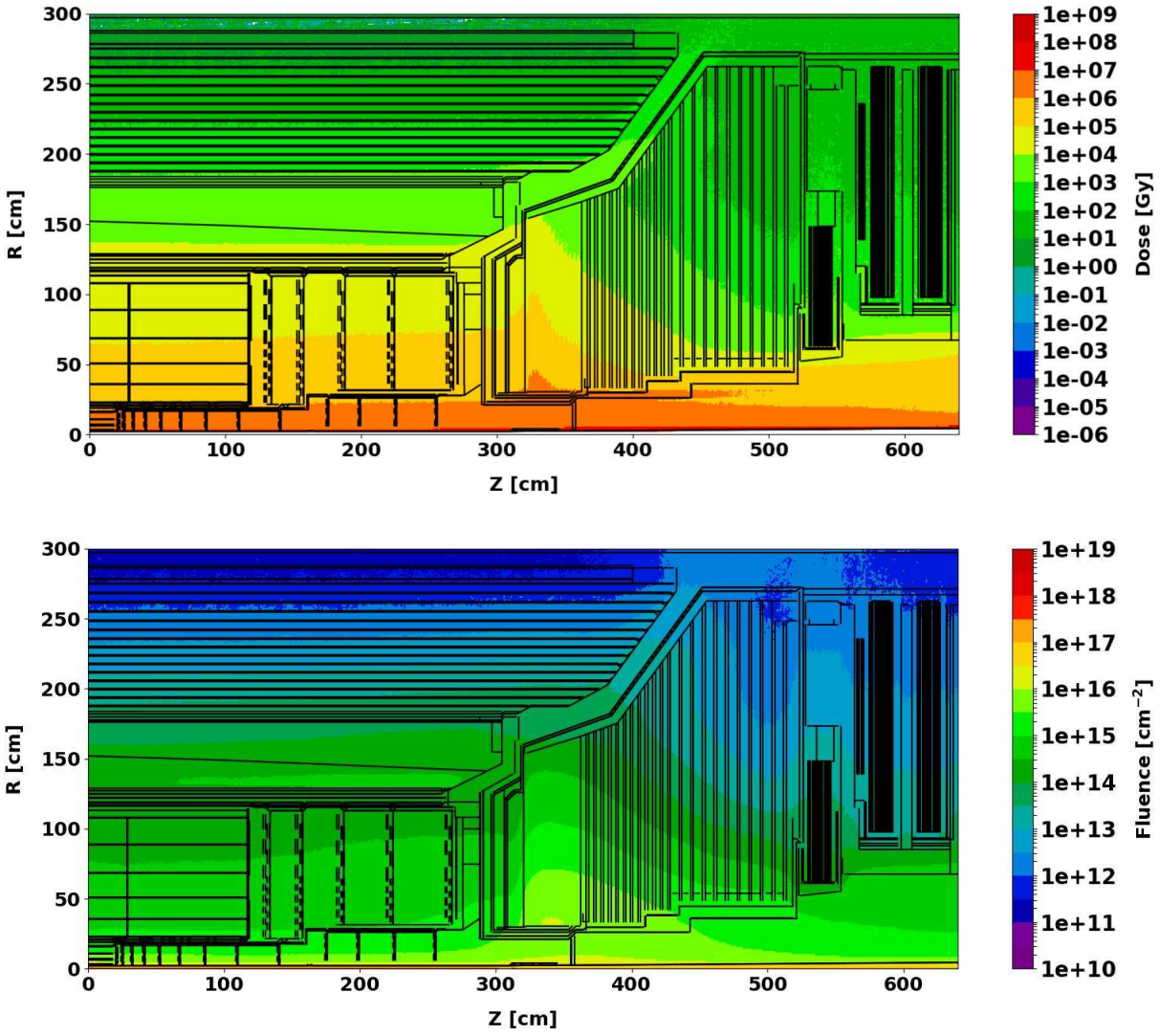

Figure 3.15: R-z projection of the distribution of the absorbed dose in Gy (top) and fluence in n_{eq}/cm^2 (bottom) for the positive endcap of the HGCAL and half the tracker, after a $4\,ab^{-1}$ exposure. Produced with the BRIL "Simulation Plotting Tool" [303] with CMS `FLUKA` geometry, version 3.7.0.0 [304, 305].

Parameter	Value
η coverage	$1.5 \lesssim \eta \lesssim 3.0$
Total area of Si sensors	620 m^2
Total area of Sci tiles	400 m^2
Endcap radial length	2.3 m
Endcap longitudinal length	2 m
Endcap weight	215 t
Temperature	−35 °C
Number of modules	30 000
Number of Si channels	6 000 000
Number of plastic tile boards	4000
Number of CE-E layers	26
Number of CE-H (hybrid) layers	21 (14)

Table 3.2: Summary of some of the properties of HGCAL.

the energy resolution is measured to be $22.1\%/\sqrt{\text{GeV}}$, with a constant term of 0.6% [307]. The term's meaning has been introduced in Section 2.2.2, when describing the electromagnetic calorimeter. The energy response is shown to be linear within 2.5%. At the highest energies, the angular resolution was measured to be 4.5 mrad, and the lateral position resolution to be 0.3 mm [307]. For charged pions, instead, the stochastic term is measured to be $\sim 130\%/\sqrt{\text{GeV}}$ [308], but improvements of a factor of 2 were shown to be achievable by employing machine learning techniques [309]. The timing resolution for positron showers was measured in 2018 to be ∼20 ps for the highest beam energies [310]. Single channels in the SKIROC2-CMS chip [311] were shown to achieve ∼60 ps, compatible with the electronics specifications. Overall, the results point to a successful operation of the HGCAL at the HL-LHC.

3.2.4 Detector Upgrades in Other LHC Experiments

For completeness, we here provide a brief overview of planned updates for other LHC experiments, which will also have to handle the future high-radiation conditions:

- **ATLAS**: In preparation for the increased LHC luminosities, ATLAS will feature a completely new IT, extending its $|\eta|$ coverage to 4, as well as the new High Granularity Timing Detector (HGTD). The TRT will be removed. The IT will include 10 times more strips and 60 times more pixels, and its innermost layer will feature 3D Si sensors, due to its superior radiation hardness [312]. The HGTD will reach resolutions of ∼30 ps, similar to the future CMS MTD. The trigger system will be fully refurbished, operating at ∼1 MHz, and making decisions within ∼10 µs. The FE and BE of the calorimeters will be replaced to increase the granularity and handle the additional radiation [291]. The FE will also be updated for the muon systems, which will be included for the first time in the ATLAS trigger decision. A new RPC layer will increase the muon trigger acceptance. Online luminosity measurements will be provided by the HGTD and by the upgraded Luminosity

Cherenkov Integrating Detector (LUCID). All these upgrades should be available from the start of the HL-LHC [313].

- **LHCb**: Following the extremely significant Run 3 detector revamp, LHCb is preparing a further upgrade for the HL-LHC phase [150, 314]. The current Shashlik EM calorimeter is being replaced by a new "spaghetti calorimeter" with scintillating fibers and with high-performance modules and extra granularity, enabling timing measurements of ∼20 ps. The RICHs are also being upgraded with very fast electronics for single Cherenkov photons. The upgrades have already been successfully tested with SPS beams.

- **ALICE**: Recent Run 3 upgrades included modifications to the Time Projection Chambers, IT, Muon Forward Tracker (MFT) and to the trigger. For Run 4, an additional update to the IT is foreseen, replacing its three innermost layers by ultra-thin layers placed closer to the IP [315], together with the addition of new EM and HAD Forward Calorimeters (FoCals) [316]. The upgrades should further constrain gluon nuclear PDFs, besides providing increased resolutions and efficiencies. Additional upgrades are planned for Run 5, after 2032. The exciting "near-massless" iris-shaped vertex tracker [317] will sit at an inner radius of 5 mm only, which requires a special retractable strategy to avoid damage from the unfocused LHC beam! The vertices will be complemented by dedicated PID TOF detectors and an additional RICH detector extending the charged high-p_{T} PID. Finally, muon detectors and the EM calorimeter will also be updated.

3.3 The Reconstruction of Trigger Primitives in HGCAL

The importance of the L1 and High-level CMS trigger systems was already highlighted in Section 2.3. The HGCAL will be integrated with the online firmware trigger system put in place by CMS, the L1 [198], which precedes the HLT running on commercial servers. In this Section we detail the hardware and software infrastructure powering the HGCAL DAQ and L1 TP generation (TPG) systems. The TPG produces TPs which are physical quantities built to encapsulate the most discriminative information in an event. TPs are used as building blocks for L1 decisions. In HGCAL, TPs consist of (η, ϕ) towers and cluster-related variables, such as energy, positions and shapes. The TPG includes all steps from data collection in the FE chips at 40 MHz to the production of TPs in the BE electronics [295]. The splitting of resources between *on-detector* and *off-detector* electronics is related to several factors, including the flexibility provided by FPGAs or other programmable boards, the cost of boards and connectors, and the high radiation levels present close to the detector [202].

The TPG must fit within a ∼5 µs latency, or 200 BXs, taken from the total 12.5 µs latency allocated to the L1 [198], and is constrained by the available bandwidth between the FE and the BE. The bandwidth is limited by several factors, including the total power consumption, the multiplicity of routing lines, the space available for electronic boards in the underground services cavern (USC), and the overall cost of hardware, among which expensive FPGAs and optical links [318]. A large number of throughput reduction techniques is thus applied as often and as soon as possible, exploiting pipelined algorithms whenever feasible. The flow of TPs can be schematically visualized in Fig. 3.16. HGCAL

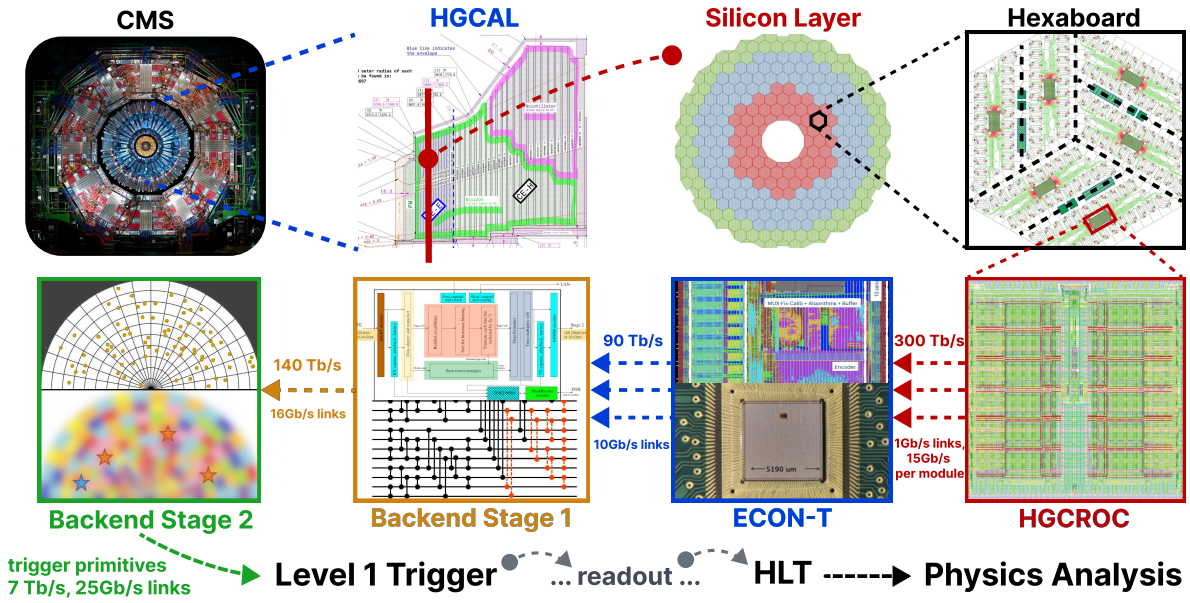

Figure 3.16: Simplified schematic of the TP dataflow in HGCAL, starting (ending) in the top left (bottom left) corner. The diagram follows the TP processing in a Si layer through the FE and BE, and up to the L1, including expected approximate bandwidths. Trigger decisions at L1 will impact the HLT and, consequently, physics analysis. Adapted from [8].

is the second subdetector in CMS with the largest fraction of allocated resources at L1, due to its high granularity.

The most granular detector elements the TPG is concerned with are TCs. As detailed in Section 3.3.2, for the Si section, TCs are arranged in a "three-fold diamond" configuration where groups of 4 or 9 Si cells form one TC. This configuration was chosen because it allows the convenient definition of uniform groups of neighboring cells to form TPs, as shown in Fig. 3.17. It also simplifies the layout of the Si module readout PCB, given the subdivision of the module into symmetric domains [295]. In the Sci section, TCs are defined as groups of 2×2 tiles, except close to boundaries, where some TCs containing less tiles exist [319]. The number of TCs and electronic links (elinks) associated to the Sci section increases with the detector's depth, reflecting the overall distribution of Si to Sci regions with increasing layer number (recall Fig. 3.11).

TPs clearly are the basis for most of the physics analyses that can be done at CMS, both now and in the future. They stem from a complex interplay of advanced software and hardware approaches, with extremely tight bandwidth and latency budgets, forming the seed for all downstream decisions. My Thesis work is deeply connected to the definition and processing of HGCAL TPs, and will hopefully grant others the basis for important extensions, with the final goal of providing robust information to enhance the sensitivity to SM and BSM physics.

3.3.1 The Infrastructure of the Backend Trigger and Data Acquisition Systems

The HGCAL BE electronics consist of the DAQ, TPG, detector control system (DCS) and detector safety system (DSS). The first two form the so-called trigger and data acquisition (TDAQ). The BE have been implemented on the Advanced Telecommuni-

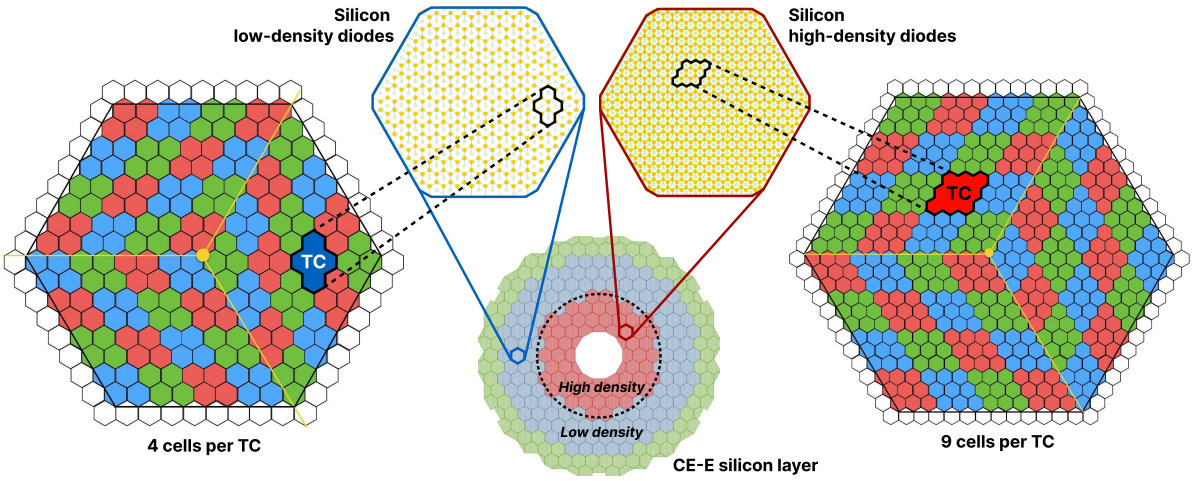

Figure 3.17: Illustration of the three-fold diamond configuration of an hexagonal 8″ module, used to associate single Si cells to TCs. Low density modules (left) associate four sensors to each trigger cell, while high density modules (right) create TCs with nine channels each. All modules have exactly 48 TCs, effectively removing a layer of complexity when processing TCs. The actual physical dimensions of the TCs vary given the boundaries of the hexagonal modules.

cations Computing Architecture (ATCA) infrastructure [320], also used for other CMS subdetectors [275]. ATCA boards will be placed in ATCA crates, which can house up to 14 boards. The DAQ system includes two parts, where the most prominent features the configuration, control and data acquisition system. The second, smaller part holds the endcap muon trigger finder (EMF) interface, which purpose is to propagate non zero-suppression (NZS) HGCAL "regions of interest", as defined by extrapolated muon tracks in the muon detectors, aiming at reconstructing MIP peaks for calibration without applying thresholds. The larger part of the DAQ consists of 6 identical hardware copies, following the three-fold polar angle symmetry of the two endcaps. Each 120° sector hosts 16 `Serenity S` DAQ boards [321] in two crates. The interface to the smaller EMF system has one `Serenity S` board per endcap, called "NZS board". Besides being responsible for data acquisition, the DAQ system has other responsibilities. Being the entry point to the FE electronics, the DAQ is also equipped with synchronous fast control, asynchronous slow control, and transporting the timing information from CMS to the FE. Concerning the HGCAL TPG system, its role is to build objects that are used as primitives by the central L1 trigger. The TPG system receives FE data for every BX at 40 MHz, and creates trigger "primitives" which are passed to the CMS central Level-1 Trigger (L1T) system to use in making the L1A trigger decisions. The DAQ and TPG systems receive data from the FE electronics through separate links, being independent systems as far as the BE electronics are concerned.

The low-power gigabit transceiver (lpGBT) application-specific integrated circuit (ASIC) [322] is the radiation-hard link driver standard to be used with HL-LHC FE electronics. It provides a two-way connection between the BE and FE, including clock distribution, control and configuration signals, and transmission of the information collected by the detector. The data is supplied to the lpGBT inputs via up to 7 elinks, each providing 32 b words at 40 MHz, or 1.28 Gb s^{-1}. This corresponds to 8.96 Gb s^{-1} data rates between the FE and the BE. In HGCAL, both the DAQ and the TPG communicate with the FE via elinks, but only the DAQ provides control and configuration. The data is then forwarded

to the BE electronics, located at a distance of ∼100 m from HGCAL. The versatile link transceiver (VTRx+) [323] module provides an optical communication interface between the lpGBTs and the BE computing boards. Additionally, CERN has developed a custom radiation resistant version of the easy-to-install `Firefly` optical transceivers [324, 325], optimized for use in the BE electronics to communicate with the VTRx+ modules.

The time and control distribution system (TCDS) interface is provided by one specialized ATCA board per crate, named DAQ and TCDS Hub (DTH400). The DTH400 also provides an interface to the Central DAQ (cDAQ), outputting ∼400 Gb s^{-1} via Ethernet data-to-surface (D2S) links. The data is then transported to the surface into the memory of commercial servers. The DTH400 boards receive the CMS-wide clock and control signals and distribute it to the remaining boards in the crate. The communication with the TCDS is two-way: this notably gives the ability to request a BE resynchronization or reset, and to throttle the entire L1 system. Communication between the BE and the cDAQ is established via 25 Gb s^{-1} fiber optics running the CMS-wide `SLinkRocket` protocol, which ensures a strict specification based on 128 b words [279]. The BE electronics will be located in the CMS USC, where space is extremely limited and has to be distributed across all subdetectors. The HGCAL TDAQ BE has been assigned 8 racks, split into 3 and 5 to DAQ and TPG, respectively. Each rack fits two ATCA crates, and sufficient power is provided to the ATCA boards.

3.3.2 Frontend Electronics

The HGCAL L1 reconstruction chain, including the TPG reconstruction, starts at the location where data is collected, namely the Si cells and Sci tiles described in Section 3.2.3. From raw energy deposits to the creation of TPs, a complex chain of electronic components and data reduction and selection algorithms is in place. The architecture surrounding the *on-detector* steps, i.e., the steps taking place very close to where the raw data is collected, constitutes the FE electronics. The entry points of the reconstruction chain are the custom chips located on the hexaboards or tileboards, depending on the detector region. They are called HGCAL read-out chips (HGCROCs) [326, 327], and are HGCAL-specific ASICs which collect, amplify and filter the produced ionization or scintillation charged currents at ∼300 Tb s^{-1} [295]. The layout of a HGCROC chip is shown in Fig. 3.18. In addition to a standard 10 b Analog-to-digital converter (ADC) charge measuring mode, the HGCROC switches to a time-over-threshold (ToT) mode as soon as a threshold on the deposited charge is reached, of the order of the preamplifier saturation threshold of ∼100 fC. The time during which the preamplifier is saturated serves as a proxy for the amount of deposited charge. During the saturation period, which can reach up to ∼200 ns, the channel is blind to new charge deposits. Once the saturation is over, the time is digitized with a 12 b time-to-digital converter (TDC). Beyond the data paths, the chip includes a phase-locked loop (PLL), which generates the clocks needed to operate the chip. An inter-integrated circuit (I2C) interface is also present, enabling the modification of all static parameters of the chip, which are triplicated to prevent SEEs, which are stochastic, localized and non-cumulative effects disrupting the chip's functioning.

Due to the similarity of the algorithms and electronics of the Si and Sci detector regions, and also taking into account differences in their development stage, we focus on the Si technology to simplify the overall description. The HGCROC trigger path aggregates the data into TCs by summing their energies, in what constitutes the first of many data reduction algorithms in the TPG. TCs are defined as energy sums of neighboring

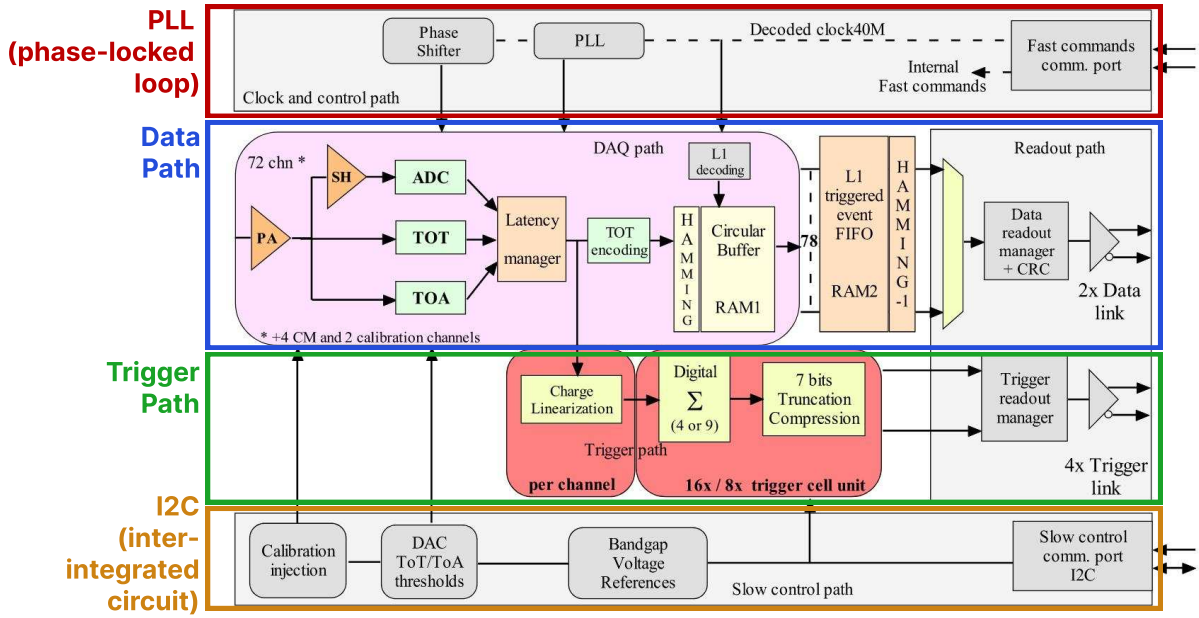

Figure 3.18: Block diagram of the HGCROC [326, 327]. It is composed of two data paths: the DAQ path (in blue), connected to the DAQ concentrator chip (ECON-D), and the trigger path (in green), connected to the TPG concentrator chip (ECON-T). It also includes a PLL, which generates the clocks needed to operate the chip, and an I2C interface, which enables the modification of all static parameters of the chip. Taken from [8].

sensor cells, and represent a simple method to reduce the prohibitive data throughput. They group 4 or 9 channels, depending on the Si module granularity. In the CE-E, only odd-numbered layers are considered for TC reduction. The summation of single channels requires all inputs to use the same energy scale, and this is not the case due to the two charge digitization modes being used, the ADC and the ToT. The produced digitized values are thus rescaled, but the rescaling procedure needs to take into account the strong non-linearity of the ToT response for medium charge values, close to the ADC regime. An approximate approach is employed to avoid an extremely demanding linearization procedure at 40 MHz. Besides the reduction in granularity, TCs also decrease the algorithms' complexity, in the sense that all modules have exactly 48 TCs, and thus HGCAL L1 algorithms can ignore differences arising from low- and high-granularities. After building TCs, the charge values to be sent to the BE are compressed by a factor of ∼3 using a floating point encoding. The compression exploits the fact that a high resolution is generally not required at L1 for particles lying well above the energy thresholds. In parallel, the full-granularity data is kept in circular buffers and is sent out via 1.28 Gb s^{-1} elinks as soon as a L1A signal arrives. Despite the chip's ability to also measure the Time of Arrival (ToA) of the charged pulses, timing information cannot be exploited in the trigger path due to bandwidth constraints.

The TPG reconstruction chain continues via the ECON-T chip, which is located very close to the HGCROCs, in the so-called "concentrator mezzanine", next to the hexaboard. The chip concentrates, selects and/or aggregates TCs within a single module, yielding one data packet per BX. Each module has either 3 or 6 HGCROCs, depending on the concerned granularity density region. The chip first calibrates the input charges coming from the HGCROC, converting them into transverse energy values [328]. It then builds *module sums*, where the energies of TCs in a module are summed without any energy

threshold being applied. The ECON-T can operate in a number of modes, of which we mention the ones most likely to be used during data-taking, also illustrated in Fig. 3.19:

- **Threshold algorithm**: Selects all TCs with an energy above a given threshold, subject to bandwidth limits. The size of the output varies event-by-event, and for different modules in the same event.
- **Best-Choice algorithm**: Selects a fixed number of TCs with the highest energy. The size of the output is fixed and thus known in advance. Requires sorting, which is implemented via batcher odd-even sorting networks [329–331]. No truncation is applied before the sorting, enabling the configuration of the selected number of TCs.
- **Super Trigger Cell algorithm**: Reduces the data granularity by summing nearby TCs. At the same time, information on the energy distribution within a super trigger cell (STC) is kept, by propagating the TC with the maximum energy. STCs will most likely be composed of 2×2 TCs in the scintillator, and will not be considered for the Si section.

The current plan envisions the usage of the Best-Choice (BC) algorithm for the CE-E and the STC algorithm for the CE-H. This combination is preferred over the threshold algorithm due to the fixed output data size, which leads to a simpler, buffer-less BE data unpacking. Several studies covered different algorithm choices, including using just one for the entire detector, or other combinations. In spite of the granularity reduction put forward by the STC algorithm, its usage is required where the available optical links are not sufficient to transmit all required information to the BE. It was found that, given the existing event-to-event rate inhomogeneities, the BC algorithm, given the number of TCs, occasionally misses an important fraction of the event [332, 333]. The effect was particularly visible for hadronic jets, where serious cost and space constraints can impose limits on the fiber optics, and thus on the number of TCs the algorithm can keep. On the other hand, the usage of STCs across the 47 layers leads to an unacceptable decrease in EM resolution.

Another flavour of concentrator chips gathers the DAQ data: the ECON-D, again one per module. The ECON-D can optionally apply zero suppression, where only channels with an energy above a certain threshold are kept, and then merges all of HGCAL's data into a single packet. One of the major challenges of the FE is the ability to deal with extremely inhomogeneous data rates across HGCAL, which may occasionally vary by almost two orders of magnitude. The ECON-D thus relies on a buffering system which supports variations in the size of the packets and in the L1A rate. Despite ensuring one sent package per L1A signal, the ECON-D cannot guarantee the package's data integrity due to the buffers being full. L1 throttling might be required in some cases.

The data is finally then sent via $1.28\,\mathrm{Gb\,s^{-1}}$ e-links to lpGBT ASICs [322] located in the FE motherboards, or engines. Each motherboard is connected with up to 6 ECON-Ts and ECON-Ds. The ASICs serialize the concentrator chip (ECON) data to $10.24\,\mathrm{Gb\,s^{-1}}$, and send it to the VTRx+ interface, which in turn distributes it to the off-detector BE via fiber optics. In total, $\sim 90\,\mathrm{Tb\,s^{-1}}$ are transferred to the BE [334].

3.3.3 Backend Electronics

The BE, located at $\sim 100\,\mathrm{m}$ from the detector, receives FE data with the goal of building cluster-shape variables within a $\sim 2.5\,\mu\mathrm{s}$ latency budget. Clusters, together with simpler

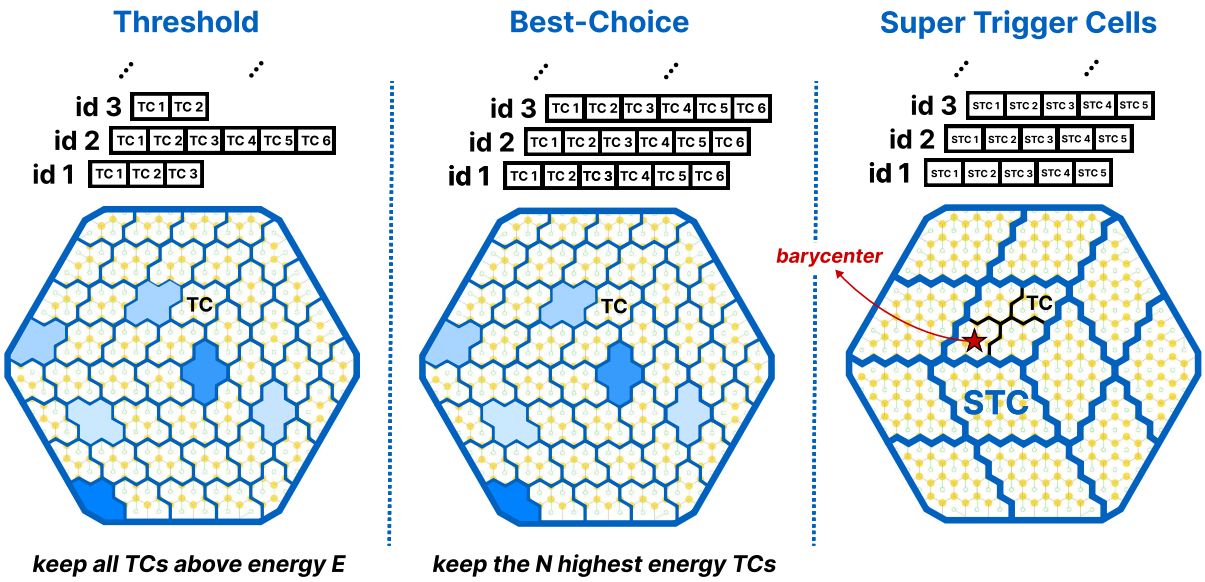

Figure 3.19: Schematic illustration of three data reduction algorithms currently implemented in the ECON-T chip. We show low-density modules, but the algorithms are identical for high-density regimes. For displaying purposes, we are assuming the maximum supported bandwidth translates to 5 STCs and 6 TCs per BX, where "id" refers to a different block of data being sent to the BE, coming from a different module or from a different event in the same module. The threshold algorithm requires a variable data size format. The STC visualization represents the scenario where each STC corresponds to 4 TCs, or 16 Si cells in a low-density module.

TTs built out of STCs and of module sums along the longitudinal direction, amount to the final HGCAL TPs to be transmitted to L1. The BE layout is split in two processing stages, called Stage 1 (S1) and Stage 2 (S2), which run on `Serenity` boards [321] with 128-link `Xilinx VU13P` FPGAs. The first stage is required to assemble data coming from multiple detector locations into a single board, and thus provide a large enough phase-space to better reconstruct clusters. Indeed, each FE optical link sends data belonging to a few modules only, which get translated into a mere 2 % of the detector per S1 FPGA. A second stage can then gather the data corresponding to a larger fraction of HGCAL to robustly build TPs. Additionally, the more data fits into a single FPGA, the less data duplication is required to handle boundaries, especially when taking into account that different BE FPGAs do not communicate with each other. The current design allots 120° of HGCAL to each S2 board, with a TMT period and hence a board multiplicity of 18, effectively representing 6 identical subsystems.

The S1 thus receives ECON-T data from multiple modules, but from a single BX, into 14 FPGAs per 120° sector, where the number of boards is driven by the existing optical link multiplicity. The data is unpacked and an energy rescaling is applied, to correct for the different encodings used by the BC and STC algorithms. The TCs are routed into projective ϕ bins. In parallel, module sums and STCs are summed into partial (η, ϕ) TTs, being formed separately for the CE-E and CE-H. Finally, the data is sent to S2 with a $\sim$140 Tb s^{-1} throughput after time-multiplexing it with a 18 BX period [328]. The TPG BE architectural layout is illustrated in Fig. 3.20, from the FE inputs to sending TPs to the central L1 system.

Before describing S2, we briefly dwell on the unintuitive fact that the data throughput between S1 and S2 is actually larger than the one between the ECON-T and S1. Where

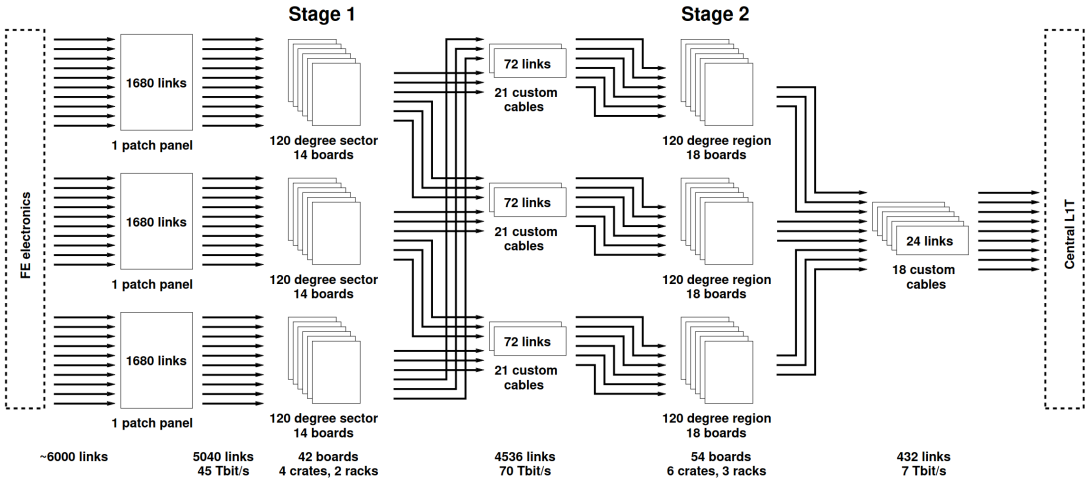

Figure 3.20: Layout of Stage 1 and Stage 2 boards for one HGCAL endcap. The 120° symmetry is used to process the data in terms of three identical and independent firmware regions. The full TPG system consists of two identical and independent copies of this layout. Taken from [328].

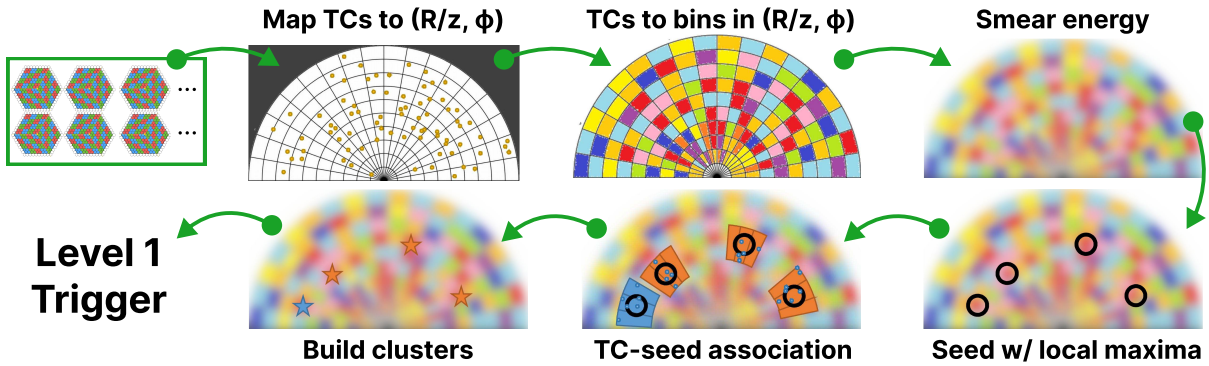

Figure 3.21: Schematic flowchart of S2's reconstruction chain. TCs from S1 are unpacked and processed in a pipelined fashion up to the creation of cluster-related variables, which are fed to the L1. The description of the steps can be found in the text, where "histogramming" refers to the first two steps in this figure. Adapted from [8].

does the additional data come from? The answer is two-fold. Firstly, one needs to account for the data duplication required to handle boundaries between 120° sectors, which is nicely illustrated in Fig. 3.20. Secondly, the data has to be inflated since:

- the memory addresses have to be encoded on a larger number of bits, because the S1 covers larger detector regions than the ECON-T;
- the energies have to be encoded on a larger number of bits to absorb different energy scales in the different detector regions used by the ECON-T;
- more bandwidth has always to be allocated to TC bins in order to absorb fluctuations and limit truncation effects in the S1. This is also true for the fixed-size BC algorithm, since it provides a fixed number of TCs per module, not per bin.

The S2 is designed to perform the main TPG reconstruction work: building clusters and TTs. Partial tower energies are accumulated into (η, ϕ) bins and clusters are built following the steps highlighted in Fig. 3.21:

- **Histogramming**: TCs are mapped to a projective (ϕ, R/z) space with (216, 42) bins, where $\mathrm{R} = (x^2+y^2)^{1/2}$ and $\tan(\theta) = \mathrm{R/z}$ (see Section 2.2.1). These coordinates are chosen since a constant R/z corresponds to a constant particle angle θ, where R is defined in the plane perpendicular to the LHC beamline. The coordinates are "projective", since 3D deposits are mapped to a 2D space. Energy deposits of neutral particles originating from the center of the detector and spanning several layers will thus lie in a single R/z bin. The binning further reduces the spatial granularity and, due to its grid-like structure, facilitates vectorized and hence parallel processing in the firmware. Each bin contains the energy sum of all its TCs, together with their $\mathrm{MIP_T}$[3]-weighted x/z and y/z positions, where $\mathrm{MIP_T}$ is defined as $\mathrm{MIP}/\cos(\theta)$, with one MIP being the energy deposited by a minimum ionizing particle [13, §34.2.3], and θ the polar angle introduced in Section 2.2.1. The weighted positions are defined as follows:

$$\left.\frac{x}{z}\right|_{\text{weighted}} = \sum_i^{\mathrm{N_{TC}}} \frac{\mathrm{MIP_T}^i\, x^i}{z^i} \quad , \quad \left.\frac{y}{z}\right|_{\text{weighted}} = \sum_i^{\mathrm{N_{TC}}} \frac{\mathrm{MIP_T}^i\, y^i}{z^i} \quad ; \tag{3.1}$$

- **Smearing**: An energy smearing step is applied to the (ϕ, R/z) bins to decrease overall variations in their energy distribution. This is meant to address biases discussed in Section 3.5. A convolutional kernel is iteratively slid along both directions, independently. For each bin, the energy of all its neigh ours covered by the finitely-sized kernel is multiplied by the corresponding kernel weight, and the energy is updated. The kernels are shown in Eq. (3.2), along ϕ (left) and R/z (right):

$$\left[\ \ldots\ \ \tfrac{1}{16}\ \ \tfrac{1}{8}\ \ \tfrac{1}{4}\ \ \tfrac{1}{2}\ \ 1\ \ \tfrac{1}{2}\ \ \tfrac{1}{4}\ \ \tfrac{1}{8}\ \ \tfrac{1}{16}\ \ \ldots\ \right] \qquad \begin{bmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix} \tag{3.2}$$

  Variations are more prominent along ϕ since the binning is finer. The length of the kernel along ϕ is R/z-dependent, as illustrated by the dots in Eq. (3.2) (left). The ϕ kernel collects the energy from more bins for lower R/z rows. The energy of each bin is normalized to its measured energy, in order to ensure that no energy is artificially added to the event.

- **Seeding**: Seeds are local $\mathrm{MIP_T}$ maxima in the histogram, and are so called since they indicate the starting TC for clustering algorithms to gather other TCs. Seeds are found via a seeding window which, for each bin, spans its immediately adjacent bins and checks whether their $\mathrm{MIP_T}$ energy is lower than the central bin. If so, and if the energy from the central bin lies above a threshold, the bin is promoted to a seed. The threshold cut limits the collection of clusters from pure noise. We define the *window size* to be k based on the number of k^{th}-order neighbors considered by the seeding window. A size of 1 will consider the central bin plus its 8 closest neighbors, a size of 2 will consider $16 + 8 + 1 = 27$ bins, and so on. The window size can also

[3]The algorithms and data flow of the trigger reconstruction chain are in constant evolution. After the work related in this Chapter had been finalized, it was decided to perform the calibration to energy values (GeV) directly in the ECON-T. This means that energy is now measured in transverse energy units, not in $\mathrm{MIP_T}$ units.

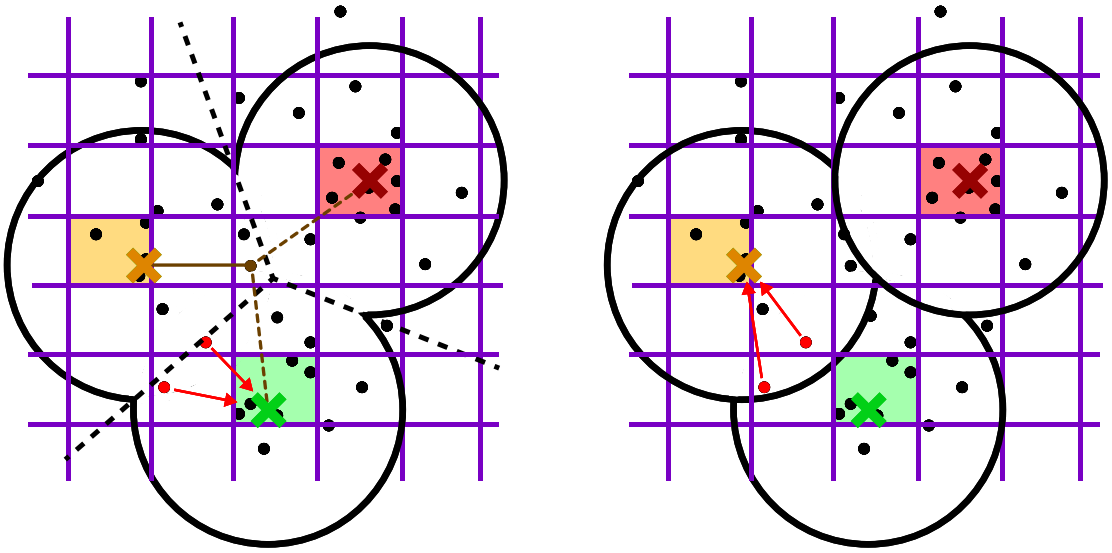

Figure 3.22: Illustration of the two clustering algorithms considered in the HGCAL TPG. The grid represents the (ϕ, R/z) bins. Black dots represent TCs. The crosses refer to the position of the seeds, ordered by color from the highest to the lowest energy: red, yellow and green. The respective colored bins corresponds to the bins where the seeds are located. The black circles represent the region of influence of a particular seed on its neighboring TCs. The two TCs in red are associated to different seeds depending on the used algorithm. The matching radii can be different for different seeds. TCs outside the three circles are not associated to any seed. (Left) The `min_distance` algorithm associates TCs based on distance. The black dashed lines represent the border between the regions where a particular seed gathers all TCs. The brown TC serves as an example: the distance to the three seeds is shown with brown lines, where the solid line shows the closest seed to that TC. (Right) The `max_energy` algorithm prioritizes instead the association based on the energy of the seeds.

be varied differently along the two directions. The default S2 reconstruction uses $k = 1$.

- **Clustering**: TCs are associated to seeds and used to calculate cluster properties. Every seed leads to exactly one cluster. Contrary to previous steps, which run on a (ϕ, R/z) space, the clustering uses the (x/z, y/z) projective space. Two different clustering algorithms are currently defined in the TPG, and illustrated in Fig. 3.22. A distance matching threshold is applied to both algorithms to ensure no TC is associated to extremely distant seeds. The distance is calculated in the same projective space. The default matching radius slowly increases with the detector's depth, from 0.015 in the first layer to 0.050 in the last CE-H layers. The first and default `min_distance` algorithm associates TCs to their closest seed, based on the 2D distance in the projective space. The second algorithm, called `max_energy`, prioritizes an association based on the seed energy, where the highest energy seed is associated to all TCs within its matching radius, the second-highest energy seed is associated to the remaining TCs within its (different) matching radius, and so forth. If no TC is left for the lowest-energy seeds, then no cluster is formed.

Once the clusters are defined, cluster-shape variables can be computed. The full list of variables is not yet defined, but they will surely include the barycenter's position and energy of the clusters. Additionally, two separate HAD and EM energy interpretations

will be defined, with possibly different parameters, such as radii or energy thresholds. We refer to "interpretations" since in the TPG no particle identification is performed.

3.4 Development of a Simplified HGCAL Geometry and Event Processing Chain

In Section 3.3 we introduced the processing steps taking place in S2. They all run a series of algorithms which are not yet finalized, being therefore subject to constant updates to improve physics performance and reduce latency. The algorithms depend on a large number of potentially tunable parameters. Several alternative algorithms might also exist for the same processing step. A mechanism is thus needed to easily compare the performance of different algorithms in the same processing chain. The mechanism should also provide an event-by-event comparison for debugging purposes. Moreover, the S2 processing chain is itself not yet finalized, and some processing steps might be added or removed according to evolving physics and computing needs. It is clear that a flexible, fast and easy-to-use framework is needed to perform all required studies and tests.

A `C++` emulation of the entire firmware processing chain is available in CMSSW. Despite not being a bit-level emulator, it reasonably replicates the future firmware processing chain, and has access to the official CMSSW geometry, including HGCAL, representing the still-evolving outcome of several years of software development within CMS. However, it would be very useful to also have access to a framework enabling the fast exploration of new algorithms and the optimization of current ones, with quick prototyping and parameter optimization. On top of that, doing it in `Python` would help newcomers, as it tends to be better known by Physics students than `C++`, and implies a smoother learning curve. To address these points, I have implemented the HGCAL S2 processing chain in `Python v3` [335] from scratch. The implementation not only ports the existing algorithms, but also rewrites potential `Python` bottlenecks, usually `C++ for` loops, using `numpy` for `C`-level performances [336]. It also exploits a modern suite of standard and HEP-focused packages, such as `uproot` [337] and `RDataFrame` [338], among others. The framework, called `bye_splits` for reasons detailed in Section 3.5, enables faster prototyping, testing and optimization. It also powers the studies presented in Section 3.5. The framework is currently being used by multiple people at Leprince Ringuet Laboratory (LLR), serving as a foundational basis for a significant portion of their work. The code has been presented at the international PyHEP 2023 Workshop [9], including a live demonstration in a web application. The application was deployed using CERN's platform-as-a-service (PaaS), which is based on `OKD4` (`OpenShift4`), and conveniently enables a project to be directly built from a `git` repository via the source-to-image (S2I) service. The PaaS also provides access to CERN's network file-system, where geometry and event data are stored. The application might become useful in years to come for sharing information, for visualization, and for tuning parameters directly from the Graphical User Interface (GUI). Indeed, since everything is implemented in `Python`, connecting the visualization to the actual data processing becomes straightforward with the definition of simple code callbacks. The callbacks are activated after a button pressing, a drop-down menu selection, or similar. The geometry and specific events are deployed via interactive `Python` visualization libraries, which I used to write 2D and 3D visualizations. Concerning 3D visualizations, I wrote a proof-of-concept that was later developed by a different student,

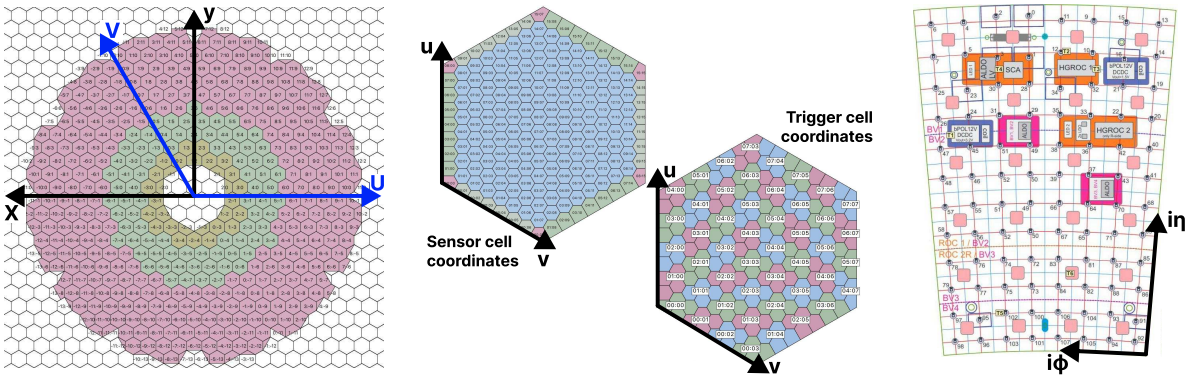

Figure 3.23: Coordinate systems in HGCAL, for the Si (left and middle) and Sci (right) sections. The Si section has two coordinate systems, one at layer level (left), and another at module level (middle). The $(u,\ v)$ coordinates in the middle figure are similar for Si and TCs cells, but the multiplicity is different. The low granularity modules are shown, but the same coordinate system is used for high granularity modules. Note that module hexagons and cell hexagons have a different orientation, flat-top and pointy-top. The Sci section uses polar-like coordinates, where each TC is uniquely identified by an index $(\mathrm{i}\phi, \mathrm{i}\eta)$ pair.

and was presented in the live demonstration mentioned above. I have also demonstrated how to potentially merge visualizations from different libraries using `Flask` [339].

The framework is conceptually split across two independent axis: geometry and event processing. Users, or clients, can develop their own algorithms transparently, i.e. without understanding the framework's internals. We start by describing HGCAL's coordinate system, which is a prerequisite for understanding the implementation of the framework.

3.4.1 HGCAL Coordinates

The HGCAL follows separate coordinates systems for the Si and Sci regions. Differences come about from the different detector configurations. The coordinates for both sections are shown in Fig. 3.23. Note that the coordinates for TCs and sensor cells are numerically different due to the different cell multiplicity, but follow exactly the same logic. Each detector element is uniquely identified by a set of two coordinates, plus the layer index where it belongs.

Silicon Coordinates

The Si region is characterized by hexagonal-shaped modules and detector cells. There are thus two sets of *axial coordinates* [340] applied at the level of one full layer, where one coordinate point corresponds to one module, and at the level of the module, where one coordinate point corresponds to a single TC (or Si cell). Each sampling layer has its own independent coordinates, for both endcaps. Given the flat-top and pointy-top relative orientations of the hexagons within their respective phase-spaces, the two coordinate systems in Si are not aligned. This further complicates *navigation* in such a coordinate system, which is by default more complex than a Cartesian system. Each TC (or cell) within a module is uniquely identified by a set of u and v indexes, and each module within a layer is uniquely identified by a set of U and V indexes. Multiple module orientations are possible, affecting the direction of the (u, v) eigenvectors. In the remainder of this Section we consider a single orientation for simplicity, as the logic is kept unchanged. Hexagonal coordinates in HGCAL are further explored in Section 3.5.4.

Scintillator Coordinates

The regions in the detector covered by scintillator material follow a simpler polar-like coordinate system, defined in terms of η and ϕ rings in the transversal plane. The indexes do not linearly map to a physical distance, given the gradual increase of the tile sizes with increased distance from the layer center. TCs comprise 2×2 groups of Sci tiles, which follow the same (adapted) coordinate system.

3.4.2 Geometry Implementation

A simplified geometry was implemented, including both Si and Sci sections. The geometry is based on the CMSSW `V11` geometry. The custom geometry is used for visualization purposes only, including visual validation, and has no impact on the implementation of the algorithms. Additionally, any final validation must be done by a bit-level emulator, or directly by the hardware system. Implementing the full detailed HGCAL geometry would thus require an effort largely exceeding the scope of the Thesis for no real benefit. The custom implementation includes only the topology of the detecting elements, i.e. Si modules and Sci tiles. Information on materials and passive components is not present. Information on low and high granularity Si modules is also not present, since the geometry concerns itself with TCs. Finer mesh details such as partial Si wafers or variations in the spacing between sensors is not included. What is included is the position and representation of every detector element across the full range of the endcaps, and its mapping to the correct geometry coordinates for later association with event information.

The Si section is implemented as a series of diamond-shaped elements representing TCs, arranged in a hexagonal fashion to mimic the Si modules. The structure replicates the three-fold configuration of Si TCs. Modules are drawn by shifting the position of a single "root" module along (U, V) HGCAL coordinates. Despite the perfect plane tesselation ensured by hexagonal coordinates, not all modules or TCs should be drawn, as shown in Fig. 3.23 (left), where the hexagons in white do not have a physical counterpart. The framework ensures nonphysical modules are skipped. The real physical dimensions of the modules are carefully taken into account, such that the positions of TCs are displayed in real data units, in cm. Due to the diamond-shaped TCs and to geometry-related shifts, the alignment of TCs relative to their true position, as provided by CMSSW, becomes ambiguous. At least two approaches exists:

1. perfectly align the TC within a hexagonal module, which leads to a worse match of each TC center relative to their true center;
2. slightly misalign module "three-folds", leading to a less appealing visualization that however better displays the true position of TCs.

The second method is used because, though less visually appealing, produces element positions closer to the real ones. This geometrical limitation can be appreciated in Fig. 3.24., where method #2 is applied.

The Sci is simpler to create, where most tile TCs are made of 2×2 cells. However, for some layers, an inner and outer η boundaries have TCs made of just 1×2 tile cells. To avoid overlaps, the size of those TCs is reduced to half using the corresponding hard-coded layer numbers. All TCs are displayed as annular wedges.

The validation of the geometry is performed by overlaying true TC positions as provided by CMSSW with custom positions, as shown in Fig. 3.24. Both module-centered

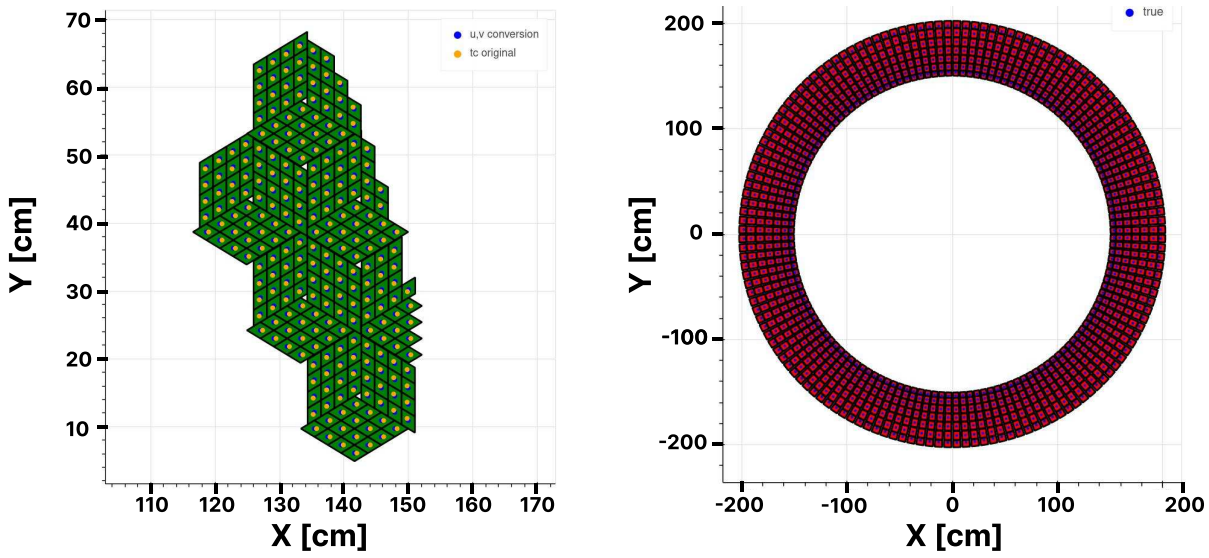

Figure 3.24: (Left) Detail of the outer boundary of a Si layer. Orange dots correspond to the original TC positions, as provided by the CMSSW geometry. Blue dots correspond to the position in the custom simplified geometry, enforcing TCs with equal sizes. The central "hole" in each module is intentionally created by shifting the "three-fold" configuration to get close matches between orange and blue dots, and thus a better visual representation of events. (Right) Entire scintillator section for one particular mixed layer, where the inner Si sensors are not represented. The true TC position is represented in blue, and placed at the center of each TC in red.

and corner-centered geometries are present in different HGCAL layers. The successful validation of the custom geometry thus requires the addition (subtraction) of a relative vertical shift to odd (even) CE-H Si layers, corresponding to approximately half the height of a module. After the correction, very good agreements are found across the entire detector: the TC positions in both the Si and Sci sections perfectly align, modulo the visual ambiguity in the Si section discussed above (and shown on the left plot of Fig. 3.24). The geometries can be compared in Fig. 3.25.

3.4.3 Event Implementation

The processing of event data starts in CMSSW, where the full reconstruction is run for different *single particle guns*, which correspond to simulations shooting single particles through HGCAL, simulating all physical interactions with the detector's material. This step produces all required TC and cluster information later required to validate the local `Python` framework. CMSSW also propagates generator-level information. In this Thesis we focus on 0 PU single-photon guns. The framework was nevertheless tested with 200 PU and found to work as expected. The large output `ROOT` files produced by CMSSW are "skimmed", i.e. only the relevant branches are kept, and some general selections are applied:

- request a photon at generator level;
- only photons that did not pair-convert before HGCAL's surface, known as *unconverted photons* (photons traverse less than $\sim$1 X_0 before reaching HGCAL [297]);
- positive endcap only for simplicity, given the endcap's mirror symmetry;

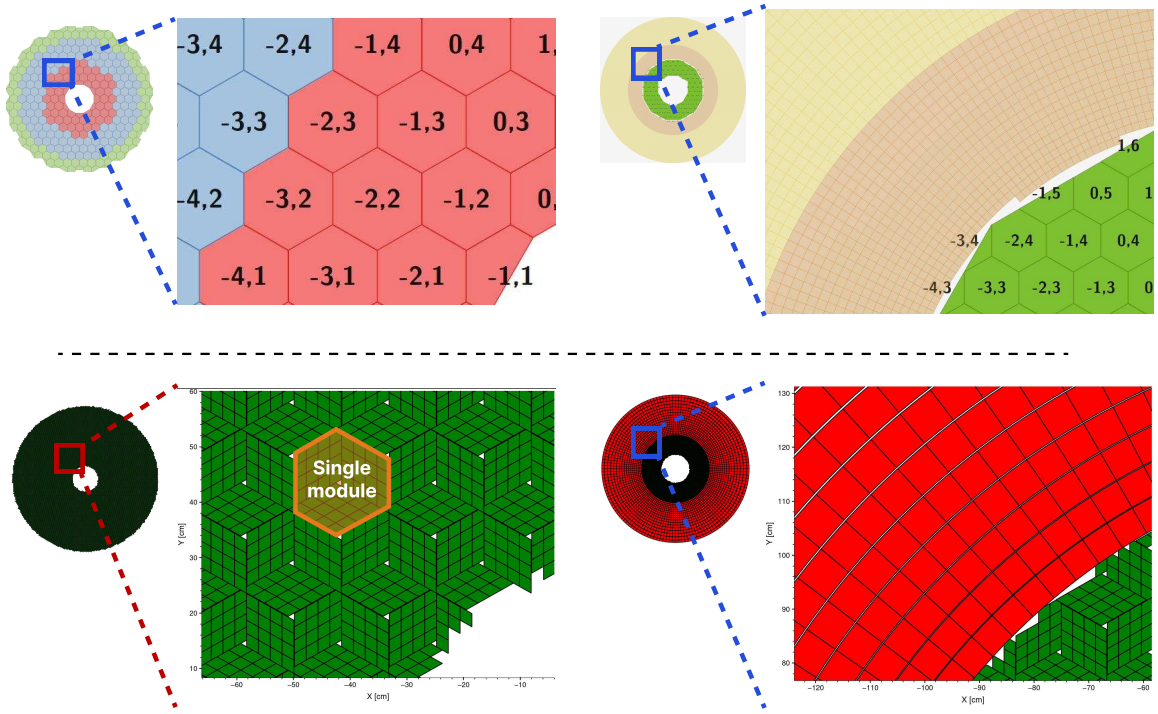

Figure 3.25: Comparison between the official CMSSW HGCAL geometry (top) and the geometry developed in this work (bottom). We show an example for CE-E (left) and CE-H (right) layers, with their zoomed-in and zoomed-out versions. Adapted from [9].

- CE-E only, as photons are not expected to leak to the CE-H;
- 0.5 MIP_T threshold to all TCs;
- cluster matching with the generated particles: $(\eta_\text{cl} - \eta_\text{gen})^2 + (\phi_\text{cl} - \phi_\text{gen})^2 < 0.05^2$

Skimmed data is then read on an event-by-event basis, is further skimmed according to user-defined variables, and is stored under the `parquet` file format [341]. From there, requested events are fed to the client-side of the framework, which might include user-defined algorithms, such as the L1 S2 algorithms, or the display applications already mentioned (see Fig. 3.26). A caching system is implemented, where `parquet` events are read and stored for later use. This avoid re-reading the same events multiple times during algorithmic processing or data display, mitigating the main latency bottleneck of the framework. The speed improvement is particularly important for 200 PU data. When absent from the `parquet` files, events are reprocessed from the skimmed data automatically. For display, event information is joined with geometry information.

In Fig. 3.27 and Fig. 3.28 we showcase 3D displays produced by the framework for particular events, for a 200 PU photon and a 0 PU hadronically decaying tau lepton, respectively. On the left, TCs belonging to the same cluster are drawn with a different color. On the right, the same event highlights color-coded energy deposits in all TCs, in MIP_T units. For the tau decay, which includes three charged and one neutral pion, we can see energy deposits in the Sci region too. Transparent TCs were not clustered. The online version of the event displays is fully interactive.

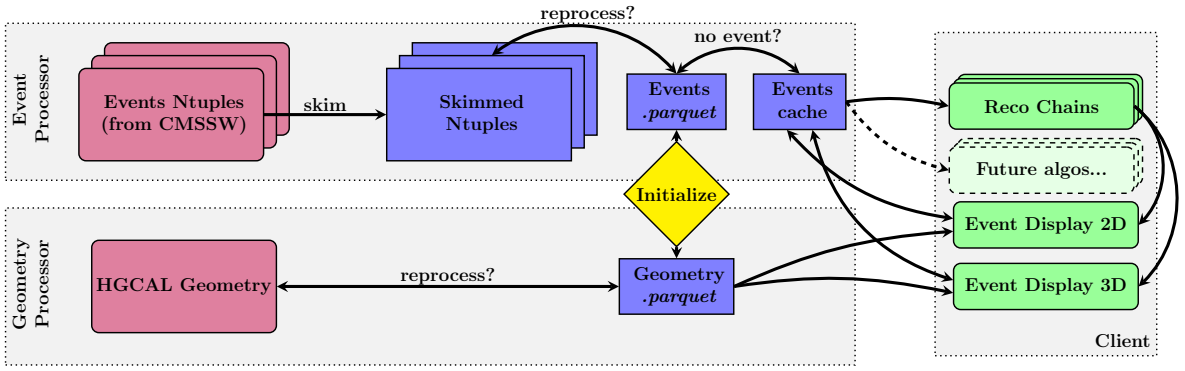

Figure 3.26: Architectural layout for the S2 reconstruction implementation in `Python`. It is roughly split in two conceptual axis: the geometry and the event processors. The two axis are used simultaneously during TPG data processing. The framework is simple and flexible enough to support the addition of algorithms in the future without much effort. A pseudo-cache mechanism is added to speed-up event processing and displaying. The user, or client, is able to use the two processors transparently. Adapted from [9].

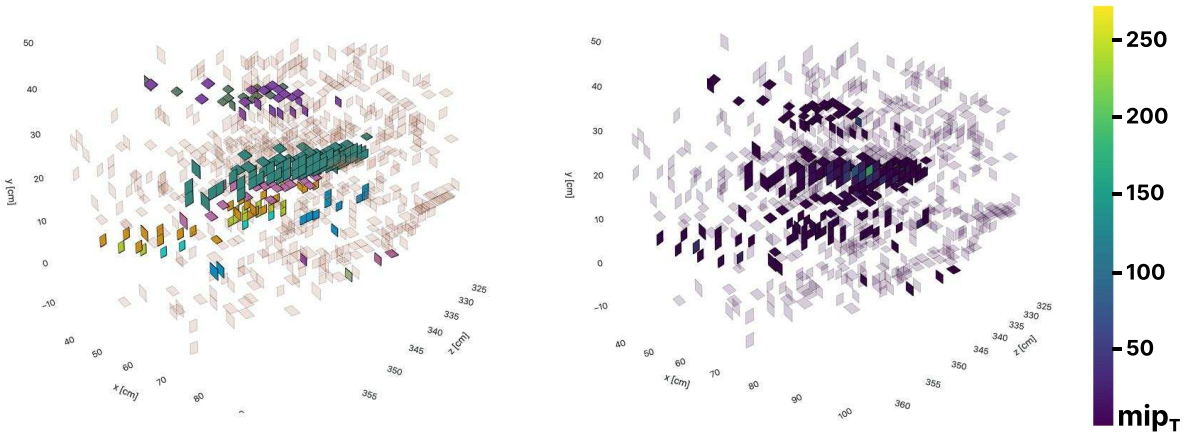

Figure 3.27: 200 PU single photon event display using this work's custom HGCAL geometry [9]. The framework supports the inspection of any event in 3D. TCs represented as transparent rectangles are not clustered. (Left) TCs associated to the same cluster have a different color. The central photon shower is clearly visible, together with some PU clusters. The default reconstruction chain was used, with the `min_dist` clustering algorithm. (Right) The same event is displayed in terms of energy deposits in MIP_T units.

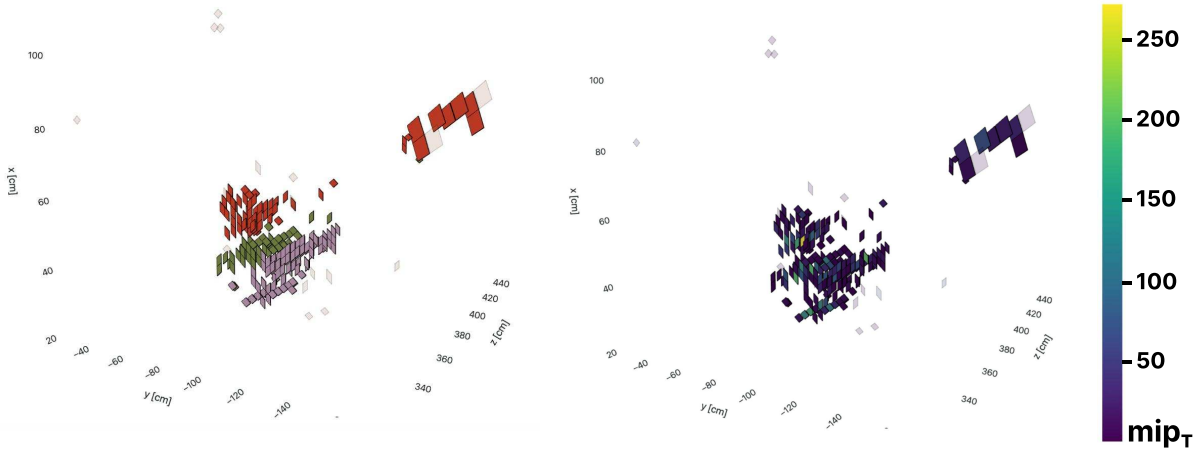

Figure 3.28: 200 PU single tau event display using this work's custom HGCAL geometry [9]. The tau particle decayed into three charged pions and one neutral pion. The framework supports the inspection of any event in 3D. TCs represented as transparent rectangles are not clustered. (Left) TCs associated to the same cluster have a different color. The neutral pion was reconstructed outside HGCAL. The default reconstruction chain was used, with the `min_dist` clustering algorithm. (Right) The same event is displayed in terms of energy deposits in MIP_T units.

3.5 Cluster Splitting

The framework described in Section 3.4 is put to use in a concrete scenario. Before the start of the Thesis here described, it had been observed that the L1 S2 chain, for unconverted photon events, occasionally identified more clusters than the number of unconverted photons originally generated in the detector. For a single photon, this corresponds to measuring two energy clusters from just one generated particle. We say that the cluster that should have been measured was "split", hence the term *cluster splits*. At the TPG level of the reconstruction, each cluster should ideally correspond to a single particle. A different behavior can degrade the detector's energy and position resolution, since:

- particles can be reconstructed with a lower energy than their true energy when only one cluster passes additional selections, since the energy distribution across the clusters can be asymmetric;
- particles can be reconstructed with a shifted position when only one cluster passes additional selections, since part of the energy deposits will not contribute to the position calculation (see Eq. (3.1));
- the multiplicity of particles in the event can be affected, when both clusters are interpreted as separate particles.

The above can negatively impact the L1 decision, and thus trigger efficiency. When studying cluster splits we require the following additional selections:

- $(E_\text{Cluster} - E_\text{Gen})/E_\text{Gen} < -0.35$, which only selects events where a split very likely happened, since on average the energy is equally split across the two clusters;
- $\eta \in [1.7, 2.8]$, to avoid unwanted reconstruction effects at HGCAL boundaries, where showers might be transversally cut.

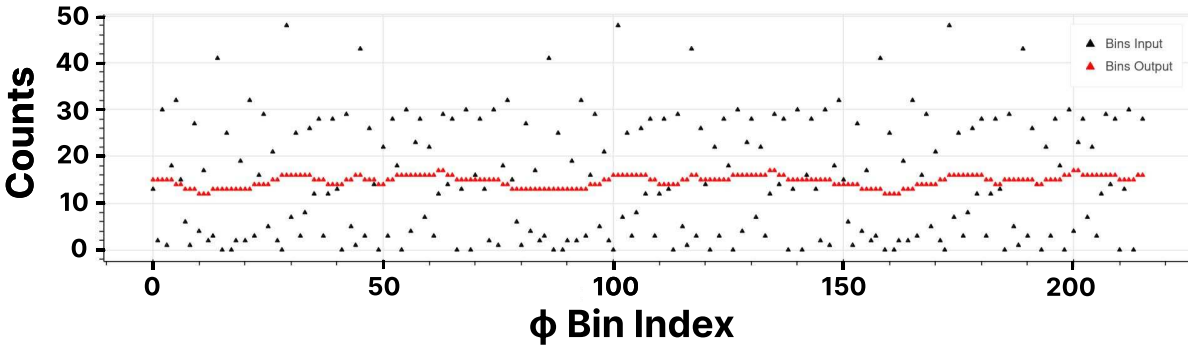

Figure 3.29: Number of TCs present as function of the ϕ bin index, for the lowest R/z bin, where differences between adjacent bins is the most extreme. Black dots show the original distribution, and red dots show the TC distribution after running the `bye-splits` algorithm with $\lambda = 0$. Related results show that the bin positions assigned to TCs, with respect to their true positions, are always less than 2 cm for all layers and R/z bins.

With these selections, we observe cluster splits in $\sim$1 % of single unconverted photon events. In the remainder of this Section we explain the causes for cluster splits, and detail the techniques developed in the context of this Thesis to mitigate or remove it.

3.5.1 The Origin of Cluster Splits

As seen in Section 3.3, the first step of the S2 consists in reducing the detector's granularity by mapping TCs into projective (ϕ, R/z) bins. The physical dimensions of the bins are strongly dependent on their location: bins closer to the beamline (low R/z) will be much smaller than bins closer to the outer boundary of the Si layers (high R/z). Given 42 R/z bins and 216 ϕ bins in a CE-E layer starting at R = 30cm and ending at R = 150 cm, the ratio of the physical areas of two far-away bins can reach a factor of 5. In spite of the varying bin area, the distribution of TCs across HGCAL is uniform. This implies a smaller number of TCs for low R/z bins. This is intended, as regions closer to the beamline expect larger energy density or occupancies, and thus the binning should be finer to increase the local resolution.

The rectangular binning grid applied over the physical hexagonal TC space inevitably introduces inhomogeneities in the distribution of TCs, since a perfect alignment of detector elements with (ϕ, R/z) bins is not possible. In other words, immediately adjacent bins, in both the R/z and ϕ dimensions, might have a significantly different number of associated TCs. Differences are extremely dependent on the bin widths, but the effect is very pronounced with current S2 parameters. The inhomogeneities are illustrated by the black dots in Fig. 3.29. The non-uniformities introduce nonphysical biases, as the distribution of deposited energy in (ϕ, R/z) bins might not closely follow the true one in the detector. Whenever the TC-to-bin mapping is such that an intermediate bin has fewer TCs compared to their neighbors, the energy deposited by single particles can follow a pattern with two maxima along ϕ. As a consequence, the seeding step, as described in Section 3.3, finds two seeds. Since each seed originates a cluster, two clusters are formed from a single particle. The degradation of the detector's energy response and position resolution ensues. The effect is overwhelmingly more present in the low R/z region, where bins are finer and TC multiplicity fluctuations larger.

Previous studies [342] have shown, unsurprisingly, that defining a R/z-dependent binning can mitigate the splits. Specifically, binning schemes with finer bins in the low R/z

regions improve the energy resolution. However, and taking into account the intended final FPGA implementation, the added complexity requires additional firmware resources, making it impractical. In what follows, we propose a solution which does not impact hardware resources.

3.5.2 The Bye-Splits Iterative Algorithm

Cluster splits are mostly located in the high-η region, where bins are smaller in the Cartesian space and TC counts are less homogeneous along adjacent bins. Splits are also present virtually along ϕ only, since R/z bins are larger, and thus less subject to TC multiplicity fluctuations. In other words, TC non-uniformity is much more prominent along ϕ. For the purpose of this Thesis, I have developed the `bye_splits` algorithm, which aims at reducing the variance of TC across all ϕ bins for a constant R/z value. The algorithm modifies the mapping of TCs to (ϕ, R/z) bins, and is run offline, fully decoupled from the online reconstruction. Its TC-to-bin output mapping can be known in advance, and therefore encoded in a Look-Up Table (LUT). No impact on firmware resources is thus expected, which represents a striking advantage over all sorts of more complex algorithms that could be included in the FPGAs. The algorithm was presented at [8].

The `bye_splits` algorithm defines a sliding window around three consecutive ϕ bins, similarly to what happens in S2's seeding step. Circular boundary conditions are taken into account. `bye_splits` computes, for each group of three bins, the differences $D_{\text{left}} = C_2 - C_1$ and $D_{\text{right}} = C_3 - C_2$ between their TC counts C, where the indexes 1, 2 and 3 refer to the left, middle and right bins in the sliding window, respectively. A pseudo-random number x is sampled from an uniform distribution $\mathcal{U}(0,1)$ to decide whether the TC position migration should occur on the left or right side of the window:

$$\text{Side} = \begin{cases} \text{left}, & \text{if } x \sim \mathcal{U}(0,1) < \frac{|D_{\text{left}}|}{|D_{\text{left}}|+|D_{\text{right}}|} \\ \text{right}, & \text{otherwise} \end{cases} \tag{3.3}$$

The randomness in Eq. (3.3) ensures that the shape of the distribution of TC counts along ϕ is kept, while turning TC migrations more likely on the side where differences is counts are larger. In fact, had we simply used the absolute value of the differences to choose the side, all TC distributions would lose some of its distinct shape properties. Our method instead "squashes" the distributions, trying as much as possible to maintain the original shape. Once a side is chosen, the shift of a TC is executed taking into account the relative distribution of TC counts in the sliding window. For a sliding window size of 3, there are four types of TC distributions, as also illustrated in Fig. 3.30:

- "valleys", where the central bin has less TC counts than its neighbors;
- "mountains", where the central bin has more TC counts than its neighbors;
- "ascent", where TC counts increase from left to right;
- "descent", where TC counts decrease from left to right.

The TC shift is always performed from the bin with more TCs to the bin with less TCs. Eq. (3.3) ensures that a side is never chosen when the difference is zero. Only one TC is moved per iteration, and the TC shift is always done relative to the bin in the center

of the sliding window. After the shift, the sliding window moves with unitary stride. The algorithm is run for all possible windows, forming one *epoch*. After each epoch, the following termination condition is checked for every ϕ bin i:

$$|D_{\text{left},i}| + |D_{\text{right},i}| \leq \max\left\{1, \lambda \times \left(\left|D^0_{\text{left},i}\right| + \left|D^0_{\text{right},i}\right|\right)\right\} \tag{3.4}$$

where $\lambda \in [0, 1]$ is a tunable parameter and D^0 refers to the differences before the algorithm was run. The max operator ensures convergence for low-λ (more aggressive) runs. As expected, running the algorithm with $\lambda = 0$ provides an essentially flat TC count distribution, as shown in red in Fig. 3.29. We note that the direction in which the sliding window moves should not impact the final mapping, since the variance of TC counts per bin is independent of ϕ. However, the iterative nature of the algorithm does not ensure this condition. We experimentally run the algorithm in both directions and observe no noticeable difference. An important constraint is brought by the number and distance of bin migrations, which should be minimized. By "distance" we mean the physical Cartesian displacement of a TC relative to its original location. Despite the wish to reduce cluster splitting, the final TC-to-bin mapping should still reflect the overall physical positions of TCs, or else other energy- or position-related biases can impact the TPs. The algorithm is therefore run with a single ϕ bin shift per iteration. For $\lambda = 0$, we verify that all TCs move less than 2 cm along ϕ, which implies they moved to their immediately adjacent bins only. We also observe that the number of moving TCs decreases with decreasing η. For the lowest R/z (highest η) row almost 50% of the TCs moved. This number drops to ∼11 % (∼8 %) for the 3$^{\text{th}}$ (6$^{\text{th}}$) R/z-row.

The algorithm removes a significant portion of cluster splits. We show one such example in Fig. 3.32, where the effect of the smearing step is also illustrated. Significant improvements in energy response and position resolution are obtained, as one can observe on the left column of Fig. 3.33. The energy response is here defined by dividing the reconstructed p_{T} by the generated one, where a value of 1 indicates that all the momentum was reconstructed. The position resolutions are instead calculated by taking the difference between reconstructed and generated η or ϕ, where a value of 0 indicates a perfect matching between generated and reconstructed positions. `bye_splits` has also been validated by verifying that it does not impact the reconstruction of samples where no cluster splits are present. The sample was obtained requiring an energy response above -0.2, instead of below -0.35. The algorithm is thus a strong candidate for the final design of the reconstruction chain.

3.5.3 Algorithmic Alternatives

The performance of `bye_splits` is very encouraging, but could be improved. Alternative approaches were therefore explored to mitigate cluster splits, trying to achieve better energy responses and position resolutions. New approaches could be easily tested thanks to the framework described in Section 3.4. Three methods are tested, and they all reduce cluster splits dramatically.

- **Seeding window increase**: The size of the seeding window along ϕ was increased to 2 bins, while keeping the size along R/z set to 1 bin. This enables the seeding step to "look further", and notice whether a particular area of the phase-space has two local maxima. The window has access to 14 neighbors instead of 8. Results are displayed on the right column of Fig. 3.33, where it becomes clear that clus-

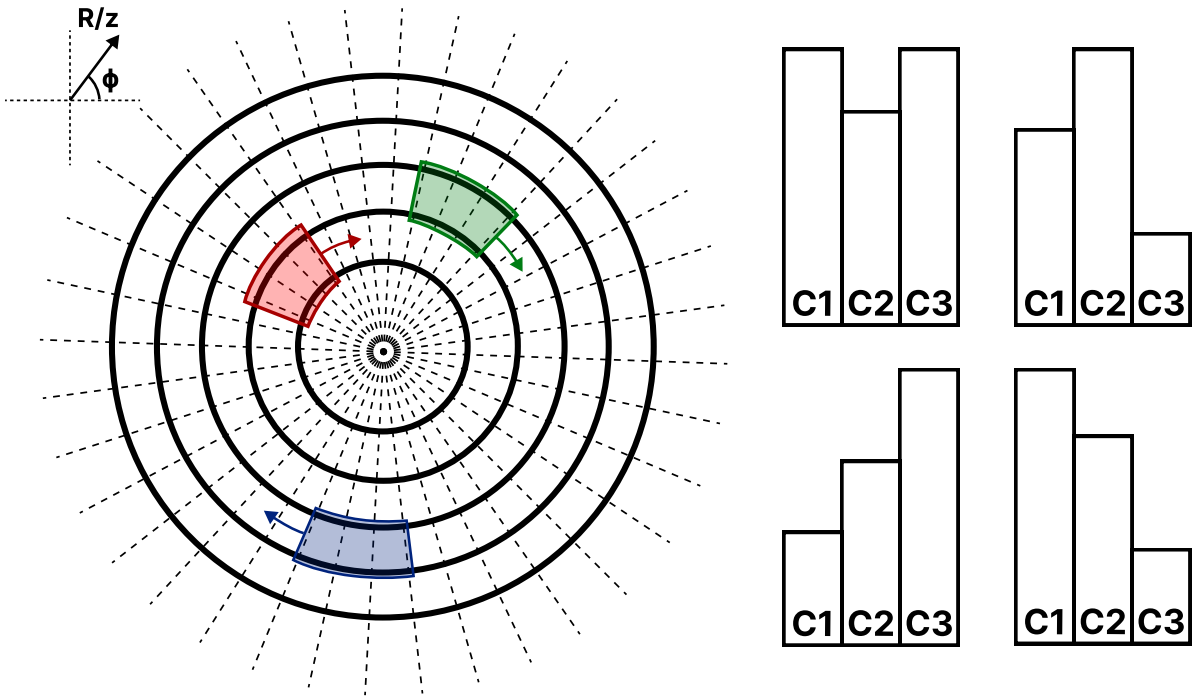

Figure 3.30: (Left) Illustration of the phase-space where the `bye_splits` iterative algorithm is run. Only the 4 lowest R/z rows are shown. The algorithm is run independently for each R/z row, with a sliding window of size 3. The arrows represent the direction followed by the sliding windows at each successive iteration. (Right) The TC migrations are executed depending on the four relative distributions of TC counts that can be present in a particular sliding window, from top-left to bottom-right: "valley", "mountain", "ascent" and "descent". The letters "C" refer to the TC counts D_{left} and D_{right} are defined with.

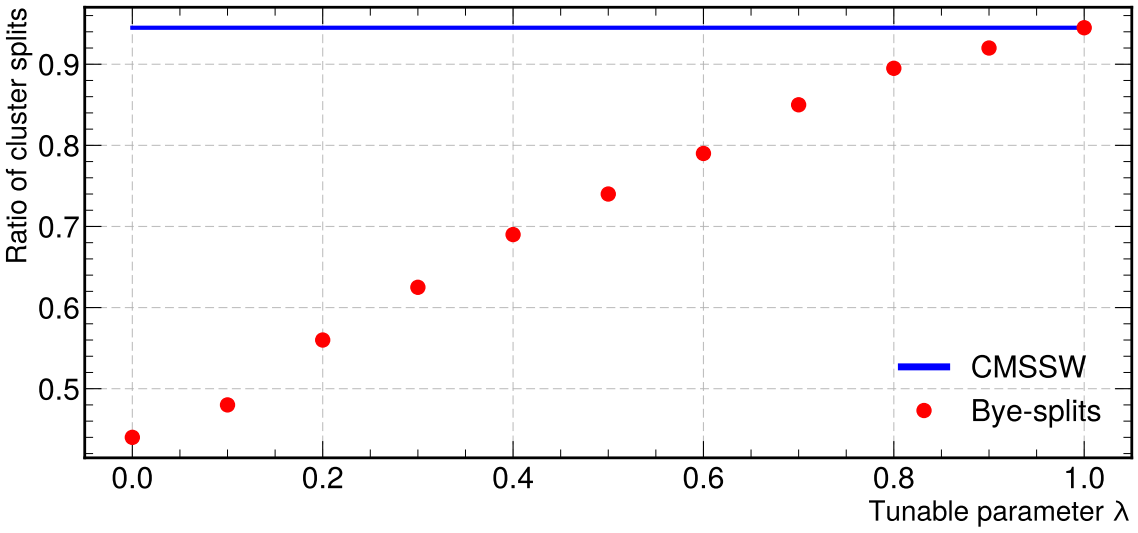

Figure 3.31: Dependence of the ratio of cluster splits with the parameter λ, which controls the stopping condition of `bye_splits` (see Eq. (3.4)). The blue line represents the fraction of cluster splits with the default chain, also implemented in CMSSW. The red dots refer to the custom framework with `bye_splits`. The more aggressive the algorithm is, the more cluster splits are removed. Following Eq. (3.4), a value of $\lambda = 0$ indicates the most aggressive stopping condition.

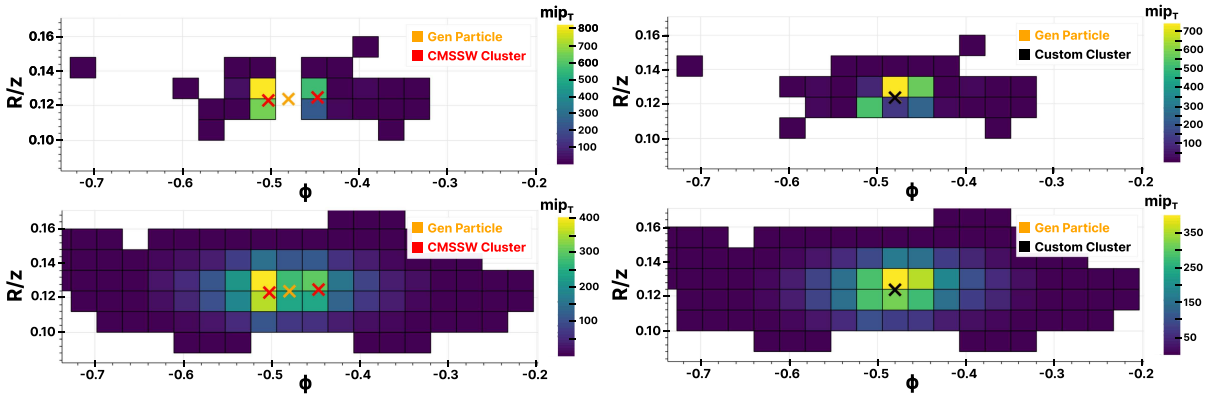

Figure 3.32: Example of a cluster split for a single photon event in the (ϕ, R/z) space, where colors represent energy deposited per bin in MIP_T units. The orange cross shows the position of the generated unconverted photon. The top (bottom) row shows the same event before (after) applying the smearing step. The left (right) column displays the event not considering (considering) the `bye_splits` algorithm, where the red (black) crosses point to the position of the reconstructed clusters. Generated and reconstructed clusters become superimposed after running `bye_splits`.

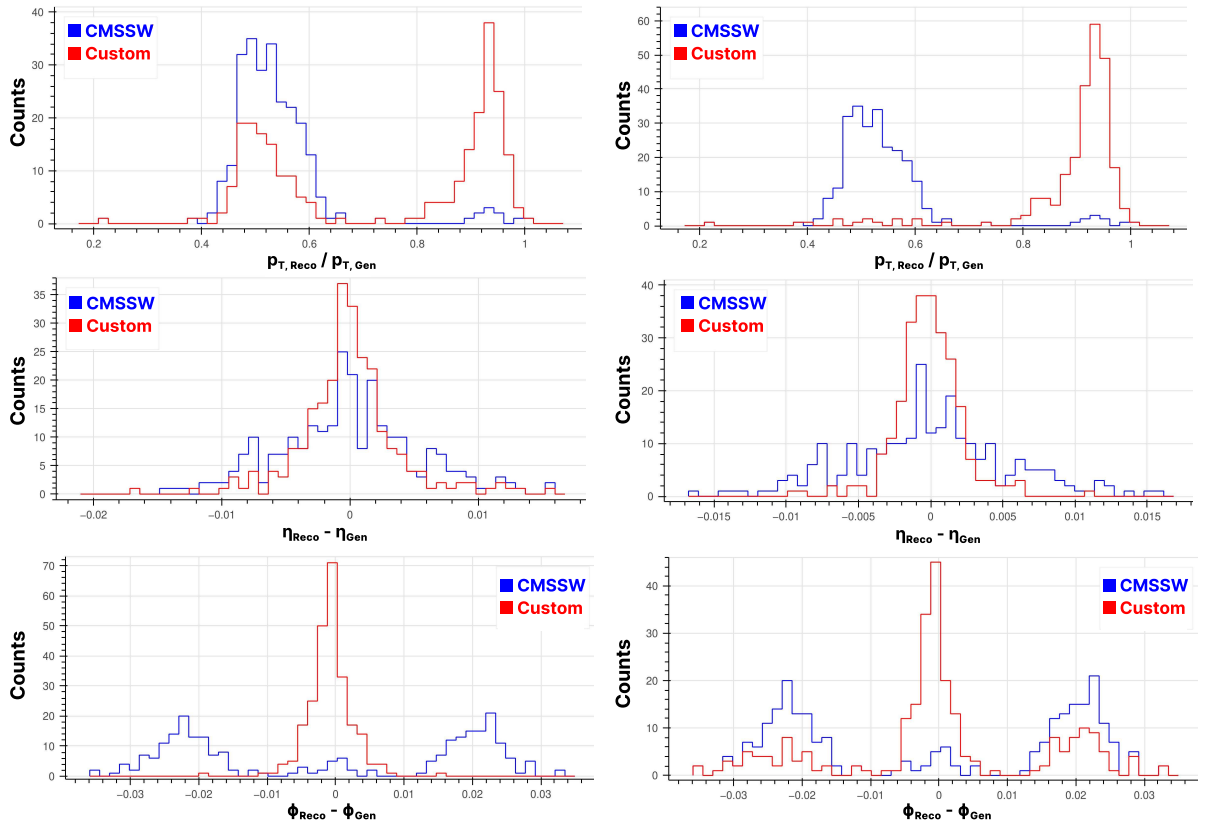

Figure 3.33: Energy response (top) and position resolution (η, middle, and ϕ, bottom). The blue curves refer to the standard CMSSW S2 chain, which is identical in both columns. The red curves shows the results of the custom `Python` framework, considering `bye-splits` with $\lambda = 0$ on the left and a seeding window of size 2 along ϕ on the right. Both approaches improve the response and resolutions. All displayed events satisfy the selections described in the text.

ter splits are virtually completely removed, with a performance clearly superior to `bye_splits`. However, each window requires six additional firmware comparators, increasing resource consumption.

- **Smearing kernel update**: Apply a "flat-top" kernel instead of the default one:

$$\left[\ \ldots \ \ \frac{1}{8} \ \ \frac{1}{4} \ \ \frac{1}{2} \ \ 1 \ \ 1 \ \ 1 \ \ \frac{1}{2} \ \ \frac{1}{4} \ \ \frac{1}{8} \ \ \ldots \ \right] \tag{3.5}$$

  The current smoothing kernel applies a larger weight to the central bin, while the updated one has a "flat" weight scheme around the central bin. This is equivalent to share the bin energy between central bins along ϕ, causing local non-uniformities in the energy deposits to become blurred out, removing cluster splits. This methods has a result almost identical to the previous one.

- **Energy prioritization**: Instead of applying the default `min_distance` clustering algorithm, we test the greedy `max_energy` algorithm (recall Fig. 3.22). Cluster splits are by construction removed, as long as the matching TC-to-seed distance is large enough, since the method concentrates TCs into the same cluster. Given that seeds issued by a cluster split are generally very close, the `max_energy` algorithm effectively removes all splits.

The final choice of the algorithm(s) to use will be strongly dependent on available resources and latency constraints. This Thesis provides however a variety of choices to consider for a L1 performance improvement.

To optimize `bye_splits` and simplify its output LUT, the algorithm was also run considering smaller regions of the detector only, namely CE-E only and the region around which the EM shower energy maxima is expected to be found, specifically between layers 8 and 15. It was observed that TCs in CE-H are irrelevant for the performance of `bye_splits` with single photon guns, and that the maximum energy region clearly dominates the performance. The latter result is expected, given that most active TCs will on average lie in that same region. These results suggest that the requirements of algorithms needing more resources can be mitigated by focusing on the most sensitive regions of the detector for EM showers.

3.5.4 Using Detector Coordinates

HGCAL's geometry uses non-Cartesian coordinates to cover the hexagonal tessellation of its Si modules and sensors, as shown in Fig. 3.23 (left and middle). The algorithms running in the BE consider instead projective and/or Cartesian coordinates. The latter are used because they enable to access the full detector in a single 2D plane, and are ideal for vectorized data processing. However, biases are created once the phase-space where measurements are taken is matched to bins defined in a different space. The cluster splits described in Section 3.5.1 represent a good example. Additionally, the current (ϕ, R/z) bins imply a complex firmware routing of TCs to bins, which is further complicated by the varying bin area.

I explored the possibility of using detector coordinates for the HGCAL TPG reconstruction. Hexagonal coordinates bring a (moderate) increase in algorithmic complexity for navigation and neighbor query across detector elements. Additionally, and as discussed in Section 3.4.1, further complications arise from the coexistence of two different

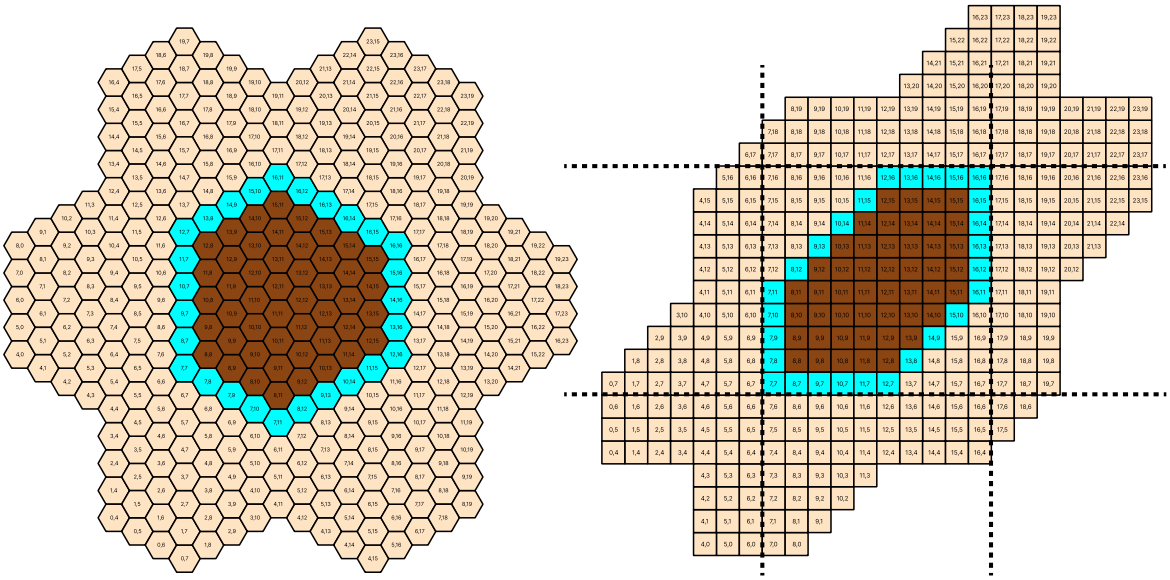

Figure 3.34: Illustration for the global hexagonal coordinate system used for the alternative seeding based on detector coordinates. Brown shows TCs belonging to the central module, while cyan shows the first-order neighbors that have to be considered when applying a size 1 seeding window to all TCs in the central module. Each TC only has 6 neighbors. (Left) Detector space. (Right) The same detector elements converted to a square grid-like space, helpful to visualize the structure of the seeding window in hexagonal coordinates. The bins contained within the dashed lines represent the bins that a naive seeding implementation would have to store in memory for each module being processed.

hexagonal coordinates, which must be integrated. Finally, detector coordinates are not projective, which implies considering groups of a few consecutive layers only. On the positive side, using coordinates that follow the geometry of the detector should immediately remove any existing bias, including cluster splits. It also simplifies the routing of TCs, and removes any need for coordinate conversion in the algorithms. The firmware resource usage remains very dependent on individual algorithms, and can actually be lowered given the removal of potential coordinate conversion steps.

I have designed an alternative seeding step for S2 which receives TCs and finds local maxima directly in detector coordinates. The step was implemented as part of an alternative processing chain in the `Python` software reconstruction. For each HGCAL layer, the seeding projects all modules into a new global hexagonal coordinate system. The coordinate translation from local $(u,\ v)$ to global (u', v') TC coordinates, using the hexagonal module coordinates (U, V), is done using the following expression:

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = -4 \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} + \begin{bmatrix} u \\ v \end{bmatrix} \tag{3.6}$$

where the constants reflect the structure of the Si modules and the hexagonal tessellation. Notably, they can be expressed as powers of 2, easing their inclusion in the firmware. The effect can be seen in Fig. 3.34, where the end result of the translation of seven modules is shown in detector and rectangular spaces. Both spaces are functionally identical, but the second helps visualizing what the alternative seeding step considers. The result of the same translation for a simulated single unconverted photon event is shown in Fig. 3.35.

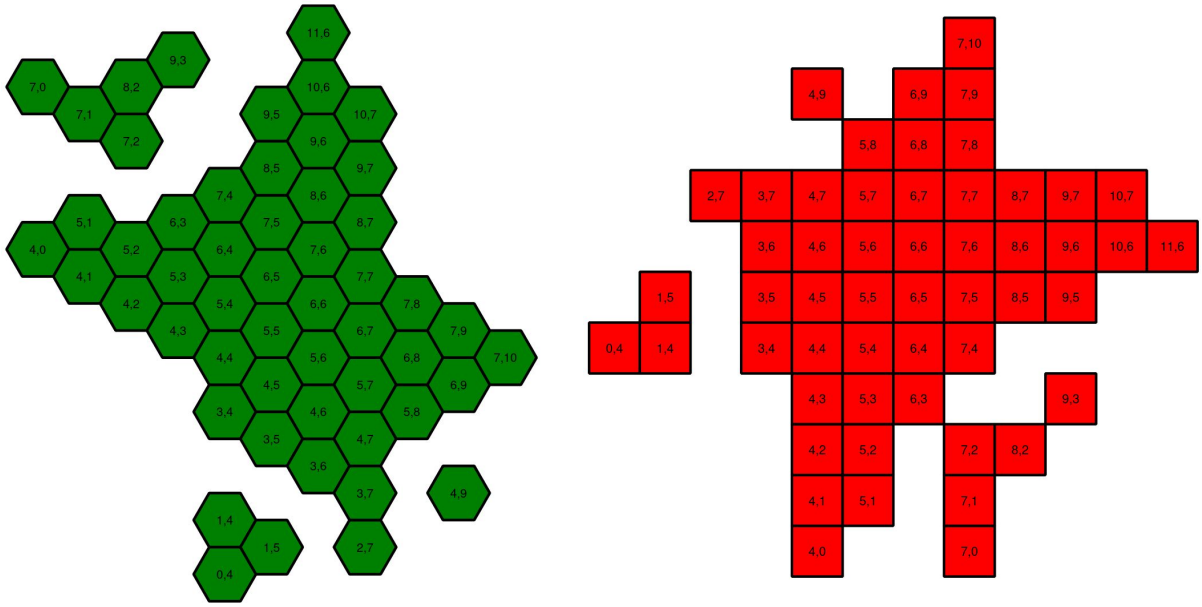

Figure 3.35: Projection of TCs in four layers (9, 11, 13 and 15) of a simulated single unconverted photon event. The layers correspond to the shower maximum region. Each TC has 6 neighbors only, not 8 as the rectangular space might suggest. (Left) Detector space. (Right) The same detector elements converted to a square grid-like space.

The internal logic of the alternative seeding is very similar to the default seeding algorithm. Once the coordinate translation in Eq. (3.6) is applied, the algorithm searches for local seeding maxima. Given the non-projective nature of the coordinates, we choose to consider the projection of (u', v') for layers 9, 11, 13 and 15, corresponding to the EM shower maximum region. Contrary to standard Cartesian coordinates, each hexagonal bin only has 6 neighbors. Less comparators are thus required at firmware level. When seeding one module, the information of nearby modules must be available to correctly compute seeds at the module's boundaries. The brown and cyan colors of Fig. 3.34 represent the central modules and the TC neighbors for a seeding window of size 1. These considerations become relevant if the seeding is run in parallel for multiple modules in the firmware. The algorithm itself can be easily parallelized at TC level, where each brown bin would correspond to a separate instance. The dashed lines contain the bins in a square-grid which would have to be kept in memory if running all instances in parallel, using some kind of shared memory in a naive implementation. Further considerations about firmware implementation are outside the scope of this Thesis. An additional advantage of the alternative seeding step is that it allows to replace the current S2 histogramming step by a simple accumulation step across layers, potentially reducing S2's latency. The smearing might also become redundant, since its main purpose was to remove cluster splits. This is indeed the case in all performed tests, but further validation would be required, as discussed in Section 3.6.

The custom processing chain is run with the alternative seeding and is compared with the default chain using (ϕ, R/z) coordinates. The energy response and position resolution remain virtually identical, as shown in the top plots of Fig. 3.36. The response is here defined as the difference between clustered and generated energy divided by the generated energy, where zero corresponds to reconstructing all generated energy. We also show in

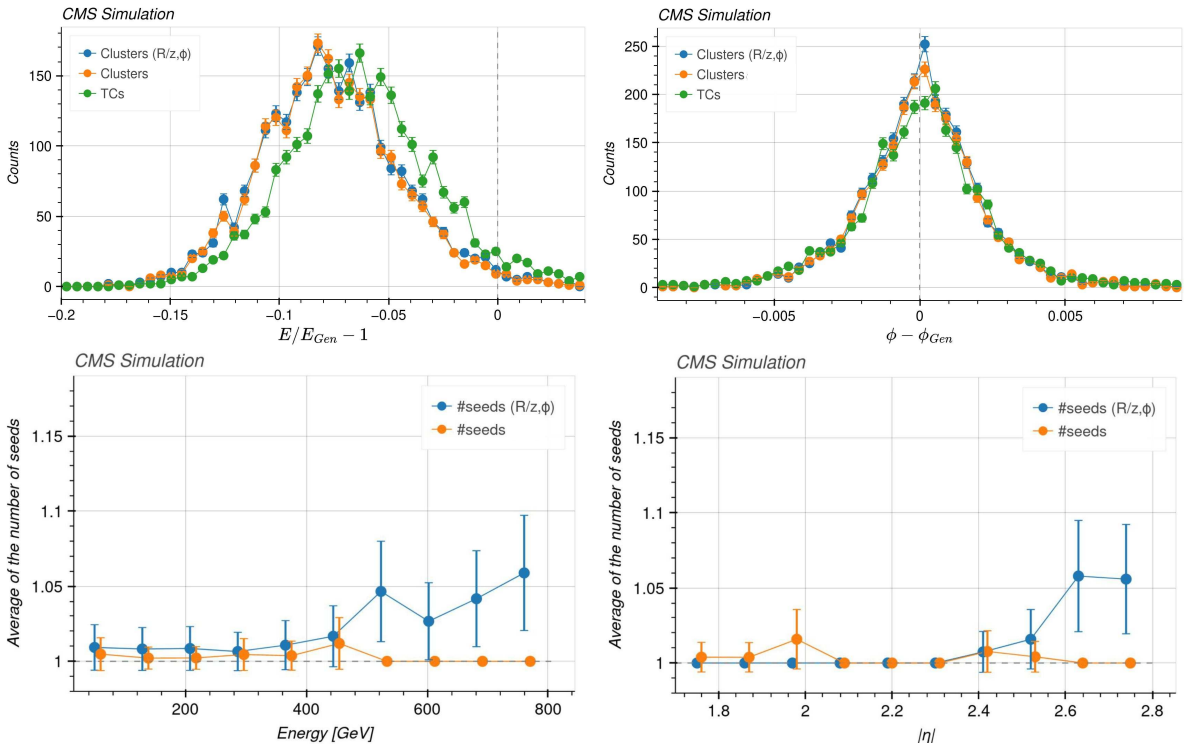

Figure 3.36: (Top) Energy response (left) and ϕ position resolution (right) for the default (blue) and alternative (orange) reconstruction chains, considering 0 PU single unconverted photons. The alternative chain drops the histogramming and smearing steps, and replaces the seeding step by the one using detector coordinates. In green we display the same quantities obtained with all TCs, where the position is weighted by the TC energy. (Bottom) Average number of seeds as a function of energy (left) and $|\eta|$ (right). Using detector coordinates naturally removes cluster splits.

green what would the response and resolution be had all TCs be considered, instead of considering only the ones filtered by the clustering algorithm. TC positions are weighted by energy. The green curve provides a more reasonable reference to which we can compare the blue and orange curves. The clustering algorithm only loses a small amount of energy compared to what is available. Interestingly, the bottom plots of Fig. 3.36 demonstrate an almost complete removal of cluster splits, by simply using detector coordinates, without running any additional algorithm or modified S2 step. The splits are removed at high η, where they are known to dominate. The additional number of splits for higher energy values can be explained by the increase of the spread of energy deposition along (ϕ, R/z). We also test a seeding window of size 2, but find no significant difference relative to the default size of 1. We believe that the rare splits still present after running the algorithm do not have a geometrical cause, but are instead the result of photon conversions that happened after the surface of HGCAL but before the layers where the EM shower maximum occurs. The algorithm is observed to have a "perfect efficiency" at 0 PU, meaning that it reconstructs at least one seed for all events.

3.6 Discussion and Future Directions

In this Chapter I introduced the HL-LHC and its upgrades, focusing mostly on the future HGCAL. The reconstruction of TPs in HGCAL was described in detail, in order

to contextualize the developments I brought forward. My contributions covered two interdependent axis: software development and algorithm design and implementation. Concerning the former, a new `Python` chain was developed in order to replicate the L1 S2 processing chain. The chain is an important piece of software for people working on TPG at HGCAL. In parallel, a simplified geometry was implemented, useful for visualization and sharing. In general, the developments above simplify testing, validation and tuning of any current and future algorithms. On the algorithmic axis, I designed and implemented from scratch two new algorithms, `bye_splits` and the hexagonal seeding. The former mitigates cluster splits without affecting firmware resources, while the latter showcases an elegant solution to the biases introduced by mismatched coordinate systems. The `bye-splits` algorithm represents a viable options to be implemented at the L1 S2, together with some alternative approaches mentioned in Section 3.5.3.

Several future research directions are available and required. Of particular note, only 0 PU single photon guns were studied in depth. Despite checks on 200 PU samples suggesting a correct data processing behavior, extensive studies are necessary. In particular, any algorithm to be added to S2 must be exposed to the challenging 200 PU environment, and compared to existing alternatives. To give an example, the presence of PU background noise might affect the capability to mitigate cluster splits. In addition, different types of particles must be tested, such as taus, hadrons and LLPs. Their inclusion requires considering the information coming from the EM and HAD sections of the detector, and might imply the adaptation and optimization of current algorithms. As an example of the issues one might encounter, tau leptons can naturally deposit their energy in single or multiple clusters via one or three pronged decays. It is also currently not clear whether seeding algorithms require information from the Sci section. Beyond different signal signatures, one must also investigate the trigger rates each algorithm brings, and ensure their compatibility with L1 requirements. Another point concerns unexpected corner cases, which might create problems that are currently not foreseen. Failures can only be addressed if caught, and this requires systematic tests with large and diverse samples.

Eventually all algorithms must be implemented and tested in the firmware before deployment. This becomes increasingly crucial as the HGCAL installation deadline approaches. A logical and much awaited extension of this Chapter's work consists in using modern tools to include the new seeding algorithm in FPGAs. The idea is to rewrite the current implementations in a pragma-enriched `C++` version, which can be converted to Register-transfer level (RTL) via `Vitis` / `Vivado` High Level Synthesis (HLS) [343]. RTL can be converted to a bit-stream by `Vivado`, which is used as configuration data to program the FPGAs. One of the challenges lies on writing the RTL interface between the modules processed by the ECON-T and the BE algorithms.

On a final note, I believe the HGCAL L1 chain represents an ideal playground to test modern optimization approaches with deep learning and/or automatic differentiation [344]. A robust parameter tuning of all algorithms would clearly fit the needs of the TPG, and would most certainly improve physics performance. However, pragmaticism should rule in the final years before HGCAL's deployment, and priorities should be such that they provide a working TPG chain within the imposed constraints and deadlines.

4 Resonant Search for $HH \to bb\tau\tau$: Setting the Stage

The study of topologies involving a pair of Higgs bosons is immediately associated to the precise determination of the shape of the Higgs potential, and the consequences such a measurement can bring to the field of HEP. There is however another way of framing these processes, namely by hypothesizing the existence of BSM particles, massive enough to decay into a Higgs boson pair: $X \to HH$. The exploration of such topologies is primarily driven by the profusion of well-motivated theoretical models predicting the existence of resonances coupled to the Higgs sector. Examples of the aforementioned models are certain SUSY flavors, WEDs, and models where one or more singlets or doublets are added, much in the vein the successful ideas introduced by the Higgs Mechanism. The observation of a new resonant state would immediately revolutionize Particle Physics, representing the first definitive evidence of BSM physics at a collider experiment.

The specific choice of the $bb\tau\tau$ decay channel can be well justified. It has consistently been observed to be one of the most sensitive HH channels, due to a balance between background contamination and BR size. Its decay topology is experimentally rather complex, consisting on two tau leptons and two jets initiated by a b quark. The tau leptons are themselves unstable, with a 2.9×10^{-13} s mean lifetime [13]. They decay via leptonic or hadronic channels. Leptonic decays include either an electron or a muon, plus neutrinos in the form of MET. Hadronic decays are more complex, featuring various combinations of neutral and charged hadrons. When considering the $H \to \tau\tau$ process, there are are thus nine possible decay combinations, as illustrated in Fig. 4.1. In our analysis, we take into account the channels where at least one hadronically-decaying τ candidate τ_h is present: $\tau_e\tau_h$, $\tau_\mu\tau_h$ and $\tau_h\tau_h$. These three dominate, given their total BR of 88%. An additional channel, $\tau_\mu\tau_\mu$, despite not being used to extract the final results given it small BR of 3%, is utilized as a control region (CR) for DY and $t\bar{t}$ backgrounds, and is used to compute MET trigger SFs, as described in Section 4.7.4. The $\tau_e\tau_e$ and $\tau_e\tau_\mu$ channels are not exploited due to their small combined BR of 9%, and because they are expected to be overwhelmed by the DY and/or $t\bar{t}$ backgrounds (DY does not affect $\tau_e\tau_\mu$). Concerning the second Higgs boson candidate, the b-jets must be discriminated against other types of particles, mostly lighter jets, using techniques discussed in Section 4.3. All the above hints at the complexities of the $bb\tau\tau$ channel, which encompasses most topologies a collider experiment can be exposed to. The event display of a $HH \to bb\tau\tau$ candidate is shown in Fig. 4.2, recorded by CMS on July 5th 2016.

Finally, our analysis is also motivated by a recent result of the ATLAS Collaboration on the $X \to HH \to bb\tau\tau$ process [345]. There, a small excess is observed at $m_X \sim 1$ TeV, with a local (global) significance of $3.1\,\sigma$ ($2.0\,\sigma$)[1]. This excess has been excluded by the latest boosted $X \to HH \to bbWW$ and $X \to HH \to bbbb$ measurements from CMS, but an additional measurement with an identical final state topology could help strengthening

[1]Local and global significances are explained in Section 5.3.4

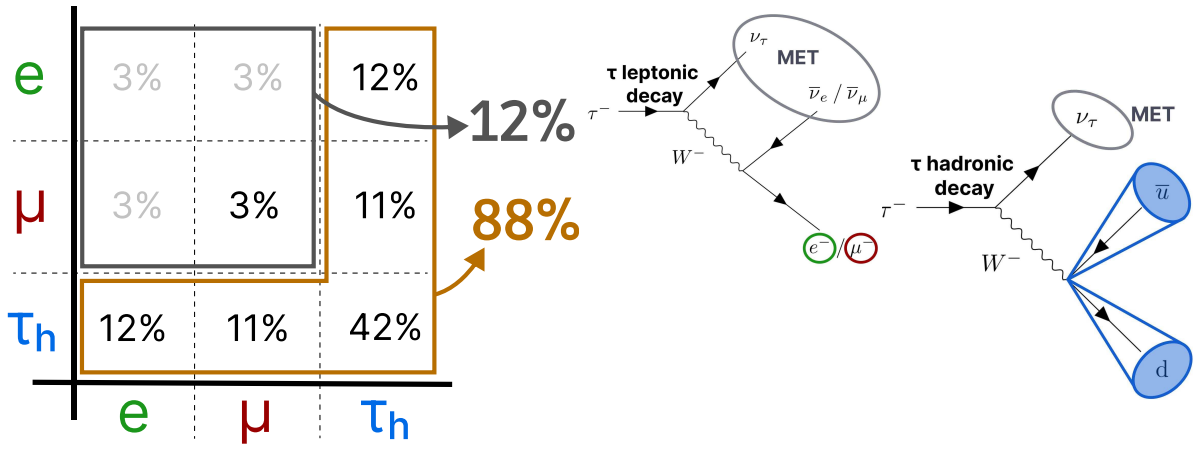

Figure 4.1: Illustration of the nine combinations two tau leptons can decay into (left), with corresponding leptonic and hadronic Feynman diagrams (right). Six decays are considered in the X → HH → bbττ analysis, highlighted in orange and corresponding to 88% of the total BR. The gray box shows the decays that were not considered in the limits, due to very large backgrounds and small BRs, which make them much less sensitive than the other channels. The $\tau_\mu\tau_\mu$ channel is however used to derive trigger scale factors (SFs) and to define a background control region. All decay channels include at least one neutrino, and thus MET.

the exclusion. This stems from the high complementarity between the bbττ channel and the other silver bullets: bbbb tends to be the most sensitive at high masses, bbγγ dominates at low masses, and bbττ shines at an intermediate range. The analysis presented in this work thus aims to replicate the excess with the CMS experiment, or instead demonstrate it to be a statistical fluctuation. It is thus fundamental to obtain a sensitivity at least as good as the one from the ATLAS result, especially close to the m_X region where the tension was observed.

This work represents an enormous improvement over the latest CMS iteration of the resonant X → HH → bbττ analysis, which used 2016 data only [110]. Besides the almost four-fold increase in luminosity, the analysis strategy has been completely revamped, and not many similarities remain. The current strategy was instead initially based on the recent CMS nonresonant bbττ publication [111]. However, a series of important upgrades has since been implemented, as summarized in Table 4.1. It is worth remarking that the resonant and nonresonant studies mutually benefit from each other. Just like the work here presented took advantage from decisions made in the past, upgrades in the analysis here presented are already being considered for Run 3 bbττ analyses to be published in the future. Our analysis exploits pp collision data recorded at a $\sqrt{s} = 13\,\text{TeV}$ center-of-mass energy by the CMS experiment during Run 2, for a total of $138\,\text{fb}^{-1}$.

It should be clear that the work reported in this Chapter and in Chapter 5 does not correspond to the efforts of a single person, but rather the collective work of an (international) analysis team. I can however claim to have been one the main analyzers, with contributions on multiple fronts. I was deeply involved in the definition and implementation of the new trigger strategy, including the computation of MET triggers SFs, the definition of trigger regions, the removal of event overlaps, and the improvement of older implementations of the legacy triggers. I have updated some of the analysis selections, either due to updates brought by the UL datasets, or in order to extend the available phase-space. The introduction or upgrade of multiple corrections was also done as a part of my Thesis: new τ identification SFs, electron, tau and jet energy corrections and resolutions, including their propagation to the definition of MET, and others. This was

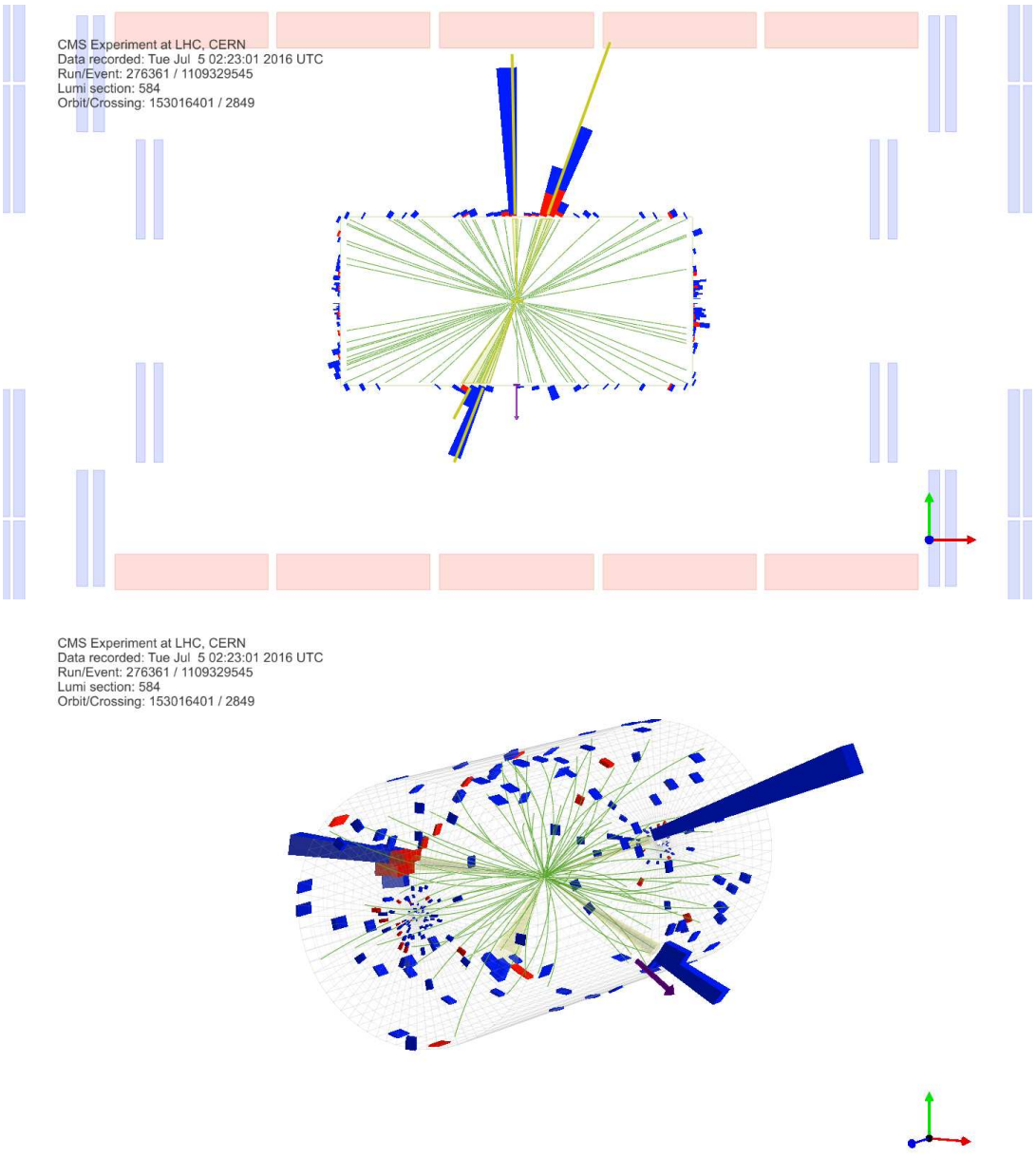

Figure 4.2: CMS event display of a $HH \to bb\tau\tau$ candidate, in 2016. Two views are shown, namely R vs z (top) and 3D Cartesian coordinates (top). Red and blue represent, respectively, ECAL and HCAL energy deposits, where the magnitude is represented by the dimension of each bar. Tracks are represented in green. The four dark green jet cones highlight the two b jets and two hadronic τs. The event passed the **res2b** selection. The selection of the analysis categories is defined in Section 4.4.3.

Stage	Section	Improvement
Samples	Section 4.1.1	Consider the latest dataset updates (UL)
	Section 4.1.2	Consider NLO DY samples (scaled to NNLO) instead of LO
Triggers	Section 4.2	Exploit MET and single-τ triggers
Selection	Section 4.4	Updates, notably a phase-space increase at low p_T and high η
	Section 4.4.2	Retraining a dedicated algorithm for b jet selection
	Section 4.4.3	Category optimization for a merged b jet topology
	Section 4.5	New $m_{\tau\tau}$ regression algorithm
Discriminant	Section 5.1	New parameterized Deep Neural Network (DNN) algorithm
Statistics	Section 5.2	Additional systematic uncertainties

Table 4.1: List of improvements adopted in the resonant $X \to HH \to bb\tau\tau$ analysis with respect to the previous CMS $bb\tau\tau$ analyses. The updates cover all stages of the workflow illustrated in Fig. 4.3. All improvements are explained in detail in this Chapter or in Chapter 5.

strongly related to the computation of systematic uncertainties, where I was also played a major role. I also ported an estimation technique for $t\bar{t}$ background SFs, developed for the $bb\tau\tau$ nonresonant analysis, which is however currently believed not to be required. More generally, I was involved in virtually all steps of the rewrite and optimization of the analysis software, in the documentation and review process of the work, and on the production of results, in the form of binned distributions and final limits.

In this Chapter we set the stage for the signal extraction and statistical analysis done in Chapter 5. We describe the data and MC samples, triggers, physics objects and selection in Sections 4.1 to 4.4, respectively. We then detail a new algorithm which performs the regression of the invariant mass of the $\tau\tau$ pair in Section 4.5. We conclude this first Chapter dedicated to the $X \to HH \to bb\tau\tau$ analysis by covering the way how backgrounds are modeled and corrected, in Sections 4.6 and 4.7. The full analysis workflow is illustrated in Fig. 4.3.

4.1 Samples

The analysis uses up to $138\,fb^{-1}$ of Run 2 UL data[2] recorded by CMS at $\sqrt{s} = 13\,TeV$. It also makes use of dozens of MC samples. These are used to model the many processes which replicate or fake the final decay topology of interest, called *backgrounds*. MC samples are also employed to simulate $pp \to X \to HH$ processes with different spin and masses, representing the analysis' *signals*. In this Section we provide details on all the above, obtained via the data processing chain described in Section 2.5.

4.1.1 Data

As seen in Section 2.3.2, CMS data is structured in PDs, each defined by a unique set of HLT trigger paths. Every event belonging to a PD fired at least one of the HLT paths defining that same PD. Each PD usually focuses on a particular type of physics object. The PDs used in the analysis are chosen based on the final leptonic products: electrons

[2]The concept of Ultra-Legacy data was first introduced in Section 2.5.

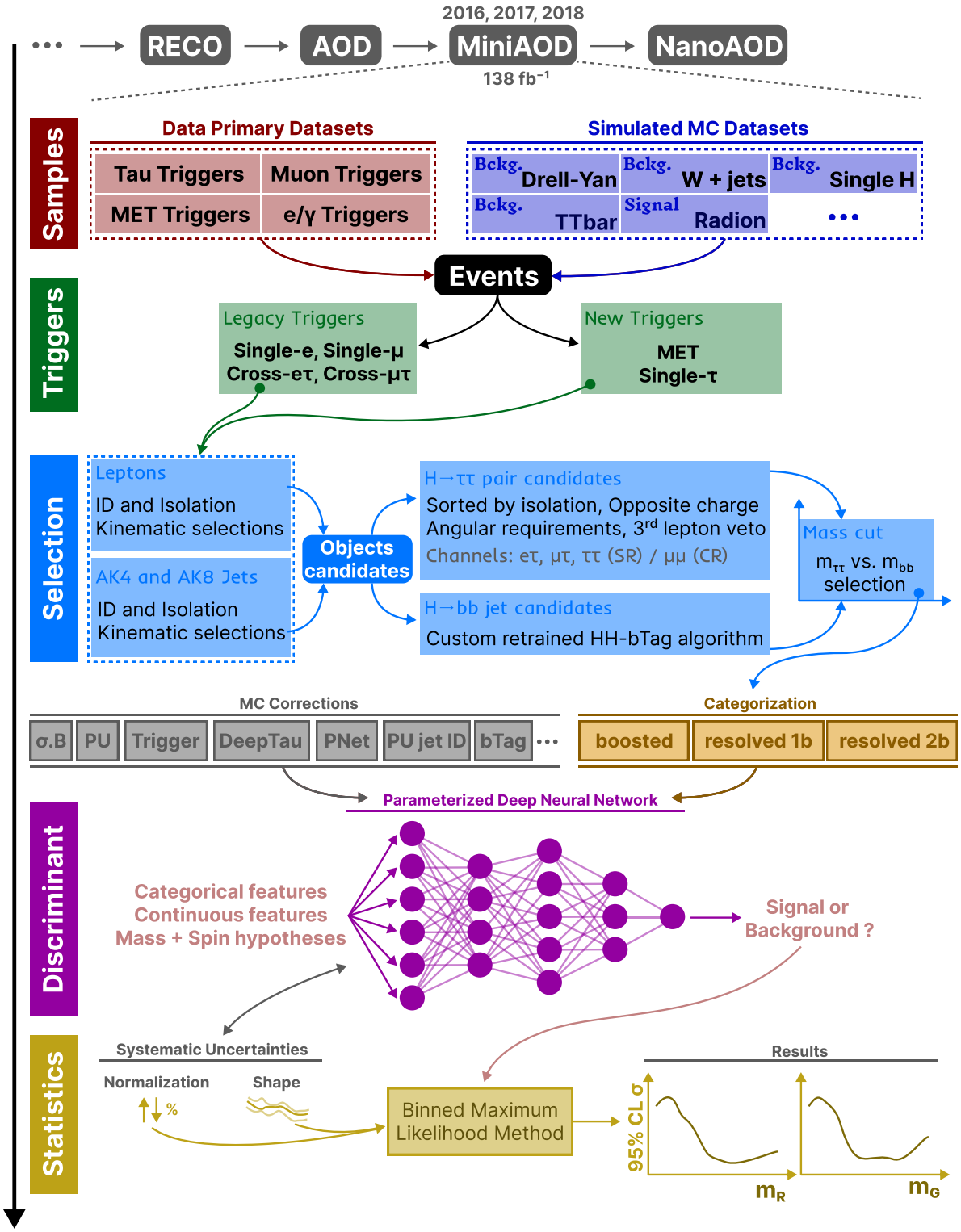

Figure 4.3: Illustration of the analysis workflow. The strategy can be visualized in different stages, each covered in detail in later Sections of this Chapter. We start with the selection of the Run 2 data and MC samples to consider. A series of triggers is then applied, in order to select events which might have a bb$\tau\tau$ topology. Selection cuts are then applied, constructing individual objects, and then pairs. Three categories are defined to enhance the analysis sensitivity. This information, together with a large number of MC SFs, is used as input to our DNN discriminant, which assigns a probability for an event to be signal or background. Such a variables is exploited in a binned maximum likelihood fit to extract 95% CLs, where systematic uncertainties are includes as nuisances. Results are obtained as a function of the masses of an hypothetical spin-0 radion (m_{R}) or spin-2 graviton (m_{G}).

from the `EGamma` PD, muons from the `SingleMuon` PD, and tau leptons from the `Tau` PD. Additionally, due to the new trigger strategy reported in Section 4.2, a dataset containing MET triggers is also considered. The names of the PDs give a strong indication on which triggers were used to collect their data. There are occasionally some mismatches; the `SingleMuon` PD, for instance, also includes double muon triggers. The usage of multiple PDs can lead to event overlaps. Their mitigation is discussed at the end of Section 4.2.3.

UL data are used across the three data-taking years. The 2016 data samples are split into two periods. This is because, in the late 2015 and early 2016 periods, the Si strip tracker experienced issues with the pre-amplifier of one of its (APV series) readout chips, leading to a loss of hits in the first period of 2016. In the remainder of this Thesis the period with issues is denoted 2016 APV, or "pre-VFP"[3], while the remaining 2016 data is referred as 2016 "post-VFP", or simply 2016. The recorded luminosity for different data-taking periods at CMS is measured to be $16.8\,\text{fb}^{-1}$ for 2016 and $19.5\,\text{fb}^{-1}$ for 2016 APV, for a total of $36.3\,\text{fb}^{-1}$ during 2016, $41.5\,\text{fb}^{-1}$ for 2017 and $59.8\,\text{fb}^{-1}$ for 2018.

4.1.2 Backgrounds

The more data one collects, the more sensitive an analysis tends to become. A sensitivity improvement can also be achieved by improving the efficiency and/or acceptance of the experiment to the signals being studied. In this work, we want to maximize the used number of hypothetical $X \to HH \to bb\tau\tau$ events present in our samples. However, by increasing the signal efficiency, the presence of backgrounds mimicking the topology of the signal will become more significant. If the backgrounds are incorrectly estimated, a part of the measured data might be wrongly classified, leading to incorrect conclusions. For instance, a missing background source can lead to the observation of a non-existing BSM signal. A careful study and modeling of backgrounds is thus required.

In particular, the $X \to HH \to bb\tau\tau$ analysis is affected by numerous sources of background, which can be qualitatively labeled as *reducible* and *irreducible*. The reducible backgrounds are those with final states which can in theory be distinguished from the final states of the signal. In practice, however, no experiment has a perfect object efficiency and identification, and some inefficiencies are present. This becomes particularly noticeable when the backgrounds have cross sections several orders of magnitude larger than the signal being studied. One of the important reducible background sources in the $X \to HH \to bb\tau\tau$ analysis is multijet production, where gluon- or quark-initiated jets are misidentified as τ or b candidates. It mostly affects the $\tau_h\tau_h$ channel, since the leptonic selections in the other two channels, as described in Section 4.4, suppress the multijet contribution. In what follows, we also refer to this source as "QCD" background. Other reducible backgrounds are the DY production of a τ pair in association with a light quark pair, $t\bar{t}$ production with the decay of at least one of its W bosons to quarks, and the production of W bosons in association with light jets, where the latter are misidentified as either leptons or b jets. Backgrounds with a single W boson plus jets are substantial for the baseline selection described in Section 4.4. However, this source is strongly suppressed by the categorization step, specifically by b-jet tagger requirements. The categorization is discussed in Section 4.4.3.

The irreducible backgrounds are instead those with a final state identical to the signal of interest. The production of $t\bar{t}$ pairs dominates, namely when the decay is fully-leptonic, $t\bar{t} \to b\bar{b}\,W^{\pm}W^{\mp} \to b\bar{b}\,\ell^{\pm}\bar{\nu}_{\ell}\tau^{\mp}\nu_{\tau}$, where ℓ can refer to electrons, muons or tau leptons.

[3] VFP refers to "Preamplifier Feedback Voltage Bias"

This background affects all analysis channels. To give an idea, as stated in Table 4.2, the $t\bar{t}$ cross section amounts to 84.1 pb in the fully-leptonic channel only, against the expected SM HH cross section of ∼31 fb. The second most prominent irreducible background is again DY lepton pair production, this time with additional b-jets: $Z/\gamma^* + b\bar{b} \to \ell\ell + b\bar{b}$. The inclusive DY cross section is 5 orders of magnitude larger than the SM HH cross section.

On top of the major sources above, several minor backgrounds are also taken into account. All sources are listed in Tables 4.2 and 4.3, together with their generators and cross sections. Most background sources, except for the ones already discussed, should be negligible given their small cross sections, and one would naively think they can be ignored. However, not only the total process rate matters, but also its shape. Some low cross section processes can peak under the $bb\tau\tau$ maximum, as for instance $t\bar{t}H$. Such behavior can be observed in the di-τ mass regression described in Section 4.5. The background samples are generated based on state-of-the-art theoretical calculations at LO and NLO, and are used to optimize the event selection, to evaluate efficiencies, and to compute systematic uncertainties. DY samples are updated to NLO, compared to the LO version used in Ref. [111]. The samples are further normalized to NNLO, shape differences being negligible. MC samples are also used to extract all sorts of SF corrections, as detailed in Section 4.7.

Process	**MC generator**	**σ [pb]**	**Ref.**
W + jets: inclusive	LO `MadGraph v2.6.5`	59375.00	[346]
W + jets: binned in H_T	LO `MadGraph v2.6.5`	–	–
DY + jets: inclusive	NLO `MadGraph v2.6.5`	6077.22	[347]
DY + jets: binned in p_T^Z	NLO `MadGraph v2.6.5`	–	–
DY + jets: binned in jet multiplicity	NLO `MadGraph v2.6.5`	–	–
EWK W^- + 2 jets, $W \to \ell\nu$	LO `MadGraph v2.6.5`	20.25	[348]
EWK W^+ + 2 jets, $W \to \ell\nu$	LO `MadGraph v2.6.5`	25.62	[348]
EWK Z + 2 jets, $Z \to \ell\ell$	LO `MadGraph v2.6.5`	3.987	[348]
$t\bar{t}$: fully-leptonic	NLO `Powheg 2.0`	84.1	[349]
$t\bar{t}$: semi-leptonic	NLO `Powheg 2.0`	347.55	[349]
$t\bar{t}$: semi-hadronic	NLO `Powheg 2.0`	359.44	[349]
single top: tW channel + antitop	NLO `Powheg 2.0`	35.85	[348]
single top: tW channel + top	NLO `Powheg 2.0`	35.85	[348]
single top: t-channel + antitop	NLO `Powheg 2.0`	80.95	[348]
single top: t-channel + top	NLO `Powheg 2.0`	136.02	[348]
single top: s-channel + antitop	NLO `Powheg 2.0`	7.104	[348]
single top: s-channel + top	NLO `Powheg 2.0`	3.549	[348]

Table 4.2: List of background samples used in the $X \to HH \to bb\tau\tau$ analysis (part 1), together with their generators and cross sections at 13 TeV. The DY and W + jets samples are stitched together and rescaled to the cross section of the inclusive sample.

Process	MC generator	σ [pb]	Ref.
ZZ	LO Pythia8	16.523	[346]
WW	LO Pythia8	118.7	[346]
WZ	LO Pythia8	47.13	[346]
WWW	NLO MadGraph v2.6.5	0.209	[348]
WWZ	NLO MadGraph v2.6.5	0.168	[348]
WZZ	NLO MadGraph v2.6.5	0.057	[348]
ZZZ	NLO MadGraph v2.6.5	0.0147	[348]
$t\bar{t}WH$	LO MadGraph v2.6.5	0.00114	[348]
$t\bar{t}ZH$	LO MadGraph v2.6.5	0.00113	[348]
$t\bar{t}W$ + jets, $W \to \ell\nu$	NLO MadGraph v2.6.5	0.2043	[348]
$t\bar{t}W$ + jets, $W \to qq$	NLO MadGraph v2.6.5	0.4062	[348]
$t\bar{t}Z$ + jets	NLO MadGraph v2.6.5	0.2529	[348]
$t\bar{t}WW$	LO MadGraph v2.6.5	0.006967	[348]
$t\bar{t}ZZ$	LO MadGraph v2.6.5	0.001386	[348]
$t\bar{t}WZ$	LO MadGraph v2.6.5	0.00158	[348]
ZH, $H \to bb$, $Z \to \ell\ell$	NLO Powheg 2.0	0.052	[350]
ZH, $H \to bb$, $Z \to qq$	NLO Powheg 2.0	0.36	[350]
ZH, $H \to \tau\tau$	NLO Powheg 2.0	0.0554	[350]
W^+H, $H \to \tau\tau$	NLO Powheg 2.0	0.05268	[350]
W^-H, $H \to \tau\tau$	NLO Powheg 2.0	0.0334	[350]
$q\bar{q} \to H$, $H \to \tau\tau$	NLO Powheg 2.0	0.237	[350]
$g\bar{g} \to H$, $H \to \tau\tau$	NLO Powheg 2.0	3.0469	[350]
$t\bar{t}H$ + jet, $H \nrightarrow bb$	LO MadGraph v2.6.5	0.17996	[350]
$t\bar{t}H$ + jet, $H \to bb$	LO MadGraph v2.6.5	0.2953	[350]
$t\bar{t}W$ + jets, $W \to \ell\nu$	LO MadGraph v2.6.5	0.2161	[350]
SM $HH \to bb\tau\tau$	LO MadGraph v2.6.5	0.0031	[62]

Table 4.3: List of background samples used in the $X \to HH \to bb\tau\tau$ analysis (part 2), together with their cross sections at 13 TeV and generators.

4.1.3 Signal

The $X \to HH \to bb\tau\tau$ signal samples refer to the ggF radion and graviton models, corresponding to spin-0 and spin-2 processes, respectively, where the Higgs bosons are forced to decay into a pair of b quarks and a pair of τ leptons. The models are introduced in Ref. [88], which defines a set of common benchmarks multiple analyses can agree on. This is crucial for comparisons and combinations across different channels. Analyses within the ATLAS Collaboration also use the same spin-2 model. The model employed for spin-0 is instead different, but equivalent when the NWA approach is considered, which is so far the case for the vast majority of HEP analyses. In Chapter 6 we discuss an alternative approach to the NWA.

The signal samples are generated at LO with `MadGraph v2.6.5`, assuming the NWA, and `Pythia8` is used for the hadronization. The samples usually have on the order of 10^5 events, depending on spin, mass, and data-taking period. The mass range considered for both spin hypotheses is identical: 250, 260, 270, 280, 300, 320, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 1000, 1250, 1500, 1750, 2000, 2500 and 3000 GeV. We do not consider lower masses to avoid the presence of extremely virtual Higgs bosons. There is no fundamental principle against it, but the width of the Higgs boson is so small as to make such a process experimentally unobservable. On the large mass side, we stop at 3 TeV since a dedicated CMS high mass analysis with the same final state is currently being developed by a different analysis team. Additionally, higher mass points are not expected to provide a better sensitivity, due to presence of boosted τs, for which our analysis is not optimized.

4.1.4 MC Reweighting

The reweighting and normalization of MC background samples is essential to accurately model the processes and to establish a sound comparison with data. The same is true for signal samples, except in what concerns the overall normalization. Indeed, when a model independent analysis is performed, the cross section of resonant signal processes is arbitrarily defined. MC samples are weighted per event, using the following factor:

$$L^{\text{period}} \times (\sigma \times \mathcal{B})^{\text{theory}}_{\text{sample}} \times \frac{\prod_{j \in \Theta^{\text{gen}}} w_j^{\text{gen}}}{\sum_{\text{events}} \left[\prod_{j \in \Theta^{\text{gen}}} w_j^{\text{gen}} \right]} , \tag{4.1}$$

where L^{period} is the integrated luminosity in a given data-taking period, σ and $\mathcal{B}$ represent, respectively, the theoretical process cross section and decay BR of a specific sample, and w_j^{gen} stands for weight j of the MC generator, belonging to the set of generator weights Θ^{gen}.

In practice, additional weights w_i must be applied to correct for discrepancies between data and MC. The full per event weight w_{event} is thus given by:

$$w_{\text{event}} = \frac{L^{\text{period}} \times (\sigma \times \mathcal{B})^{\text{theory}}_{\text{sample}} \times \prod_{j \in \Theta^{\text{gen}}} w_j^{\text{gen}} \times \prod_{i \notin \Theta^{\text{gen}}} w_i}{\sum_{\text{events}} \left[\prod_{j \in \Theta^{\text{gen}}} w_j^{\text{gen}} \right]} . \tag{4.2}$$

The total number of events corresponds to the sum of all w_{event} weights. Different data-taking periods can have different sets of weights, depending on the detector's evolving conditions. In our analysis, the w_i weights consist of stitching weights for DY samples, L1 prefiring weights, trigger SF weights, weights for objects faking taus, PU jet identification weights, reshaping b-tagging weights and b-jet discriminator weights. In Θ^{gen} we include the PU reweighting and the NLO reweighting weights. All MC weights are described in Section 4.7.

4.2 Triggers

The offline data events are recorded with a set of L1+HLT triggers directly related to the analysis decay channels. The goal is to minimize the number of lost $\text{HH} \to \text{bb}\tau\tau$

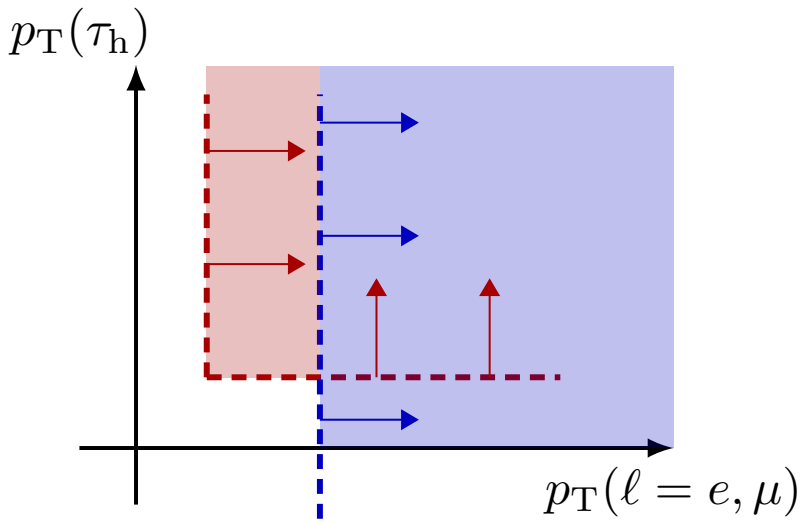

Figure 4.4: Simplified illustration of the increased acceptance brought by a cross trigger. We shows the $p_{\mathrm{T}}(\ell)$ vs. $p_{\mathrm{T}}(\tau_{\mathrm{h}})$ phase-space, with ℓ representing an electron or a muon. In blue we depict the region covered by the single lepton trigger, while in red we highlight the region added by the cross trigger. The dashed lines and arrows indicate the regions covered by the respective triggers. The additional tau leg in the cross trigger enables to decrease the lepton leg threshold, relative to the single trigger. Triggers also depend on other variables, making a realistic multi-dimensional picture more complex.

events, extending the analysis acceptance and maximizing the sensitivity to BSM signals. This is why analyses occasionally consider a large set of different HLT trigger paths, since all sources of increased acceptance matter, especially for processes as rare as double Higgs production. More event equates to more stringent exclusion limits, or luckily to the observation of some previously hidden resonance. However, it is also true that one expects diminishing returns for each additional trigger considered. Taking also into account the complexities associated to data processing and trigger corrections, a balance must be found.

The trigger strategy employed in this work extends the strategy used in the previous CMS Run 2 nonresonant bbττ result [111], which in itself already represents an extension of the strategy considered for the 2016 analysis [110]. In the following, we detail the new triggers used in the resonant HH → bbττ analysis for the first time, and explain in which phase-space regions they are applied. The logical-`OR` is always intended whenever several HLT paths are applied simultaneously to the same events.

4.2.1 Legacy Triggers

In the 2016 analysis, a single HLT path targets each channel: a trigger requiring one electron for the $\tau_e\tau_{\mathrm{h}}$ topology, another trigger requiring a muon for $\tau_\mu\tau_{\mathrm{h}}$, and a di-$\tau$ trigger for the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel. Cross trigger paths are then added to the semi-leptonic channels, following what was done in the nonresonant bbττ analysis of Ref. [111]. By "cross" we refer to triggers applying a logical-`AND` on two different object types, so-called *trigger legs*: electron and tau for $\tau_e\tau_{\mathrm{h}}$, muon and tau for $\tau_\mu\tau_{\mathrm{h}}$. These new triggers enlarge the analysis' phase-space, as shown in Fig. 4.4, in terms of the p_{T} thresholds considered. We call *legacy triggers* the set of triggers that has already been used in previous analysis iterations, and list them for the three analysis channels in Tables 4.4 to 4.6. The region

covered by those triggers is accordingly named *legacy region*. Additionally, the $\tau_\mu\tau_\mu$ channel, despite not being exploited for the final results, is nevertheless used for defining a CR and for extracting MET trigger SFs as explained in Section 4.7.4. The $\tau_\mu\tau_\mu$ triggers are thus shown in Table 4.7, corresponding to a subset of the triggers used for the $\tau_\mu\tau_h$ channel.

Year	Trigger requirement	Lumi [fb^{-1}]
2016	One electron, $p_T > 25$ GeV, $\|\eta\| < 2.1$, Tight Iso	35.9
2017	One electron, $p_T > 32$ GeV, Tight Iso, L1 Double-e/γ	41.5
	One electron, $p_T > 35$ GeV, Tight Iso	41.5
	One electron, $p_T > 24$ GeV, $\|\eta\| < 2.1$, Tight Iso + One τ_h, $p_T > 30$GeV, $\|\eta\| < 2.1$, Loose PF Iso	41.5
2018	One electron, $p_T > 32$ GeV, Tight Iso	59.8
	One electron, $p_T > 35$ GeV, Tight Iso	59.8
	One electron, $p_T > 24$ GeV, $\|\eta\| < 2.1$, Calo. Iso + One HPS τ_h, $p_T > 30$ GeV, $\|\eta\| < 2.1$, Loose PF Iso	42.1

Table 4.4: Trigger paths used in the $\tau_e\tau_h$ channel for the three years of data-taking and the corresponding integrated luminosity. Within each year, the logical OR of the HLT paths is always intended. The paths requiring one τ_h were updated during the 2018 data-taking after the implementation of the HPS algorithm (see Section 4.3.3). The Isolation criteria (Calo. or Tracker Iso) reported in the table are explained in the text.

Concerning the $\tau_e\tau_h$ channel, which triggers are listed in Table 4.4, isolation selections are applied to the electron legs, using a Tight WP. The isolation is defined based on the amount of calorimetric energy surrounding the e/γ L1 candidate. Additionally, an $|\eta| < 2.1$ cut is occasionally applied to avoid the high rates close to the beamline. The electron can be joined to a tau leg, where analogous but tighter cuts are applied, given the higher rates. Concerning Table 4.5, which covers the $\tau_\mu\tau_h$ channel, we can see that two types of muon isolation are defined. The first uses energy information from the two CMS calorimeters; the second exploits reconstructed tracks close to the muon. The muon can be also connected to a hadronic tau leg, forming another cross-trigger. The difference in p_T thresholds between the electron and muon triggers is an indication of the excquisite muon ID capabilities in CMS. Finally, for Table 4.6, we observe similar requirements as in the tau legs above, but with necessarily tighter cuts. Indeed, tau leptons are often faked by other physics objects, and their rates are thus harder to control, given the available trigger bandwidth. This explains why no single-τ trigger had been so far considered in the legacy region: p_T thresholds are too large, and cannot reach the p_T regions covered by the di-τ triggers. A description of the algorithms used to reconstruct tau leptons and of how their isolation is computed is available in Section 2.4.5.

4.2.2 Additional Triggers

Triggers which had not been present in previous analysis' iterations have been added to our analysis. The aim of the inclusions is to extend the available phase-space, increasing signal acceptance. The first novelty is the usage of an unprescaled MET$_{\text{no-}\mu}$ trigger. This

Year	Trigger requirement	Lumi [fb^{-1}]
2016	One muon, $p_T > 24\,GeV$, Tracker Iso	35.9
	One muon, $p_T > 24\,GeV$, Calo. Iso	35.9
	One muon, $p_T > 24\,GeV$, Calo. Iso + One τ_h, $p_T > 20GeV$, Loose PF Iso	35.9
2017	One muon, $p_T > 27\,GeV$, Calo. Iso	41.5
	One muon, $p_T > 20\,GeV$, $\lvert\eta\rvert < 2.1$, Calo. Iso + One τ_h, $p_T > 27GeV$, $\lvert\eta\rvert < 2.1$, Loose PF Iso	41.5
2018	One muon, $p_T > 24\,GeV$, Calo. Iso	59.8
	One muon, $p_T > 27\,GeV$, Calo. Iso	59.8
	One muon, $p_T > 20\,GeV$, $\lvert\eta\rvert < 2.1$, Calo. Iso + One τ_h, $p_T > 27\,GeV$, $\lvert\eta\rvert < 2.1$, Loose PF Iso	59.8

Table 4.5: Trigger paths used in the $\tau_\mu\tau_h$ channel for the three years of data-taking and the corresponding integrated luminosity. Within each year, the logical OR of the HLT paths is always intended. The paths requiring one τ_h were updated during the 2018 data-taking after the implementation of the HPS algorithm (see Section 4.3.3). The Isolation criteria (Calo. or Tracker Iso) reported in the table are explained in the text.

Year	Trigger requirement	Lumi [fb^{-1}]
2016	Two τ_h, $p_T > 35\,GeV$, $\lvert\eta\rvert < 2.1$, Medium PF Iso	27.2
	Two τ_h, $p_T > 35\,GeV$, $\lvert\eta\rvert < 2.1$, Medium PF Comb. Iso	8.7
2017	Two τ_h, $p_T > 35\,GeV$, $\lvert\eta\rvert < 2.1$, Tight PF Iso	41.5
	Two τ_h, $p_T > 40\,GeV$, $\lvert\eta\rvert < 2.1$, Medium PF Iso	41.5
	Two τ_h, $p_T > 40\,GeV$, $\lvert\eta\rvert < 2.1$, Tight PF Iso	41.5
2018	Two τ_h, $p_T > 35\,GeV$, $\lvert\eta\rvert < 2.1$, Medium PF Iso	59.8

Table 4.6: Trigger paths used in the $\tau_h\tau_h$ channel for the three years of data-taking and the corresponding integrated luminosity. Within each year, the logical OR of the HLT paths is always intended. The paths requiring one τ_h were updated during the 2018 data-taking after the implementation of the HPS algorithm (see Section 4.3.3). In 2016, the τ isolation requirement was modified, creating the Combined (Comb.) isolation, which also considers the p_T sum of neutral candidates, in addition to the number of hits in the tracker.

Year	Trigger requirement	Lumi [fb^{-1}]
2016	One muon, $p_T > 24\,GeV$, Tracker Iso	35.9
	One muon, $p_T > 24\,GeV$, Calo. Iso	35.9
2017	One muon, $p_T > 27\,GeV$, Calo. Iso	41.5
2018	One muon, $p_T > 24\,GeV$, Calo. Iso	59.8
	One muon, $p_T > 27\,GeV$, Calo. Iso	59.8

Table 4.7: Trigger paths used in the $\tau_\mu\tau_\mu$ channel for the three years of data-taking and the corresponding integrated luminosity. Within each year, the logical OR of the HLT paths is always intended. The paths requiring one τ_h were updated during the 2018 data-taking after the implementation of the HPS algorithm (see Section 4.3.3). The Isolation (Iso) criteria reported in the table are explained in the text.

variable closely follows Eq. (2.11), but reconstructed muons are explicitly removed from its definition:

$$\mathrm{MET}_{\text{no-}\mu} = -\sum_{i}^{\mathrm{N_{particles}}} \vec{p}_{\mathrm{T},i} + \sum_{j}^{\mathrm{N_{muons}}} \vec{p}_{\mathrm{T},j} \,. \tag{4.3}$$

The above quantity is introduced for the computation of $\mathrm{MET}_{\text{no-}\mu}$ trigger SFs, as described in Section 4.7.4. Following the same logic, a similar $\mathrm{MHT}_{\text{no-}\mu}$ quantity can be defined by extending Eq. (2.12), and removing contributions from PF muons. The $\mathrm{MET}_{\text{no-}\mu}$ trigger requires the $\mathrm{MET}_{\text{no-}\mu}$ and $\mathrm{MHT}_{\text{no-}\mu}$ of the event to lie above 120 GeV. The new trigger is used in all channels, exploiting the low lepton p_{T} kinematic region not covered by legacy triggers. Since no $|\eta|$ cut is enforced, the $\mathrm{MET}_{\text{no-}\mu}$ trigger also enables to reach high-$|\eta|$ events that would otherwise not be kept by the cross lepton triggers. Even leptonic triggers without explicit $|\eta|$ cuts are limited by the kinematic range of their corresponding trigger SFs. Furthermore, given the additional p_{T} thresholds one has to apply on top of the HLT thresholds, as described in Section 4.4, the $\mathrm{MET}_{\text{no-}\mu}$ trigger brings more events than one might initially think. We also expect the $\mathrm{MET}_{\text{no-}\mu}$ trigger to become more and more significant with increasing m_{X}. For low resonance masses, the two Higgs bosons lie close to a back-to-back topology in the laboratory frame, and the MET (and $\mathrm{MET}_{\text{no-}\mu}$) largely cancels. As m_{X} increases, the bosons become more and more collimated, leading to a MET increase. Taking into account the hard p_{T} thresholds of the di-τ trigger when compared to legacy triggers in other decay channels, and the fact that the acceptance increase applies to two τ_{h}s rather than one, we expect the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel to benefit the most from the added trigger. Finally, we have imposed a $\mathrm{MET}_{\text{no-}\mu}$ threshold of 150 GeV for all data-taking periods, to take into account the turn-on of the trigger. This choice is detailed in Section 4.7.4, where the corresponding SF are derived.

A single-τ trigger is also added for the first time to our analysis, requesting a hadronic tau lepton p_{T} above 180 GeV for 2017 and 2018, and above 140 GeV for 2016. The trigger covers events with exactly one high-p_{T} τ_{h}. When both leptons have a high p_{T}, the legacy triggers are used, since they have a larger efficiency, as discussed in Section 4.2.3 and shown in Fig. 4.6. We note that the $\mathrm{MET}_{\text{no-}\mu}$ trigger also extends the acceptance for high-p_{T} τ_{h}s, given the $|\eta| < 2.1$ of the single-τ trigger.

The most significant selections included in the newly added $\mathrm{MET}_{\text{no-}\mu}$ and single-τ triggers are listed in Table 4.8. Trigger SFs are applied to correct for discrepancies between data and MC trigger efficiencies.

4.2.3 Trigger Regions

Three non-overlapping trigger regions are defined for each channel, in order to simultaneously exploit different triggers. The regions are associated to the three groups of trigger presented in previous Sections: legacy, $\mathrm{MET}_{\text{no-}\mu}$ and single-τ. The definition of the regions does not depend on the trigger being fired, but only on kinematic properties of the events. The full logical-`OR` is not explored for reasons discussed in Section 4.2.4. The regions have a very similar structure for all years and channels, and are illustrated in Fig. 4.5 for 2018, showing only p_{T} cuts for simplicity. The $\tau_e\tau_{\mathrm{h}}$ channel in 2016 represents an exception, as no cross trigger was available. There, the legacy region is defined by $p_{\mathrm{T}} > 26$ GeV and $|\eta| < 2.5$ only. The $|\eta|$ cut comes from hard limits from the definition of trigger SFs.

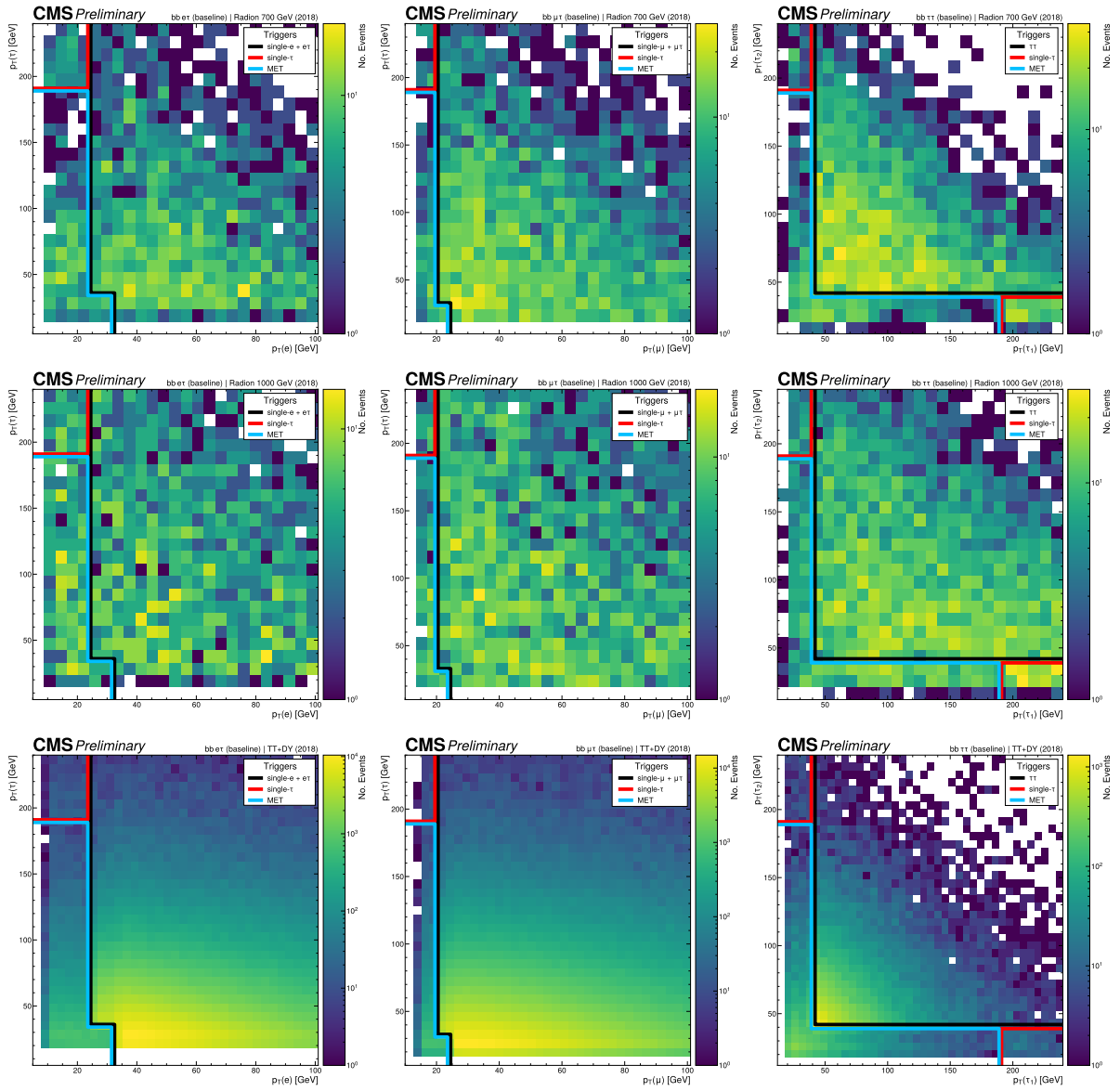

Figure 4.5: Lepton-p_{T} kinematic regions where different triggers of the analysis are applied, for 2018. The regions are defined based on the triggers' HLT cuts plus thresholds, and the colored thick lines represent the boundaries of those regions. The three $\tau\tau$ decay channels are represented by the left, middle and right columns. Single-$e(\mu)$ and cross-$e\tau(\mu\tau)$ triggers are used for high-p_{T} leptons in the $\tau_e\tau_{\mathrm{h}}$ ($\tau_\mu\tau_{\mathrm{h}}$) channels, associated to the black region. The $\mathrm{MET}_{\text{no-}\mu}$ trigger covers low-p_{T} leptons, in the blue region. The single-τ trigger covers the red regions where one of the leptons has high-p_{T}, and the other p_{T} is lower than the single- and cross-trigger thresholds. The three regions are explicitly defined in the text. The two top rows represent a signal radion with masses of 700 and 1000 GeV, while the bottom row shows the distribution of the two most important background components: DY and $\mathrm{t\bar{t}}$.

Year	Trigger requirement	Lumi [fb^{-1}]
2016	$\text{MET}_{\text{no-}\mu} > 120\,\text{GeV}$, $\text{MHT}_{\text{no-}\mu} > 120\,\text{GeV}$	35.9
	One τ_h, $p_\text{T} > 120\,\text{GeV}$, $\lvert\eta\rvert < 2.1$, ${p_\text{T}}^{\text{track}} > 50\,\text{GeV}$, VLoose Iso	29.1
	One τ_h, $p_\text{T} > 140\,\text{GeV}$, $\lvert\eta\rvert < 2.1$, ${p_\text{T}}^{\text{track}} > 50\,\text{GeV}$, VLoose Iso	33.3
2017	$\text{MET}_{\text{no-}\mu} > 120\,\text{GeV}$, $\text{MHT}_{\text{no-}\mu} > 120\,\text{GeV}$	40.6
	$\text{MET}_{\text{no-}\mu} > 120\,\text{GeV}$, $\text{MHT}_{\text{no-}\mu} > 120\,\text{GeV}$, $\text{H}_\text{T} > 60\,\text{GeV}$	36.7
	One τ_h, $p_\text{T} > 180\,\text{GeV}$, $\lvert\eta\rvert < 2.1$, ${p_\text{T}}^{\text{track}} > 50\,\text{GeV}$, Medium Iso	41.5
2018	$\text{MET}_{\text{no-}\mu} > 120\,\text{GeV}$, $\text{MHT}_{\text{no-}\mu} > 120\,\text{GeV}$	59.8
	One τ_h, $p_\text{T} > 180\,\text{GeV}$, $\lvert\eta\rvert < 2.1$, ${p_\text{T}}^{\text{track}} > 50\,\text{GeV}$, Medium Iso	59.8

Table 4.8: Trigger paths used in the $\tau_e\tau_\text{h}$, $\tau_\mu\tau_\text{h}$ and $\tau_\text{h}\tau_\text{h}$ channels for the three years of data-taking and corresponding integrated luminosity values. Within each year, the logical-OR of all the paths is always intended, also including the triggers in the legacy region. The paths requiring one τ_h were updated during the 2018 data-taking after the implementation of the HPS algorithm. The Isolation (Iso) criteria reported in the table are explained in the text.

The p_T and $|\eta|$ cuts defining the regions are based on the corresponding HLT thresholds in Table 4.5, plus additional standard thresholds applied on top of p_T cuts:

$$p_\text{T}^{\text{offline}} \geq p_\text{T}^{\text{HLT}} + \text{threshold}\,, \tag{4.4}$$

where $p_\text{T}^{\text{offline}}$ is the transverse momentum of the offline selected lepton, p_T^{HLT} is the p_T threshold applied at HLT level, and "threshold" depends on the lepton type: 1 GeV for electrons, 2 GeV for muons, 5 GeV for tau leptons firing the di-τ trigger, and 10 GeV for tau leptons firing the single-τ trigger. The thresholds are chosen to be conservative with respect to the trigger turn-on curves. For events passing more than one trigger, the loosest thresholds among those two triggers are applied.

The legacy region takes precedence over all other regions, since it is where the triggers more closely target the analysis topology and cover a larger signal acceptance. The second region to be defined is the one associated to the single-τ trigger. In the $\tau_\text{h}\tau_\text{h}$ channel, the single-τ region is actually composed of two disjoint regions, one for each tau lepton. Finally, all events not belonging to the two kinematic regions above are associated to the $\text{MET}_{\text{no-}\mu}$ region, on the condition that they meet the $\text{MET}_{\text{no-}\mu}$ trigger requirements. The complete selections are shown in Table 4.9, where their five-dimensional nature becomes clear. For $\tau_e\tau_\text{h}$ and $\tau_\mu\tau_\text{h}$, the lepton is by definition considered the first lepton; in the $\tau_\text{h}\tau_\text{h}$ channel the first lepton is the most isolated one.

We additionally study whether giving precedence to the single-τ region over the legacy region increases the analysis acceptance. In practice, this amounts to associating events with two high p_T leptons to the single-τ trigger region. In terms of Fig. 4.5, this scenario corresponds to a red band covering the top of each plot, plus the right side for $\tau_\text{h}\tau_\text{h}$, using 190 GeV as the p_T threshold for both leptons. We find this change to be detrimental since, as shown in Fig. 4.6, the signal acceptance is lower for intermediate m_X values. The m_X region at $\sim$1 TeV is particularly relevant, given the bb$\tau\tau$ excess observed by ATLAS [351]. The results imply that the single-τ trigger is less efficient than the legacy triggers, for the three channels. This is surprising for the $\tau_\text{h}\tau_\text{h}$ channel, where the legacy trigger is a double τ trigger. We believe the difference is due to the additional track p_T cut at 50 GeV present only in the single-τ HLT path.

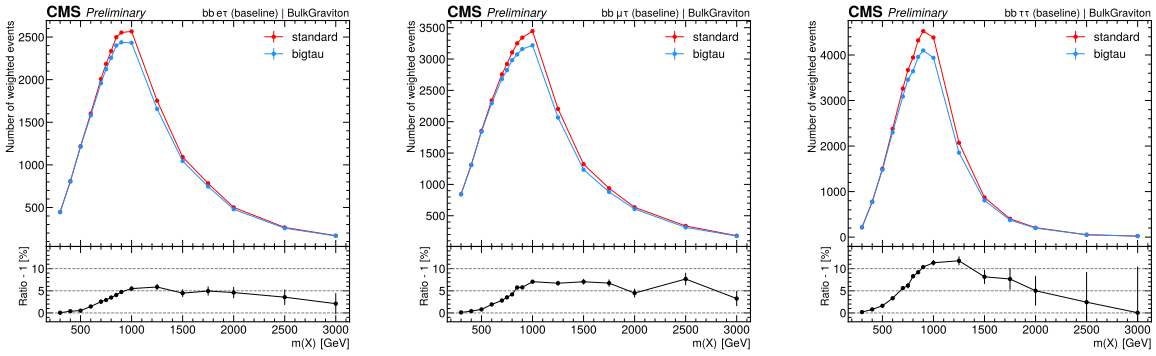

Figure 4.6: Comparison of the weighted number of signal events obtained after applying two scenarios for the trigger regions, for the three analysis channels. In all channels, the region giving precedence to the legacy triggers provides an improvement for intermediate masses. Equivalently, the legacy triggers are more signal efficient than the single-τ trigger. This unexpected result can be due to a p_{T} cut on the tau track, not present in the legacy triggers. The plots refer to the spin-2 hypothesis, but the results for spin-0 are nearly identical.

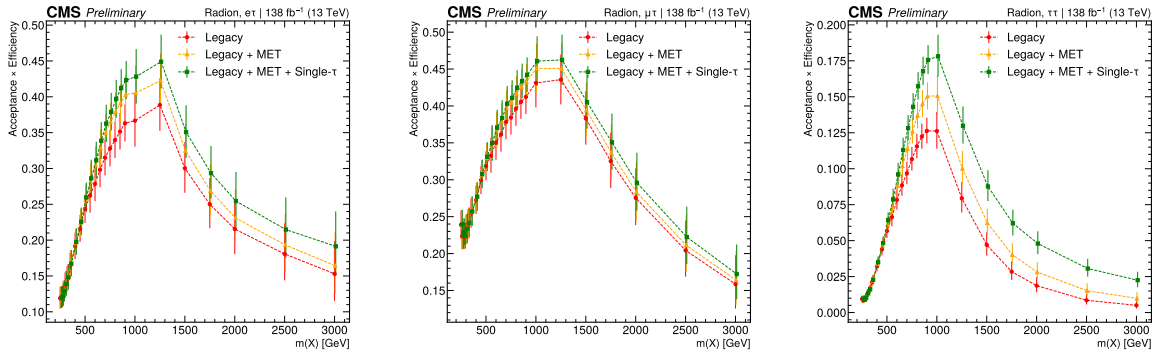

Figure 4.7: Acceptance times efficiency for the trigger selections, in the four analysis channels, considering only events in the trigger regions of the analysis. The denominator refers to the loosest possible kinematic cuts, before performing any of the selection steps detailed in Section 4.4. The improvement brought by the new triggers is clear.

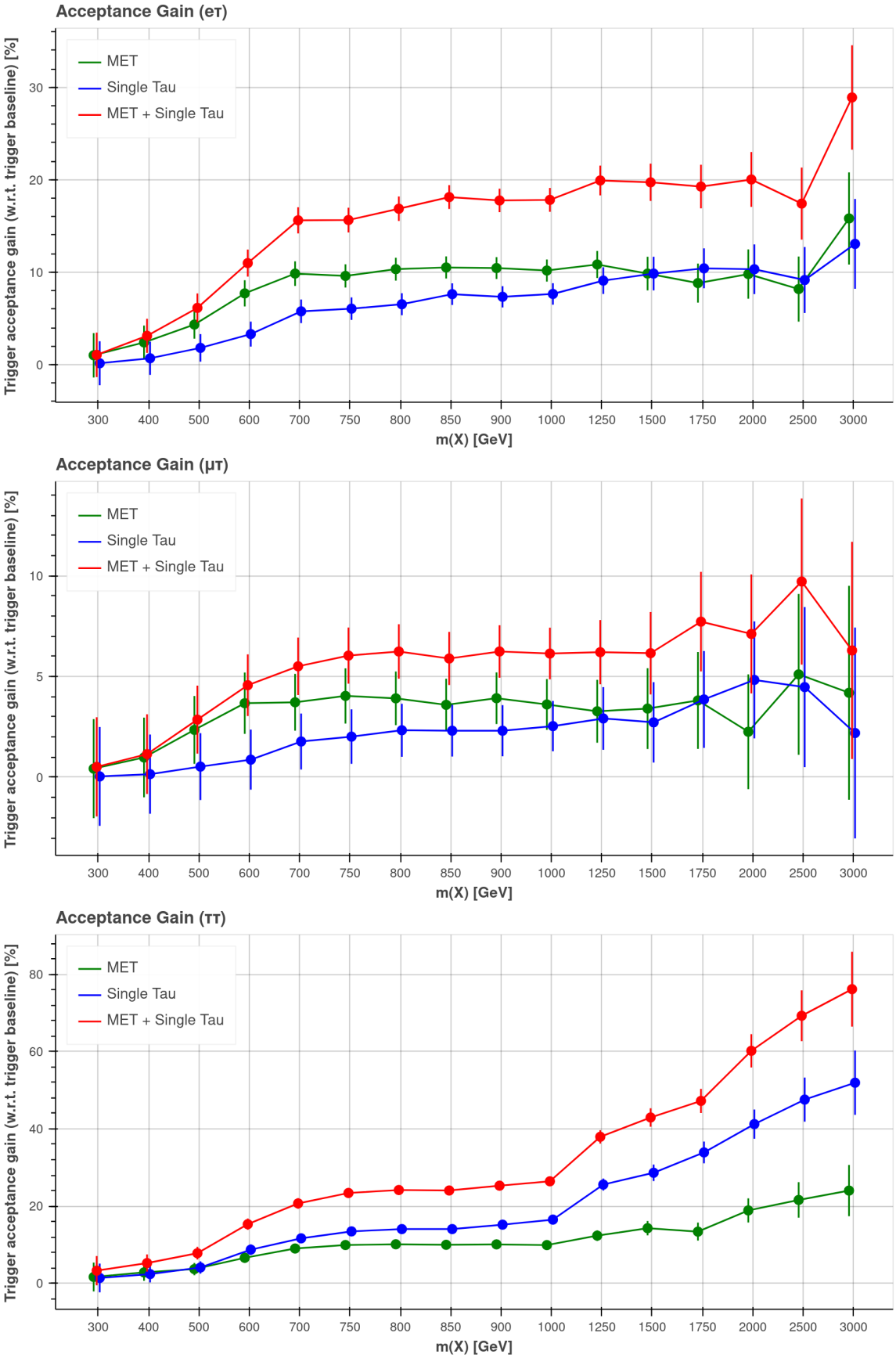

Figure 4.8: Trigger acceptance gain for 2018 in the $\tau_e\tau_h$ (top), $\tau_\mu\tau_h$ (middle) and $\tau_h\tau_h$ (bottom) channels, with respect to the legacy triggers, in percentage, when considering the addition of the $\text{MET}_{\text{no-}\mu}$ and single-τ triggers. The analysis trigger regions are taken into account. The expected statistics improvement is significant, especially for high m_X values. We show the spin-0 hypothesis, but the result with the spin-2 hypothesis is very similar.

Chn.	Region	Trigs.	First Lepton		Second Lepton	$\text{MET}_{\text{no-}\mu}$
$\tau_e\tau_h$	Legacy	cross-$e\tau$	$p_T > -/25/25$, $\lvert\eta\rvert \leq 2.1$	&&	$p_T > -/35/35$, $\lvert\eta\rvert \leq 2.1$	–
		single-e	$p_T > 26/33/33$, $\lvert\eta\rvert < 2.5$	–	–	
	Tau	single-τ	–	–	$p_T > 130/190/190$, $\lvert\eta\rvert \leq 2.1$	–
	MET	$\text{MET}_{\text{no-}\mu}$	–	–	–	$> 150/150/150$
$\tau_\mu\tau_h$	Legacy	cross-$\mu\tau$	$p_T > 21/22/22$, $\lvert\eta\rvert \leq 2.1$	&&	$p_T > 25/32/32$, $\lvert\eta\rvert \leq 2.1$	–
		single-μ	$p_T > 26/29/26$, $\lvert\eta\rvert < 2.4$	–	–	
	Tau	single-τ	–	–	$p_T > 130/190/190$, $\lvert\eta\rvert \leq 2.1$	–
	MET	$\text{MET}_{\text{no-}\mu}$	–	–	–	$> 150/150/150$
$\tau_h\tau_h$	Legacy	di-τ	$p_T > 40/40/40$, $\lvert\eta\rvert \leq 2.1$	&&	$p_T > 40/40/40$, $\lvert\eta\rvert \leq 2.1$	–
	Tau	single-τ	$p_T > 130/190/190$, $\lvert\eta\rvert \leq 2.1$	||	$p_T > 130/190/190$, $\lvert\eta\rvert \leq 2.1$	–
	MET	$\text{MET}_{\text{no-}\mu}$	–	–	–	$> 150/150/150$

Table 4.9: Kinematic definition of the three trigger regions considered in this analysis, for the $\tau_e\tau_h$ channel. The slashes separate the thresholds applied in 2016, 2017 and 2018, in order. Notice that no cross-$e\tau$ was present for 2016. The p_T and $\text{MET}_{\text{no-}\mu}$ values are in GeV units.

The acceptance times efficiency gains from adding the $\text{MET}_{\text{no-}\mu}$ and single-τ triggers are displayed in Fig. 4.7, for the three channels, using the spin-0 samples, and without applying any selection, except for requiring the trigger regions described above. The improvement is obvious, and larger in the $\tau_h\tau_h$ channel. Similarly, the expected relative signal acceptance gains with the new triggers are shown in Fig. 4.8 for the three analysis channels, after applying the analysis selection detailed in Section 4.4, and after taking into account the analysis weights. The graphs show the impact on the spin-0 signal, as a function of m_X. The results are observed to be comparable for spin-2 samples. We require the τs to be matched with generated objects, identified based on their leptonic or hadronic τ decays. As expected, the $\tau\tau$ channel is the one mostly benefiting from the new triggers. The expected signal acceptance increase is very significant, especially at high masses. The difference between the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels comes from the higher p_T cut of muons at 15 GeV, compared to 10 GeV for electrons, which is needed due to the validity range of the respective identification and isolation SFs.

We have additionally tested a non-isolated $p_T(\mu) > 50$ GeV trigger to ensure no additional gain could be obtained in the $\tau_\mu\tau_h$ channel. We applied the trigger for all muons with a p_T above the 50 GeV HLT cut plus the offline threshold. We observe that this alternative trigger brings no increase in signal acceptance, and we therefore do not add it to the analysis trigger strategy. In other words, for the relatively high $p_T(\mu) > 50$ GeV values considered, the HLT isolation does not impact trigger efficiency. This verification was performed after removing events with an isolation larger than 0.15, as part of the analysis selection presented in Section 4.4.

Dataset event overlap

Event overlaps across analysis channels are impossible by construction, since an event can only be assigned to one channel, as described in Section 4.4.1. Overlaps can instead occur across PDs, because an event can, and often does, fire triggers associated to different PDs. Ideally, one would take into account individual event identifiers and make sure each

event is only used once. This method is however impractical, especially when processing samples in parallel. Alternatively, as done in this work, a set of rules is agreed upon:

- An event firing only the $\text{MET}_{\text{no-}\mu}$ trigger can only be accepted if it belongs to the `MET` dataset, and vice-versa, an event from the `MET` dataset can only be accepted if it fired the $\text{MET}_{\text{no-}\mu}$ trigger;
- An event firing only the single-τ trigger can only be accepted of it belongs to the `Tau` dataset, and vice-versa, an event from the `Tau` dataset can only be accepted if it fired the single-τ trigger. This rule is however ignored if the concerned channel is $\tau_{\text{h}}\tau_{\text{h}}$, since in that case the `Tau` dataset is used both for the legacy and for the single-τ trigger regions.

The above rules ensure that no event overlap takes place.

We provide an example to illustrate how the rules work. Take a $\tau_\mu\tau_{\text{h}}$ event with large enough $\text{MET}_{\text{no-}\mu}$. Such event would be present in both `SingleMuon` and `MET` datasets. Given certain p_{T} and η values of the two leptons, it could be assigned to the $\text{MET}_{\text{no-}\mu}$ trigger region (blue region in Fig. 4.5). The rules above guarantee that only the event present in the `MET` dataset would be selected, avoiding event duplication.

4.2.4 A Note on the Inclusion Method

For completeness, we provide a brief description of a general method for estimating SFs when considering the logical-`OR` of any number of triggers. The implementation and testing of the method represented a significant part of my PhD, but ended up not being fully exploited by reasons detailed below. We nevertheless describe the underlying idea, as it can serve as basis for future work.

The probability P that at least one out of N_{items} trigger items i accepts an event j with trigger item bit x is given by [352]:

$$\text{P}_j = 1 - \prod_i^{\text{N}_{\text{items}}} [1 - x_{ij}\epsilon_i(q_j)] \tag{4.5}$$

where ϵ refers to the trigger efficiency, depending on variables q, defined as the number of events passing trigger item i and some reference trigger, divided by the number of events passing the same reference trigger. In Eq. (4.5) we use the fact that the triggers we consider are not prescaled and different data runs have the same conditions. Correlations might exist between triggers, so we rewrite $x_{1j}\epsilon_{1j}x_{2j}\epsilon_{2j} \to x_{1j}\epsilon_{1j}x_{2j}\epsilon_{2|1j}$, and similar for higher order terms with $\text{N}_{\text{items}} > 2$, where $\epsilon_i(q_j) \equiv \epsilon_{ij}$ simplifies the notation, and the symbol "|" stands for a condition probability. Note the commutativity relationship $\epsilon_{1j}\epsilon_{2|1j} = \epsilon_{2j}\epsilon_{1|2j}$. Using Bayes' theorem, we express the efficiency product as an *intersection efficiency* $\epsilon_{1\cap 2j}$, representing the fraction of events passing triggers 1 and 2, plus some reference trigger, when compared to all events passing the same reference trigger. The probability becomes:

$$\begin{aligned}\text{P}_j = {} & x_{1j}\epsilon_{1j} + x_{2j}\epsilon_{2j} + x_{3j}\epsilon_{3j} + \cdots \\ & - x_{1j}x_{2j}\epsilon_{1\cap 2j} - x_{1j}x_{3j}\epsilon_{1\cap 3j} - x_{2j}x_{3j}\epsilon_{2\cap 3j} - \cdots \\ & + x_{1j}x_{2j}x_{3j}\epsilon_{1\cap 2\cap 3j} + \cdots\end{aligned} \tag{4.6}$$

For each intersection, we must define the variables q. We should consider the ones used in the definition of trigger items, since efficiencies will mostly depend on the variables defining their corresponding HLT trigger paths. For instance, $q = p_{\mathrm{T}}(\mu)$ is a reasonable choice for the single-μ trigger. Finally, we can correct our MC distributions event-by-event, using SFs defined as follows:

$$\mathrm{Event}_{\mathrm{MC,\ Corrected}} = \mathrm{Event}_{\mathrm{MC}} \times \frac{\mathrm{P}_{\mathrm{Data}}}{\mathrm{P}_{\mathrm{MC}}} \tag{4.7}$$

This is called the *inclusion method*, and its implementation is challenging. Many intersections might have to be considered. Their calculation is done as a function of different variables q (potentially 1D, 2D or 3D ϵ distributions), for all intersections. We nevertheless note that the terms in Eq. (4.5) with more triggers will naturally be the ones which contain less events overall, and can be dropped when statistics lie below a tunable threshold. Importantly, orthogonal reference triggers have to be found for each intersection, where the latter should have enough statistics. Finally, the inclusion method represents a new approach, thus requiring extensive validation.

The method also brings important advantages. Contrary to common division-like methods, all data are used. It may become very advantageous when significant overlaps exist between triggers. Indeed, most analysis do not consider the logical-OR between triggers due to the increased complexity. I developed an analysis-independent framework to measure all intersection efficiencies. A preliminary test was carried using the $\tau_\mu\tau_{\mathrm{h}}$ channel, exploiting its large statistics. For simplicity, we used the muon p_{T} as q across all intersections. We consider the logical-OR between the single-τ and $\tau_\mu\tau_{\mathrm{h}}$ triggers, and apply the standard analysis selection. We observe that the SFs calculated via Eq. (4.5) provide better Data/MC agreement in the analysis phase-space than the SFs provided centrally by CMS. The improvement is observed only for the variable being used, namely the lepton p_{T}. This encouraging result can be explained, as our custom SFs are calculated with kinematics specific to $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$, while central SFs must be general enough to be used by multiple analyses. For this test, variables not correlated with the p_{T} are not affected.

The inclusion method is not applied to the $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis, despite being fully implemented and despite a preliminary validation. This can be explained by the fact that the complete validation would require more time, which would not be then dedicated to work on the analysis as a whole. There were also significant challenges in defining appropriate reference triggers and datasets for each intersection combination. Furthermore, the correct definition of possibly multi-dimensional bins for each variable q, together with the integration of custom SFs with SFs centrally provided by CMS, was no simple task. Additionally, the project was experimental, and there was no guarantee the final result would converge into robust and usable corrections. For all the reasons above, we decided to create a trigger strategy with non-overlapping regions, as described in Section 4.2.3, so that the logical-OR between all triggers would not have to be considered. However, the developed framework was not left unused. Given its flexibility, it was utilized with one single trigger, namely the MET trigger studies presented in Section 4.7.4. It was also used for on-going discussions concerning the definition of the Run 3 $\mathrm{bb}\tau\tau$ trigger strategy. Estimates have shown an expected acceptance gain of the order of $\sim$10 % when using the full logical-OR for the Run 3 nonresonant $\mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis, compared to the

trigger strategy employed in this work. Exploiting the logical-OR of multiple trigger is thus one of many possible approaches to increase the performance of future HH searches.

4.3 Physics Objects

In this Section, we define the physics objects and related quantities which are used for event selection in Section 4.4. We focus on the objects used by the analysis, which happen to cover a wide variety: electrons, muons, tau leptons and MET for the $H \to \tau\tau$ decay, and jets for the $H \to bb$ decay. Within the CMS Collaboration, dedicated teams called Physics Object Groups (POGs) are responsible for the definition, calibration and validation of specific objects. As such, some selection and calibration procedures are often common to a broad range of physics analyses. In the context of our analysis, we had interactions with the EGamma, Tau, Muon, BTV (b-tagging and vertexing) and JetMET, including during review steps internal to CMS. We follow POG recommendations whenever possible.

4.3.1 Muons

The muon selections applied in this work were originally based on the selections used by the CMS $H \to \tau\tau$ analyses [57–59], but have evolved according to updated analysis needs, further corrections, and sample updates. Muons are required to be reconstructed by the Tracker or the Global muon reconstruction algorithms, which were described in Section 2.4.2. The muon selections later presented in Table 4.12 apply kinematic, isolation and ID thresholds.

One of the most efficient ways to reject backgrounds for lepton candidates is the usage of isolation quantities, a generic class of discriminating variables that are constructed from the p_T sum of PF particles inside a cone around the lepton, relative to its p_T.

$$\mathcal{I}^{\ell}_{\text{rel}} = \frac{\sum p_T^{\text{charged}} + \max\left[0, \sum p_T^{\text{neutral}} + \sum p_T^{\gamma} - \frac{1}{2}\sum p_T^{\text{PU}}\right]}{p_T^{\ell}} \text{ , with } \ell = \mu, e, \tag{4.8}$$

where the $\sum p_T^{\text{charged}}$, $\sum p_T^{\text{neutral}}$, and $\sum p_T^{\gamma}$ are the scalar sums of the transverse momenta of charged hadrons originating from the PV, neutral hadrons and photons, respectively, and the $\sum p_T^{\text{PU}}$ is the sum of transverse momenta of charged hadrons not originating from the PV. Notice that Eq. (4.8) is valid both for muons and for electrons. For the case of muon candidates, the cone is defined with $\Delta R < 0.4$. An additional isolation criteria $\mathcal{I}^{\text{Trk}}_{\text{rel}}$ is used for high p_T muons, using exclusively the energy measured in the tracker, in a $\Delta R < 0.3$ cone. The isolation criteria help disentangling prompt muons from muons originating from heavy-flavor quark decays, and from charged particles in jets misidentified as muons. In this analysis, we required $\mathcal{I}^{\mu}_{\text{rel}} < 0.15$.

Three ID selections for PF muons are centrally recommended by CMS, in the form of `Loose`, `Medium` and `Tight` WPs. In this analysis, the signal muon candidates are required to pass the `Tight` ID, which uses global muons. On top, requirements are applied on the number of hits in the muon chambers, strip tracker, and pixel detectors, plus a quality condition on the χ^2 fit of the track. The selections aim at decreasing the misidentification of other particles as muons. The main background sources are cosmic rays, punch-through hadrons[4] and products of in-flight secondary decays.

[4] Hadrons escaping the calorimeter and leaving energy deposits in the muon system.

A correction factor is applied to MC simulations to take into account differences with respect to data in the isolation and ID efficiencies of muons. These factors are derived from $Z \to \mu\mu$ events selected with a Tag & Probe technique (see the note below), and are provided by the CMS Muon POG. For vetoed muons, $\mathcal{I}^{\ell}_{\text{rel}} < 0.3$ is required separately for the logical-`OR` for two WPs, `Medium` and `Tight`. The logical-`OR` because since `Tight` is not a subset of the `Medium`, and therefore there will be a small number of events that fail `Medium` and pass `Tight`. The veto selection on leptons is described in Section 4.4.1.

An additional WP is available, dedicated to muons with a p_{T} above 120 GeV. Given the fraction of such muons in our analysis, as listed in Appendix A.3, a decision was made not to consider SFs for the high-p_{T} WP. Moreover, binned distributions indicate a good agreement between muon data and MC. Such distributions are shown at the end of this Chapter, in Section 4.8.

Brief note on the Tag & Probe

This technique is often used in HEP to measure the efficiency of a given process directly from data. In a nutshell, for resonances decaying to two leptons, such as $Z \to \mu\mu$ or $Z \to ee$, one of the leptons, the *tag*, is identified with tight selection criteria, while the other lepton, the *probe*, has to satisfy very loose criteria only. The probe is matched to the tag by verifying the invariant mass of the tag + probe system, which should match the mass of the decaying resonance. The tag is used to trigger the resonance, while the probe is exploited to make an unbiased estimate of the efficiency, since very little selections were applied on it. The efficiency is measured by computing the ratio between passing probes and all available probes.

4.3.2 Electrons

Just like for muons, the electron selections applied in this work were originally based on the CMS $H \to \tau\tau$ analyses [57–59]. The standard CMS electron reconstruction algorithm is used [223], combining ECAL and tracker information. Electron candidates are reconstructed from clusters of energy deposits in the ECAL, which are then matched to tracks in the inner silicon tracker, as discussed in Section 2.4.3. Given that the electron reconstruction heavily relies on the deposits in the ECAL, electrons in the the $1.44 < |\eta| < 1.57$ transition region between the ECAL barrel and endcaps are excluded from this analysis, as they are typically of low reconstruction quality.

When defining the isolation for electrons, a $\Delta R < 0.3$ cone is defined around the electron, with respect to the electron direction, and Eq. (4.8) is used. In this analysis, we require $\mathcal{I}^{e}_{\text{rel}} < 0.1$ for electrons in the $\tau_e\tau_{\text{h}}$ channel, and $\mathcal{I}^{e}_{\text{rel}} < 0.3$ for vetoed electrons.

The electron ID uses a multivariate analysis (MVA) which has been updated and improved for CMS Run 2 analyses. The discriminant is based on a BDT [353], combining several observables sensitive to bremsstrahlung along the trajectory of the electron, to PF isolation components and the energy density within the isolation cone, to shower-shape variables, to variables related to electron conversion, and to the matching between the trajectory of the electron and its related clusters. The MVA is trained on all electrons, regardless of whether they pass the trigger requirements or not, and is then tuned for electrons with $p_{\text{T}} > 10$ GeV. Three categories are defined based on electron η: two in the barrel and one in the endcap. Signal electrons require the `Tight` WP, which has a ∼80 % signal efficiency, while veto electrons must pass the `Medium` WP, which has a ∼90 %

efficiency. The full veto selection is described in Section 4.4.1. A correction factor is applied to the MC to take into account differences with respect to data in ID efficiencies of electrons. These factors are derived from $Z \to ee$ events, also selected with a Tag & Probe technique.

The energy of prompt electrons in data does not match what the simulations provide. Energy scaling and smearing corrections are applied to genuine electrons, following the recommendations of the E/Gamma POG. The correction is applied to MC using a variable which represents the combined ECAL and tracker electron energy after applying scale and smearing corrections. The corrections are taken into account both in the selection of genuine electrons and in their veto. They also lead to several systematic shape uncertainties, as described in Section 5.2.2.

4.3.3 Hadronic Tau Leptons

Hadronically-decaying tau leptons are reconstructed using the HPS algorithm, which was described in Section 2.4.5. In our analysis, as well as in others, hadronically-decaying τs are the most important τ decays, given their large BR. There is thus a strong interest in improving the τ reconstruction performance, leading to the implementation of dedicated algorithms. In Run 2, `DeepTau` [354] has demonstrated good performances, and is used in our analysis on top of HPS. As a consequence, there is no need to use more traditional quantities, like the ones described above for muon and electron candidates. The goal of the `DeepTau` algorithm is to disentangle τ_hs from quark- and gluon-initiated jets, and also from electrons and muons, which can occasionally mimic a hadronic tau lepton decay. It uses information from all CMS subdetectors, including variables used by HPS. It also considers information on candidates reconstructed within the HPS tau signal and isolation cones, such as track and cluster properties and kinematics. The algorithm exploits a DNN architecture, where the final discriminants D, against electrons, muons and jets, are computed as follows:

$$y_\alpha = \frac{e^{x_\alpha}}{\sum_\beta e^{x_\beta}} \quad , \quad D_\alpha = \frac{y_\tau}{y_\tau + y_\alpha} \tag{4.9}$$

with $\alpha \in \text{jet}, \mu, e$, and x representing the four output nodes: x_{jet}, x_μ, x_e and x_{τ_h}. The discriminants are also known as `DeepTauVSjet`, `DeepTauVSe` and `DeepTauVSmu` for D_{jet}, D_μ and D_e, respectively. The expected τ_h ID efficiencies are obtained with validation samples. The efficiencies for a particular `DeepJet` WP are defined using genuine τ_hs in a $H \to \tau\tau$ sample, where the τs are reconstructed as τ_hs in a $30 < p_T < 170\,\text{GeV}$ range, and have passed that same WP. As shown in Table 4.10, the efficiencies range from 40 to 98% for jets, from 60 to 99.5% for electrons, and from 99.5 to 99.95% for muons, depending on the WP. The jet $\to \tau$ misidentification rate varies jet p_T and quark flavor. It has been estimated to be 0.43% for a genuine τ ID efficiency of 70%, using simulated W + jets events. The same rate for electrons and muons is 2.60(0.03)% for a genuine τ_h ID efficiency of 80(> 99)%. Significant updates are being put in place, mostly for Run 3 analyses, including using newer and extended data for training, improved training techniques, and optimized hyper-parameter tuning [355].

	VVTight	VTight	Tight	Medium	Loose	VLoose	VVLoose	VVVLoose
D_e	60%	70%	80%	90%	95%	98%	99%	99.5%
D_μ	–	–	99.5%	99.8%	99.9%	99.95%	–	–
D_{jet}	40%	50%	60%	70%	80%	90%	95%	98%

Table 4.10: ID efficiencies of τ_hs for all DeepTau WPs, considering its three classes. The efficiencies are measured with $H \to \tau\tau$ samples for τ_hs in a $30 < p_T < 70$ GeV range [354].

4.3.4 Jets

The CMS PF algorithm creates a list of particle candidates which account for all tracker and muon tracks, and for all energy deposits in the calorimeters above a certain threshold. This information is assembled into jets using the anti-k_T clustering algorithm, described in Section 2.4.6, with distance parameters of 0.4 for AK4 jets and 0.8 for AK8 jets. AK4 jets are required to satisfy $p_T > 20$ GeV and to not overlap with the two leptons from the $H \to \tau\tau$ decay, with an angular selection of $\Delta R(\text{jet}, \tau) < 0.5$. Since tracking information is only available in the central region of the CMS detector and the b-tagging process heavily relies on it, all b-jet candidates are required to have $|\eta| < 2.5$ for the 2017 and 2018 datasets, while $|\eta| < 2.4$ is required in 2016. The difference in η coverage between different years stems from the new CMS pixel detector installed during the Phase 1 upgrade [356]. A more detailed description of jets coming from b quarks and identified as b-jets follows below. The recommended set of jet energy corrections are applied to both AK4 and AK8 jets in data and MC, as described in Section 4.3.4.

Some jets must occasionally be vetoed due to their low reconstruction quality, or because they originate from electronic noise. A PF jet ID criterion is available to CMS analyzers, and all AK4 jets in our analysis are required to pass its Tight WP. The criterion is based on many jet observables, including the multiplicity of charged hadrons, the energy fraction deposited in ECAL by hadrons, and the fraction of hadrons clustered within the jet. The efficiency is around 98/99% or more for all η values, with a background rejection above 98% at $|\eta| < 2.7$.

Jets are also frequently produced by PU, being unrelated to the PV. These jets often result from the overlap of many low-energy jets, being thus broader than PV jets. To avoid such background jets, AK4 jets satisfying $p_T < 50$ GeV are required to pass the Loose WP of a dedicated PU jet discriminant. The discriminant uses a BDT to find an optimized decision boundary, using information related to jet shape, object multiplicity and compatibility with the PV.

Jets from b-quarks originating from the decay of high p_T Higgs bosons are often close enough to be merged into a single large radius jet by the anti-k_T algorithm, forming an AK8 jet. In our analysis, the Graph Neural Network (GNN)-based PNet algorithm [266] is used to discriminate $H \to bb$ decays from the multijet background, as detailed below. We require AK8 jets to satisfy $p_T > 250$ GeV, and to not overlap with the two analysis leptons: $\Delta R(\text{jet}, \tau) < 0.8$. The jets must also have a SoftDrop mass above 30 GeV, where SoftDrop [357] is a boosted jet grooming algorithm which removes soft and wide-angle radiation, aiming at mitigating the effects from contamination of ISR, UE and PU.

Identification of B-jets

Jets originated by the hadronization of b quarks distinguish themselves from other jets, inasmuch as they contain particles known to be relatively long-lived. Such b mesons and hadrons can thus decay with a displacement of a few millimeters with respect to the PV, defining the so-called *secondary vertices*. Additionally, b hadrons decay into electrons or muons with a probability of ∼20 %. Distance parameters and displaced leptons can thus be exploited for discriminative purposes [358].

During Run 1, the b-jet reconstruction algorithms available within CMS worked by manually building discriminative variables. The most advanced, the Combined Secondary Vertex (CSV) algorithm, used the secondary vertex mass and the number of tracks in a jet, among other variables. Deep learning techniques first appeared in Run 2, starting with `DeepCSV` [359], and later `DeepJet` [360, 361], which is based on Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). Further improvements, particularly the widespread adoption of GNNs, have lead to PNet [266], and finally to Particle Transformer (ParT) [120]. The latter exploits state-of-the-art transformer technology [362], and should start being used in Run 3.

In our analysis, AK4 jets originating from b quarks are identified using the `DeepJet` algorithm. In order to separate b-jets from other jets, `DeepJet` combines secondary vertex properties, track-based variables and PF jet constituents (neutral and charged candidates) in a DNN. It then classifies jets into three main categories: light, charm or b jets. Bottom jets are further categorized into jets with at least two b hadrons, jets with exactly one b hadron decaying hadronically, and jets with exactly one b hadron decaying leptonically. Light jets are also split into quark and gluon jets. In this Thesis, whenever we refer to `DeepJet`, we imply its b jet discrimination capabilities compared to all other jet classes. Comparisons between other specific jet types are referred to as `DeepJetCvsB` and `DeepJetCvL`, when discriminating c jets against b jets, and c jets against light jets, respectively. The latter are used as input features for the $m_{\tau\tau}$ regression in Section 4.5.

Year	**DeepJet**			**PNet**		
	WP	Eff. [%]	Cut	WP	Eff. [%]	Cut
2016	`Loose`	86.3	0.0408	`LP`	80	0.9137
	`Medium`	71.4	0.2489	`MP`	60	0.9735
	`Tight`	54.7	0.8819	`HP`	40	0.9883
2016APV	`Loose`	87.3	0.0508	`LP`	80	0.9088
	`Medium`	73.3	0.2598	`MP`	60	0.9737
	`Tight`	57.5	0.8819	`HP`	40	0.9883
2017	`Loose`	91.0	0.0532	`LP`	80	0.9105
	`Medium`	79.1	0.3040	`MP`	60	0.9714
	`Tight`	61.6	0.7476	`HP`	30	0.987
2018	`Loose`	91.5	0.0490	`LP`	80	0.9172
	`Medium`	80.7	0.2783	`MP`	60	0.9734
	`Tight`	65.1	0.7100	`HP`	40	0.988

Table 4.11: `DeepJet` and mass decorrelated PNet $X \to b\bar{b}$ thresholds for different data-taking periods, with associated WPs and Run 2 $H \to bb$ signal jet efficiencies. LP, MP and HP refer to Low, Medium and High purities, respectively.

AK8 jets originating from merged $H \to bb$ decays are instead tagged by the PNet algorithm. This algorithm is able to identify hadronic decays of highly Lorentz-boosted top quarks and W, Z, and Higgs bosons, and classify different decay modes, such as $b\bar{b}$, $c\bar{c}$ or $q\bar{q}$ pairs. The tagger is trained with $X \to b\bar{b}$, $X \to c\bar{c}$ and $X \to q\bar{q}$ signal jets, where X is a spin-0 scalar, and with QCD multijet background samples. It accordingly outputs four discriminant scores, each representing the probability P for one of the four following processes to occur: $X \to b\bar{b}$, $X \to c\bar{c}$, $X \to q\bar{q}$ and QCD. We use a mass-decorrelated version of PNet. The decorrelation is achieved by reweighting the training samples into uniform jet p_T and jet `SoftDrop` mass distributions. The $X \to b\bar{b}$ discriminant is given by:

$$\frac{P(X \to b\bar{b})}{P(X \to b\bar{b}) + P(QCD)} . \tag{4.10}$$

Three WPs are defined with $H \to bb$ signal jets at efficiencies of 40%, 60%, and 80%: Low Purity (LP), Medium Purity (MP), and High Purity (HP), respectively. In order to select the most performant WP, the full analysis workflow is run once per WP, and the LP WP is found to provide the most stringent results. It is however important to note that discrepancies between data and MC require the application of dedicated SFs to all jets passing the PNet WPs. AK8 analysis jets must thus be corrected, in a procedure described in Section 4.7.7. The thresholds on the `DeepJet` and PNet discriminant values, and corresponding efficiencies, are listed in Table 4.11.

Jet Energy Scale and Resolution Corrections

The measured jet energy can significantly differ from the underlying true hadron energy. Differences can arise due to detector noise, PU or a non-linear calorimetric response. The precise understanding of jet energy corrections (JECs), scales and resolutions, is of crucial importance for multiple analyses, also entering as an important component in their systematic uncertainties. The energy of jets must therefore be appropriately corrected, in order to match the true particle-level deposited energy [363, 364]. In Fig. 4.9 we illustrate the approach adopted by CMS in Run 2. It consists of sequential steps, where each step is responsible to independently correct a different effect. Each data-taking period has its own set of corrections. The first step addresses the spurious energy deposits from PU interactions. For each type of PF candidate an offset energy is subtracted from the jet energy. In the second step, detector response corrections are applied, in order to fix its non-uniformity across the jet p_T and η phase-space. Next, remaining differences between data and MC are corrected by accounting for PU effects, which also depend on the p_T and η of jets. Finally, optional flavour dependent corrections can be applied. For all jet types, the energy scale uncertainties are smaller than 3% for $p_T > 50\,\text{GeV}$ in the $|\eta| < 3.0$ region, increasing to 5% for $3.0 < |\eta| < 5.0$.

Since measurements show that the jet energy resolution in data is worse than in the simulation, resolution corrections must be applied to MC jets. The latter are smeared to describe the data. The smearing procedure uses a "hybrid" approach, recommended to all CMS analyses, and composed of two methods. If a matched generator-level jet exists, then the four-momentum of the corresponding reconstructed jet is rescaled, with a factor which depends on the p_T of the reconstructed and generated jet:

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1)\,\frac{p_T - p_T^{\text{Gen.}}}{p_T} \tag{4.11}$$

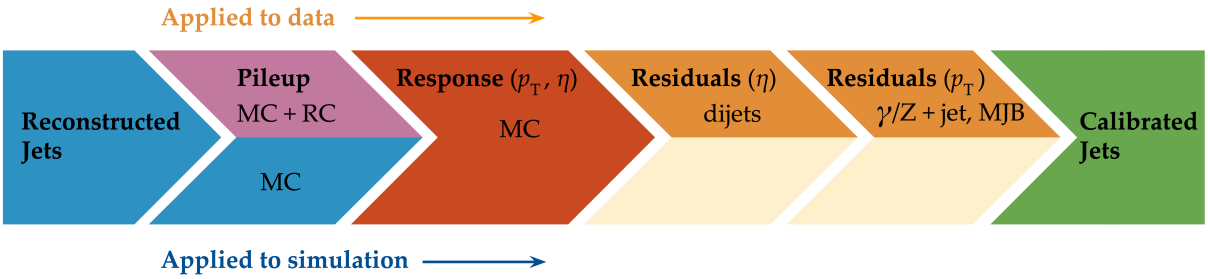

Figure 4.9: Illustration of the jet energy correction stages that must be sequentially applied in order to obtain a calibrated jet, as done for Run 2 in CMS. Taken from [364].

where s_{JER} is the data-to-simulation core resolution scale factor. If the jet was not matched (and thus $p_{\text{T}}^{\text{Gen.}}$ is not available), then a stochastic smearing is applied, performing the four-momentum rescaling using a different factor:

$$c_{\text{JER}} = 1 + \mathcal{N}(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)} \tag{4.12}$$

where σ_{JER} is the relative p_{T} resolution in simulation, and $\mathcal{N}(0, \sigma)$ denotes a random number sampled from a normal distribution with zero mean and standard deviation σ. The resolution corrections are computed after applying the above jet energy corrections. The data/MC SFs usually vary between 1 and 1.2, but are larger in the transition region between the endcaps and the forward detectors. No significant dependencies on the p_{T} and η of the jets are observed, except in the transition region [365].

4.3.5 Missing Transverse Energy

As discussed in Section 2.4.7, MET is the negative vector sum of all PF reconstructed particles in an event. Despite being well defined, the "raw", uncorrected MET is systematically different from the transverse momentum actually carried by invisible particles. This happens due to a variety of detector effects, most notably the non-compensating nature of the CMS calorimeters, which was explained in Section 2.4.1, and due to detector misalignments. In this analysis, we apply corrections as instructed by the CMS JetMET POG, turning the measured $\vec{p}_{\text{T}}^{\text{miss}}$ into a better estimate of the "true" MET.

Measurements show that the jet energy resolution (JER) in data is worse than in the simulation. As discussed in Section 4.3.4, jets in simulation should thus be smeared to achieve a better agreement with data. Given that jets are one of the building blocks of MET, their smearing should be propagated to the MET. The corrections replace the vector sum of transverse momenta of particles clustered as jets by the vector sum of the transverse momenta of the jets to which JECs are applied. Corrections are applied to AK4 jets.

It has been observed that uncorrected MET features a modulation in the azimuthal ϕ coordinate. The modulation roughly follows a sinusoidal curve with a 2π period. The distribution of true MET should instead be independent of ϕ because of the collisions' rotational symmetry along the transverse axis. The modulation can be due to anisotropic detector responses, to inactive calorimeter cells and/or tracking regions, to the detector misalignment, and even to the displacement of the beam spot. The amplitude of the modulation increases roughly linearly with the number of PU interactions. In this analysis, we reduce the amplitude of the ϕ modulation by shifting the origin of the x and y

coordinates in the transverse momentum plane, as a function of the run number and of the number of PVs.

We also apply MET quality filters provided by the JetMET POG, in order to improve the quality of the reconstructed MET:

- Events where the PV is not of good quality are rejected.
- A beam halo filter is used, to reduce the non negligible probability of high-energy halo muons to interact in the calorimeters. Such interactions can create clusters of up to several hundreds of GeV.
- Events with problematic dead cell TP energy recovery are removed.
- Events where a large nonphysical MET is erroneously reconstructed due to the presence of additional muons are rejected.
- Additional filters are applied to reject events with high HCAL or ECAL noise.

4.4 Selection and Categorization

The observation of $X \to HH \to bb\tau\tau$ naturally requires the capability to precisely reconstruct the two pairs of b jets and tau leptons. Having introduced all single physics objects included in the $bb\tau\tau$ topology, we can now describe in detail the full event selection, which aims at providing the subset of events processed that match the signal of interest. However, the selection should not be too strict, or equivalently lead to a very high $bb\tau\tau$ purity; that could hinder analysis steps to come, and specifically affect the final discriminant, by reducing its training data size. Given the rarity of HH processes, it is imperative to measure as many $HH \to bb\tau\tau$ events as possible; the goal of the selection is to maximize signal acceptance.

The selection starts from the data passing the analysis triggers, and is performed in various stages, as illustrated in Fig. 4.3. In the following, we detail the selection cuts applied to leptons and b jets in Sections 4.4.1 and 4.4.2. All events that satisfy the selection on tau leptons and b-quark pairs are said to belong to the **baseline** selection. We then proceed by introducing the categorization step in Section 4.4.3 and the invariant mass cut in Section 4.4.4. Finally, we provide the definitions of the two CRs used in the analysis, which are used to train the discriminant and are visualized in Section 4.4.5.

4.4.1 Tau Lepton Pair

This step aims at identifying the visible decay products of one of the 125 GeV Higgs bosons decaying to a τ pair. Selected signal events are required to have at least one τ candidate decaying hadronically and that has been reconstructed by the HPS algorithm. The `DeepTau` discriminant [354] identifies τ_hs, distinguishing them from jets, electrons and muons. Due to charge conservation, HPS decay modes with two prongs are explicitly rejected. Furthermore, all events assigned the $h^{\pm}\pi^0\pi^0\nu_\tau$ topology are analyzed as though they were $h^{\pm}\pi^0\nu_\tau$. This happens because the former is much rarer than the latter, and the HPS algorithm is not tuned to reconstruct decays containing two π^0s. This same assignment is also taken into account when applying `DeepTau` SFs, as described in Section 4.7.5 The identification requirement applied to τ_hs is the `Medium` WP of the `DeepTauVSjet` algorithm, the `Tight DeepTauVSmu` WP, and the `VVLoose DeepTauVSe`

WP. The choice of WPs is related to the WPs used by the Tau POG to derive SFs for genuine τ_hs, which have a dependency on `DeepTauVSmu` and `DeepTauVSe`. The reasoning is to use corrections as close to POG recommendations as possible, in order to reach a good data/MC agreement in all relevant parts of the parameter space.

Object	p_T **[GeV]**	$\|\eta\|$
Electron	> 10	< 2.5
Muon	> 15	< 2.4
Tau lepton	> 20	< 2.3

Table 4.12: List of the minimum p_T and maximum $|\eta|$ thresholds considered for electron, muon and tau lepton candidates. These value are a consequence of hard limits imposed by object-specific SFs,as provided by the relevant POGs. Leptons firing triggers with tighter p_T cuts must apply the thresholds according to Eq. (4.4).

All events lying above the minimum p_T or below the maximum $|\eta|$ values reported in Table 4.12 are considered, as long as they satisfy the triggers described in Section 4.2, and as long as the offline objects are geometrically matched to the online HLT objects firing the paths. The values in Table 4.12 come from hard limits imposed by the isolation and ID trigger SFs used in the analysis. We look for lepton candidates passing simple kinematic, isolation and ID criteria, as summarized in Table 4.13, where the Δ_{xy} and Δ_z distances are also taken into account, measuring the compatibility with the PV. Additionally, the two objects in a pair should have opposite charges, and should pass p_T thresholds based on the triggers they fired. An exception occurs for the $\text{MET}_{\text{no-}\mu}$ trigger, since it imposes no p_T cut, and can thus explore the full phase-space presented in Table 4.12.

The analysis channels are defined according to specific priority rules. An event is classified as $\tau_\mu\tau_h$ if a single muon candidate passes all criteria. If two muons pass the criteria, where the second muon can have a looser isolation than the first, the event becomes $\tau_\mu\tau_\mu$. In case no muon is selected, we check whether any electron passing the criteria was found. In case a single electron candidate is present, the event is associated the $\tau_e\tau_h$ channel. If two electrons are present, the event would be assigned the $\tau_e\tau_e$ channel, but it is instead discarded, as we are not taking this channel into consideration. If neither muons or electrons are found, the event is classified as $\tau_h\tau_h$, as long as a two τ_hs are present, and all pair candidates are kept. Within each pair, leptons are ordered as follows, depending on the channel:

- $\boldsymbol{\tau_e\tau_h}$ **and** $\boldsymbol{\tau_\mu\tau_h}$: The first position is assigned to the leptonic leg, either a muon or an electron.

- $\boldsymbol{\tau_h\tau_h}$: All selected pairs are first sorted according to the `DeepTau` score of the first lepton. If the two first legs have the same isolation, the highest p_T leg is used to order the pairs. If the p_T is also identical, i.e. the pairs share the same first leg, then the pair with the most isolated second leg is preferred. If ambiguity is still present, priority is given to the pair with the highest p_T of the second leg.

This channel classification strategy was chosen because it was seen to maximize signal purity, while removing event overlaps among the three different final states.

Some events can include multiple leptons of the same kind, and their choice becomes a matter of some ambiguity. We mentioned above that $\tau_\mu\tau_h$ events take precedence over $\tau_e\tau_h$ events; no ambiguity is present at the decay channel level. However, nothing forbids

an event to include, on top of a τ_h, two muons or two electrons, as long as they satisfy the requirements in Table 4.13. For those cases, there would be multiple ways to choose the "correct" lepton, i.e. the lepton truly associated to the Higgs boson decay. To simplify the selection and avoid taking the wrong decision, the *third lepton veto* is implemented, were events with a third lepton are rejected. This veto also helps removing background events with two leptons and a fake τ_h, as for instance di-boson processes, or fully leptonic decays of $t\bar{t}$. The third lepton selection criteria are very similar to the ones for the "standard" leptons, just slightly looser, in order to remove most situations where ambiguity might be present.

Chn.	$\lvert\eta_1\rvert$	**ID$_1$**	**Iso$_1$**	$\lvert\eta_2\rvert$	**ID$_2$**	**Iso$_2$**	Δ_{xy}	Δ_z	$\Delta R(\ell_1, \ell_2)$
$\tau_e\tau_h$	< 2.5	`Tight`	< 0.1	< 2.3	`DeepTau`	`DeepTau`	< 0.45	< 2.0	> 0.4
$\tau_\mu\tau_h$	< 2.4	`Tight`	$< 0.15^\dagger$	< 2.3	`DeepTau`	`DeepTau`	< 0.45	< 2.0	> 0.4
$\tau_h\tau_h$	< 2.3	`DeepTau`	`DeepTau`	< 2.3	`DeepTau`	`DeepTau`	–	< 2.0	> 0.4
$\tau_\mu\tau_\mu$	< 2.4	`Tight`	$< 0.15^\dagger$	< 2.4	`Tight`	< 0.30	< 0.45	< 2.0	> 0.4
3rd e	< 2.5	`Medium`	< 0.3	–	–	–	< 0.45	< 2.0	–
3rd μ	< 2.4	`Medium`	$< 0.3^\dagger$	–	–	–	< 0.45	< 2.0	–

Table 4.13: Selections defining the analysis channels, including the third lepton vetos. Besides the cuts listed in the table, opposite charges are requested between any two leptons, and the p_T thresholds follow Eq. (4.4), except when the the leptons fired only MET, in which case no specific cuts are requested, and Table 4.12 is instead considered. The $|\eta|$ cut on the second lepton in $\tau_e\tau_h$ and $\tau_\mu\tau_h$ can be 2.1 if the event only fires the cross trigger. With the exception of not applying a Δ_{xy} cut on τ_hs, the Δ_{xy} and Δ_z cuts below are identical for any two objects in a pair, are are provided in mm. The isolation cuts marked with † are applied to the isolation considering all PF muon candidates, but also to an isolation considering only muon tracks.

We remind the reader that the phase-space is always divided into three regions, according to the triggers being used, as described in Section 4.2.3.

4.4.2 B Quark Pair

We now turn to the $H \to bb$ process, where jets coming from a Higgs boson must be selected, following at least one of the below criteria:

- Two AK4 jets with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.5$ for 2017 and 2018, with a $\Delta R > 0.5$ distance between each jet and each selected τ candidate. For 2016, $|\eta| < 2.4$ is instead used.
- One AK8 boosted jet, with the distance between the jet and both selected τ candidates of $\Delta R > 0.8$.

Additionally, the $H \to bb$ selection is improved by applying a discrimination algorithm to identify b jets, dubbed `HH-bTag`. The algorithm is based on studies done in the context of Ref. [110]. At its core, a DNN architecture assigns a score between 0 and 1 to all possible AK4 b jet candidates, and the two jets with the highest score are selected. The score stands for how confident the model is that a particular b jet originated from a $H \to bb$ decay. The architecture of the model is characterized by the following elements:

- Five concatenated Long Short-Term Memory (LSTM) layers [366], using a sigmoid activation function. After each layer, a batch normalization step is applied.

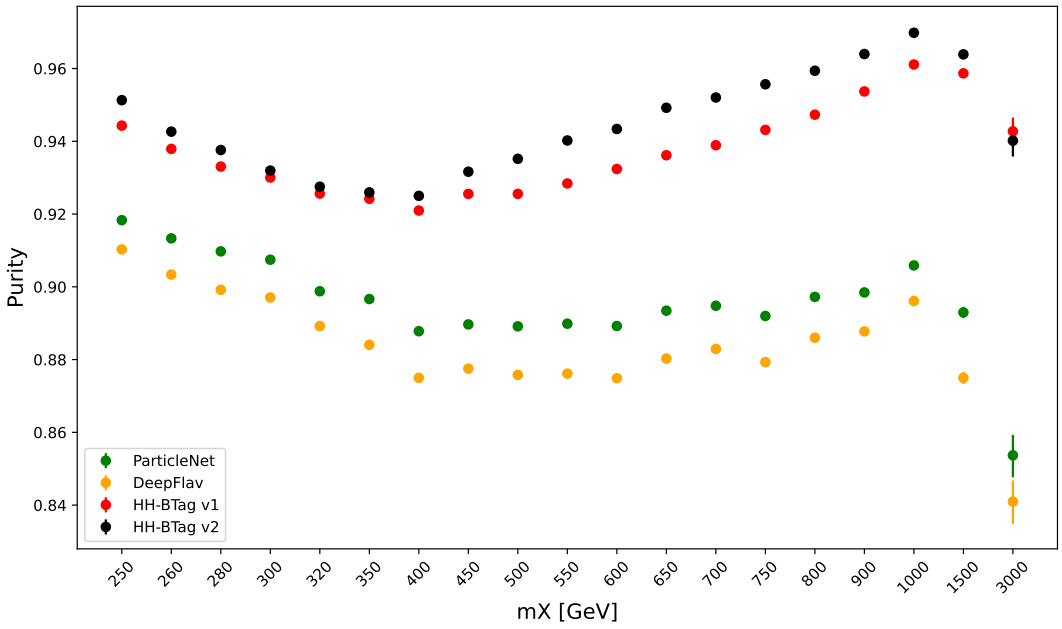

Figure 4.10: Comparison of the purity for the spin-0 bb$\tau\tau$ resonant signal, as a function of m_X, between the original (`v1`) and retrained (`v2`) `HH-bTag` versions, and two other CMS algorithms, namely PNet and `DeepJet` (also known as `DeepFlavour`). Purity is defined in Eq. (4.13). The retrained algorithm clearly provides the best performance across the entire mass range.

- Ten Time Distributed Dense (TDD) layers [367], with sigmoid as activation function. A batch normalization step is also applied between each layer.
- Binary cross-entropy as the loss function [368], minimized with the `AdamW` algorithm [369].
- A final TDD layer with only one unit and with a sigmoid as activation function, providing the final score.

The algorithm has been retrained with UL data, using both nonresonant and resonant ggF bb$\tau\tau$ signal samples, in order to provide the best performance possible. The algorithm is trained via cross-validation with two folds, a procedure described in Section 5.1. A total of 14 input features are considered, including the score of `DeepJet`, several kinematic variables, and MET. Categorical variables are also used, such as the data-taking year and decay channel. The performance of the algorithm is evaluated and compared to the previous version, and also to other b jet ID algorithms, as shown in Fig. 4.10. The retrained version provides better results than all alternatives, across the full probed m_X range. The so-called *purity* is used as comparison metric:

$$\text{purity}^{\text{classifier}} = \frac{\text{N}^{\text{classifier}}_{\text{true}}}{\text{N}^{\text{classifier}}} \,, \tag{4.13}$$

where $\text{N}^{\text{classifier}}_{\text{true}}$ is the number of events in which the selection of the b jet pair candidate matched the ground truth's definition, and $\text{N}^{\text{classifier}}$ represents the total number of events where a candidate is reconstructed. The matching uses a $\Delta\text{R} < 0.5$ cut around the direction of the reconstructed b jet.

For AK8 jets, the `HH-bTag` algorithm is not employed; the PNet discriminant [266] is used instead, as explained in Section 4.4.3.

4.4.3 Categorization

With the selection fully defined, we proceed to split selected events in orthogonal categories. The categories are meant to boost the analysis' sensitivity. That happens because, during the extraction of the final results, each category is separately fitted. In the limit where all categories have identical event topologies, the result of the fit should be equal to a fit performed on the selected events taken together. If the categories possess some distinguishing features, as is the case in our analysis, a per-category fit improves the result, since individual background sources can be better constrained.

The categorization scheme follows the angular radius parameters adopted in the reconstruction of jets within CMS.

- $\Delta\mathrm{R}(\mathrm{b},\mathrm{b}) > 0.8$: each b-quark is reconstructed as a AK4 jet;
- $0.4 < \Delta\mathrm{R}(\mathrm{b},\mathrm{b}) < 0.8$: the two b-quarks are reconstructed both as two AK4 jets and as one large AK8 jet;
- $\Delta\mathrm{R}(\mathrm{b},\mathrm{b}) < 0.4$: the two b-quarks are reconstructed as an AK8 jet only.

The so-called *resolved* topologies refer to the first scenario, while *boosted* topologies refer to the third. The events matching the second scenario can be classified as one or the other, depending on the analysis' strategy. In this work, two resolved categories and one boosted category are defined, as follows:

- Events with a reconstructed AK8 jet having $m_{\mathrm{SoftDrop}} > 30\,\mathrm{GeV}$, $p_{\mathrm{T}} > 250\,\mathrm{GeV}$, $\Delta\mathrm{R}(\mathrm{jet},\tau) > 0.8$ for both τs, and a PNet discriminant score passing the LP WP are assigned to the **boosted** category.
- Events with two AK4 jets and no AK8 jets, where only one of its b jet candidates passes the `Medium` WP of `DeepJet`, are assigned to the **res1b** category.
- Events with two AK4 jets and no AK8 jets, where both its b jet candidates pass the `Medium` WP of `DeepJet`, are assigned to the **res2b** category.

The **res2b** category provides the most sensitive measurements for resonance masses below ∼700 GeV, while the **boosted** category drives the analysis sensitivity for resonance masses above ∼700 GeV. The categories are attributed following a specific precedence order, as illustrated in Fig. 4.11, where **res2b** has precedence over the other categories. Other possibilities were tried, namely giving the top precedence to the **boosted** category, but the choice used here provided the best final results.

4.4.4 Invariant Mass Cut

Events classified as resolved are required to have reconstructed visible masses of the b and τ pairs within a rectangular window. The goal is to maximize signal acceptance, requiring at least 98% of signal presence. These cuts remove the tails of the mass spectrum, and potential outliers, easing the task of discriminants further down in the analysis chain. The mass cut also allows to define CRs with low signal contamination, useful to assess the proper modeling of some of the main analysis' backgrounds.

In order to define the mass window interval, the ggF spin-0 and spin-2 signal samples are utilized. The samples are merged, considering all mass and spin configurations at once. The $\tau_e\tau_\mathrm{h}$, $\tau_\mu\tau_\mathrm{h}$ and $\tau_\mathrm{h}\tau_\mathrm{h}$ channels have been separately considered to estimate the

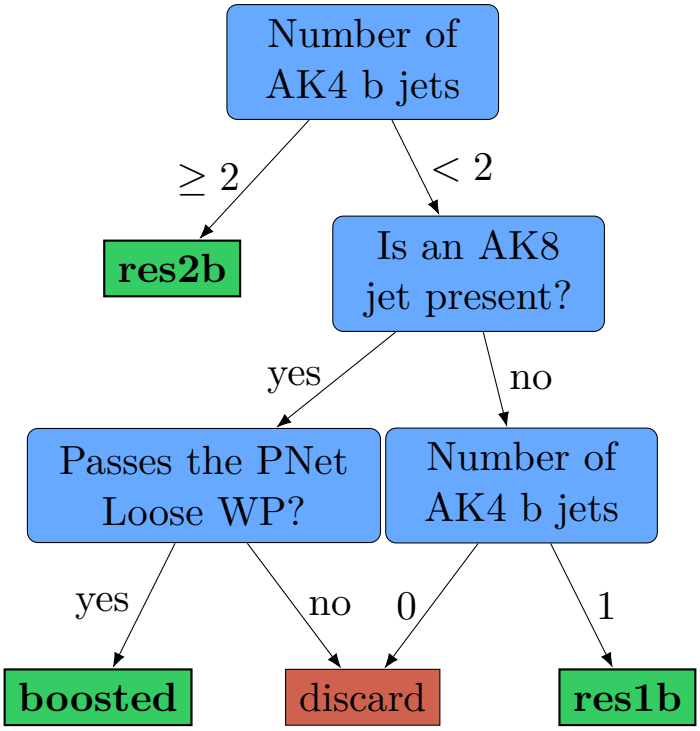

Figure 4.11: Flowchart of the analysis categorization. Three categories are defined: **res1b**, **res2b** and **boosted**, based on the number of AK4 and AK8 jets, and on the `Loose` WP of the PNet discriminant.

rectangular cuts. The event selection, in addition to the **baseline** requirements, includes the following conditions:

- presence of two resolved b jet candidates passing the `Loose` b-tag WP;
- b jets matched to a generated b quark.

The maximum and minimum values of the $m_{\mathrm{b\bar{b}}}$ and $m_{\tau\tau}$ visible masses are calculated from their 99.5% and 0.5% quantiles. To define the mass window interval, the limits for $m_{\mathrm{b\bar{b}}}$ are calculated first. An additional requirement is then added while computing the limits for $m_{\tau\tau}$: to consider $m_{\mathrm{b\bar{b}}}$ only within the limits calculated in the previous step. The values obtained are:

- visible $m_{\tau\tau}$ between 20 GeV and 130 GeV;
- $m_{\mathrm{b\bar{b}}}$ between 40 GeV and 270 GeV.

The cuts ensure a very high signal efficiency. It has been shown that tighter cuts, although providing a larger S/B ratio, result in a poorer limit when compared to a DNN discriminant [111]. The boundaries of the cut are thus kept very loose, focusing on acceptance and not purity. The two-dimensional distribution of $m_{\mathrm{b\bar{b}}}$ versus $m_{\tau\tau}$ is displayed in Fig. 4.12, where a red rectangle highlights the computed mass interval. We note that the visible mass signal distributions are similar for all mass points, including the ones not displayed.

4.4.5 Control Regions

It is often useful to understand if single background sources are being correctly modeled. Additionally, one often wants to perform cross-checks on data outside the signal region

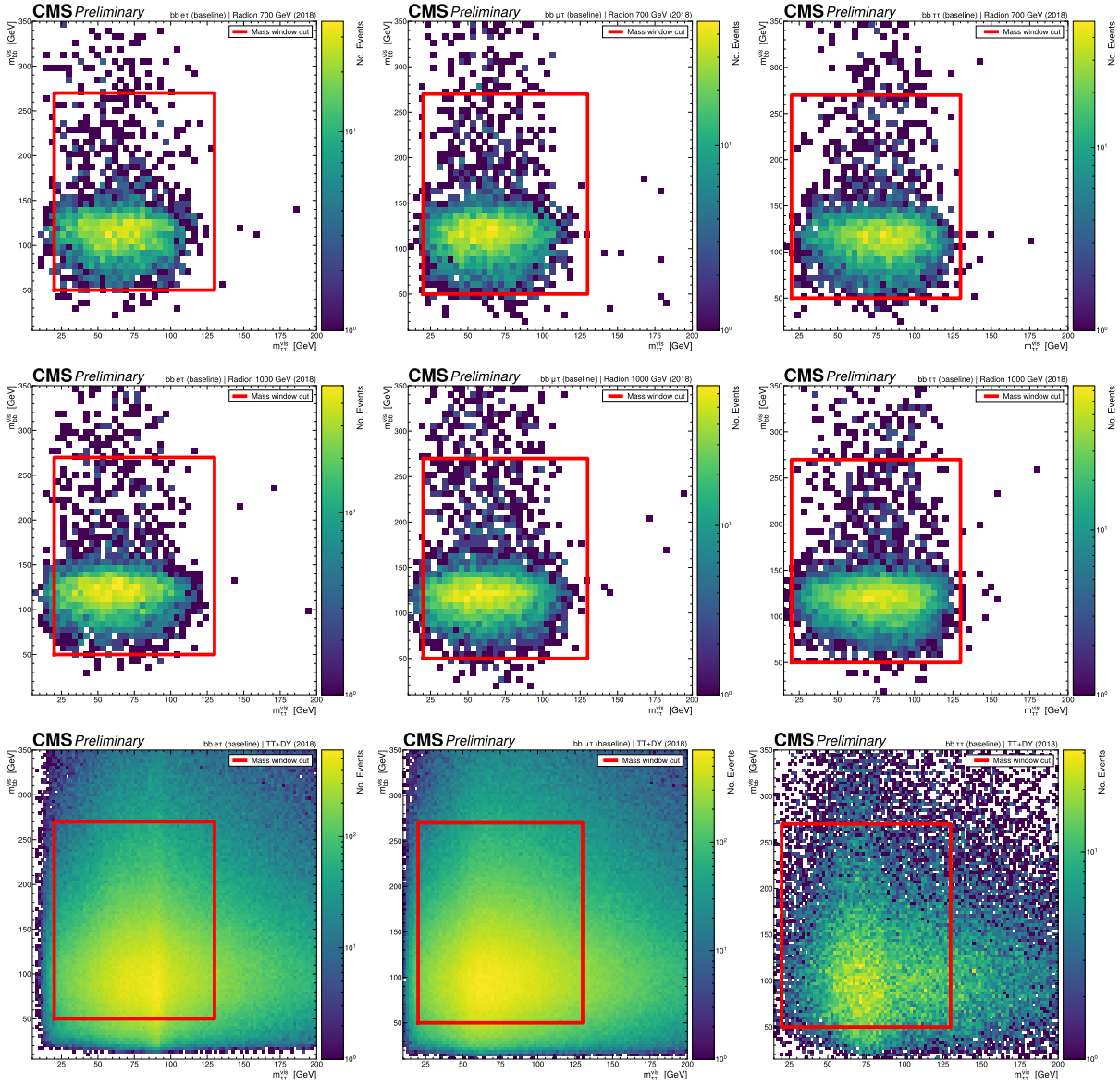

Figure 4.12: Illustration of the rectangular window mass cut (in red) on top of signal (700 GeV and 1 TeV for, respectively, the top and middle rows) and background (bottom row). The three analysis channels are represented in the left, middle and right columns.

(SR), to avoid creating biases, but with a similar topology. An example would be to determine correction weights using a data-driven approach. CRs are thus introduced, one focused on $t\bar{t}$ and another designed to be DY-dominated:

- **$t\bar{t}$ CR**: Events satisfy the **baseline** selection, and must be associated to the $\tau_e\tau_h$ or $\tau_\mu\tau_h$ channels. They additionally do not include any AK8 jets, thus having a resolved topology. The two AK4 jets must pass the `Medium DeepJet` WP, just like events in the **res2b** category. This requirements increases the fraction of $t\bar{t}$ events, compared to other backgrounds associated with lighter jets. Finally, the mass of the $\tau\tau$ system must lie above 130 GeV, in order for the CR to be orthogonal to the rectangular mass cut, and to remove DY events.

- **DY CR**: Events satisfy the **baseline** selection, and must be associated to the $\tau_\mu\tau_\mu$ channel. They additionally do not include any AK8 jets, thus having a resolved topology. Either one or two AK4 jets must pass the `Medium DeepJet` WP, mimicking the **res1b** or **res2b** categories, depending on the use case. The fraction of DY events decreases with **res2b** cuts, and $t\bar{t}$ increases. Finally, an invariant mass cut is requested around the mass of the Z boson, to remove any non-DY background source. Notice that there is no need to ensure orthogonality with respect to the rectangular mass cut, since the $\tau_\mu\tau_\mu$ channel is not added to the final analysis fit.

In the $X \to HH \to bb\tau\tau$ analysis, CRs regions are used to inspect the agreement of data and MC in different kinematic and categorical distributions. In particular, they are quite useful to determine whether additional corrections are required for specific background sources.

4.5 Mass Regression of the Tau Lepton Pair

A sizable fraction of the energy in the $H \to \tau\tau$ system is lost from a detector's point of view, due to the neutrinos in the decays of the tau lepton. Experimentally, the neutrinos are measured as MET. Such a signature severely limits the reconstruction of the invariant mass of the di-τ system, $m_{\tau\tau}$, which has an impact in the sensitivity of any analysis relying on tau lepton pairs. In our analysis, the correct reconstruction of the full $H \to \tau\tau$ leg is essential for separating signal from background events. Additionally, in order to obtain an accurate assessment of the properties of the HH system, a good estimate of the $H \to \tau\tau$ momentum is necessary.

In order to reconstruct the energy carried by the neutrinos, CMS analyses with the $H \to \tau\tau$ topology, including the previous $bb\tau\tau$ iteration [111], were so far using the Secondary Vertex Fit (SVFit) algorithm [370]. This algorithm reconstructs $m_{\tau\tau}$ based on a likelihood function. The function optimizes the reconstruction of the visible momenta of the τ decay products plus the MET in the event. As a consequence, the SVFit algorithm improves the $m_{\tau\tau}$ reconstruction, so that the $H \to \tau\tau$ system lies closer to the expected 125 GeV mass, with an improved resolution.

However, SVFit faces two important issues. Firstly, it is a computationally expensive approach, leading to bottlenecks in the analysis workflow. Secondly, it does not work optimally when the MET is close to zero, which happens for HH systems at low masses, since the two τs tend to be emitted back-to-back. In this work, we introduce an alternative approach to fix such issues.

A DNN is trained in order to *regress* the neutrino momenta. A generic DNN represents a function with a given, usually large number of parameters, mapping input features into a set of continuous variables, in the case of *regression*, or into class probabilities, in the case of *classification*. Its goal is to perform *inference*, which corresponds to the estimate of some unknown quantity or quantities, based on the observed data [371]. The parameters of the network are optimized, or learned, such that the network is better able to predict the ground truth. This is achieved via a *training* procedure, which is in this case *supervised*, i.e. it uses labeled data. In a nutshell, the optimization procedure is performed via *back-propagation* [372], where the parameters of the network are iteratively adjusted. The procedure minimizes a *loss function*, which quantifies the difference between the target prediction and true values. In this context, the target corresponds to neutrino momenta, and the true values are obtained with simulated samples. The parameters of the network are updated in the direction of the negative gradient of the loss function. A factor called *learning rate* scales the magnitude of the adjustments, influencing the time the algorithm takes to converge.

For each of the four data-taking periods, the following samples are used for the training:

- ggF spin-0 and spin-2 signals in the mass range of 250 GeV to 3 TeV;
- DY samples binned in the p_{T} of the Z boson;
- $\mathrm{t\bar{t}}$: fully-hadronic, semi-leptonic and fully-leptonic;
- $\mathrm{t\bar{t}H}$ ($\mathrm{H} \to \tau\tau$).

The backgrounds above cover the two dominant background sources, namely DY and $\mathrm{t\bar{t}}$. The $\mathrm{t\bar{t}H}$ process is also considered, since it has the same final state as the signal, including a peaking structure.

The input features used to train the network can be continuous or categorical, and are described in Table 4.14. The continuous features are passed through a normalization layer ensuring zero mean and unitary standard deviation. The categorical features are encoded as integers, deprived of physical meaning, which are passed through an embedding layer, with output dimension 10, before being concatenated with the continuous features. Some categorical features are occasionally set to a default value depending on the resolved or boosted topology of an event. The concatenation is used as an input to the first layer of the network.

The network consists of a common block with 5 dense layers and two "heads", as illustrated in Fig. 4.13. Each layer has 128 nodes and uses the Exponential Linear Unit (ELU) activation function [373]. After each layer, batch normalization [374] is applied. One of the heads is used for regressing the neutrino momenta, and the other performs classification, with 4 dense layers each. The classification head is not actively used, but was added since it improved the performance of the regression network. We can intuitively reason that the classification loss forces the network's weights to better extract the underlying properties of the input samples, which in turn affects the regression head, given the significant sharing of weights in the inital common block. The regression head has 6 outputs, corresponding to the three-momenta of the neutrinos from the two τ legs. For leptonic τ decays, the two neutrino momenta are summed. The classification head has 4 outputs. The elements of the architecture are listed in Table 4.15.

The loss function includes a mean squared error term for the regression head, and a cross-entropy term for the classification head. Both terms are added. On top, a

Type	Variable
Continuous	`DeepJetCvsB`: Charm versus bottom `DeepJet` score
	`DeepJetCvL`: Charm versus light flavour `DeepJet` score
	`DeepJet`: b-tagging score
	`HH-bTag`: Prob. that the jets come from a $H \to bb$ decay
	MET†: x and y MET projections
	Cov(MET): MET covariance matrix
	4-momenta†: visible 4-momenta of τs, b jets, $\tau\tau$ and $b\bar{b}$
Categorical	τ decay channel: $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$
	τ_h decay mode
	Presence of two AK4 jet
	Presence of one AK8 jet
	Data-taking period
	Spin hypothesis
	Mass hypothesis

Table 4.14: Lists of continuous and categorical input variables used by the DNN. The items marked with a † correspond to vectors that were rotated in the transverse plane with respect to the visible di-τ system. The rotation is performed to establish a common reference, which helps the training.

regularization term is included, with the sum of the L2 norm $(\sum_i^n |x_i|^2)^{1/2}$ of all the n weights i in the network. This term is multiplied by a factor which depends on n, to avoid a dependency on the network's size. The factor is chosen such that the regularization term makes up $\sim$20 % of the total loss.

When training a DNN, one should ensure that the network can generalize to unseen data. This is crucial, as DNNs can suffer from *overfitting*, a phenomenon in which sufficiently large networks, after enough time, perfectly describe the data they were trained on, but cannot make correct predictions on different datasets. This is a consequence of fitting the dataset's variance, or noise, which is inevitably present, in addition to the relevant model structures. The noise is usually irrelevant for the understanding of a particular system and, due to its stochastic nature, is detrimental to the algorithm's capacity for generalization. To counteract such effects, the dataset is split into training and *validation* subsets. The latter is used for evaluating the performance of the network in an unbiased way, using data it has never seen, thus preventing overfitting. In turn, the introduction of a validation subset has the disadvantage of holding out a portion of the data from the training, leading to an increase of the statistical uncertainty associated to each prediction. A strategy is therefore chosen such that all samples can be used, called *cross-validation*. This procedure defines k DNN discriminants associated to k subsets of the full dataset, or *folds*. Each DNN is trained on $k-1$ folds, with the remaining fold being held out for validation. No two DNNs are validated with the same fold, and the full data is exploited. In the end, the predictions of the k discriminants are averaged. For this work, we use five folds ($k = 5$).

The training is performed in *batches*, with a size of 4096, and exploits the `AdamW` optimizer [369] with an initial learning rate of 0.003. If the validation loss does not

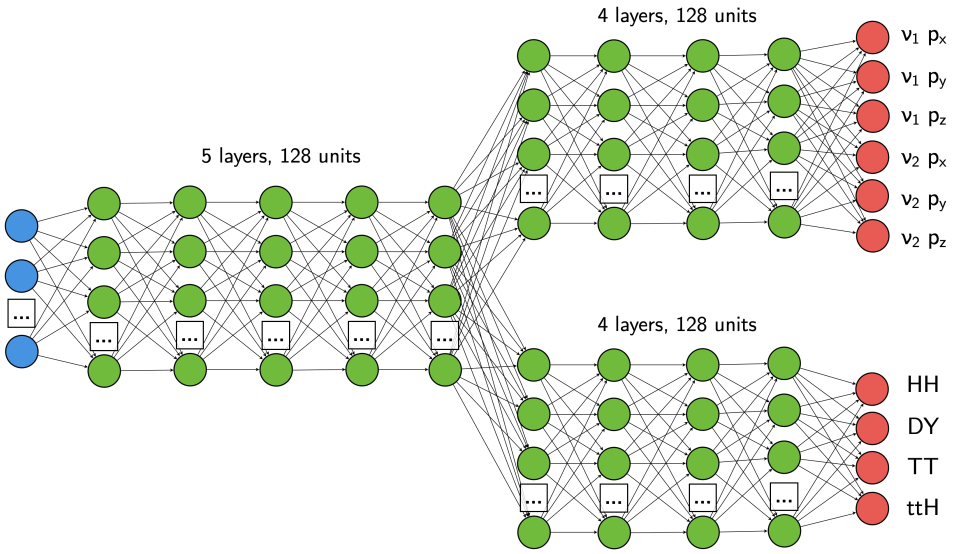

Figure 4.13: Visual representation of the $m_{\tau\tau}$ regression network with its two "heads", for regression and classification. The regression estimates the three-momenta $\vec{p} = (p_x, p_y, p_z)$ of the neutrinos ν_1 and ν_2 produced in the $H \to \tau\tau$ decay. Depending on the τ decay channel, the ν notation can refer to one or two neutrinos. The classification assigns every event a score, representing the probability to belong to one of four classes. Details are provided in the text. Courtesy of Tobias Kramer.

improve within 10 epochs, the learning rate is halved. Once no effect on the validation loss is observed, the adaptation of the learning rate is stopped. As soon as the validation loss has not improved within 15 epochs, the training ends.

Event weights are considered, to account for different cross sections, selection efficiencies and other differences between individual processes. Instead of applying such weights directly in the calculation of the loss of each batch, the batch composition directly depends on the weights. As a first step, the events of the 4 classes are distributed across a batch in equal proportions. This is done to prevent the samples with large MC statistics, such as DY and $t\bar{t}$, from driving the learning in the batch, when compared to samples with less events, such as the signal. Within the sample classes, two different approaches are chosen. For the signal class, each mass or spin hypothesis is equally weighted, whereas for the background classes, the events are distributed based on the event weights. The network is implemented using the `Keras` library [367] with a `Tensorflow` backend [375].

DNN section	Layers	Nodes/layer	Loss function	Number of Outputs
Common block	5	128	–	–
Regression head	4	128	Mean Square Error	6: p_T of 2 ν's
Classification head	4	128	Cross-Entropy	4: HH, DY, $t\bar{t}$ and $t\bar{t}H$

Table 4.15: Overview of the architecture of the $m_{\tau\tau}$ regression network. A batch size of 4096 is employed. ELU activations and batch normalization are applied after each layer.

The performance of the regression is dramatically superior to what SVFit achieves. Results vary depending on the class being probed, but lead to $m_{\tau\tau}$ widths at least two

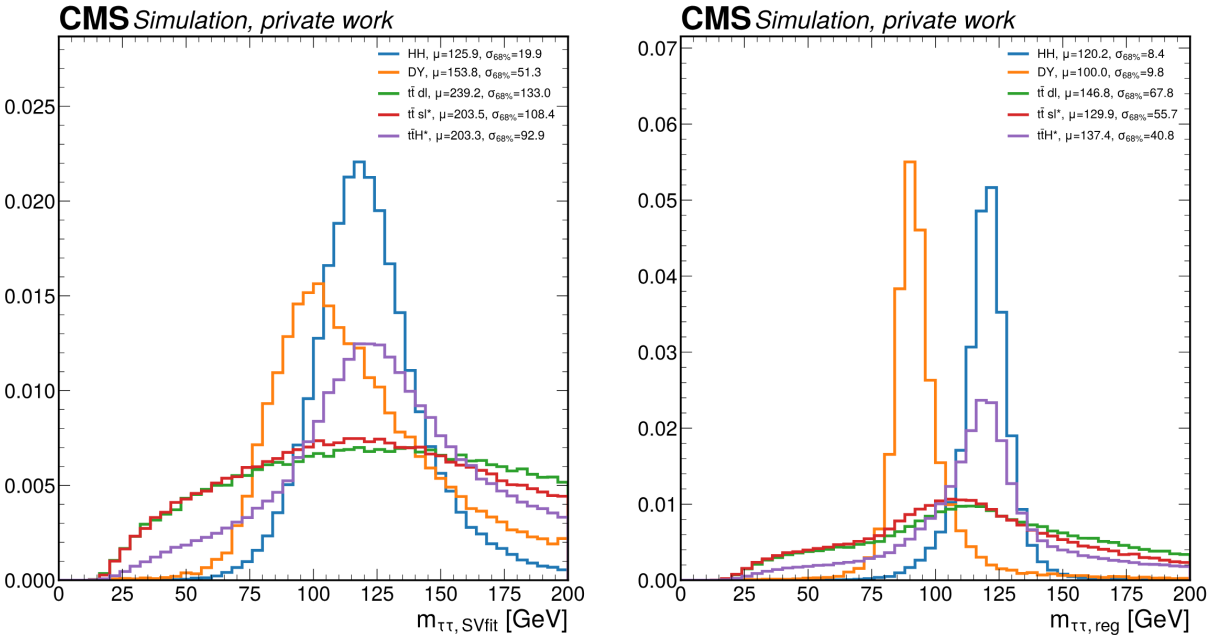

Figure 4.14: Comparison of the binned $m_{\tau\tau}$ distribution after regressing $m_{\tau\tau}$ with SVFit (left) and the new DNN algorithm (right). The distributions are shown for the four samples which are encoded as a class by the network: HH, DY, $t\bar{t}$ and $t\bar{t}H$. $t\bar{t}$ is split into fully-leptonic ("dl") and semi-leptonic ("sl") samples. The HH signal represents a merged sample with all signal hypotheses. The resolution improvement is very large. The HH and DY peaks become clearly separate. Courtesy of Tobias Kramer.

times narrower. The width of the DY peak is reduced by a factor of 5. The regressions are shown in Fig. 4.14, both for the sum of all HH signal samples and for the three backgrounds considered in the classification head. The DY and signal peaks can be fully disentangled, contrary to what happens with SVFit. These results are achieved with only a minor $t\bar{t}$ background sculpting. We notice that the method here described has the potential to be applied to other decay topologies, given appropriate training data and input features. The network is further exploited in the definition of the final discriminant, detailed in Section 5.1.

4.6 Modeling of Background Processes

A complete understanding of all backgrounds that could mimic the signal of interest is a complex but essential step, and the only way to observe rare phenomena and make precise measurements. If a given MC sample is accurate enough, *MC-driven* methods are used to perform background estimates. If instead, despite all efforts, significant discrepancies are known to exist between simulation and data, the analysts might resort to *data-driven* methods, where real data outside the SR is instead exploited, as long as it is kinematically similar to the SR. In this work, we consider the first approach for all background sources except for the multijet background, where the data-driven ABCD method is employed. Residual discrepancies are corrected *a posteriori*, via dedicated SFs, detailed in Section 4.7. In the following, we describe the modeling of all background sources, given emphasis to multijet, DY and $t\bar{t}$, which are the dominant background processes in the $X \to HH \to bb\tau\tau$ analysis.

4.6.1 Multijet

Collision events produce large numbers of jets, mostly coming from gluons or light jets. Such jets can be occasionally misreconstructed as τ_h particles, or also as jets issued from heavier quarks. The bb$\tau\tau$ topology can thus be faked whenever one or two jets are misidentified as τ_h candidates, and two additional jet candidates are present. The multijet background affects in particular the $\tau_h\tau_h$ channel, where no lepton can be used as a background-suppressing handle. The modeling of QCD background from simulations presents two major challenges. Firstly, large samples are needed, especially in the distribution tails of the SR, given the relatively low probability (10^{-2} to 10^{-3}) for a quark or gluon jet to be identified as a τ_h candidate, together with the presence of two jets passing the b jet selections. Secondly, the misidentification of quark and gluon jets as τ_h candidates suffers from imperfect modeling, given that detector effects are hard to quantify. As a consequence, the QCD multijet background is estimated via a data-driven method, exploiting jet-enriched regions.

The basic idea is to find fully uncorrelated variables upon which the SR selection depends on, and define new regions by inverting the variables' selection cuts, removing as much non-multijet processes as possible. These regions can be used to estimate both the shape and the normalization of the QCD background in the SR, without using the SR directly. The method is dubbed ABCD, since it explicitly divides the analysis phase-space into four regions: A, B, C and D. In our analysis, the regions are based on the isolation of the τ_h lepton and on the electric charges of the lepton pairs. The A region corresponds to the SR, as defined in Section 4.4 for each category, while the B, C and D regions are sidebands used to estimate the background in the SR. In Fig. 4.15 we illustrate how the phase-space is divided. The tau pairs are classified as SS when they have the same charge, or OS, when they have opposite charge signs. In region B, the pair charge requirement is inverted, thus including SS pairs of isolated leptons. In regions C (OS) and D (SS), the `DeepTau` selection is inverted by requiring the τ_h candidate to fail the `Medium` WP of the `DeepTauVSjet` discriminator, but still pass the `VVVLoose` threshold (see Table 4.10). These regions thus include non-isolated leptons. In the $\tau_h\tau_h$ final state, only the `DeepTau` selection of the τ_h candidate with the lowest isolation is inverted.

The shape of the multijet QCD contribution in any given variable can be estimated from the B or C regions. Yields from all other MC-based backgrounds are subtracted from the data. Given some inevitable correlations between the lepton pair charge and the τ isolation, the remaining number of events has to be corrected by a multiplicative factor given by the ratio between the two control regions left. Hence, if region B is chosen to define the shape of the QCD background, the correction to its yield will be given by the C/D ratio, or *transfer factor*. Equivalently, if region C is chosen, the yield will be corrected by B/D:

$$N_A = N_B \times \frac{N_C}{N_D} \quad \text{or} \quad N_A = N_C \times \frac{N_B}{N_D}\,, \tag{4.14}$$

where, despite being identical, the two expressions are experimentally obtained in a different way, as one component (the ratio) provides the event yield and the other provides the shape of the SR. Both the shape and the ratios are calculated after subtracting all other MC-based backgrounds. The two alternative shapes, from regions C and B, are used as the upper and lower templates for the QCD shape uncertainty, and the nominal value is obtained by averaging the two options in Eq. (4.14). This effectively symmetrizes the

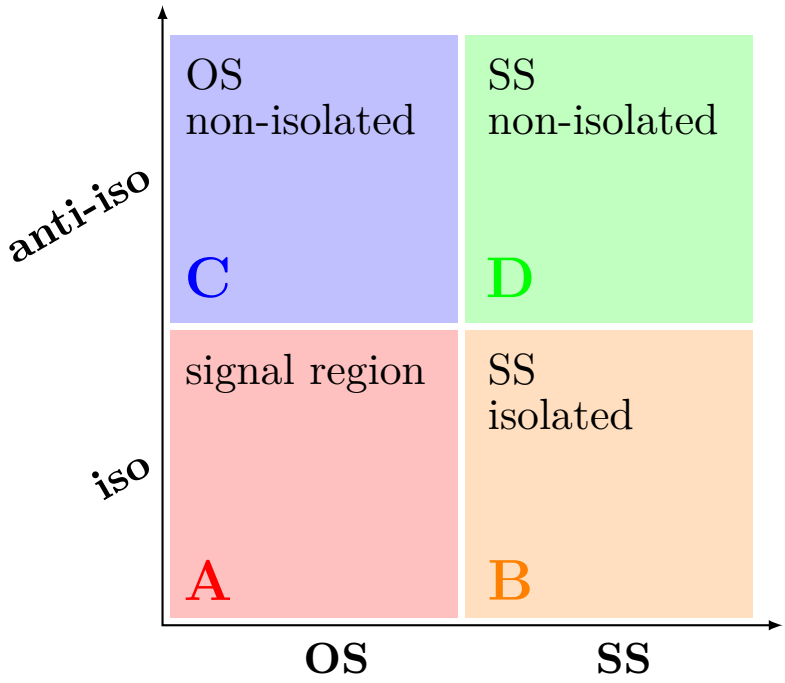

Figure 4.15: Sketch of the four ABCD regions used to estimate the multijet background. Region A corresponds to the SR, while the other regions are sidebands used to estimate the multijet background in the SR. The phase-space division is based on the isolation of the τ_h and on the relative sign of the τ lepton charges. More details are provided in the text.

corresponding systematic uncertainties, discussed in Section 5.2, ensuring the nominal value always lies between the up and down variations.

Several tests are performed to evaluate the validity of the ABCD method for QCD estimation. In the first test, the stability of the estimated yield is evaluated by modifying the definition of the C and D regions, using four different `DeepTau` WPs. The C and D regions are then defined as regions where the τ_h candidate passes a WP $\mathcal{A}$ from the `DeepTauVSjet` discriminator, but not a WP $\mathcal{B}$, where $\mathcal{B}$ is tighter than $\mathcal{A}$. Four alternatives are tested:

- $\mathcal{A}$: `VVVLoose`, $\mathcal{B}$: `VVLoose`
- $\mathcal{A}$: `VVLoose`, $\mathcal{B}$: `VLoose`
- $\mathcal{A}$: `VLoose`, $\mathcal{B}$: `Loose`
- $\mathcal{A}$: `Loose`, $\mathcal{B}$: `Medium`

In these four regions, the ratio of C/D yields is computed and compared to the value of the C/D yield ratio obtained with the standard definitions of the C and D regions. The level of compatibility between the measurements is satisfactory. In the second test, the ABCD estimation is compared to a direct MC-subtracted data sideband region where signal presence is negligible. This sideband region has been defined by inverting the $m_{\tau\tau}$ and $m_{b\bar{b}}$ mass cuts from the **res1b** category. Again, the QCD estimation obtained with the ABCD method is found to be in good agreement with the MC-subtracted data in this sideband, validating the QCD estimations obtained in the analysis. In the third and final test, it is verified that the shapes extracted from the B and C regions are compatible. Being all tests successful, the ABCD method is used for estimating the

multijet background in the $X \to HH \to bb\tau\tau$ analysis, as it had already been the case with previous analysis iterations. Examples of the distributions of some kinematic variables in the B, C and D regions are provided in Appendix A.5.

Some alternative methods have historically been explored for multijet estimation. Parametric fits of data sidebands are probably the easiest to implement, but require narrow resonances and QCD-pure sidebands. Some variations of the ABCD method have instead been explored, adopting more than four regions. The idea is that some residual correlations can exist between the two variables defining the ABCD phase-space. By adding an "E" and "F" regions, one can leverage the additional information from the added sidebands to better estimate the multijet background. This method is appropriately called "alphabet", and has been extended into higher-dimensionality spaces [376]. Furthermore, the transfer factors can be learned by DNNs or BDTs, such as in the resolved CMS $HH \to bbbb$ analysis [113]. Despite all possible improvements, it is important to state that the main drive for the proposals above is the possibility to obtain a very precise estimate, leading to smaller systematic uncertainties. However, the $X \to HH \to bb\tau\tau$ measurement is currently limited by statistical uncertainties, and so the reduction of systematic uncertainties is generally not a crucial item. On top, the definition of additional phase-space regions tends to increase statistical uncertainties. Overall, taking also into account the systematics discussed in Section 5.2.1, the ABCD method is deemed appropriate for the analysis here reported.

4.6.2 Drell-Yan

The contribution of the DY $Z/\gamma^* \to \ell\ell$ plus jets events is estimated using MC simulations. NLO samples are used to model it. Within the CMS Collaboration, DY samples are generated in three separate ways:

- as an inclusive sample;
- in slices of $p_T^{Z,Gen}$ (GeV units):]0; 50],]50;100],]100, 250],]250;400],]400,650],]650;∞[;
- in slices of the number of jets N_{jet}: 0, 1 and 2 jets.

In order to maximize the number of available MC events for such a significant background, all three types are added together, using a simple technique called *stitching*. There, the samples being added are scaled back down to the total DY inclusive cross section, listed in Table 4.2.

A *stitching weight* of 1/3 is applied to all DY events, given the three sample types above, to avoid event duplication. There is however an exception, in which events with $p_T^{Z,Gen} = 0\,\text{GeV}$ are weighted by 1/2, since the first bin of the p_T-sliced samples does not include events with zero generated Z p_T. The quality of the DY MC simulation is checked in the **res1b**-like DY CR, defined in Section 4.4.5. Fig. 4.16 shows the distribution of two kinematic variables in the $\tau_\mu\tau_\mu$ channel for the 2018 data-taking period, namely the p_T of the first b jet and the η of the first lepton. There, the shaded gray bars represent the statistical uncertainty of the background samples, and the statistical uncertainties of the data are too small to be seen. A good data/MC agreement is obtained.

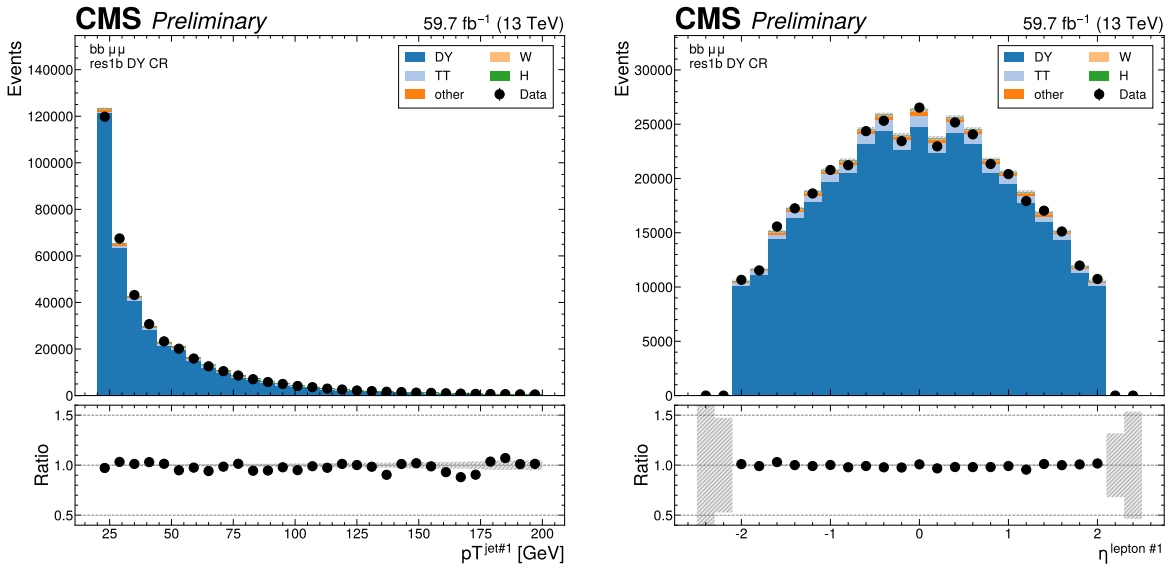

Figure 4.16: Distribution of the p_{T} of the first b jet (left) and the η of the first lepton (a muon) in the $\tau_\mu\tau_\mu$ channel, for the 2018 data-taking period, using the **res1b**-like DY CR. The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

4.6.3 Top/anti-top pairs

The contribution of the $\mathrm{t\bar{t}}$ background is modeled using MC simulations. Virtually all top quarks decay to a b quark and a W boson, which can in turn decay leptonically or hadronically. Samples are thus divided into fully-hadronic, semi-leptonic (or, equivalently, semi-hadronic), and fully-leptonic samples. Fig. 4.17 shows the distribution of two kinematic variables in the $\tau_\mu\tau_{\mathrm{h}}$ channel for the 2018 data-taking period, namely the p_{T} of the first b jet and the η of the first lepton, using the $\mathrm{t\bar{t}}$ CR. There, the shaded gray bars represent the statistical uncertainty of the background samples, and the statistical uncertainties of the data are too small to be seen. The agreement with data is found to be sufficient.

For completeness, we analyzed the purity of $\mathrm{t\bar{t}}$ samples in terms of the presence of at least one fake τ_{h}. The results can be seen in Table 4.16. The more modest presence of fakes in the leptonic channels comes from the requirement of an electron or a muon, which removes the fakes present in fully-hadronic $\mathrm{t\bar{t}}$ events. There is also a higher fraction of fakes in the **boosted** category with respect to the resolved categories. This happens because τ reconstruction algorithms perform worse in boosted regimes. Finally, we note the additional requirement on the b jet of the **res2b** category slightly improves the purity, as expected.

MC Mismodeling

Contrary to what is observed at the time of writing, during the initial stages of the PhD work a $\mathrm{t\bar{t}}$ MC mismodeling was observed. The issue was predominant in the $\tau_e\tau_{\mathrm{h}}$ and $\tau_\mu\tau_{\mathrm{h}}$ channels, just as it had been observed in the previous nonresonant iteration of the $\mathrm{bb}\tau\tau$ analysis. The discrepancies did at some point disappear, due to the implementation of several corrections that were missing at the time. Still, we here report a method to

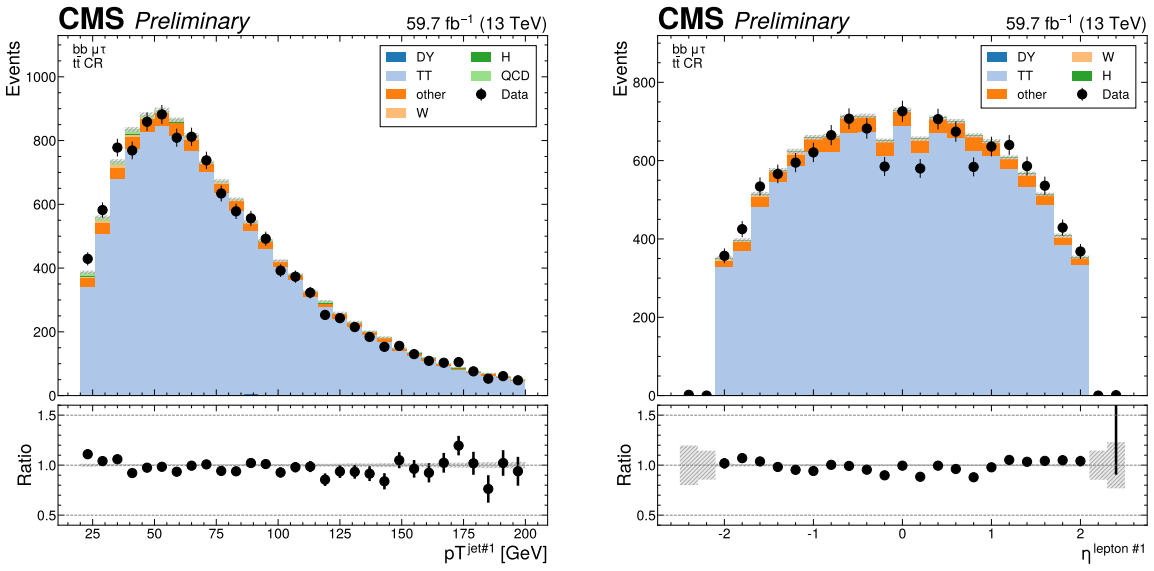

Figure 4.17: Distribution of the p_T of the first b jet (left) and the η of the first lepton (a muon) in the $\tau_\mu\tau_h$ channel, for the 2018 data-taking period, using the $t\bar{t}$ CR. The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

Channel	baseline	res1b	res2b	boosted
$\tau_e\tau_h$	0.329 [77094.9]	0.380 [22362.4]	0.309 [19105.5]	0.596 [130.6]
$\tau_\mu\tau_h$	0.332 [115723.8]	0.379 [35168.1]	0.310 [30247.2]	0.597 [134.2]
$\tau_h\tau_h$	0.422 [2509.0]	0.477 [522.2]	0.415 [408.1]	0.79 [9.7]

Table 4.16: Fraction of $t\bar{t}$ events with at least one fake τ_h, in all analysis channels and categories. The numbers between parenthesis refer to the total number of weighted events across fully-leptonic, semi-leptonic and fully-hadronic $t\bar{t}$ samples. The baseline selection refers to Section 4.4 before applying any categorization. Categories are defined in Section 4.4.3.

fix such issues, following Ref. [111], since it might become useful if residual mismodelings appear in future iterations.

We set to fix the discrepancy by rescaling the normalization of the $t\bar{t}$ background to the data. The $t\bar{t}$ CR is used. Assuming the region has no signal, we fit the CR by setting the normalization of $t\bar{t}$ as a free-floating parameter. No systematics are introduced, as they can partially encode the $t\bar{t}$ normalization. The result of the fit provides a constant factor which can be used to multiply the $t\bar{t}$ yield in the SR. In Fig. 4.18 we show how the method worked when it was needed. The data selection there performed is outdated.

The result of the fit can be validated by repeating it on the CR and SR put together, where a cut is applied to the analysis' final discriminant, so that only background-like events are used, avoiding biases from looking at the signal-sensitive region. The cut ensures the usage of a signal depleted region, avoiding biases. The result of the CR and CR+SR fits should match.

4.6.4 Other backgrounds

All remaining processes are modeled, in shape and normalization, using solely MC samples. Cross sections can be inspected in Section 4.1.2. The inclusive W + jets sample

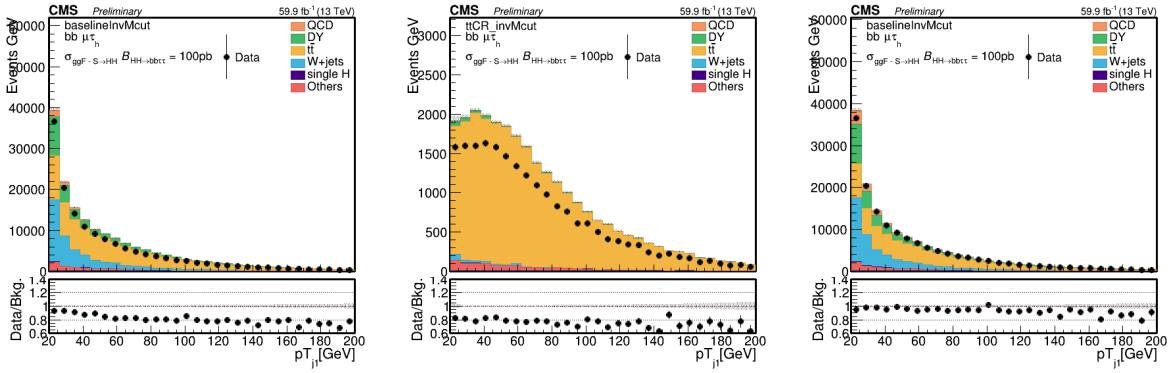

Figure 4.18: Example of the $t\bar{t}$ mismodeling correction in the $\tau_\mu\tau_h$ channel, in 2018. The $t\bar{t}$ background appears in yellow. (Left) p_T of the b-jet with the highest `HH-bTag` score before applying the $t\bar{t}$ SF. (Middle) p_T of the same b-jet in the $t\bar{t}$ CR. (Right) p_T of the same b-jet after applying the $t\bar{t}$ SF.

is used in the $0 \leq H_T < 70\,\text{GeV}$ only for simplicity, in order to avoid overlaps with the samples binned in H_T. The impact of the W + jets sample in the analysis becomes very small once b-tagged jets are required, especially in the **res2b** category. Processes involving the presence of two or three vector bosons, like ZZ or WWZ, are modeled, in shape and normalization, through inclusive samples. For the contributions arising from EW processes with one W or Z boson plus 2 jets, from single top both in the s- and t-channels, or from $t\bar{t}$ pairs produced in association with a single boson or a pair of vector bosons, the cross sections are extremely small, but are nevertheless taken into account. The single Higgs production cross section is also tiny, when compared to the dominant backgrounds. However, such processes have the potential to have very similar m_{HH} topologies as the signals of interest, and are also considered. This is the case for ggF and VBF Higgs production, but also when a Higgs boson is produced in association with a vector boson or a pair of top quarks. Finally, the nonresonant SM Higgs boson pair production is also added as a background source, assuming $\lambda_{HHH} = 1$.

4.7 Monte Carlo Corrections

Despite the occasional existence of alternative data-driven techniques, many physics analyses extensively rely on MC simulations for most of its steps, as is the case for the analysis covered in this work. From studying the selection efficiency of various thresholds, to training discriminants and performing the fits for extracting the signal, MC samples are absolutely crucial. However, they do not always perfectly match the observed data. To avoid introducing biases in the final result, discrepancies between data and MC should be understood, and ultimately corrected.

The correction is done with the application of SFs, often defined as the ratio between data and MC in dedicated phase-space regions. SFs can sometimes be applied to the entire event, corresponding to a multiplicative factor that changes the event's importance when compared to other events. The SFs can also be applied to particular objects. Residual differences are handled by the addition of systematic uncertainties, covered in the next Chapter.

In this Section, we detail all MC corrections applied in our analysis. Some corrections are centrally provided by the CMS Collaboration, so that multiple analyses

can benefit from them, while other corrections have to be computed specifically in the $X \to HH \to bb\tau\tau$ context. We mention here the corrections that are straightforward to describe:

- **Pileup Jet Identification**: As discussed in Section 4.3.4, jets with $p_T < 50\,GeV$, which do not pass the `Loose` WP of the PU jet discriminator, are discarded. A set of centrally provided SFs must then be applied.
- **Jet Energy Smearing**: The jet smearing procedure is applied to all jets, both AK4 and AK8, as discussed in Section 4.3.4. The smearing is also propagated to MET. The smearing SFs are centrally provided.
- **Electron and Muon SFs**: Specific SFs are applied to correct possible data/MC discrepancies in the reconstruction and ID of electrons and muons, in the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels. The corrections are centrally provided, binned as a function of the p_T and η of the leptons.
- **Tau Energy Scale Corrections**: The Tau Energy Scale (TES) corrections are centrally provided. They are binned in four τ decay modes, based on the number of neutral and charged decay products, as explained in Section 2.4.5. For genuine τs, we consider the scenarios with one prong, one prong and one neutral, three prongs, and three prongs plus one neutral. For electrons misidentified as taus, only the first two scenarios are considered, while for muons misidentified as taus, no energy scale correction is recommended, due to the rarity of such process.
- **Pileup Reweighting**: The PV is fit by a deterministic annealing algorithm [377]. The same algorithm also obtains the total number of vertices and defines an assignment of clustered tracks to different collisions, effectively identifying all interaction vertices. In a first step, tracks passing certain quality criteria are clustered mostly based on the their z coordinate of closest approach to the beamline. The procedure is fully equivalent to gradual cooling in statistical mechanics, hence its name. In a second step, the clustered tracks are fitted three-dimensionally using the entirety of the available track information. The vertices are sorted according to the sum of p_T^2 of their associated tracks, and the vertex with the highest value is selected as the PV. The other vertices in the event are assumed to come from PU collisions.

  During the LHC Run 2 there were an average of 27, 38 and 37 pp interactions per bunch crossing, for 2016, 2017 and 2018, respectively. Unfortunately, the distribution of the number of interaction vertices in MC events does not exactly matches the one in data. To achieve a better match, a standard reweighting procedure is applied to MC events, as centrally recommended by the CMS Collaboration for Run 2. The provided PU weights use a nominal Minimum Bias, inelastic pp cross section of 69.2 mb, which is used as an input to produce the correction weights. The SFs are centrally provided. The cross section includes a 4.6% uncertainty, which must be taken into account when computing systematic uncertainties.

In what follows, we cover the corrections which require a more detailed explanation.

4.7.1 Level-1 Trigger Prefiring

The time measurement of ECAL was observed to shift in 2016 and 2017, leading to ECAL TPs being associated to the wrong BX. Unfortunately, this *prefiring* effect was

not propagated to the L1 TPs, which caused a significant loss of events. Indeed, TPs with large η values were mistakenly associated to the previous BX, and because at L1 two consecutive bunch crossings are not allowed to fire, the correct BX was vetoed if enough ECAL energy was found in an adjacent BX. Even if the event was kept, TPs in the BX following the one which prefired are ignored. The effect strongly depends on p_{T} and η, reaching values larger than 50% for high p_{T} jets, in a $2.75 < |\eta| < 3$ window.

A similar situation is present in the Muon Chambers, where the assignment of muon candidates to a particular BX can be wrong. This happens given the modest time resolution of the muon detectors. The effect is most evident in 2016, but cannot be neglected for the other years. The prefiring rate remains constant for $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, but impacts almost the full available η range. The magnitude of the effects varies from 0% to 3%. We note that UL samples are the first set of CMS Run 2 samples for which this effect is taken into account.

The prefiring shifts are not described by the simulations. MC simulations are thus corrected via a reweighting procedure. The final MC event weight w is obtained as the product of the non-prefiring probability P for all affected objects, measured using efficiencies ε, computed with events that did not prefire:

$$w = 1 - \mathrm{P}(\text{prefiring}) = \prod_{i=\text{photons, jets, muons}} \Big(1 - \varepsilon_i^{\text{pref}}(\eta, p_{\mathrm{T}})\Big). \tag{4.15}$$

As an example, if in a particular bin 5% of the events prefired, MC events in that bin will be scaled down by a factor of $1 - 0.05 = 0.95$. This corresponds to less events being considered, as expected.

4.7.2 Lepton Trigger Scale Factors

Whenever an event belongs to a phase-space region where more than one trigger is active, the computation of the corresponding trigger efficiency must take into account the logical-`OR` between the triggers that could have fired. The efficiency for events firing more than one trigger would otherwise be incorrect, since those events would be taken into account multiple times; twice for two triggers, thrice for three, and so on. This idea was already discussed in Section 4.2.4, and the probability that at least one out of all considered triggers fires is given by Eq. (4.5).

In our analysis, the $\tau_e\tau_{\mathrm{h}}$ and $\tau_\mu\tau_{\mathrm{h}}$ final states include both a single-lepton and a cross-lepton trigger, where the latter is composed of two trigger legs. By leg we mean an element of a HLT path applying specific selections on a physics object, as already discussed in Section 4.2.1. Legs within a full path follow a logical-`AND`: the HLT path fires only if all independent legs fired. The HLT trigger paths used in this work have been presented in Section 4.2.1.

The SFs associated to lepton triggers must take into account the efficiency of the logical-`OR` between the two triggers used in the leptonic channels. Assuming the efficiencies of the two legs of the cross trigger to be independent, the efficiency of the logical-`OR` can be factorized and computed from the efficiencies of the single objects, following closely Eq. (4.5):

$$\text{efficiency} = \varepsilon_{\mathrm{L}} + \varepsilon_\ell\,\varepsilon_\tau - \varepsilon_\ell\,\varepsilon_\tau\,\varepsilon_{\mathrm{L}|\ell}\,, \tag{4.16}$$

where ε_{L} is the single-lepton trigger efficiency, ε_ℓ represents the cross-lepton trigger efficiency for the τ_e or τ_μ leg, and ε_τ stands for the cross-lepton trigger efficiency for the

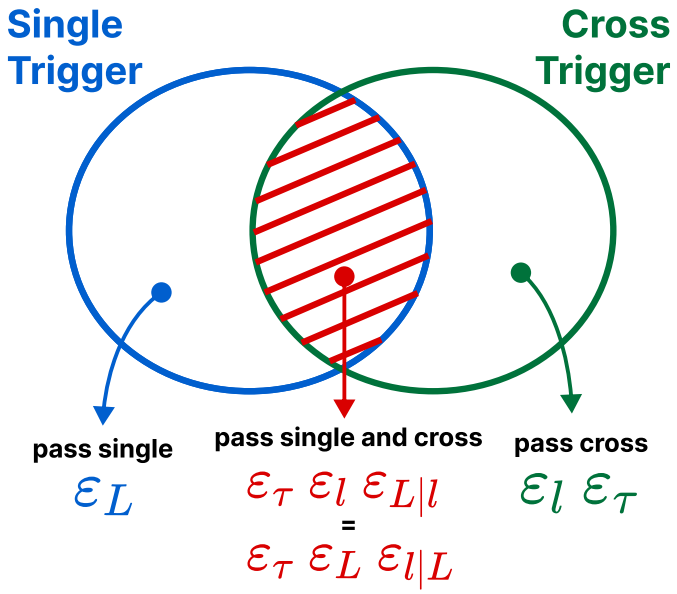

Figure 4.19: Venn diagram illustrating the single- and cross-trigger phase-spaces together with their intersection, as considered for the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels. The meaning of the different efficiency terms ε is described in the text. Eq. (4.16) is obtained by summing the two separate efficiencies and subtracting their intersection, as dictated by Eq. (4.5). The result represents the probability for an event to pass the logical-`OR` between the single- and cross-lepton triggers. The fact that the L and ℓ triggers are essentially the same, modulo their p_T threshold, enables to use the simplified variant shown in Eq. (4.17).

τ_h leg. $\varepsilon_{L|\ell}$ stands for the single-lepton efficiency given that the lepton leg of the cross trigger fired. We write the latter explicitly given the extreme correlation between those two trigger elements. Note that we could shift from the probabilities in Eq. (4.5) to actual efficiencies because the latter are being computed in phase-space bins, and not for single events, where the notion of efficiency is meaningless. Eq. (4.16) can be understood by taking into account the three terms shown in Fig. 4.19, which again follow Eq. (4.5).

Eq. (4.16) is correct, but the Tau POG does not provide $\varepsilon_{L|\ell}$ out of the box. That would imply knowing beforehand which triggers each analysis would choose, and store all possible combinations, which is highly unpractical. Instead, the analyzers are left with the task of deriving the analysis' SFs themselves. We can avoid doing so by exploiting the extreme similarity of the single-lepton trigger and the lepton leg of the cross trigger. Indeed, assuming noise-free triggers, $\varepsilon_{L|\ell}$ should be exactly one, as the two triggers are identical modulo their p_T threshold, which is lower for the cross-lepton trigger leg. For those cases, which we expect to completely dominate, the last term of Eq. (4.16) reduces to $\varepsilon_\ell\,\varepsilon_\tau$. In other words, one trigger is a subset of the other. However, due to (admittedly rare) trigger inefficiencies, it could happen that the single-lepton trigger fires and the other does not. For such cases, $\varepsilon_L > \varepsilon_\ell$, and thus $\varepsilon_{\ell|L} = 1$ (but $\varepsilon_{L|\ell} \neq 1$, in general). Using Bayes' theorem, we can express $\varepsilon_\ell\,\varepsilon_{L|\ell}$ as $\varepsilon_L\,\varepsilon_{\ell|L}$, which simplifies to ε_L. Finally, using a `min()` operand, we can choose whichever expression is correct for both situations above. We thus finally obtain the following formula, which is used for the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels:

$$\text{Eff} = \varepsilon_L + \varepsilon_\ell\,\varepsilon_\tau - \min(\varepsilon_L, \varepsilon_\ell)\,\varepsilon_\tau\,, \tag{4.17}$$

The formula is equivalent to Eq. (4.16). Avoiding the calculation of an additional efficiency term also removes a possible concern of larger trigger SF uncertainties. The

single-μ trigger and τ-legs efficiencies are provided by the CMS POGs, while the remaining lepton triggers are kindly provided by the authors of the $\gamma\gamma \to \tau\tau$ analysis [378]. The SFs depend on the η and p_{T} of the object. In Fig. 4.20 we provide examples for the used data efficiency maps in 2017 and 2018 for the $\tau_e\tau_{\mathrm{h}}$ and $\tau_\mu\tau_{\mathrm{h}}$ channels.

For the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state, di-τ trigger efficiencies and SFs are provided by the Tau POG in the context of the SM $\mathrm{H} \to \tau\tau$ analysis. The SFs are measured using $\mathrm{Z} \to \tau\tau \to \mu\nu_\mu\nu_\tau\tau_{\mathrm{h}}\nu_\tau$ events selected with the Tag & Probe technique[5], and cover the logical-`OR` of the three trigger paths used. The SFs also depend on the η and p_{T} of the objects.

4.7.3 Single-τ Trigger Scale factors

A single-τ trigger has been exploited for the first time in CMS bb$\tau\tau$ analyses, to the best of our knowledge. More details are given in Section 4.2.2. Flat SFs for the single-τ trigger are provided by the Tau POG, which recommends their usage in the region where the trigger efficiency plateaus, defined to be 10 GeV above their trigger threshold (130 GeV for 2016 and 190 GeV for 2017 and 2018). The recommended SFs can be inspected in Table 4.17. They are used to correct the MC event-by-event. The corrections are applied only to events within the single-τ region, as detailed in Section 4.4.

Year	**Single-τ SF**
2016	0.88 ± 0.08
2017	1.08 ± 0.10
2018	0.87 ± 0.11

Table 4.17: Single-τ trigger SFs as recommended by the CMS Tau POG. The corresponding HLT paths are defined in Table 4.8.

4.7.4 MET Trigger Scale Factors

A $\mathrm{MET}_{\text{no-}\mu}$ trigger has been exploited, inspired by the past high-mass resonant bb$\tau\tau$ analysis [101]. More details are given in Section 4.2.2. Contrary to what happens for the single-τ trigger, no SFs are centrally available for MET triggers; they must be derived in the context of this analysis. The efficiency of MET triggers is in general challenging to calculate given that, by construction, MET contains all objects present in the event. No dataset can therefore be used as an orthogonal reference, or denominator in the efficiency computation, against which to measure the MET trigger efficiency. However, by removing the contribution of muons in the definition of MET, as shown in Eq. (4.3), events triggered by muon triggers become orthogonal to the $\mathrm{MET}_{\text{no-}\mu}$ trigger, which is used in this work. We thus measure the efficiency ε of the $\mathrm{MET}_{\text{no-}\mu}$ trigger in data and MC, using a `SingleMuon` PD, independently for the four data periods under consideration (2016, 2016 APV, 2017 and 2018):

$$\varepsilon(\mathrm{MET}_{\text{no-}\mu}) = \frac{\text{Analysis Selection \&\& Single-}\mu\text{ Trigger \&\& MET}_{\text{no-}\mu}\text{ Trigger}}{\text{Analysis Selection \&\& Single-}\mu\text{ Trigger}}\,, \quad (4.18)$$

[5] See the brief note on the Tag & Probe technique at the end of Section 4.3.1.

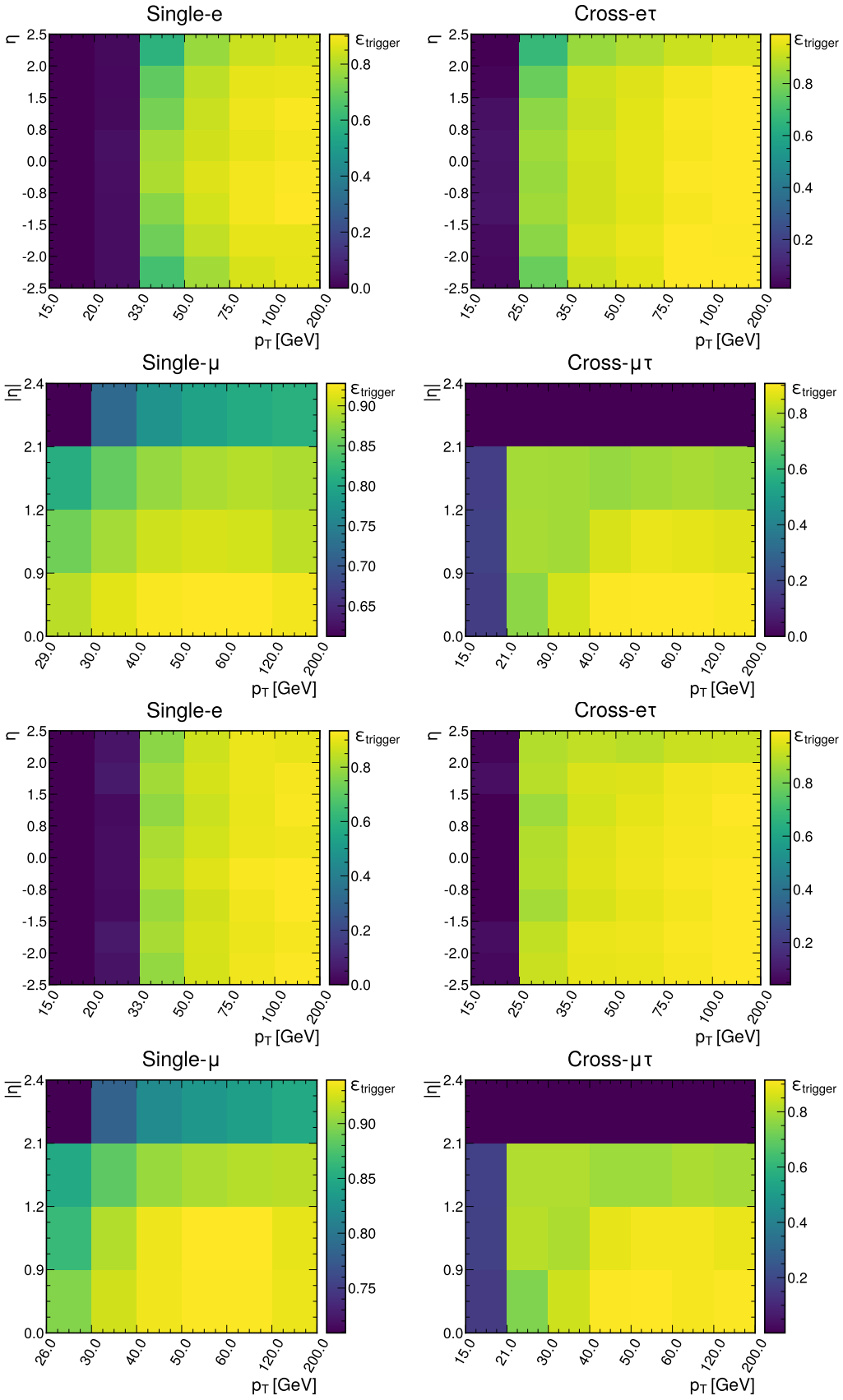

Figure 4.20: (p_T, η) trigger efficiency maps. Notice the occasionally different axis ranges. (Top row) Single-e (left) and cross-$e\tau$ (right) for 2017. (Middle-top row) Single-μ (left) and cross-$\mu\tau$ (right) for 2017. (Mid-bottom row) Single-e (left) and cross-$e\tau$ (right) for 2018. (Bottom row) Single-μ (left) and cross-$\mu\tau$ (right) for 2018.

where "Analysis Selection" refers to the **baseline** selection described in Section 4.4, plus the presence of two b jet candidates without `DeepFlavour` requirements. We explicitly enforce the single-μ trigger to be fired, ensuring a robust definition of the efficiency.

We apply a selection similar to the ones detailed in Section 4.4.1, but considering the $\tau_\mu\tau_\mu$ channel. Since the $\tau_\mu\tau_\mu$ channel is not part of the three analysis channels, we can use all its events while keeping orthogonality to the analysis. No additional selection is needed to define orthogonal phase-space regions, avoiding a decrease in statistics. We require two muons with $p_T > 15\,\text{GeV}$ each, and other selections as defined in Table 4.13. The three most important sources of background in the $\tau_\mu\tau_\mu$ channel are taken into account for the MC efficiency computation: $t\bar{t}$, DY and W+jets. Note that the multijet background is mostly absent in $\tau_\mu\tau_\mu$. A custom binning is set so to sufficiently sample the efficiency curves, especially in the turn-on region. To smooth out the fluctuations, a sigmoid function is fitted to both the data and the MC efficiency curves in their turn-on regions. The sigmoid function depends on three parameters:

$$f(x, a, b, c) = \frac{c}{1 + e^{-a(x-b)}} \,. \tag{4.19}$$

Four sets of SFs are calculated, one per data period, as the ratio of the data and MC sigmoid curves, as shown in Fig. 4.21. In order to obtain the best possible fit result, the range of the sigmoid fit is varied, and multiple values are tested. We find that a good result is obtained for all data periods by starting the fit at 150 GeV and ending it at 350 GeV. Values after 350 GeV can be fit by a horizontal line. Multiple starting values are tried and compared, and we find that they do not significantly impact the result, except when using the full $\text{MET}_{\text{no-}\mu}$ range, as illustrated in Fig. 4.23 (left), for 2018. For validation purposes, we also derive $\text{MET}_{\text{no-}\mu}$ SFs using the $\tau_\mu\tau_h$ channel, in order to make a comparison with the $\tau_\mu\tau_\mu$ SFs, following the selection described in Tables 4.12 and 4.13. They are found to be compatible within statistical uncertainties, as shown in Fig. 4.23 (right), for 2018. For completeness, we also compare the used $\tau_\mu\tau_\mu$ $\text{MET}_{\text{no-}\mu}$ SF curves across the four data periods in Fig. 4.24. Differences can arise due to changes in data-taking conditions across years. Plots for all the periods can be inspected in Appendix A.4.

We observe that, in 2017, the $\text{MET}_{\text{no-}\mu}$ trigger does not become fully efficient for high $\text{MET}_{\text{no-}\mu}$ values. Such inefficiency could be seen as a $\sim$5 % drop in the efficiency curve, at the plateau. This happens because the trigger was not active in the last runs of 2017. To recover the missing luminosity, we decided to consider instead, for 2017 only, a logical-`OR` between the "standard" $\text{MET}_{\text{no-}\mu}$ trigger used for other years, plus a trigger identical to the "standard one", but with an additional $H_T > 60\,\text{GeV}$ cut. We can see in Fig. 4.22 that the new trigger collects more data during some of the last few runs in 2017. The additional trigger enables to fully recover the lost efficiency.

The SFs are used to correct the MC event-by-event, only for events within the MET region, as discussed in Section 4.2.3, and after applying a turn-on cut. The cut is set to 180 GeV for all eras. The value is chosen based on the control distributions shown in Figs. 4.25 to 4.27. Whenever an event has a $\text{MET}_{\text{no-}\mu}$ value above 350 GeV, the SF is taken to be exactly 1 for all eras. Uncertainties are calculated using the uncertainties from the sigmoid fit and applying error-propagation for the ratio. The uncertainty values of the sigmoid functions at the upper limit of the fit range are used whenever the event has a $\text{MET}_{\text{no-}\mu}$ value lying above the fit validity range. The turn-on cut at 180 GeV prevents

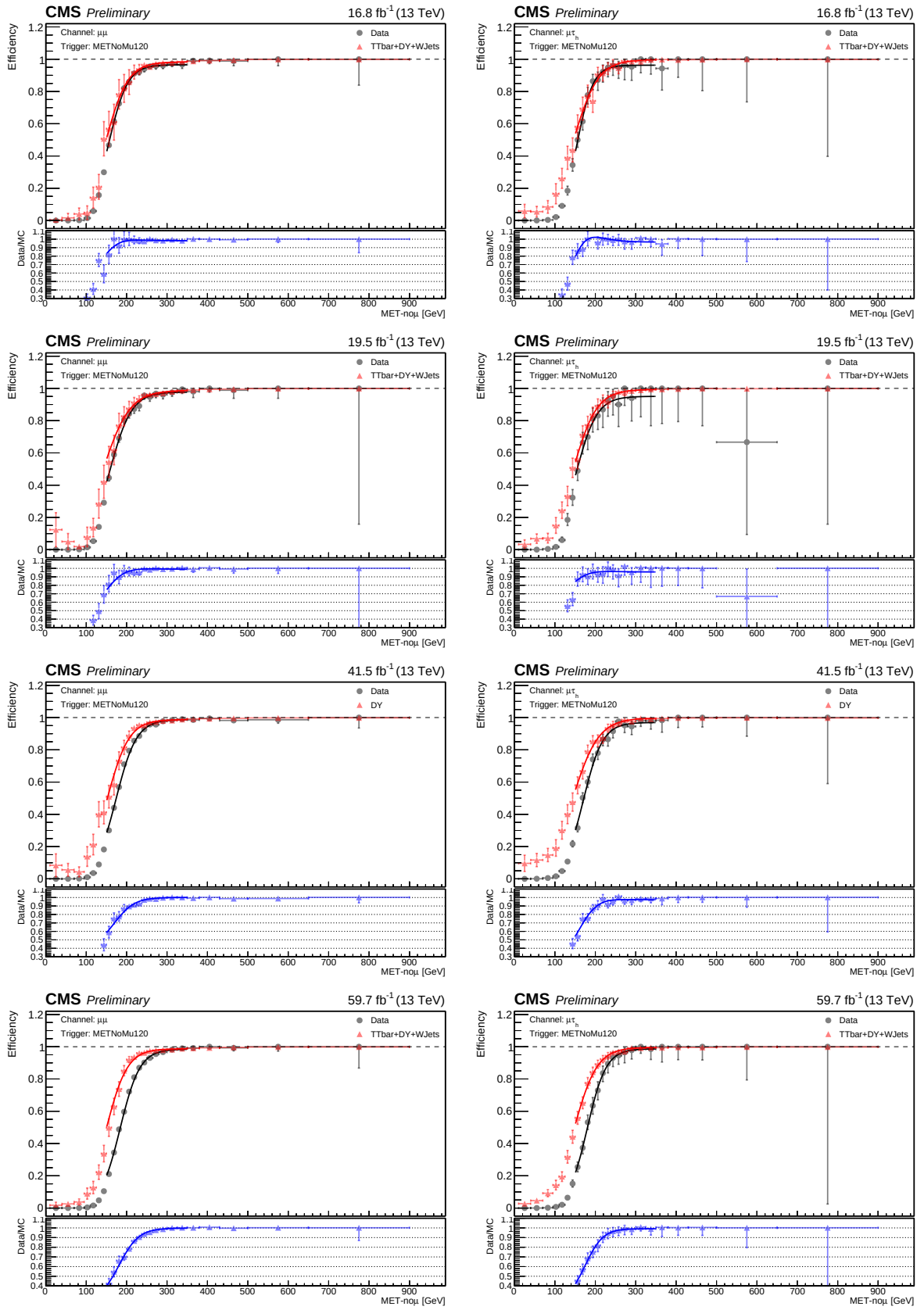

Figure 4.21: $\mathrm{MET}_{\text{no-}\mu}$ data and MC trigger efficiencies (top panels) and corresponding SFs (lower panels), for different years. The left (right) row refers to the $\tau_\mu\tau_\mu$ ($\tau_\mu\tau_\mathrm{h}$) channel. The $\tau_\mu\tau_\mathrm{h}$ channel is used for validation, while $\tau_\mu\tau_\mu$ is used to extract the analysis SFs. SFs are extracted from the ratio of the data and MC sigmoid fits, implemented to smooth out the SF's distribution. The SFs are taken to be one for $\mathrm{MET}_{\text{no-}\mu}$ values above 350 GeV. From the top to the bottom row, we show the 2016, 2016 APV, 2017 and 2018 periods.

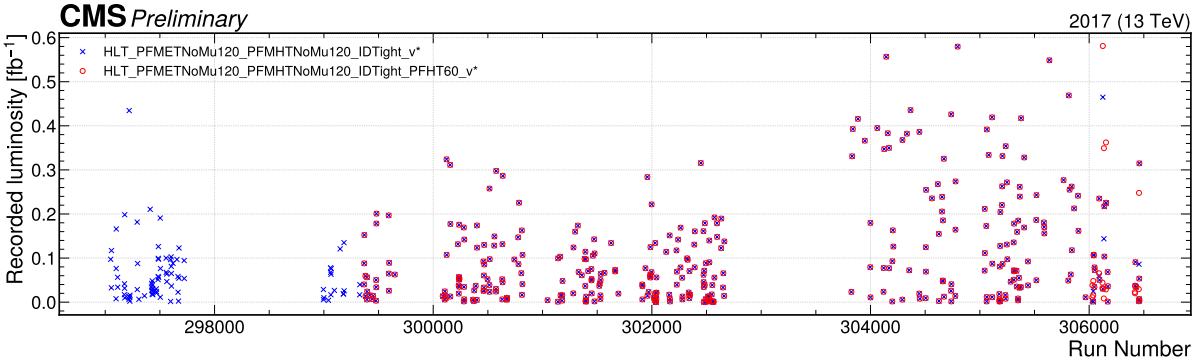

Figure 4.22: Recorded luminosity per run as a function of the unique LHC run number, for the 2017 data-taking period. The two $\text{MET}_{\text{no-}\mu}$ triggers considered in 2017 are shown. While the trigger with the additional H_T cut (empty red circles) was not active in the first runs, it collected all available luminosity once it was on. This enables to recover the luminosity lost by the "standard" $\text{MET}_{\text{no-}\mu}$ trigger, shown as blue crosses. One should notice the discrepancies in some of the last few runs.

this from happening for values below the fit validity range. Despite the low statistics involved, one can see that the $\text{MET}_{\text{no-}\mu}$ SFs improve the description of the observed data.

4.7.5 DeepTau Scale Factors for Hadronic Tau Leptons

Data/MC discrepancies in the ID efficiency of the hadronically-decaying tau leptons must be corrected. Different WPs of the `DeepTau` algorithm are employed for the selection of the $\tau\tau$ pair, as described in detail in Section 4.3.3. The SFs are centrally provided, and are implemented using the following logic:

- For genuine taus, the SFs are provided per data-taking period, in bins of the tau lepton decay mode, and the p_T dependency is fitted using linear functions in the $[20; 140]\,\text{GeV}$ range. τs with $p_\text{T} > 140\,\text{GeV}$ have separate corrections binned in p_T: $]140; 200]$ and $]200; \infty[\,\text{GeV}$. The SFs used here represent an update by the Tau POG over what was previously available, leading to a significant data/MC improvement for 2016, as shown in Fig. 4.28.
- For genuine electrons misidentified as tau leptons, the SFs are provided in barrel and endcap categories.
- For genuine muons misidentified as tau leptons, the SFs are provided binned as a function of η.

4.7.6 B-Tag Reweighting

To account for discrepancies in the MC b-tag performance, its full distribution is corrected to match the one in data, following the shape calibration procedure centrally recommended. For each MC event with a given jet configuration, the event weight ω is computed as:

$$\omega = \prod_{i}^{\text{N}_\text{jets}} \text{SF}\left(\text{D}^i,\, p_\text{T}^i,\, \eta^i\right) \tag{4.20}$$

where the SFs are provided by the CMS BTV POG as a function of the discriminator score D, the p_T and the η of the jets. In our analysis, D refers to `DeepJet`. The event

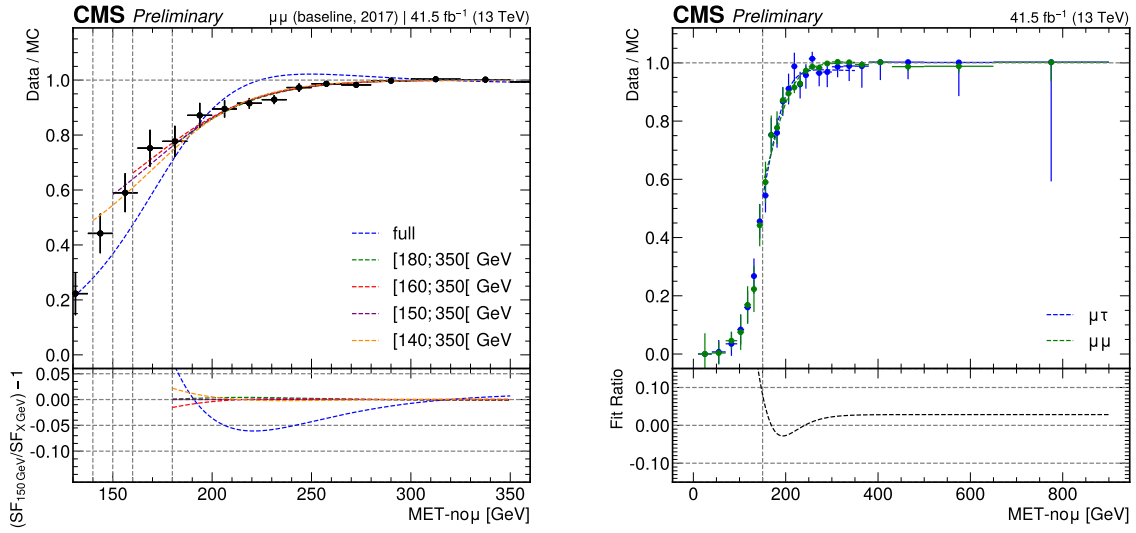

Figure 4.23: Data/MC SFs of $\text{MET}_{\text{no-}\mu}$ trigger efficiencies, in 2017. SFs are extracted from the ratio of the sigmoid fits of data and MC efficiency curves, implemented to smooth out the SFs' distributions. (Left) Five different fit ranges were tested, and zoomed in the turn-on region to better display differences. All fits are reasonably compatible, except for the full range fit, which cannot describe the data. We decided to use the fit starting at 150 GeV, for all data periods. (Right) The SFs are observed to be compatible between the $\tau_\mu\tau_\text{h}$ and $\tau_\mu\tau_\mu$ channels, within statistical uncertainties.

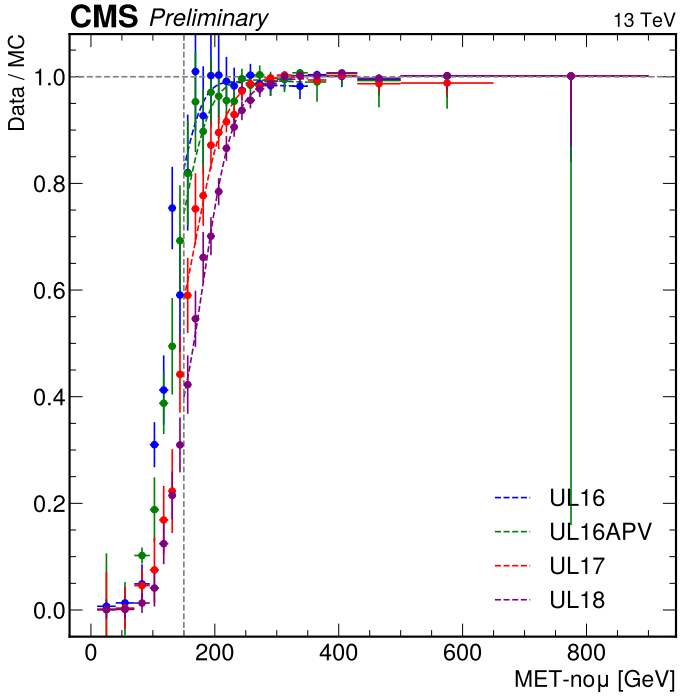

Figure 4.24: Comparison of the $\text{MET}_{\text{no-}\mu}$ SFs used in the analysis, for all data periods. SFs are extracted from the ratio of the sigmoid fits of data and MC efficiency curves, implemented to smooth out the SFs' distributions. All triggers become fully efficient starting from $\text{MET}_{\text{no-}\mu} \sim 300$ GeV.

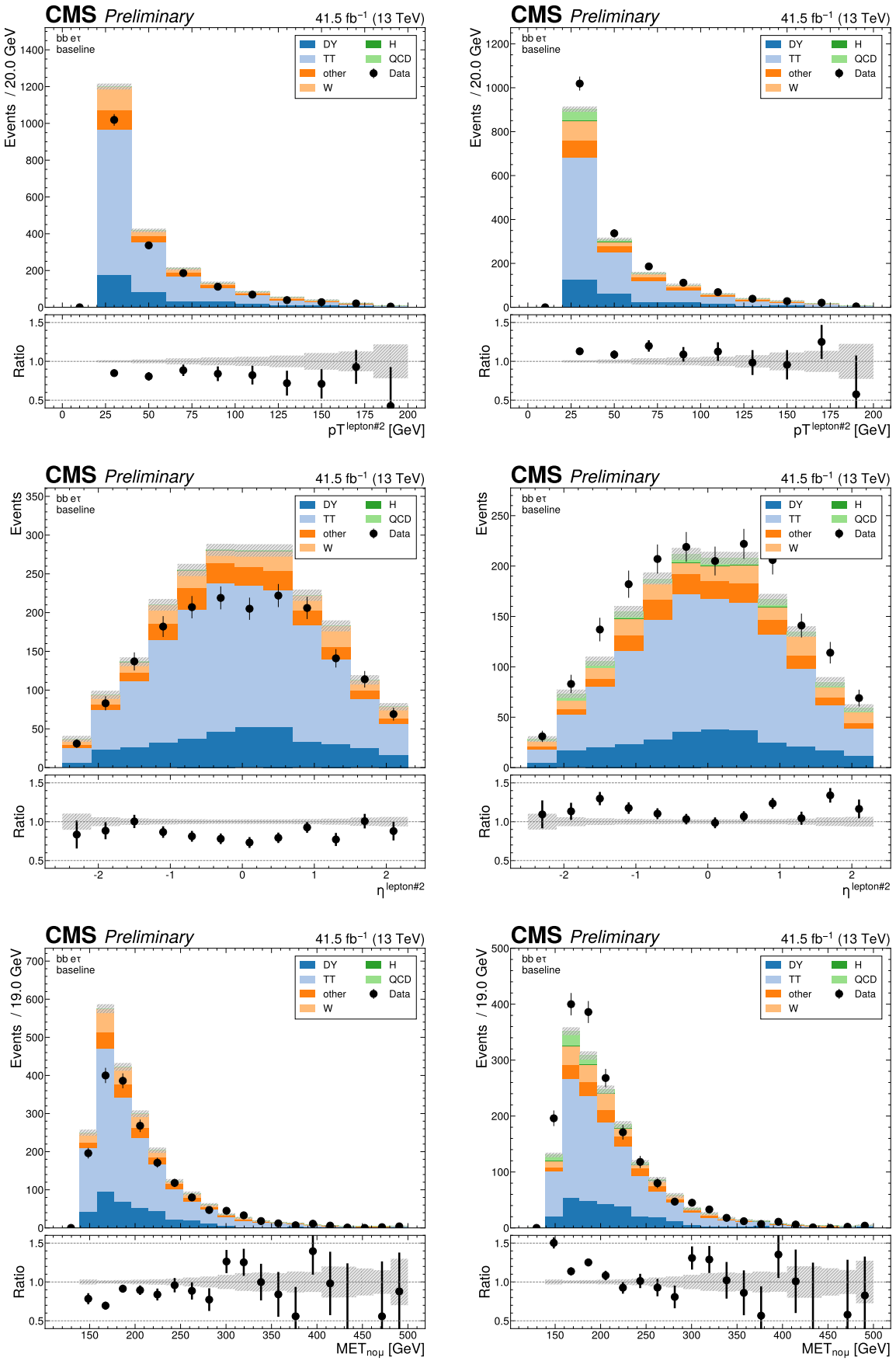

Figure 4.25: Comparison of chosen distributions without (left) and with (right) $\text{MET}_{\text{no-}\mu}$ SFs, for events triggered only by the $\text{MET}_{\text{no-}\mu}$ trigger, in 2017. We display the $\tau(p_{\text{T}})$ (top), $\tau(|\eta|)$ (middle) and $\text{MET}_{\text{no-}\mu}$ (bottom) for the $\tau_e\tau_{\text{h}}$ channel. The $\text{MET}_{\text{no-}\mu}$ SFs decrease the data to MC mismatch. Events triggered by MET with $\text{MET}_{\text{no-}\mu}$ below 180 GeV are removed from the SR.

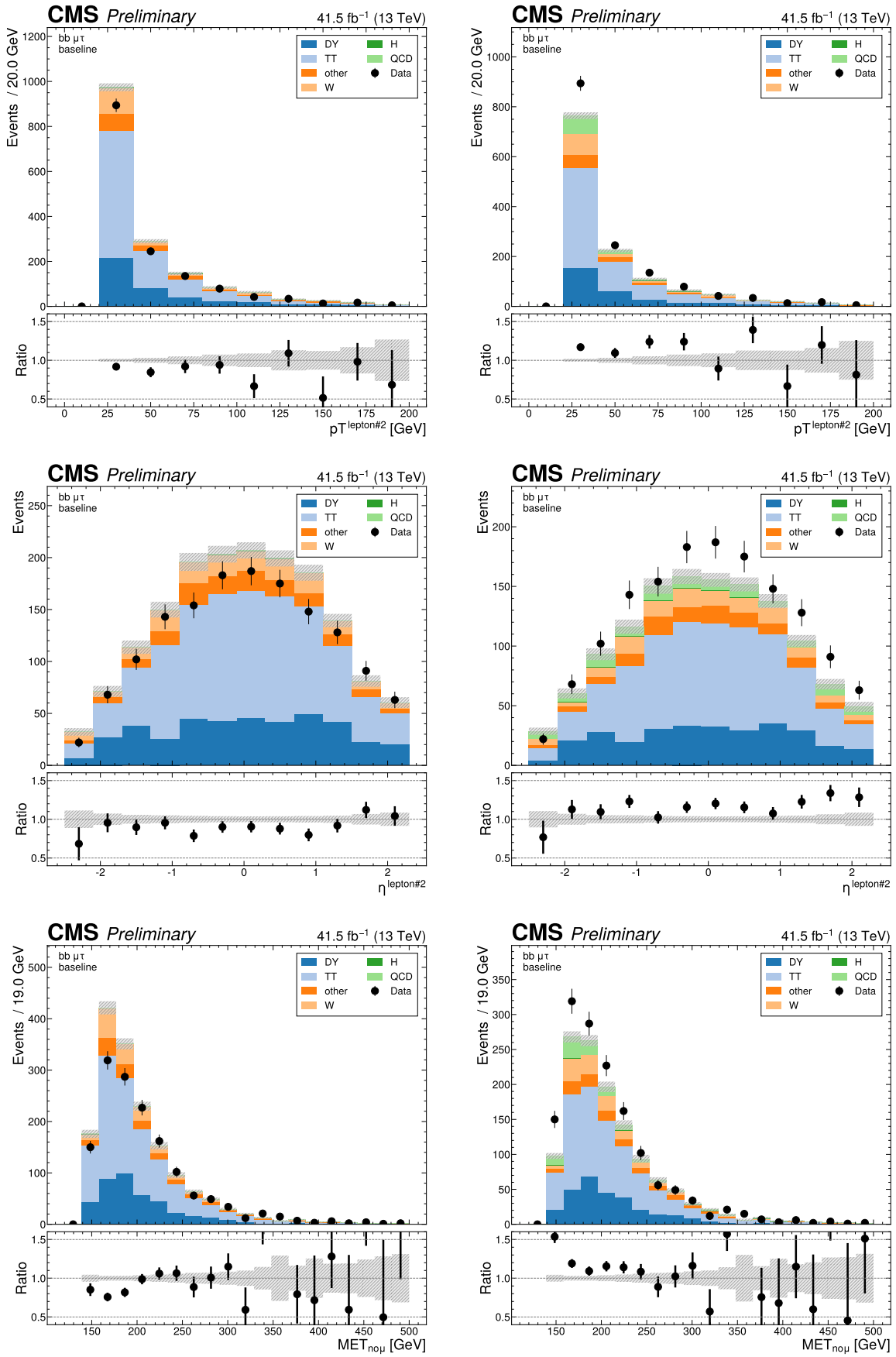

Figure 4.26: Comparison of chosen distributions without (left) and with (right) $\mathrm{MET}_{\text{no-}\mu}$ SFs, for events triggered only by the $\mathrm{MET}_{\text{no-}\mu}$ trigger, in 2017. We display the $\tau(p_{\mathrm{T}})$ (top), $\tau(|\eta|)$ (middle) and $\mathrm{MET}_{\text{no-}\mu}$ (bottom) for the $\tau_\mu\tau_{\mathrm{h}}$ channel. The $\mathrm{MET}_{\text{no-}\mu}$ SFs decrease the data to MC mismatch. Events triggered by MET with $\mathrm{MET}_{\text{no-}\mu}$ below 180 GeV are removed from the SR.

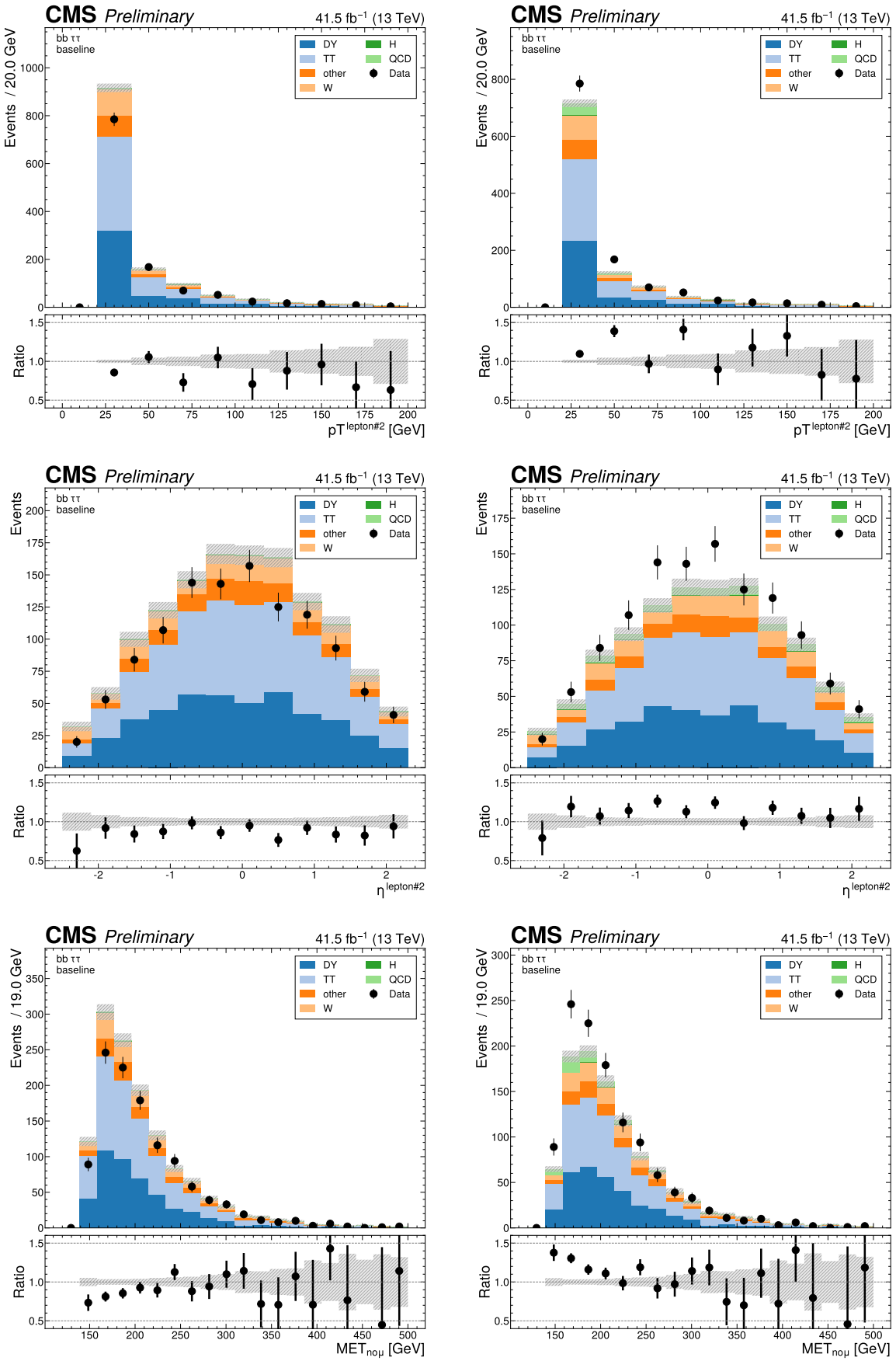

Figure 4.27: Comparison of chosen distributions without (left) and with (right) $\mathrm{MET}_{\text{no-}\mu}$ SFs, for events triggered only by the $\mathrm{MET}_{\text{no-}\mu}$ trigger, in 2017. We display the $\tau(p_{\mathrm{T}})$ (top), $\tau(|\eta|)$ (middle) and $\mathrm{MET}_{\text{no-}\mu}$ (bottom) for the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel. The $\mathrm{MET}_{\text{no-}\mu}$ SFs decrease the data to MC mismatch. Events triggered by MET with $\mathrm{MET}_{\text{no-}\mu}$ below 180 GeV are removed from the SR.

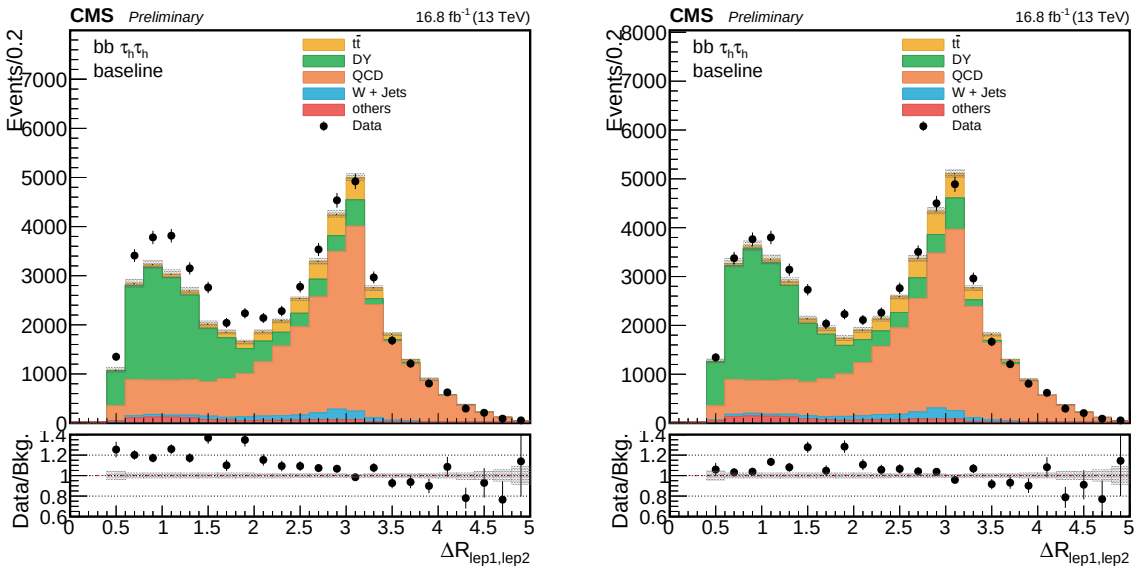

Figure 4.28: Comparison of the ΔR distribution with the baseline selection between the two leptons in the $\tau_h\tau_h$ channel before (left) and after (right) updating the `DeepTauVSjet` scale factors for 2016, as instructed by the CMS Tau POG. The agreement improves significantly.

weights computed with the method in Eq. (4.20) should change only the shape of the b-tag discriminant. In other words, before applying any b-tag selection criteria, expected event yields should be preserved: this means that the number of events (i.e. the sum of event weights) before and after applying b-tag weights should remain constant. In order to ensure this, the sum of event weights before and after applying b-tag event weights, without requiring any b-tag selection, is computed. The ratio $r = \sum \omega_{\text{before}} / \sum \omega_{\text{after}}$ is multiplied by the b-tag event weight. The values of these r factors are reported in Table 4.18.

Year	Decay Channel	r factor
2016	$\tau_\mu\tau_h$	1.0081
	$\tau_e\tau_h$	1.0068
	$\tau_h\tau_h$	1.0103
2017	$\tau_\mu\tau_h$	0.9993
	$\tau_e\tau_h$	0.9949
	$\tau_h\tau_h$	0.9547
2018	$\tau_\mu\tau_h$	1.0039
	$\tau_e\tau_h$	1.0040
	$\tau_h\tau_h$	0.9795

Table 4.18: Values of the r factors used to correct the b-tag event weights and preserve the normalization of the b-tagging discriminant.

4.7.7 Particle Net SFs

Our analysis considers the mass-decorrelated PNet $X \to b\bar{b}$ algorithm for its boosted category, as explained in Section 4.3.4. In particular, a selection cut is applied on the Low Purity WP of the algorithm's score, defined in Eq. (4.10). Since the jet tagger is

trained on MC samples only, and the latter do not perfectly agree with data, cutting on the tagger score inevitably leads to data/MC mismodelings. As usual, discrepancies must be corrected with appropriate SFs. Corrections vary depending on the MC sample considered, since the jets are generated by different physics processes. Given the development timescale of the PNet algorithm, no centrally provided SFs are yet defined for Run 2 background samples; they are only available for signal-like signatures. A custom derivation of SFs for all backgrounds samples is thus necessary.

The procedure developed to compute background PNet SFs starts from the observation that the analysis is dominated by DY and $\mathrm{t\bar{t}}$ backgrounds. In DY (plus jets), the "fat", or merged $\mathrm{b\bar{b}}$ jet most likely comes from random gluon or quark jets, misidentified as b jets. The SFs are derived in the DY CR, within a **boosted** topology. For the case of $\mathrm{t\bar{t}}$, one of the b quarks most likely comes from a top decay, and the second jet has a combinatorial origin. To derive the $\mathrm{t\bar{t}}$ SFs, the $\tau_e\tau_\mathrm{h}$ and $\tau_\mu\tau_\mathrm{h}$ SRs are combined in a region of high $\mathrm{t\bar{t}}$ purity, by considering events with a di-τ mass above 130 GeV. This region is very similar to the $\mathrm{t\bar{t}}$ CR detailed in Section 4.4.5, but without any requirement on $m_{\mathrm{b\bar{b}}}$. The DY and $\mathrm{t\bar{t}}$ CRs obtained suffer from a relatively small number of events in the **boosted** category. The SFs are obtained in p_T distributions with three p_T bins, as follows:

$$\varepsilon_{\mathrm{PNet}}(p_\mathrm{T}) = \frac{\text{Boosted CR \&\& Score}_{\mathrm{PNet}} > \text{Loose}}{\text{Boosted CR}}\,, \tag{4.21}$$

where "Boosted CR" refers to the CRs described above with events having at least one AK8 jet, and the year-dependent PNet scores can be inspected in Table 4.11 (right). The SFs are then simply calculated as:

$$\mathrm{SF}_k = \frac{\varepsilon_{\mathrm{PNet}}\left(\mathrm{Data} - \sum_{j\neq k}^{\mathrm{N_{MC}}} \mathrm{MC}_j\right)}{\varepsilon_{\mathrm{PNet}}(\mathrm{MC})}\,, \tag{4.22}$$

where $k \in$ DY, $\mathrm{t\bar{t}}$, and $\mathrm{N_{MC}}$ is the number of MC samples our analysis considers. The equation explicitly states that all backgrounds are removed from the data, except the one for which the SFs are being computed.

Considering now signal-like processes with a $\mathrm{b\bar{b}}$ decay, methods to derive SFs are already available within the CMS Collaboration, and the development of a custom method is therefore not necessary. Available methods always use "proxy jets", since it is experimentally very difficult to isolate a pure region of $\mathrm{H}\to\mathrm{bb}$ jets from data [379]. In particular, the SFs are here computed with the "sfBDT" method, which uses as proxy jets a large collection of multijet $\mathrm{g}\to\mathrm{b\bar{b}}$ events with additional selections. To ensure that the proxy jets are similar to the target signal-like jets, a BDT is developed to select a subset of multijet events exhibiting similar characteristics to the bb signal. The BDT, from which the SF method derives its name, was originally developed for the $\mathrm{VH}(\to\mathrm{c\bar{c}})$ analysis [380].

Three sets of SFs are thus defined, two for "DY-like" and "$\mathrm{t\bar{t}}$-like" backgrounds, and one for signal-like topologies. Each separate MC background is associated to one of these sets, depending on its topology:

- processes with vector bosons and potentially jets are DY-like: W+jets and EW processes in association with a vector boson;

- processes enriched with top quarks are $t\bar{t}$-like: tW, single top, $t\bar{t}H$, TTW, TTZ, TTWW;
- processes with $H \to bb$ or $Z \to bb$ signatures are signal-like: ZH, WZ, ZZ, WWZ, WZZ, ZZZ, TTWZ, TTZZ, TTWH.

More than one association is possible for some of the backgrounds, especially those including more particles. At the same time, processes with lower cross sections do not significantly impact final results. The chosen SF set is therefore not particularly important for those cases.

4.8 Binned Distributions

Having defined the entire analysis worflow from datasets to MC corrections, we can finally provide some examples to illustrate the kinematic phase-space being covered, and the relative importance of different background sources. Distributions are available in Figs. 4.29 to 4.31, where the three analysis decay channels and three categories are represented. In Fig. 4.32, we additionally provide examples of similar distributions in the two CRs introduced in Section 4.4.5. These distributions are used to evaluate the overall MC agreement with data, and are fundamental to understand whether additional corrections are required. The MC agreement with data is in general good, modulo some phase-space regions with lower number of events, notably in the **boosted** category or in the $\tau_h\tau_h$ channel. The statistical procedure for extracting the final results uses instead a parameterized DNN discriminant, a topic covered in the next Chapter, where the discriminative potential developed so far is fully realized.

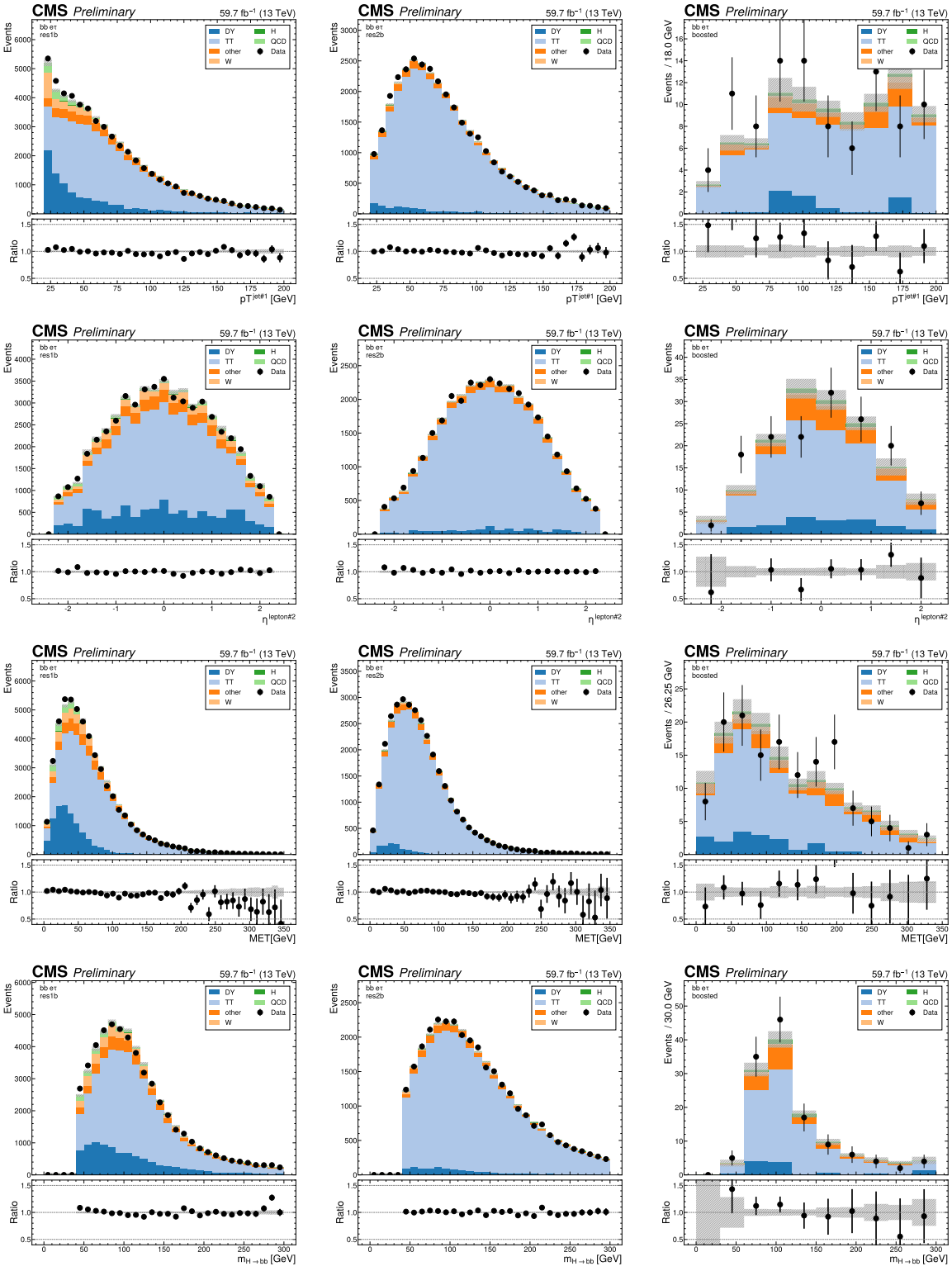

Figure 4.29: Binned distributions for the $\tau_e\tau_h$ channel in 2018. The four rows show, from top to bottom, the p_T of the first jet, the η of the second lepton, the MET, and the mass of the $b\bar{b}$ coming from one of the Higgs bosons. The three categories are shown, **res1b**, **res2b** and **boosted** in the left, middle, and right columns, respectively. Minor backgrounds are grouped together as "others". The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

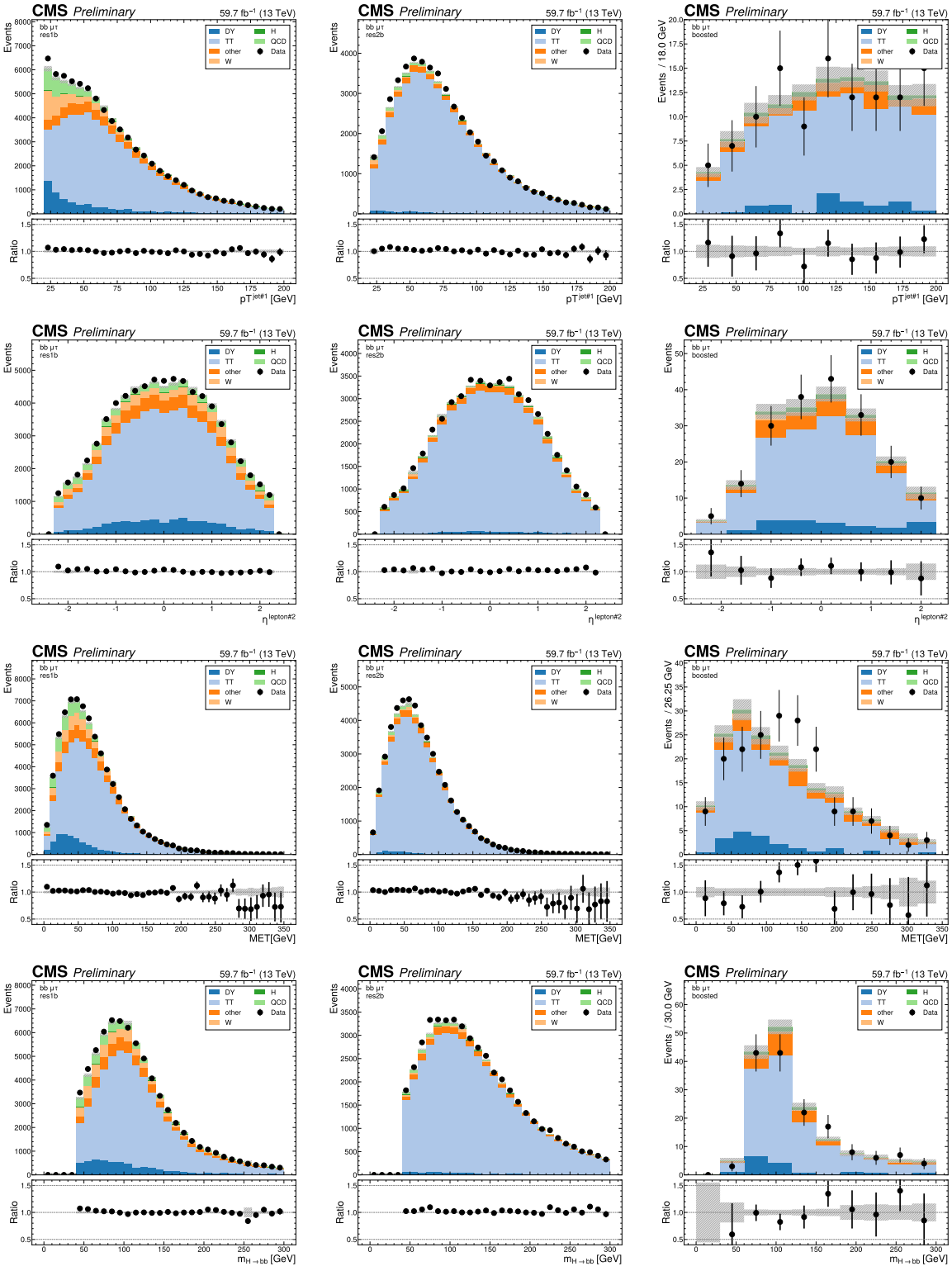

Figure 4.30: Binned distributions for the $\tau_\mu\tau_h$ channel in 2018. The four rows show, from top to bottom, the p_T of the first jet, the η of the second lepton, the MET, and the mass of the $b\bar{b}$ coming from one of the Higgs bosons. The three categories are shown, **res1b**, **res2b** and **boosted** in the left, middle, and right columns, respectively. Minor backgrounds are grouped together as "others". The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

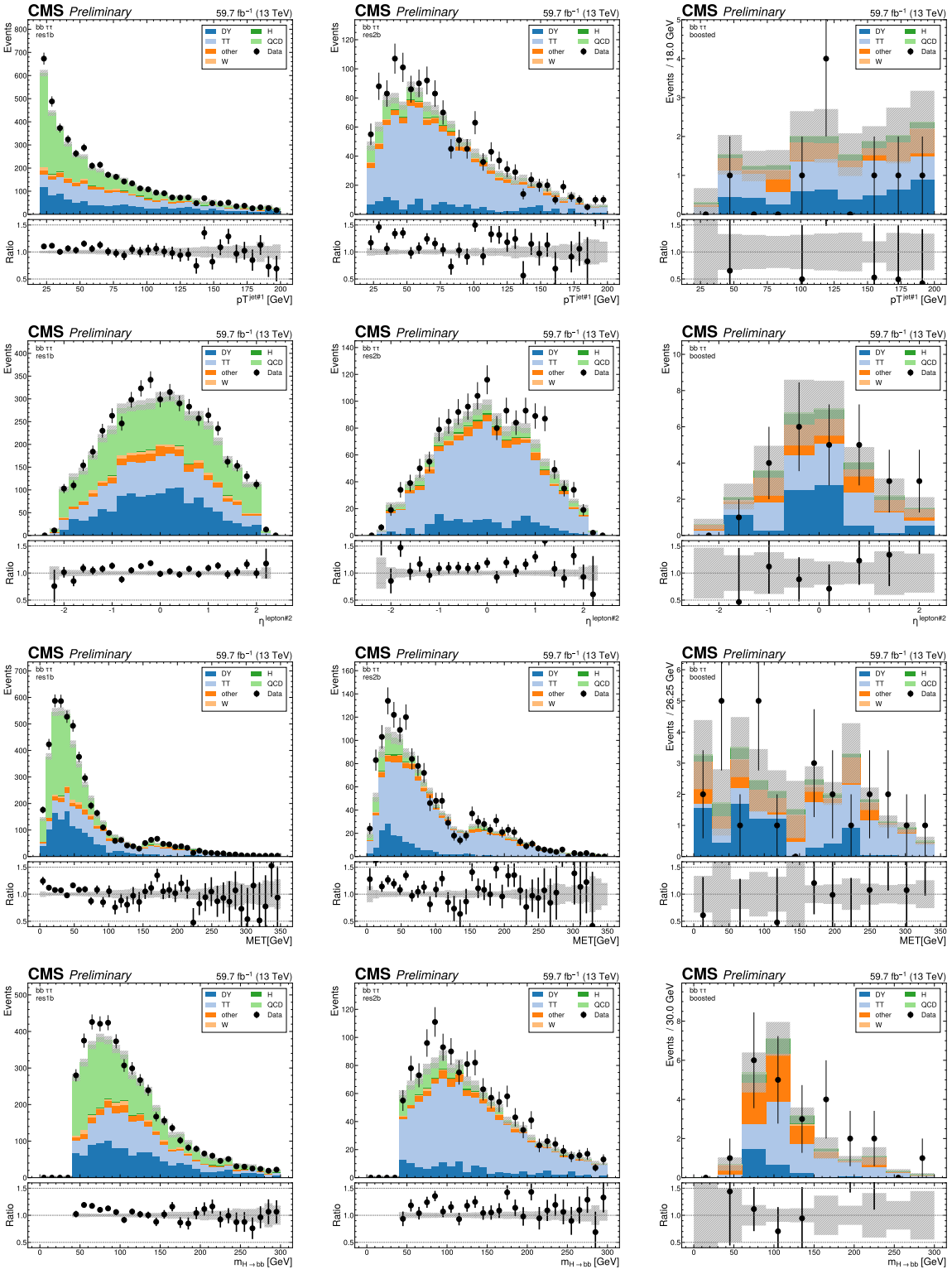

Figure 4.31: Binned distributions for the $\tau_h\tau_h$ channel in 2018. The four rows show, from top to bottom, the p_T of the first jet, the η of the second lepton, the MET, and the mass of the $b\bar{b}$ coming from one of the Higgs bosons. The three categories are shown, **res1b**, **res2b** and **boosted** in the left, middle, and right columns, respectively. Minor backgrounds are grouped together as "others". The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

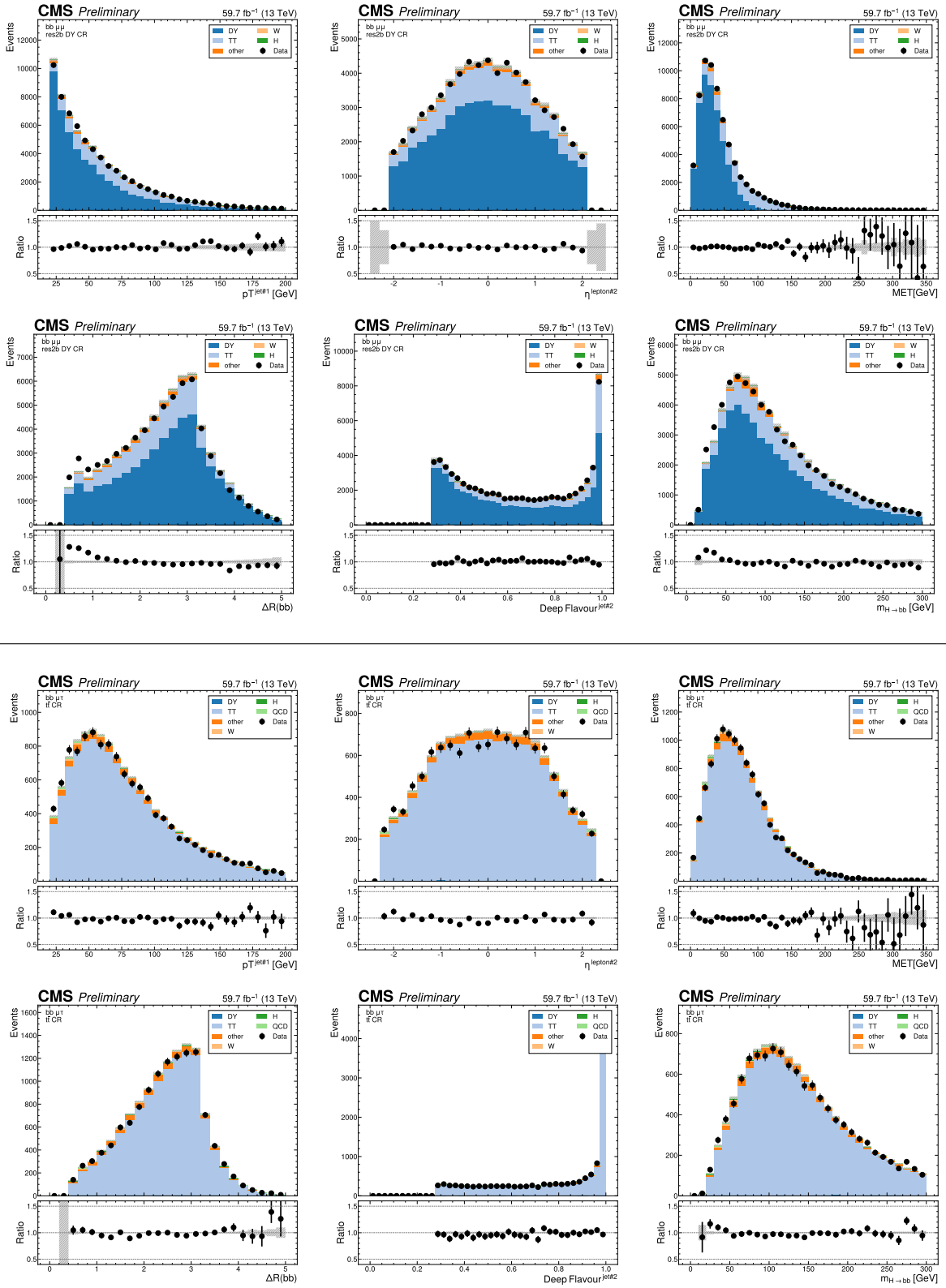

Figure 4.32: Binned distributions for the DY (top panel) and $t\bar{t}$ (bottom panel) CRs, in 2018. For each CR we show, from left to right and top to bottom, the p_T of the first jet, the η of the second lepton, the MET, the ΔR between the two b-jets, the `DeepJet` discriminant (also known as `DeepFlavour`) for the second jet, and the invariant mass of the $b\bar{b}$ pair. The categories are defined in Section 4.4.5. Minor backgrounds are grouped together as "others". The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

5 Resonant Search for HH $\to$ bb$\tau\tau$: Extracting the Limits

We have so far described in detail what in simple words amounts to a complex data processing step. Techniques that have been repeated and refined over the years were covered, including selections, estimates and related validations, with one goal only: increase the analysis sensitivity. The material presented in this Chapter finally substantiates all efforts into a quantifiable measurement. We want to be able to state whether the observed data matches the backgrounds, and thus the SM, or if it can alternatively accommodate one or more signal excesses. To draw such conclusion, we make use of a *discriminating variable*, sensitive to the underlying probability distributions of data and backgrounds, and which optimally encodes the analysis information so far discussed. The higher the discrimination power of such observable is, the more sensitive the analysis becomes, equating to measurements of smaller and smaller BSM cross sections. Conversely, would such observable be poorly defined, the analysis would not be able to detect NP, even if abundant. Therefore, the design of a discriminating variable is of dramatic importance, representing the nexus of all analysis decisions so far taken.

Historically, HEP discriminants were built from kinematical and topological variables related to the events, as these are expected to encode the most information regarding the processes under study[1]. Similarly, for many resonant analysis, as was the case for the analysis here reported, the combined invariant mass of the final state particles, or a modified version thereof, was usually considered. Such a choice naturally stems from the "bumpy" signature expected, i.e. the presence of a localized excess over a broadly distributed background.

In recent years, a paradigm shift took place in HEP: machine learning techniques, and in particular deep learning (DL), became widespread, due to their exponential improvement in classification and regression problems.In Despite being based on established principles [372, 386, 387], DNNs rose to prominence only about 15 years ago, but are today used in most industries, companies and academic institutions around the world, for tasks as varied as language generation, visual recognition, machine translation and fraud detection, among many, many others. The main reasons that lead to such a technological shift were the increased computational power, the introduction of new high-performant NN architectures, and crucially the vast amounts of available data, which lie at the center of most DL techniques. HEP has accompanied the trend, exploiting such algorithms to improve the analyses' performances. Even triggers have been included in the process, achieving improvements that surpass all expectations.

The results here presented are shown in the form of expected 95% CL upper limits. Observed upper limits will be computed once the CMS internal analysis review process is completed, which should happen by early 2025 the latest, in preparation for the 2025

[1] Matrix element discriminants are occasionally employed, as they have been shown to provide the maximal information for the description of the underlying physics processes. They can only be successfully applied for "clean" backgrounds, without jets, for instance. The approach is applied via the MELA package [381–385].

Moriond conference. Until then, analyzers do not have permission to show real data in the final discriminant's distributions, or to consider real data when extracting limits.

In Section 5.1, we start by explaining how DL techniques are leveraged to define a robust discriminating variable. We then proceed, in Section 5.2, with a description of all systematic uncertainties taken into account, arising from variations to the inputs of the DL procedure, and also due to shifts to the several MC correction weights considered in our analysis. The quantification of the differences between two alternative scenarios is the task of the statistical treatment presented in Section 5.3. Specifically, we formulate the hypothesis that the data is not described by the SM, and verify if the measurements agree with such a claim. The 95% CL expected upper limits are presented in Section 5.4.

5.1 Signal Extraction

The adoption of DL in the X → HH → bbττ analysis naturally arises from the wish to extract the best possible limits with current technology. The latest nonresonant HH → bbττ analysis [111] has already exploited such techniques, following the approach introduced in Ref. [388], achieving significant improvements with respect to previous publications. However, while in the nonresonant analysis one wished to extract the SM HH signal alone, in the resonant analysis here presented such discrimination must be done for each signal sample, associated to a given mass and spin hypothesis. In order to create an optimal discriminant for N signal samples, one could define N separate discriminants, and optimize all of them separately. This approach is however highly impractical with N = 50 (2 spin hypothesis combined with 25 resonance mass points), leading to an unsustainable complexity in data processing, and to prohibitively large resource needs. Instead, using a single network is more convenient, as it encodes all the information for different hypotheses in the same location and, if needed, enables the properties of under-populated signal samples to be extrapolated from other samples.

A parameterised DNN (pDNN) [389] is exploited in this Thesis, where the parameters refer to the hypotheses being tested, namely the mass and spin values. Parameterized learning has already been tested with the HH → bbττ topology, and was found to be useful in Ref. [390]. Additionally, we have observed that the pDNN here developed is more performant than a non-parameterized version with similar architecture and size. Despite the performance improvement, the parameterized approach introduces some complexities. During the training of the model, the signal events are passed to the network together with their mass and spin values. For the background, the mass and spin are randomly sampled from all hypothesis being tested. During inference, for the signal, a prediction is generated for the sample's mass and spin values. Instead, for data and background, the mass and spin are fixed to all possible combinations, such that to each specific hypothesis a score is attributed. This equates to generating N predictions per event. Table 5.1 summarizes the approach.

The parameterized approach also introduces some challenges in the determination of systematic uncertainties. Given that N = 50, and that each of those predictions has an independent set of associated systematic uncertainties, the total number of systematic uncertainties to be considered increases enormously. For each variation affecting the input features of the pDNN, a corresponding systematic uncertainty must be computed for all 50 combinations. A significant part of this Thesis was devoted to process, distribute

	Sample Type	Description
Training	Background	Randomly sample m and s from all possible values.
	Signal	Use the m and s of the signal sample.
Inference	Data and Background	Generate a prediction for each event N times.
	Signal	Generate a prediction using the m and s of the signal.

Table 5.1: Treatment of resonance mass and spin values during the pDNN training and inference steps. N = 50 stands for the number of signal hypotheses, a combination of 25 mass values and 2 spin values, while m and s refer to specific mass and spin values, respectively.

and store all the data necessary to correctly compute systematic uncertainties within a parameterized approach.

5.1.1 Architecture

The network's architecture is illustrated in Fig. 5.1. It has a densely connected structure [391], prepended by a Lorentz Boost Network (LBN) [392] and the regression head of the network described in Section 4.5. The total set of input features is split into the subset of features for the $m_{\tau\tau}$ regression network and the subset of features fed into the LBN, which receives all four-momenta of the two reconstructed leptons and of the two reconstructed b-jets in the event. Depending on the resolved or boosted topology of an event, default values are assigned to some input nodes. The outputs of these two networks are concatenated in an intermediate layer and fed into a densely connected DNN, which consists of 8 hidden layers, each containing 128 nodes. This architecture has shown to smooth out the surface of the loss function [393], helping the training procedure when using large networks [394]. A dropout [395] of 5% is considered. Batch-normalization layers are added after the concatenation of the two heads and before every dense layer [374]. Swish-1 activation functions are used for hidden layers [396]. For both the $m_{\tau\tau}$ regression network and the LBN, categorical features are embedded using entity embeddings [397]. At the end of the network, each event is categorized according to one of three possible classes: resonant ggF bb$\tau\tau$ signal, DY background or t$\bar{\text{t}}$ background. For simplicity, we are currently exploiting the signal score only.

5.1.2 Training

In our analysis, in order to avoid overfitting, the available samples are divided into five folds, using the same cross-validation technique that was described in Section 4.5. Each DNN is initialized with an independent "He-Uniform" weight-initialization scheme [398]. After the training, the predictions are averaged out across all discriminants.

The training data consist of all background resonant signal samples, for all 4 data-taking periods. Three classes of input features are considered:

- **continuous**: all continuous features listed in Table 4.14, already used by the $m_{\tau\tau}$ regression model, and additional inputs to the LBN, namely the $\tau\tau$ and bb four-momenta;

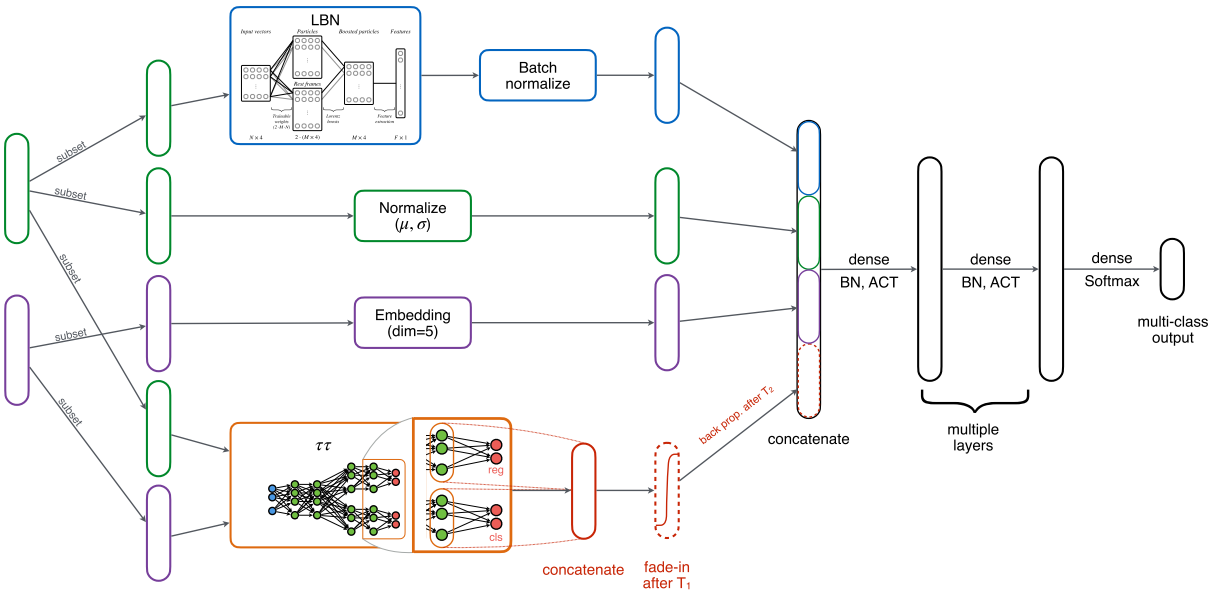

Figure 5.1: Architecture of the X → HH → bbττ final discriminant. The green and purple colors refer to continuous and categorical inputs, respectively. The layout corresponds to a combination of the $m_{\tau\tau}$ regression network, the LBN and a densely connected DNN. The weights of the regression network are initially fixed to the values from previous independent trainings, and are slowly activated as the training time passes, which we refer to as "fade-in". "BN" stands for batch-normalization, and "ACT" stands for activation function. Courtesy of Marcel Rieger.

- **categorical**: decay channel, the charges of the reconstructed τs, the data-taking period, and a boolean value standing for whether an AK8 jet is present in the event;

- **parameterized**: the mass and spin values representing the hypotheses being tested.

The data is processed in batches, each with a size of 4096 events. The `Adam` optimizer [399] is considered, starting with a learning rate of 0.003. As the training progresses, the learning rate is reduced so that the minimum of the loss function is not overshoot. A dynamic learning rate scheduler steadily decreases the learning rate based on the validation loss. The training is stopped as soon as 10 epochs have passed without the improvement of the validation loss.

Since the network utilizes the weights of the $m_{\tau\tau}$ regression head that were already optimized, care must be taken during the final discriminant optimization to avoid hurting the performance. For this reason, the weights of the $m_{\tau\tau}$ regression network are initially kept fixed, and are introduced only after 150 training steps. Additionally, the connection is established through a *fade-in* layer. The layer slowly introduces the output of the $m_{\tau\tau}$ regression network to the densely connected DNN, such that the weights are not changed too abruptly. This is achieved by multiplying the outputs of the connection between the two networks by a factor which is linearly increased over 20 epochs.

Finally, the signal, DY and $t\bar{t}$ classes are represented in equal fractions within each batch. This is required to prevent the discriminant to ignore signal events, given the multiplicity disproportion between signal and background. Within each fraction, the $t\bar{t}$ and DY samples are distributed according to MC weights, to reflect the true background source composition.

5.1.3 DNN Score Distributions

The distributions can be seen in Fig. 5.2, where a signal sample is also included for reference. The cross section of the signal is arbitrary, and in the plot it is scaled for visualization purposes to the final upper limits, shown in Section 5.4.1, multiplied by the bb$\tau\tau$ BR. Some data points are included, despite belonging to the SR, in order to demonstrate compatibility with the MC. The region being shown has a DNN score below 0.8, to avoid biasing the results by looking at the data in a signal-rich region.

Occasionally, the QCD background contribution is so small that statistical fluctuations with the ABCD method can lead to a negative bin content. In such cases, the QCD contribution is removed from the histograms.

The binning follows a "flat-signal" approach, where the number of signal events per bin is, as much as possible, required to be the same, and certain requirements are imposed on the number N of background events in each bin:

- $N_{DY} > 1$;
- $N_{t\bar{t}} > 1$;
- $N_{DY} + N_{DY} > 4$;

where the goal is to avoid bins with too little background, preventing fit instabilities, while ensuring that the two most dominant background sources are always present. For simplicity, the bin with the lowest DNN score is left with less signal events, due to the termination conditions of the binning algorithm. It does not affect the final results, since that bin provides no sensitivity.

5.2 Systematic Uncertainties

Measurements in all branches of Science consist of estimating one or more unknown parameters from observed data, which are later compared to existing models. This process is called *statistical inference*, or just inference for short [371]. The parameters one wishes to access can be categorized in two classes:

- **Parameters of Interest**: Representing the end goal of the experiments, Parameters of Interest (POIs) define the underlying probability distributions analyses wish to measure. In our case, we want to know whether the observed data can be described by our signal models, on top of the expected background. This corresponds to the definition of a single POI which scales the signal, the *signal strength modifier* $\mu \equiv \sigma_{\text{obs.}}/\sigma_{\text{signal}}$. The signal is normalized to an arbitrary value, $\sigma \times \mathcal{B} = 1\,\text{pb}$ in our case. Measuring $\mu \neq 0$ would indicate the presence of a BSM signal. In Section 5.3.3 we explain how this POI is encoded in the maximum likelihood approach to extract the final results.
- **Nuisance parameters**: The probability distributions connected to the observed data often depend on various effects not directly related to the POIs. The effects arise from the experimental apparatuses being used, such as for instance the limited resolution of a detector, miscalibrations, the existence of discrepancies between background models and real data, or the lack of sufficient events in simulations or data sidebands. The POIs can also depend on related theoretical predictions lacking

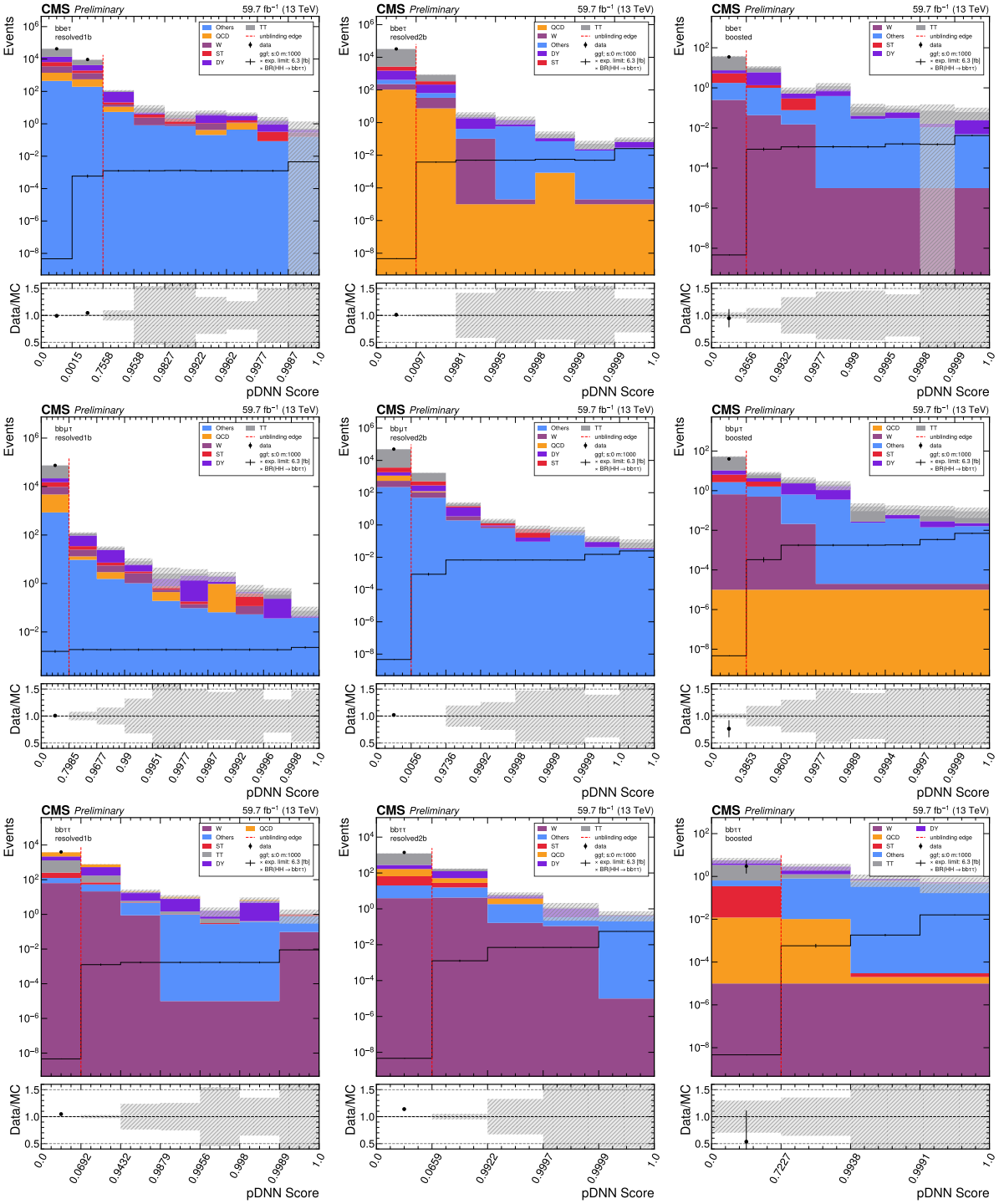

Figure 5.2: Distribution of the pDNN score for the $\tau_e\tau_h$ (top row), $\tau_\mu\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) channels, in 2018. The three columns refer to the **res1b** (left), **res2b** (middle) and **boosted** (right) categories. The signal distribution is mostly flat, and is scaled for visualization to the product of the bb$\tau\tau$ BR with the expected limits shown in Section 5.4.1. The plots are partially unblinded (see Section 5.3.5), for all background dominated bins with a pDNN score below 0.8. Details are provided in the text. Minor backgrounds are grouped together as "others". The shaded gray bars represent the statistical uncertainty of the background samples. The statistical uncertainties of the data are shown as black vertical error bars.

sufficient precision. In order to measure the POIs, one has therefore to also measure additional unknown parameters, appropriately called *nuisances*. The integration of nuisances on the final result is also done with a dedicated statistical treatment based on the maximum likelihood approach.

Uncertainties related to the POIs caused by data fluctuations alone are denoted *statistical uncertainties*. Instead, the propagation of the imperfect knowledge of all the elements of an analysis to the final result corresponds to *systematic uncertainties*, which are encoded via nuisances. The careful measurement of systematics plays a central role in any physics analysis, since systematics can dramatically affect the sensitivity of the result. In this section, we list the many sources of systematics considered. Systematic uncertainties affecting only the yield of a given signal or background process are detailed in Section 5.2.1, and systematics also affecting the shape of the final discriminating variable are listed in Section 5.2.2.

5.2.1 Normalization Uncertainties

All systematic uncertainties should be assumed to have an impact on the shape of related distributions, and not simply on their rate. However, if for a specific systematic source it is verified that shape modifications are negligible when compared to the expected sensitivity of the analysis, downgrading an uncertainty to be normalization-only has two important advantages. The first relates to the much lower analysis complexity involved. For each additional shape systematic, an extra set of variables representing the corresponding variations must be computed, processed and stored. This is also valid for the inference step of the final discriminant, as long as the distributions being varied are used as a pDNN input. In that case, one must assess how much the input's changes affect the discriminant output, for each mass point and spin hypothesis. The second advantage comes from the time taken to run the binned maximum likelihood fit. A shape systematic is naturally more time-consuming, given that the binned distributions include more parameters than one single rate parameter.

The specific uncertainties are often centrally recommended, or are instead devised by the analysts themselves, and then approved by the Collaboration during the many review steps the work is subject to. We list them in what follows:

- **Luminosity**: The uncertainty on the luminosity measurement is estimated by the CMS Lumi POG. Different uncertainty sources are considered, and their correlations are reported in Table 5.2. These values are obtained from dedicated Van-der-Meer scans during data-taking. The uncertainties are applied to the signals and to all the background processes estimated from MC simulation. The normalization of the QCD background is obtained from data and hence not subject to luminosity uncertainties.

- **Electron and muon isolation and identification efficiencies**: The uncertainties on electron and muon ID and isolation efficiencies are determined from the uncertainties on the SFs describing data/MC discrepancies. A value of 1% for both electrons and muons is obtained.

- **QCD estimation uncertainty**: The QCD background, which is estimated from data using the ABCD method, can be affected by different sources of systematics. We define three uncertainty sources:

Uncertainty [%]	2016	2017	2018
2016, uncorrelated	1.0	0.0	0.0
2017, uncorrelated	0.0	2.0	0.0
2018, uncorrelated	0.0	0.0	1.5
2017-2018, correlated	0.6	0.9	2.0
2016-2017-2018, correlated	0.0	0.6	0.2

Table 5.2: Luminosity uncertainties for all data-taking years, in percentage, explicitly separating correlated and uncorrelated sources. The uncertainties are centrally recommended by the CMS Collaboration, and are used by all physics analyses.

- **Shape**: Since the shape of the QCD contribution can be estimated from two regions, B and C, a normalization uncertainty can be obtained by comparing the two shapes.
- **Yield correction factor**: The statistical uncertainty on the B/D and C/D correction factors is defined as the sum in quadrature of the statistical uncertainties on the event yields in the concerned regions. The value of this uncertainty ranges from 5% to 100% depending on the category, channel and year considered.
- **Additional term**: The first validation procedure described in Section 4.6.1 uses different `DeepTau` WPs to ensure that the multijet estimate does not depend on the τ isolation. In the few cases where a dependency is observed, an additional uncertainty is taken into account.

Table 5.3 shows the QCD normalization uncertainties for all years, channels and categories. The cell is empty if no QCD could be estimated (as either regions B, C or D have negative yield).

Category	Channel	2016	2016 APV	2017	2018
boosted	$\tau_h\tau_h$	–	62.0	–	57.0
	$\tau_\mu\tau_h$	–	–	–	–
	$\tau_e\tau_h$	–	–	44.6	
res1b	$\tau_h\tau_h$	8.8	6.7	6.4	4.8
	$\tau_\mu\tau_h$	7.4	5.3	6.3	3.8
	$\tau_e\tau_h$	17.2	9.2	10.9	11.2
res2b	$\tau_h\tau_h$	55.5	35.0	36.2	25.6
	$\tau_\mu\tau_h$	24.1	16.1	16.0	10.0
	$\tau_e\tau_h$	–	–	–	

Table 5.3: QCD normalization uncertainties, taking into account all uncertainty sources, in percentage. A cell is empty if no QCD background could be estimated due to negative B, C or D region yields.

- **Final state branching fractions**: Two different normalization systematics are obtained by propagating the theoretical uncertainties on the Higgs decays to $b\bar{b}$ and $\tau\tau$, assuming $m_H = 125\,\text{GeV}$. We here list the values with their total uncertainties

[49]: $\mathcal{B}(\mathrm{H} \to \mathrm{bb}) = 0.5824 \pm^{1.25\%}_{-1.27\%}$ and $\mathcal{B}(\mathrm{H} \to \tau\tau) = 0.06272 \pm 1.65\%$. These uncertainties are included when converting the channel-specific upper limits to the inclusive $\mathrm{X} \to \mathrm{HH}$ process.

- **Cross sections of simulated processes**: For $\mathrm{t\bar{t}}$, W + jets, single top, single Higgs and multi-boson backgrounds, uncertainties due to the imperfect knowledge of the normalizations of the simulated processes are considered, with the values reported in Table 5.4.

Process	Uncertainty	Value [%]
ZH	QCD scale	+3.8/−3.0
	PDF	±1.3
	α_S	±0.9
WH	QCD scale	+0.5/−0.7
	PDF	±1.7
	α_S	±0.9
ttH	QCD scale	+5.8/−9.2
	PDF	±3.0
	α_S	±2
ggH	QCD scale	±3.9
	PDF	±1.9
	α_S	±2.6
qqH	QCD scale	+0.4/−0.3
	PDF	±2.1
	α_S	±0.5
$\mathrm{t\bar{t}}$	QCD scale	+2.4/−3.5
	PDF	±4.2
single top	QCD scale	+4.2/−3.5
tW	QCD scale	±5.4
W	QCD scale	+0.8/−0.4
EWK	QCD scale	+2.0
VV	QCD scale	+10.0
VVV	QCD scale	+10.0

Table 5.4: Theoretical uncertainties on the background production cross sections.

5.2.2 Shape Uncertainties

Significant shape dependencies might exist between the final discriminant and some of the variables for which uncertainties are defined. Whenever such correlations exist, a single normalization factor is not sufficient, and the full binned shape is taken into account. The evaluation of a single uncertainty is achieved by shifting its template up and down, i.e. by estimating an upward or downward variation for what can be reasonably expected to cover the most common values. For continuous variables, this can for instance correspond to $1\,\sigma$ intervals around the nominal value. The kinematic properties of objects depending on those variations must also be accordingly shifted. For instance, when the energies of τ_hs are shifted due to jet energy scale (JES) or JER corrections, the MET in the event must also take the changes in energy into account.

Concerning the final discriminant, all input features must also be shifted up and down, independently for each systematic source, leading to numerous pDNN evaluations. If the same systematic source shifts multiple distributions, those variations are considered together during the pDNN inference, leading to a single up or down template. For example, when shifting the energy scale of resolved jets, we shift the energy of the two

AK4 jets together, and also propagate the shift to the MET. In the following, we provide a list of all shape systematics considered in the $X \to HH \to bb\tau\tau$ analysis.

- **L1 prefiring**: The prefiring issue described in Section 4.7 is addressed via two independent shape uncertainties, for the ECAL and Muon Chambers. The ECAL uncertainty contains a fixed 20% contribution, to which one adds in quadrature a statistical uncertainty dependent on η and p_T. The uncertainty of the Muon Chambers is very similarly defined, but its statistical component also includes a dependence on ϕ. While the ECAL prefiring uncertainty is applied to 2016 and 2017 MC simulations only, the muon one is applied to all years.

- **Pileup reweighting uncertainty**: The systematic uncertainty on the PU reweighting technique is estimated by applying up and down variations on the PU weights. The variations correspond to using an inelastic pp cross section shifted by 4.6%, leading to 66 mb for the downward variation and to 72.4 mb for the upward variation, compared to the nominal cross section value of 69.2 mb, for Run 2. We recall that PU variations are additionally used to correct the normalization used to weight MC samples.

- **Energy scale of hadronic tau leptons**: The uncertainties are derived by combining two different measurements: the low and high-p_T measurements of $Z \to \tau\tau$ and off-shell $q\bar{q} \to W^* \to \tau\nu$ events[2], respectively. The following p_T-binned scheme is applied:
  - $p_T(\tau) \leq 34$ GeV: from the low-p_T measurement;
  - $34 < p_T(\tau) < 170$ GeV: linearly interpolated between the uncertainties in the low- and high-p_T measurements;
  - $p_T(\tau) \geq 170$ GeV: from the high-p_T measurement.

  The 34 and 170 GeV boundaries represent the mean p_T values from the two measurements.

  Four different uncertainties are provided by the Tau POG to take into account the different τ decay modes in this analysis. When considering the uncertainty for a particular decay mode, the shift is applied only to the truth-matched τ_h candidates that are reconstructed with that particular decay mode; all other τ_h candidates are left unchanged. The uncertainty sources are considered to be decorrelated between data-taking years and tau decay modes.

- **Energy corrections for prompt electrons**: A set of uncertainties is associated to the electron energy corrections, as a change in electron energy can affect the final result in several ways. Some events with electrons can for instance be excluded if they were previously part of the analysis, or vice-versa, based on some kinematic cut. The energy scaling and smearing are considered separately, leading to four variations in total. In practice, the energy of the electrons are corrected by applying the variations on the original, uncorrected electron four-momenta. The uncertainties are only applied to genuine electrons. Uncertainties on genuine electrons reconstructed as taus are instead computed using the `DeepTau` uncertainties covered in the next bullet point.

[2] The $q\bar{q} \to W^* \to \tau\nu$ process produces high-p_T τs via highly virtual W bosons with little hadronic activity.

Equivalent shape uncertainties for muons are known to be negligible and are not implemented. If they were included, they would be structured in the same way as described for electrons.

- **Energy scale corrections for electrons and muons faking taus**: Occasionally, electrons and muons are misidentified as τ_hs. For electrons, this might happen when there is one charged hadron and zero or none neutral hadrons. When only one charged hadron is present, with a BR of 11.5%, the electron can be misidentified as the charged hadron. When a neutral hadron is additionally present, with a BR of 26.0%, the fake electron can be wrongly reconstructed with additional bremsstrahlung photons. Muons can be misidentified just like the electrons, but only when no neutral hadron is present, as they radiate less.

  The uncertainties are centrally provided. They are treated as uncorrelated across different decay modes and data-taking years. They depend on the two decay modes mentioned above, and are only applied to truth-matched electrons and muons. The muon energy scales were added for completeness only, as the recommended approach is to simply assign a 1% flat uncertainty.

- **Jet energy scales and resolutions**: For the JESs, a reduced set of 11 sources is used, as recommended. The sources take into consideration dependence effects on ISR, final state radiation (FSR), jet flavour, kinematics, detector regions and miscalibrations. The uncertainties are the same for AK4 and AK8 jets. They are propagated to the definition of MET, replacing all nominal jets in an event by the corrected ones. This is repeated 11 times, leading to 22 variation templates. Concerning the JERs, up and down variations are centrally provided for all jets, taking into account their clustering radii (AK4 or AK8). The corrections are then applied before performing any selection on the jets and MET.

- **Scale factors for τ identification**: The uncertainties arising from the application of the `DeepTau` tau lepton ID SFs do not modify single objects, but rather the overall event weight. All SFs are centrally provided.
  - **Discrimination against jets**: The SFs are given in bins of the τ decay mode, and the p_T dependence is fitted using linear functions. Two statistical uncertainties, decorrelated across decay mode and data-taking period, appear from the linear fit to the p_T distributions. Several systematic components are defined, based on the correlation between decay mode and data-taking periods. τs of $p_T > 140\,\text{GeV}$ bring two additional systematic contributions. An additional systematic is included to correct for extrapolations when the τ_h has $p_T > 300\,\text{GeV}$. In total, 18 systematic sources are considered, per data-taking period.
  - **Discrimination against electrons**: The SFs are provided as a function of the data-taking years and $\tau(\eta)$, being uncorrelated across both. Two different uncertainties are identified per data-taking period, one for the CMS barrel and one for the endcaps.
  - **Discrimination against muons**: The uncertainties are provided in five $|\eta|$ bins, but are found to be negligible, and are thus not considered.

- **Trigger scale factors**: The uncertainties arising from the application of the trigger SFs do not modify single objects, but only the overall weight of each event.

  - **Legacy triggers**: Four different uncertainties are centrally provided, binned in the decay mode of the τ_hs, and they take into account the hadronically decaying legs of the $\tau\tau$ pair. Two more trigger uncertainties are used to cover the lepton legs in the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels.
  - **MET trigger**: $\text{MET}_{\text{no-}\mu}$ SFs are extracted from the fitted parameters of the sigmoid function in Eq. (4.19), and propagated using the error propagation of a ratio. The maximum allowed value of the sigmoid function is used whenever the event has a $\text{MET}_{\text{no-}\mu}$ value beyond the fit validity range.
  - **Single tau trigger**: For the single-τ trigger, we use as uncertainties the ones listed in Table 4.17, which are centrally provided.

- **B-tagging scale factors**: Uncertainties on b-tag SFs are centrally provided. The impact on the purity of b-jets is estimated by varying the contamination from light + c jets and b + c jets in heavy and light flavor regions, respectively, by ±20%. Multiple statistical uncertainties are also taken into account. In total, 8 systematic uncertainties are considered, per data-taking period. These uncertainties do not modify single objects, but rather the overall event weight.

- **Pile-up jet identification**: Uncertainties on the PU jet identification SFs are centrally provided as a function of jet p_T and η, being applied on an event-by-event base. This uncertainty does not modify the single objects, but only the overall event weights.

- **Monte Carlo finite sample size**: The limited number of simulated events in the bins of the discriminant distributions in the final fit also carry an uncertainty, independent for each bin. The Barlow-Beeston approach [400] is used to take such uncertainties into account, introducing a set of nuisance parameters which multiply the expected number of events in each bin, for each background source.

5.3 Statistical Treatment

Having defined a final discriminant and the systematic uncertainties that might impact the analysis, all ingredients are in place to extract the final result. This is the task of the statistical treatment, which in this Thesis is the now standard *modified frequentist approach*, a statistical method originally developed for Higgs boson searches at LEP [401]. Since then, it has been extensively used by searches and combinations from the LHC experiments, notably for the combination paper reporting the observation of the Higgs boson [402, 403]. Central to the approach is the much older concept of *hypothesis testing*, which is introduced in Section 5.3.1. We then provide a short description of the Neyman-Pearson Lemma in Section 5.3.2 which, together, with the concept of likelihood maximization described in Section 5.3.3, forms the basis of the definition of the test statistics used in HEP, and which is described in Section 5.3.4. We conclude by explaining in Section 5.3.5 the difference between expected and observed results, and why the most signal-sensitive data regions are not analyzed until well into the final stages of a HEP study.

5.3.1 Hypothesis Testing

In statistics, the term *hypothesis* refers to a statement concerning the agreement of the observed data with a given predictive model [404]. The hypothesis being assessed is traditionally called the *null hypothesis*, or H_0, against which a series of *alternative* hypotheses can be compared: H_1, H_2, H_3, and so on. Hypotheses denote PDFs $f(x;\theta)$ which depend on the measured data $x = (x_1, x_2, x_3, ...)$ and potentially on free parameters $\theta = (\theta_1, \theta_2, \theta_3, ...)$ which are estimated from data. To measure the agreement between a given hypothesis and the data, one constructs a function of the variables being measured, called *test statistic*, or $t(x;\theta)$. Each hypothesis implies a different PDF for the test statistic, denoted as $g(t|H_0)$, $g(t|H_1)$, $g(t|H_2)$, etc. One usually tries to construct the simplest test statistic enabling the largest discrimination possible between the hypotheses being compared. In order to accept or reject a given H_0, one has to define a test statistic *cut* t_{cut}, establishing an acceptance and a rejection region. The decision is made by comparing the observed value $t^{\text{obs.}}$ with the t_{cut} defined beforehand. When $t^{\text{obs.}}$ is within the acceptance region, the null hypothesis is said to be *accepted*; otherwise, it is *rejected*. Naturally, t_{cut} can be arbitrarily chosen, affecting the sizes of the acceptance and rejection regions. In general, it is defined as to provide a significance level α according to some common threshold, oftentimes 5% or 10%:

$$\alpha = \int_{t_{\text{cut}}}^{\infty} g(t|H_0)dt \, . \tag{5.1}$$

The larger α is, the harder it is to reject H_0, and the less frequent *false negatives* become, where H_0 is rejected despite being true. Conversely, the probability for H_0 to be accepted given that H_1 is true is called a *false positive*, and is given by:

$$\beta = \int_{-\infty}^{t_{\text{cut}}} g(t|H_1)dt \, , \tag{5.2}$$

where $1 - \beta$ is referred as the *power* of the test. The smaller t_{cut} is, the smaller β will become, creating a "powerful" test in what concerns the discrimination against alternative hypotheses.

The level of agreement between the data and H_0 can be expressed by the *p-value*. The p-value represents the probability, under the assumption of the null hypothesis, of obtaining a result equal to or more extreme (less compatible) than the one observed, given the measured value of a particular test statistic $t^{\text{obs.}}$. The formal definition is given by:

$$p = \int_{t^{\text{obs.}}}^{\infty} f(t|H_0)\, dt \, , \tag{5.3}$$

which can be converted into a significance Z, using a one-sided Gaussian integral, often used to report results in HEP publications:

$$p = \int_{\text{Z}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \, . \tag{5.4}$$

Two levels of significance are traditionally defined. An *evidence* corresponds to Z = 3, or equivalently $p = 1.3 \times 10^{-3}$, while an *observation* sets the bar higher at Z = 5, or $p = 2.8 \times 10^{-7}$.

5.3.2 The Neyman-Pearson Lemma

When associating an event to a particular null hypothesis, the *efficiency* $1-\alpha$ of such selection is controlled by t_{cut}. A larger t_{cut} implies a larger probed phase-space, and in the $t_{\text{cut}} \to \infty$ limit, all events will be associated to H_0. At the same time, less events will be assigned to an alternative H_1. It therefore exists a compromise between the efficiency and the amount of contamination in a sample, or *purity*, defined as the fraction of the number of truth-matched events. By construction, a scenario with the highest purity corresponds to a scenario with the highest power.

Defining the optimal balance as the one providing the maximum purity for a given selection efficiency, the *Neyman-Pearson lemma* states that such balance, in the space of the test statistic t, is given by:

$$\frac{g(t|H_0)}{g(t|H_1)} > c\,, \tag{5.5}$$

where c can be defined based on the efficiency one wishes to obtain. While Eq. (5.5) is rather obvious for a one-dimensional test statistic, the lemma works for multi-dimensional test statistics too. The functions g are not always simple to compute. Methods have nevertheless been devised to provide good estimates, such as the Fisher discriminant for linear test statistics, and NNs for nonlinear scenarios. The left hand-side of Eq. (5.5) gives a ratio that is guaranteed to provide the highest possible power, with an acceptance region corresponding to values above c. The quantity is known as the *likelihood ratio*, and is described in the next Section.

5.3.3 The Maximum Likelihood Method

Given a finite data sample $x = (x_1, x_2, \ldots, x_n)$, being n the number of measurements, one often wishes to estimate the parameters $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$ defining the underlying $f(x;\varphi)$ PDF hypothesis. Assuming the hypothesis to be true, the probability to find the first measurement in the interval $[x_1, x_1 + dx_1]$ is simply given by the product between $f(x;\varphi)$ and the interval dx_1. If all measurements x are independent, the probability to find each measurement i in a $[x_i, x_i + dx_i]$ interval is given by the product of all individual probabilities. If $f(x;\varphi)$ truly describes the underlying data, a high probability is expected for the data to be measured where the PDF predicts it to be. In other words, we can define the *likelihood function* L:

$$L(\varphi) = \prod_{i=1}^{n} f(x_i;\varphi)\,, \tag{5.6}$$

where the dependence on the intervals was removed since dx_i do not depend on the parameters. Note that L does not depend on the data x, which is treated as fixed, or equivalently, the data collection is assumed to be over. The likelihood can be seen as the probability of observing the data, given a set of (unknown) model parameters, where we set $L(\varphi) \equiv L(x|\varphi)$ for brevity. Therefore, the parameters that maximize such probability are the ones which maximize $L(\varphi)$. We can thus search for the *maximum likelihood estimators* of the parameters, by finding the highest maximum of the likelihood.

It is often the case that n is itself a Poisson random variable with mean ν. This is true in HEP, given the stochastic nature of any given physics process. For these cases, the *extended likelihood function* is employed:

$$L(\nu, \varphi) = \frac{e^{-\nu}}{n!} \prod_{i=1}^{n} \nu \, f(x_i; \varphi) \, . \tag{5.7}$$

In practice, the *log-likelihood* $\log L$ is usually considered, since it is mathematically much easier to treat:

$$\log L(\nu, \varphi) = -\nu + \sum_{i=1}^{n} \log \left[\sum_{j=1}^{m} \nu \, \varphi_j \, f_j(x_i) \right] , \tag{5.8}$$

where it was assumed that ν does not depend on φ, and all terms not depending on the parameters were dropped. As the logarithm is a monotonically increasing function, the maxima of L will also maximize $\log L$.

The analysis reported in this Thesis uses binned distributions, and thus one can see the PDFs as vectors, representing the event yield in all the bins of the distributions satisfying the analysis' selections. A binned approach is here preferred for simplicity and computation efficiency. For n_{tot} observations, the expectation values $\nu = (\nu_1, \nu_2, \ldots, \nu_N)$ of the number of entries $n = (n_1, n_2, \ldots, n_N)$ in N bins is given by:

$$\nu_i(\varphi) = n_{\text{tot}} \int_{x_i^{\min}}^{x_i^{\max}} f(x; \varphi) dx \, , \tag{5.9}$$

where $x_i^{\min}$ and $x_i^{\max}$ represent the edges of bin i.

Following the parameter classification discussed in Section 5.2, in HEP the parameters φ are usually split into the POI μ and nuisances θ. The estimation of nuisance parameters is denoted $\tilde{\theta}$. We express our degree of belief that the real values θ are correctly represented by the estimates $\tilde{\theta}$ with a Bayesian probability density function $\rho(\theta_k|\tilde{\theta}_k)$, with k referring to a given nuisance parameter out of all n_k. Using Bayes' theorem, one can convert ρ into a frequentist probability $p(\tilde{\theta}_k|\theta_k)$, by using a uniform prior PDF representing our full ignorance before the experiment takes place. For the case of the resonant analysis here reported, the functional form of $\rho(\theta_k|\tilde{\theta}_k)$ is given by the log-normal function:

$$\rho(\theta_k|\tilde{\theta}_k; \kappa) = \frac{1}{\sqrt{2\pi} \ln \kappa} \exp \left[- \left(\frac{\ln (\theta_k)/\tilde{\theta}_k}{2 \ln \kappa} \right)^2 \right] \frac{1}{\theta_k} \, , \tag{5.10}$$

which is the recommended choice for multiplicative corrections, like all the ones used in $X \to HH \to bb\tau\tau$. For small uncertainties, Eq. (5.10) with $\kappa = 1 + \varepsilon$ is identical to a Gaussian distribution of width ε, but can describe positively defined observables by going to zero at $\theta_k = 0$.

In HEP, it is customary to write the expected signal event yield as s and the expected background event yield as b. This enables to explicitly encode a dependence on the signal strength modifier μ as follows:

$$\nu_i(\mu, \theta) = \mu \, s_i(\theta) + b_i(\theta) \, , \tag{5.11}$$

where ν_i refers to the expected bin counts in Eq. (5.9). Often, the null hypothesis corresponds to $\mu = 0$, while the alternative hypothesis stands for the presence of a BSM signal with $\mu \neq 0$. Putting together all ingredients above, one can write the full expression of the likelihood in Eq. (5.7), as considered for the binned $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis, including systematic uncertainties:

$$L(\mu, \theta) = \prod_{i=1}^{n} \frac{e^{-[\mu\, s_i(\theta) + b_i(\theta)]}}{n_i!} [\mu\, s_i(\theta) + b_i(\theta)]^{n_i} \times \prod_{k=1}^{n_k} p(\tilde{\theta}_k | \theta_k) \,, \tag{5.12}$$

where the last term can be determined using Eq. (5.10). Note that nuisance parameters make the likelihood broader, reflecting the information lost due to systematic uncertainties. When combining multiple channels, categories or eras, the left side of Eq. (5.12) can be extended by performing further multiplications with the corresponding poissonian distribution functions.

5.3.4 Finding a Signal in HEP Experiments

We can now use the Neyman-Pearson lemma to rewrite the optimal test statistic using the likelihood function. To test a given value of the signal strength modifier μ, we define the *profile likelihood ratio (PRL)* [402]:

$$\lambda(\mu) \equiv \frac{L(\mu, \hat{\theta}_\mu)}{L(\hat{\mu}, \hat{\theta})} \,, \tag{5.13}$$

where $\hat{\mu}$ and $\hat{\theta}$ in the denominator are the maximum likelihood estimators of μ and θ, respectively, and $\hat{\theta}_\mu$ denotes the value of θ that maximizes the likelihood for a given μ. In other words, the denominator is the unconditional maximized likelihood function, obtained when the minimization is performed simultaneously on μ and θ, while the numerator is conditioned on the value of μ being probed. $\lambda(\mu)$ is by construction bounded between 0 and 1, and higher values correspond to better compatibilities between the data and the probed μ. The profile likelihood ratio satisfies the conditions required by *Wilks' theorem*, which states that, given a large dataset and assuming some regularity conditions on the likelihood, $-2 \ln \lambda(\mu)$ is asymptotically distributed according to a χ^2 distribution with k degrees of freedom, where k is given by the number of POIs [405]. Importantly, the square root of the χ^2 distribution at its minimum provides an approximate estimate of the significance, avoiding the generation of large quantities of pseudo-data to simulate $\lambda(\mu)$. To give an idea, a $5\,\sigma$ p-value represents a $\sim 10^{-7}$ probability, thus requiring $\sim 10^9$ samples for a precise determination.

In order to quantify an excess, the following test statistic is defined:

$$q_0 \equiv \begin{cases} -2 \ln \lambda(0) & \text{if } \hat{\mu} \geq 0 \\ 0 & \text{if } \hat{\mu} < 0 \end{cases} . \tag{5.14}$$

Instead, in the case of the $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis, we are looking for a signal that has never been observed. In such cases, *upper limits* are instead quoted, since an excess is not expected. When setting upper limits on the strength parameter μ, the following test statistic q_μ is used:

$$q_\mu \equiv \begin{cases} -2 \ln \lambda(\mu) & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} > \mu \end{cases} . \tag{5.15}$$

Due to the negative sign, larger values of q_μ indicate that the probability for observing the data given μ is smaller, i.e. the data and parameters become increasingly incompatible. The test statistic is set to zero for $\hat{\mu} > \mu$ because, in the specific case of an upper limit, that situation would not be interpreted as a rejection of the null hypothesis, as we would be probing a value of μ that lies closer to the background-only hypothesis than the maximum likelihood estimator $\hat{\mu}$.

The test statistic in Eq. (5.15) can be exploited to define exclusion limits via the modified frequentist confidence level criterion, or "$\mathrm{CL_s}$ method" [401, 406]. Given $q_\mu^{\mathrm{obs.}}$, we calculate the probabilities for q_μ to be equal or larger under the null or alternative hypotheses:

$$p_{s+b} = P(q_\mu \geq q_\mu^{\mathrm{obs.}} | \, s+b) = \int_{q_\mu^{\mathrm{obs.}}}^{\infty} f(q_\mu | \, s+b) \, dq_\mu \, ,$$
$$p_b = P(q_\mu \geq q_\mu^{\mathrm{obs.}} | \, b) = \int_{q_\mu^{\mathrm{obs.}}}^{\infty} f(q_\mu | \, b) \, dq_\mu \, . \tag{5.16}$$

A given signal strength μ is said to be excluded at a confidence level $\mathrm{CL} = 1 - \alpha$ if:

$$\mathrm{CL_s}(\mu) \equiv \frac{p_{s+b}}{p_b} < \alpha \, , \tag{5.17}$$

where α is the significance level. The inclusion of the denominator protects against cases where $s \ll b$, for which signal models can be excluded when there is no sensitivity due to under-fluctuations of the background. A value of 5% is commonly chosen for the significance level, leading to the 95% CL results quoted in this work. Would the experiment be repeated a large number of times, we would expect to observe μ within the CL in 95% of cases. The limits are obtained on the POI μ, but are rescaled to the cross section of the signal process.

When looking for resonances as a function of multiple hypotheses, as done for instance for the $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis in terms of the resonance masses m_X, the p-value underestimates the chances of observing fluctuations when jointly considering all signals being probed. In other words, the *local* significance, as computed for a fixed value of a parameter, systematically overestimates the more correct, *global* significance, in a phenomenon dubbed the *look-elsewhere effect*. This effect takes place because q depends on several hypotheses $m = (m_1, m_2, m_3, ...)$, and not just on a single m_i, and thus one has to take into account background fluctuations at any hypothesis value in the relevant domain. The local p-value is thus converted into a global one through a new test statistic, computed by taking the maximum value of test statistics over the hypotheses' domain [403]:

$$q(\mu) = \max_i q(\mu; m_i) \, . \tag{5.18}$$

We remark that the effect becomes stronger for worse detector resolutions, since it becomes more likely to observe multiple mass shifts accumulating at the same point, thus mimicking an excess over the background [371].

5.3.5 Expected and Observed Limits

The statistical treatment is centered around the usage of a (hopefully discriminant) test statistic, ultimately defining how much sensitive analyses become. A potentially minor adjustment in the analysis' selection, or a slight change in the training of the final discriminant, for instance, might very well modify the test statistic, and thus affect the final upper limits. What would happen if one would compute limits over and over again, each time iterating on the analysis strategy to obtain tighter limits? The final result would become biased, since decisions would be made given the final results. This might in turn lead to the appearance of spurious excesses. Instead, HEP analyses start *blinded*: the data in the SR is not looked at until the analysis strategy has been considered stable. Data in CRs can be used, together with MC samples.

During the blinded stage, the analysis' sensitivity is assessed via *expected results*. Such results are determined thanks to a *Asimov dataset* [402], a representative dataset where all observed quantities, POI and nuisances, are set to their expected values, and statistical fluctuations are not considered. The data thus derived can be used to define an "Asimov likelihood" L_A, and the PRL becomes:

$$\lambda_A(\mu) \equiv \frac{L_A(\mu, \hat{\theta}_\mu)}{L_A(\hat{\mu}, \hat{\theta})} = \frac{L_A(\mu, \hat{\theta}_\mu)}{L_A(\mu', \theta')} , \tag{5.19}$$

where μ' and θ' represent the expected values of the POI and nuisances. Once all possible issues in the analysis have been addressed and understood, the *unblinding* step can proceed, in which the maximum likelihood fit provides the *observed* result.

Given the timeline of our analysis, we present expected $X \to HH \to bb\tau\tau$ upper limits in the next Section. The current publication plan has been defined so that the unblinded analysis can be presented at the "59th Rencontres de Moriond" conference, which takes places in March 2025.

5.4 Results

We are finally able to condense all the work so far presented into quantifiable results. We show the results in Section 5.4.1, and establish comparisons with related CMS and ATLAS results in Section 5.4.2.

5.4.1 Exclusion Limits

We extract expected 95% CL upper limits on the cross section of the $X \to HH$ process, assuming spin-0 and spin-2 hypotheses. The result is obtained by rescaling the $X \to HH \to bb\tau\tau$ upper limits to the inclusive cross section. The limits are shown in Fig. 5.3, split per data-taking period and combined, and in Fig. 5.4, for the full Run 2 data with $1\,\sigma$ and $2\,\sigma$ bands. The combination is also available in tabular format in Appendix A.2. The results are similar between the two spin hypothesis. The overall trend of the curves shows a maximum sensitivity close to $m_X = 1\,\text{TeV}$. We recall that one of the goals of the analysis is to assess the tension observed by ATLAS at precisely 1 TeV.

At larger masses, the topology of the event becomes extremely boosted, both for b quarks and tau leptons. HPS is not optimized to reconstruct boosted tau leptons. This leads to the loss of many events. The effect becomes progressively worse, as more and

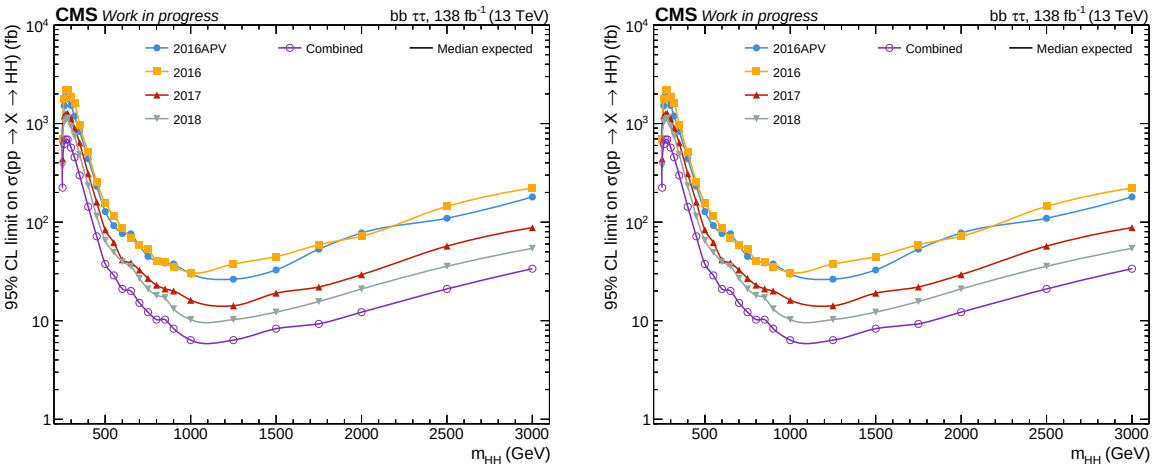

Figure 5.3: Expected $X \to HH$ cross section 95% CL median upper limits, for spin-0 (left) and spin-2 (right). The full Run 2 combination is shown in purple, while the other curves represent separate contributions for all data-taking periods, scaled to the total Run 2 luminosity. The markers refer to the probed mass points, while the lines are the result of an interpolation.

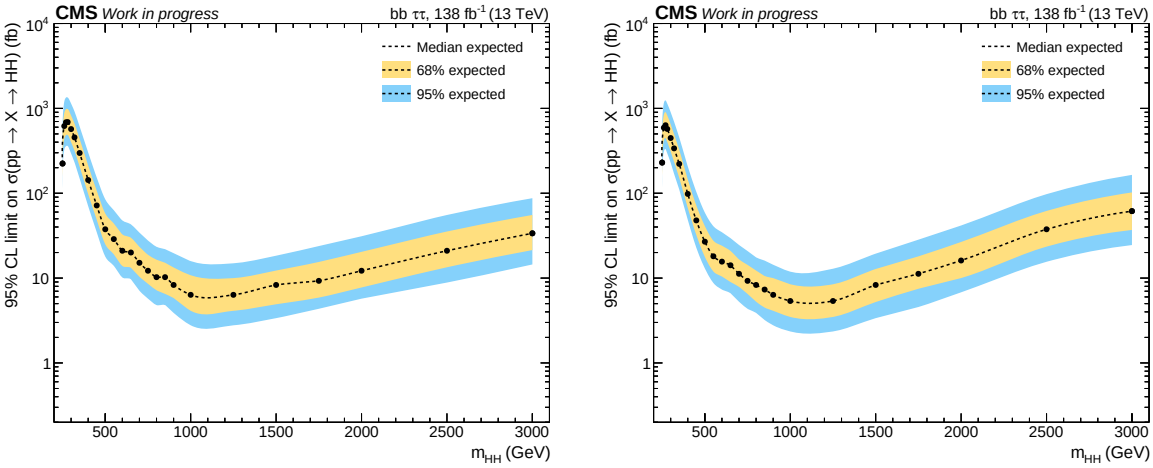

Figure 5.4: Expected $X \to HH$ cross section 95% CL upper limits, for spin-0 (left) and spin-2 (right), using full Run 2 data. The dashed line refers to the median result, while $1\,\sigma$ and $2\,\sigma$ bands are show in yellow and blue, respectively. The black points refer to the probed mass points.

more tau leptons merge into a single boosted object. This issue can be addressed by using dedicated τ reconstruction algorithms at high m_X. Such analysis improvement lies outside the scope of this Thesis, since a different CMS analysis is currently looking at such topologies, following the work performed in Ref. [407].

At lower m_X, we observe two effects. Firstly, the sensitivity decrease is due to the increasing presence of the sum of all backgrounds. Secondly, the peak at ~280 GeV is a consequence of the $t\bar{t}$ background shape, which features a peak at that location. The same overall trends were present in the previous iteration of the resonant analysis [110].

We draw attention for the relative importance of the different data-taking periods. Assuming the standard naive $s/\sqrt{b}$ scaling, where the number of signal events s stands above the background Poisson noise $\sqrt{b}$, the sensitivity increases with $\sqrt{L}$, being L the integrated luminosity. For upper limits on the cross section of interest, in the absence of a signal, the scaling will be inversely proportional to the statistical uncertainty of the background. Upper limits are thus expected to scale with $1/\sqrt{L}$. The Run 2 luminosities were measured to be $16.8\,\text{fb}^{-1}$ for 2016, $19.5\,\text{fb}^{-1}$ for 2016 APV, $41.5\,\text{fb}^{-1}$ for 2017 and $59.8\,\text{fb}^{-1}$ for 2018. The analysis strategy is identical for different periods, but some exception exist for specific thresholds and triggers. The differences between results from different periods roughly follows a $1/\sqrt{L}$ scaling, but some fluctuations are present.

The robustness of the statistical approach must be verified. Analysis usually monitor the impact that nuisance parameters have on the results with the calculation of *pulls* and *impacts*. The pull is defined as the difference between the nominal value of a parameter θ before the fit, and its estimated value after the fit, divided by the θ's uncertainty. Its distribution is expected to follow a Gaussian centered around zero, with a standard deviation of one, according to the Central Limit Theorem. If the width of the pull is smaller than one, then the analysis is able to constrain that particular nuisance, which indicates that enough sensitivity is available, or equivalently, that the nominal values of the nuisances are too conservative. The impact is instead defined as the variation induced on the POI when a nuisance is fixed to $\hat{\theta} \pm 1\,\sigma$, keeping all other nuisances fixed to their maximum profiled likelihood estimate $\hat{\theta}$.

We show a preliminary version of the pulls and impacts for the 25 leading nuisances in Fig. 5.5, for $m_X = 300\,\text{GeV}$. The pulls are all centered at zero since, for expected limits, the difference is computed with respect to the Asimov dataset.

5.4.2 Comparison with Other Results

We perform a comparison of the Run 2 spin-0 results obtained in the previous Section with the most stringent ATLAS and CMS results to date. The expected upper limits are shown in Fig. 5.6 (top). The limits from ATLAS were published in Ref. [345], where a small tension is observed at $m_X \sim 1\,\text{TeV}$, with a local (global) significance of $3.1\,\sigma$ ($2.0\,\sigma$). The CMS result comes instead from a $X \to YH$ analysis [101], where $m_Y = m_H\,\text{GeV}$ was not considered. We thus compare our result to their $m_Y = 120\,\text{GeV}$ limit, which was shown to be comparable to their $m_Y = 130\,\text{GeV}$ result. This same analysis was made part of the combination results presented in Section 1.4.1. The results are available in numerical format in Appendix A.2.

Our result improves the previous best CMS result across the entire m_X range by a very significant margin. This is expected, given that we are targeting a specific topology, while the $X \to YH$ work covers a much larger mass phase-space. Concerning the ATLAS result, we achieve a better expected performance across most of the m_X range. In particular, we

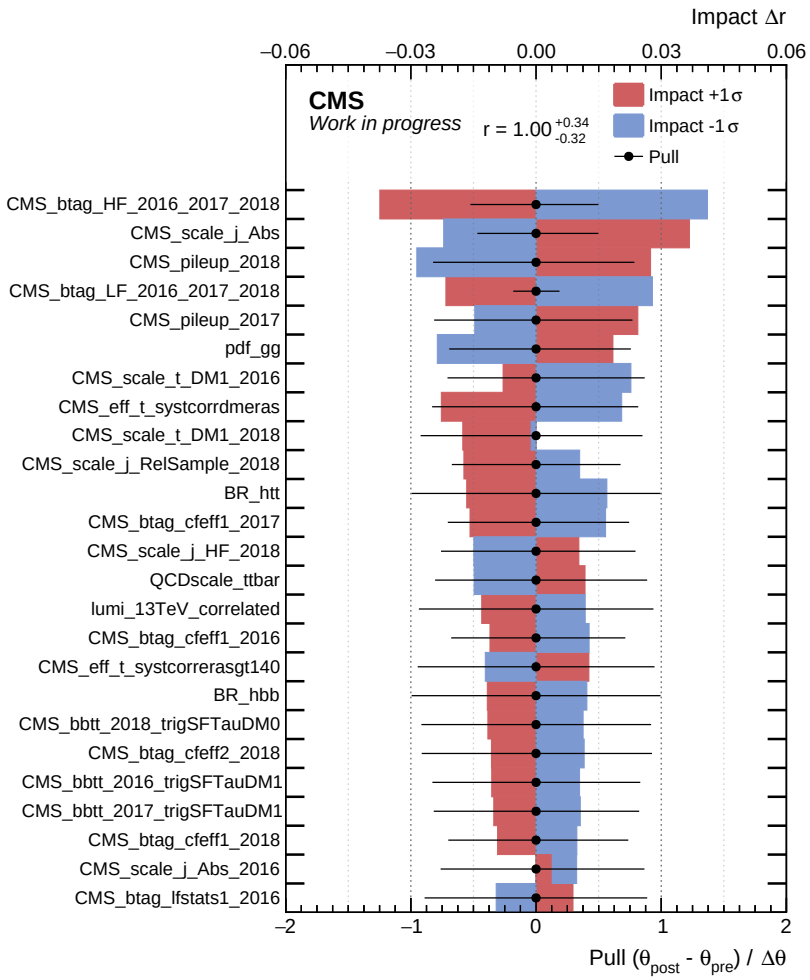

Figure 5.5: Leading 25 systematic uncertainties of the $X \to HH \to bb\tau\tau$ analysis. The pulls are shown by the black points and the black horizontal error bars, which indicate the median and the $1\,\sigma$ bands, respectively. The pulls are centered at zero because an Asimov dataset is being used. The "pre" and "post" labels refer to the maximum likelihood fits. The red and blue bars indicate the impacts on the POI, and its direction. The plot was obtained for $m_X = 300\,\text{GeV}$. The signal was injected with a cross section of 1 pb.

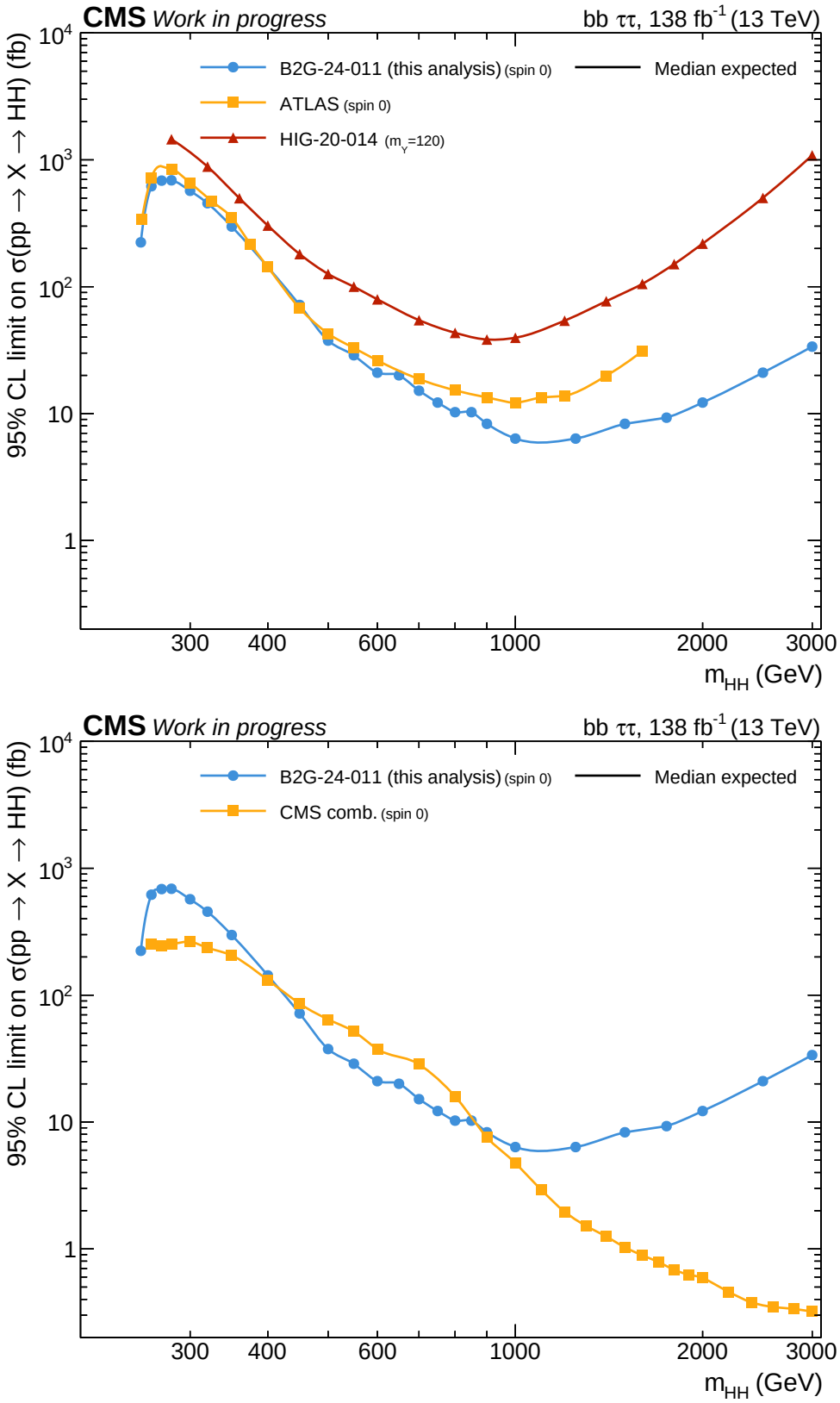

Figure 5.6: Comparison of the expected $X \to HH$ cross section 95% CL upper limits for spin-0 in this work with other results. (Top) Comparison with equivalent expected limits from ATLAS, and with the most stringent CMS $X \to HH \to bb\tau\tau$ upper limits to date. (Bottom) Comparison with the full HH CMS combination. The results in this Thesis should drive the sensitivity of HH upper limits in CMS in the $450 \leq m_X \leq 800\,\text{GeV}$ range.

are almost a factor of 2 more sensitive at 1 TeV, and can thus conclude that the analysis should be able to confirm or reject the tension observed by ATLAS.

It is also extremely relevant to compare our result with the current best CMS combination of resonant HH analyses. We show such comparison in Fig. 5.6 (bottom). Given the results, we expect our analysis to drive the CMS HH sensitivity in the $450 \leq m_X \lesssim 800\,\text{GeV}$ range.

We start by quantifying the sensitivity improvement within CMS over the years. Comparing the full Run 2 luminosity with the luminosity used by Ref. [110], which used 2016 data only, our analysis has access to ∼4 times more data, which would naively correspond to an improvement of a factor of ∼2. Looking at the limits obtained for the spin-0 hypothesis, we can instead report an improvement of ∼8.5 at $m_X = 300\,\text{GeV}$, ∼15 at $m_X = 500\,\text{GeV}$, and ∼10 at $m_X = 900\,\text{GeV}$ (only $m_X \leq 900\,\text{GeV}$ values were considered). This once more demonstrates the excellent performances obtained by CMS, which continue to improve over past results by exploiting novel techniques spanning all the steps of a physics analysis.

5.5 Prospects for HH Searches

The results presented in the previous Section used the full CMS Run 2 data. We here discuss the future of Higgs boson pair production, considering future data-taking periods at the LHC and HL-LHC.

5.5.1 Run 3

At the time of writing, Run 3 has already collected more data than Run 2 ever did, and should reach about $300\,\text{fb}^{-1}$ by the end of 2025. Multiple analyses are being prepared. Concerning the $\text{bb}\tau\tau$ channel, both nonresonant and resonant searches are being explored, and should be published within the next few years. The sensitivity of HH analyses is expected to increase significantly. Firstly, the center-of-mass energy will slightly increase, leading to a HH cross section increase of approx. 10 %. Secondly, and most importantly, the luminosity will increase by a factor of about 3. Using the naive luminosity scaling, we would expect an improvement of the order of 70%. We should however realistically expect larger improvements [408]. On one hand, the increase in luminosity, besides adding more events in known regions, opens up a way to observables that were not previously available. Examples are final state regions with large invariant mass or p_T. On the other hand, new reconstruction techniques have the potential to enhance analyses' sensitivities, as recent years show. Furthermore, the data increase often also enables to better control systematic uncertainties. The significance of such a fact will depend on the relative importance of all uncertainties, knowing that HH studies will still remain statistically dominated. To recap, we should expect reasonably larger improvements than what the naive scaling may provide, without any modification on current analysis strategies.

On top, one has to consider improvements stemming from improved methods. New machine learning techniques applied to all sorts of discriminants have the potential to make current results quickly obsolete. The most prominent example is the mass-decorrelated PNet [266]. Beyond its current AK4 an AK8 jet-tagging capabilities, notably reconstructing merged $\text{b}\bar{\text{b}}$ and $\text{c}\bar{\text{c}}$ topologies, PNet has been extended with jet flavour classification, jet mass regression, and even $\tau\tau$ tagging. As an example, PNet has already been used for energy regression in the context of jet energy scale calibrations, improving energy

response resolution by around 15 % [409]. Additionally, based on performance measurements, it is likely that some top or boson tagging with PNet will be slowly replaced by updated versions based on the state-of-the-art transformer technology [120]. Needless to say, the impact on HH analyses, including bbττ topologies, is expected to be significant.

Improvements due to novel trigger strategies are also to be expected, delivered both by improved data scouting and parking streams [212], and also by the implementation of PNet b-tagging directly at HLT. The improved tagger enables to lower HH → bbbb and HH → bbττ jet p_T thresholds and H_T selections. Completely new triggers have also been added; for the bbττ analysis we can mention new 4j+2b and 4j+1b+1τ_h paths, the latter only from 2024 onward. Using a different approach, the introduction of VBF and HH parking streams has already been shown to massively impact HH studies (see Fig. 2.19). Finally, the inclusion of PNet τ-tagging at trigger level is being envisaged, and might still be done during Run 3. When considering all the improvements above, trigger efficiencies are expected to increase by more than 60% for bbττ, and more than 80% for bbbb.

Further improvements on bbττ topologies are also certainly conceivable. There is currently an on-going effort to explore extremely boosted bbττ topologies, considering m_X values of at least 1 TeV. Such topologies require significantly different analysis strategies. The analysis is dominated by the presence of MET, given the large Higgs boson boosts. Importantly, the τ pair is also extremely boosted, and the conventional HPS algorithm fails; different tools must thus be developed, PNet being a possibility. A future combination between the results produced in this Thesis and potential high mass results should improve the limits at large m_X values.

We also draw a quick note on ZZ and ZH processes, since they represent standard candles to validate HH analyses, given their cross sections 31 and 8 times larger than the SM cross section, respectively, and their similar decay topologies. A recently published resolved ZZ/ZH → bbbb analysis [267] uses a series of advanced background estimation techniques that might be useful in future HH studies. A particularly promising example is the usage of synthetic datasets to successfully validate QCD estimates, using the "hemisphere mixing technique" [410]. Pseudo-data might be exploited in the future also to measure QCD-related systematics directly in the analysis SR. A similar strategy, given appropriate modifications, would be applicable to bbττ analyses. The analysis here reported was also developed within the same time period as a similarly defined ZZ → bbττ analysis, enabling helpful discussions. Future exchanges between such similar topologies might enrich the field, and ultimately provide better sensitivities in multiple channels at once.

5.5.2 HL-LHC

The future of HH studies can also be envisaged beyond Run 3, namely during the next phase planned for the LHC [132]. This gigantic update should commence data-taking in 2029, delivering more than 3000 fb^{-1} pp collisions, at a center-of-mass energy of 14 TeV. The measurement of the Higgs boson self-coupling will arguably become the nexus of all HEP research until the 2040s, should no interesting excess or deviation be measured by then. Tighter constraints on κ_λ and EFT couplings will certainly be achieved, also with the help of yet unexplored HH production modes and decay channels. However, HH studies will continue to be statistically-dominated, despite a possible increase in the importance of theoretical uncertainties. On top of the recent $\kappa_{2V} = 0$ exclusion, and assuming $\kappa_\lambda = 1$, we believe nonresonant Higgs boson pair production will be measured via

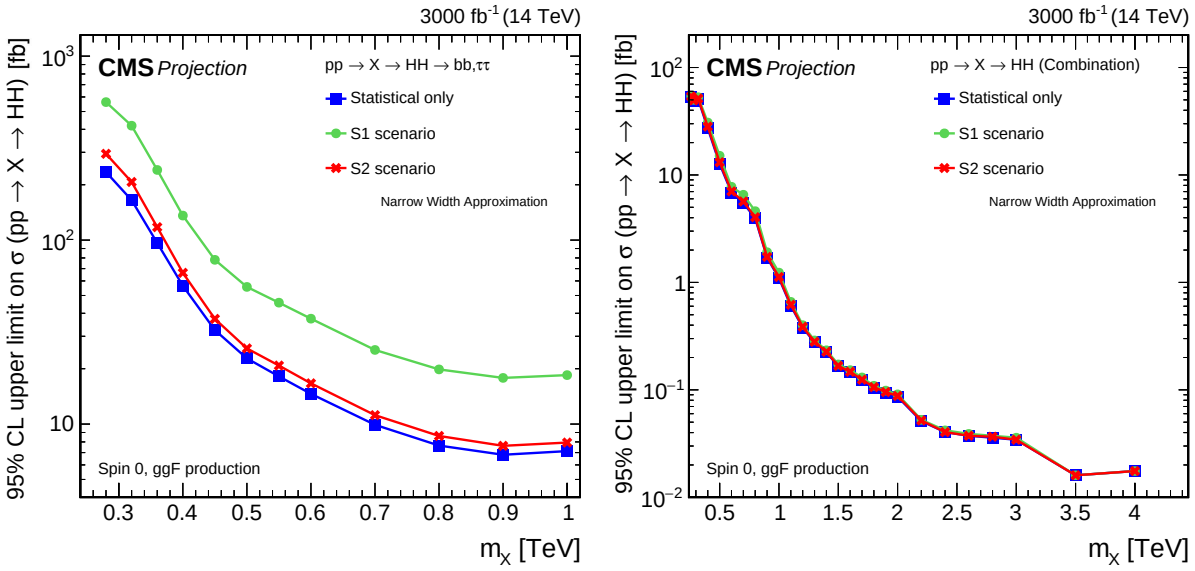

Figure 5.7: Expected upper limits at 95% CL on the product of the cross section for the production of a spin-0 resonance X and the branching fraction $\mathcal{B}(\mathrm{X} \to \mathrm{HH})$, as a function of m_X, for an integrated luminosity of 3000 fb^{-1}. Shown are the effects of different systematic uncertainty scenarios. (Left) bb$\tau\tau$ decay channel [101]. (Right) Combination of the three analyses, including bb$\tau\tau$. Taken from [7].

a multi-channel combination by the end of the HL-LHC [38], perhaps after a combination with ATLAS.

Strong prospects are also envisaged for future resonant HH results. We here describe an overview for their projections at the HL-LHC [7]. The studies, based on the CMS full Run 2 resonant HH and YH searches, focus on the most sensitive HH decay channels, namely bb$\gamma\gamma$ [107], bb$\tau\tau$ [101], and boosted bbbb [108]. Individual channels are statistically combined to take advantage of their complementary sensitivity to different BSM phase-space regions. The expected cross section upper limits at 95 % CL are provided as a function of the masses of the BSM scalars, m_X and m_Y. Where appropriate, the signal cross sections are scaled to the HL-LHC 14 TeV center-of-mass energy. The efficiency in the reconstruction and identification of photons, leptons, jets and b-jets, as well as the resolution of their energy and momentum measurements, are assumed to remain unchanged with respect to Phase 1. The experimental sensitivity expected at the HL-LHC is derived using three systematic uncertainty scenarios, detailed in Ref. [7]. Ordered from the most to the least conservative, they are called "S1", "S2" and "statistical only". Given constant updates in analyses, it is reasonable to believe the three scenarios are very conservative. For the bb$\tau\tau$ channel, 95% CL cross section upper limits are derived for m_X within 300 and 1000 GeV. The expected upper limits at 95% CL on the $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ cross section, projected to 3000 fb^{-1}, are shown for bb$\tau\tau$ channel and for the combination of the three most sensitive HH channels in Fig. 5.7, for spin-0.

Comparing the combination result with this Thesis' spin-0 result, one can unsurprisingly conclude that significant improvements are expected for the HL-LHC period. However, for some m_X values in the bb$\tau\tau$ channel, the comparison shows similar results, despite our analysis having approx. 22 times less luminosity that what will be available at the HL-LHC. Including our limits in the next CMS projection studies should thus improve limits by a good margin, in principle similar to the increased sensitivity in the $450 \leq m_\mathrm{X} \leq 800$ GeV range for the current combination. In parallel, $\mathrm{X} \to \mathrm{YH}$ analyses,

where the Y scalar decays to $b\bar{b}$, are also projected into the future, considering the same three scenarios. Interestingly, the bb$\tau\tau$ final state is shown to drive the sensitivity when both X and Y scalar have a low mass, and also when they both have a high mass, for the reasons discussed in Section 1.4.1.

In conclusion, it should by now be clear that HH studies, either in their resonant or nonresonant forms, have vast potential for exploration up to the end of the HL-LHC. Despite not being yet clear which ground-breaking accelerators and experiments will follow the HL-LHC, it is also hard to imagine a future in HEP where the pair production of Higgs bosons does not take center stage.

6 Going Beyond the Narrow Width Approximation in HH Resonant Searches

The majority of studies involving the SM Higgs boson or BSM particles considers the Narrow Width Approximation, where the production via mode i and decay f of a particle P can be factorized as:

$$(\sigma \cdot \mathcal{B})(i \to \mathrm{P} \to f) = \frac{\sigma_i \cdot \Gamma_f}{\Gamma_\mathrm{P}} \ , \tag{6.1}$$

where σ_i is the cross section of the production mode i, $\mathcal{B}$ is the BR defined as $\Gamma_f/\Gamma_\mathrm{P}$, Γ_f is the partial decay width into the final state f, and Γ_P is the total width of particle P. The approximation most notably holds when the width Γ of a particle is much smaller than its mass M, $\Gamma \ll \mathrm{M}$, hence its name. Additionally, four other conditions are required to justify the NWA applicability [411]: *i)* the mass of the daughter particles is much smaller than the mass of the resonance, *ii)* the center-of-mass-energy is much larger than the mass of the resonance, *iii)* at the diagram level, the propagator of the resonance must be separable from the corresponding matrix element, and *iv)* no significant *interferences* with nonresonant processes exist. The validity of all four points strongly depend on the process being probed. The last point is particularly relevant in the work here described, and is later shown not to be applicable in the full probed phase-space.

Experimentally speaking, the NWA typically equates to using MC signal samples where the width Γ of the resonance is neglected, by setting it to zero or to an arbitrarily small value, much smaller than the particle's mass and smaller than the detector's resolution. For a single Higgs boson, taking into account the predicted SM Γ_H of $\sim$4.1 MeV [49, App.A] at $m_\mathrm{H} = 125.08 \pm 0.12\,\mathrm{GeV}$[1] [42], the approximation holds. Instead, when considering resonant processes with a pair of Higgs bosons, one cannot *a priori* argue that the detector resolution will always be lower than the width of any BSM resonance $\mathrm{X} \to \mathrm{HH}$. Especially for large resonance masses, Γ_X values are expected to be large, potentially invalidating the NWA. Despite the above, the NWA has been repeatedly used in past CMS HH resonant analyses, including the analysis presented in this Thesis. A sizable signal width can create interferences between the resonant and nonresonant components of the signal, which are absent under the NWA. The interference could be either destructive or constructive, respectively reducing or enhancing the HH production rate [414]. The change can be significant, both overall and as a function of m_HH, altering the di-Higgs *lineshapes*. Some examples are shown in Figs. 6.7 and 6.8.

I have been the main responsible for a CMS phenomenological study that assesses the impact of the NWA in different phase-space regions of the inclusive ggF $\mathrm{X} \to \mathrm{HH}$ process, and am the sole author of all CMS results presented in this Chapter. To the

[1] The value of the PDG average [13] at the time of writing is slightly different, and based on older CMS and ATLAS measurements: [44, 412, 413]

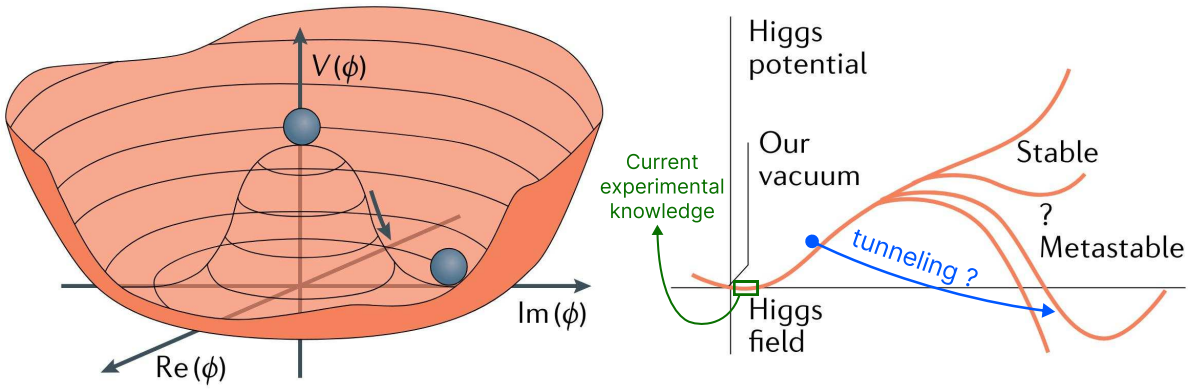

Figure 6.1: (Left) Illustration of the Higgs potential $V(\phi)$ for the scalar field ϕ. Choosing any of the minima in the complex circle breaks the symmetry. (Right) Quantum corrections can lead to changes in the shape of the Higgs potential, including the development of a second minimum implying metastability. These effects are not verifiable with our current experimental sensitivity. The minimum is taken at $v/\sqrt{2}$. Adapted from [417].

best of my knowledge, this is the first time such a complete study on the validity of the NWA is performed within the CMS Collaboration. The study was originally motivated by my involvement in the Run 2 $\mathrm{HH} \to \mathrm{bb}\tau\tau$ analysis, the latter representing an ideal setup to test new HH models. In the end, the work done settled on a more general approach, not specific to any particular decay channel. The ongoing Run 3 should see the first appearances of non-NWA results, since large deviations with respect to the NWA approach have recently drawn a significant interest from the community. For simplicity, we here consider the resonant model with the smallest number of additional free parameters: the singlet model [415]. Similar studies could be repeated with other, more complex models, such as models containing two singlets [416]. As a future step, we would like to extend the same study to the $\mathrm{bb}\tau\tau$ channel, namely measuring the impact of the NWA on the final upper limits of the analysis. The content of this Chapter was published as part of a review paper covering the Higgs boson production searches through decays of heavy resonances [7].

6.1 The Stability of the Electroweak Vacuum

As discussed in Section 1.3.1, despite the impressive success of the SM, we know that it must be incomplete. One of the big questions still left unanswered relates to the nature of the EWSB of the Higgs field. When the Universe was very young, at extremely high temperatures, it was energetically favorable for the Higgs field to be zero everywhere. Only around a hundred picoseconds after the Big Bang, with a stark temperature drop, was the symmetry broken. It is currently unknown how EWSB takes place, specifically which kind of phase transition drives it. According to the SM, the crossover should be smooth [418], but such scenario is considered unnatural by some, due to the extreme distance between EW and Planck scales. A strong first-order electroweak phase transition (SFO-EWPT) would point to BSM effects, producing signatures that can be detected in next-generation space interferometers, such as LISA [419] (already validated by the LISA Pathfinder [420]) or the similar Decigo experiment [421]. Some BSM models, including the singlet model discussed in Section 6.2, can create the required conditions to enable a SFO-EWPT, which can notably also explain baryogenesis. In particular, new states can

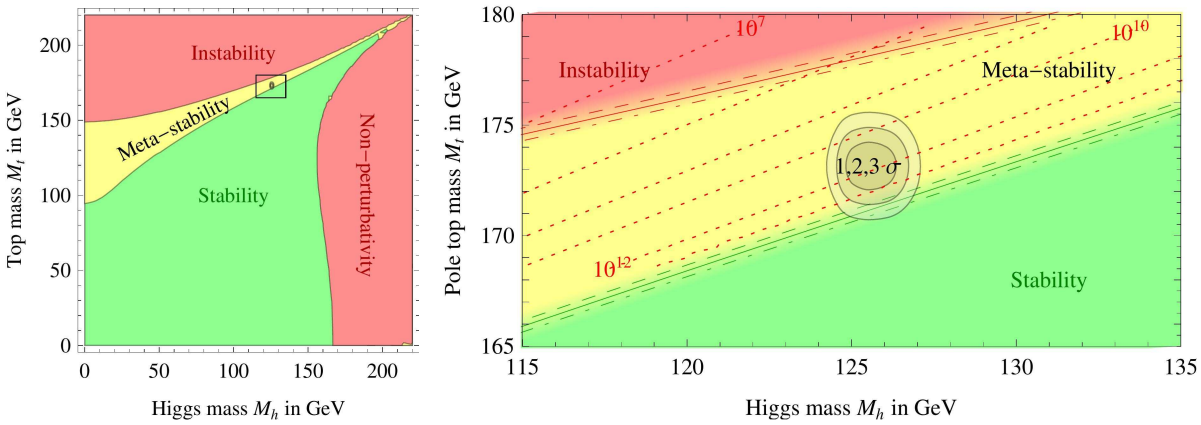

Figure 6.2: Regions of absolute stability (green), meta-stability (yellow) and instability (red) of the SM vacuum in the $m_t - m_H$ plane. Current measures tend towards a meta-stable Universe. (Left) Zoomed-out version. (Right) Zoomed-in version, also displaying the measured m_H and m_t experimental range, where the gray areas denote the allowed regions with their 1, 2 and 3 σ contours. The color gradient indicates the size of the theoretical uncertainties. Taken from [425].

modify the Higgs potential and its dependence on temperature, potentially introducing false, metastable vacua, as illustrated in Fig. 6.1. The false potential minima would allow sudden phase transitions via the sci-fi-esque coalescence of vacuum bubbles[2]. It is however important to note that, based on current m_t and α_S constraints [424], our Universe could be metastable without additional states. The stability conditions are ruled by the running of the self-coupling with energy, in a function that also depends on the Higgs boson and top quark masses. For λ_{HHH} at a certain energy scale, quantum tunneling from the current potential minimum to a lower minimum becomes possible. In Fig. 6.2, we can see measurements of m_t and m_H overlayed on stability regions. The current measurements seem to favor a scenario where the tunneling probability timescale is comparable to or longer than the age of the Universe, while remaining meta-stable. Despite the absence of a one-to-one correspondence between λ_{HHH} and the vacuum instability, the measurement of the self-coupling could nevertheless shed light on whether the long-term existence of the electroweak vacuum is not guaranteed in Nature, or if a yet unknown mechanism ensures stability [83]. We again point out the importance of HH production in the possible observation of variations of the Higgs potential, which might be connected to the EWPT, elucidating the matter-antimatter asymmetry experienced in our Universe [63, 424].

6.2 The Singlet Model

An important point under investigation is to know whether the scalar sector realized in Nature corresponds to the SM. Despite the remarkable success of the SM so far, many BSM physics predicts the existence of additional scalars, which could be singlets, doublets, or any other multiplets under the EW gauge group. These extensions typically come with additional BSM particles. Such models nevertheless have to obey the constraints set from both theory and experiments [426].

In this Chapter we focus on the SM extended by a real singlet scalar field S. This corresponds to the simplest possible extension to the SM, with the smallest addition of new

[2]Vacuum decay, with the creation of vacuum bubbles, might destroy the Universe as we know it [422, 423]!

free parameters, and thus serves as an ideal test case against which to validate ideas and new techniques. The singlet model also has no direct gauge interactions, making it harder to detect experimentally. Despite its simplicity, the model can generate a SFO-EWPT sufficient to explain EW baryogenesis, and provides a complementary approach with respect to using exotic Higgs decays [427]. The most general form of the Higgs scalar potential V becomes (compare with Eq. (1.39)):

$$V(\mathrm{H},\mathrm{S}) = \mu^2(\mathrm{H}^\dagger\mathrm{H}) + \frac{1}{2}(\mathrm{H}^\dagger\mathrm{H})^2 + c_1(\mathrm{H}^\dagger\mathrm{H})\mathrm{S} + \frac{c_2}{2}(\mathrm{H}^\dagger\mathrm{H})\mathrm{S}^2 + \frac{m_\mathrm{X}^2}{2}\mathrm{S}^2 + \frac{\kappa}{3}\mathrm{S}^3 + \frac{\lambda_\mathrm{S}}{2}\mathrm{S}^4 \ , \quad (6.2)$$

where the constants c_1 and c_2 represent the connection between the SM and the new singlet scalar. At this point, some authors impose a $\mathbb{Z}_2$ symmetry on the real singlet field, which is then softly broken by the VEV of the singlet field, leading to the mixing of the eigenstates of the H and S fields. This has the advantage of reducing the number of free parameters to only 3, usually parameterized as the mixing angle, the ratio of VEV's, and the mass of the new resonance. For this specific study, we consider instead a version of the singlet model where no $\mathbb{Z}_2$ symmetry $S \to -S$ is imposed on the scalar field. This is motivated by the fact that the $\mathbb{Z}_2$ symmetry precludes terms of odd powers of the additional singlet scalar field S, such as the ones multiplied by c_1 and κ in Eq. (6.4). The odd terms in S are known to be often responsible for a SFO-EWPT [415, 428], as discussed in Section 6.1.

Exploiting EW symmetry breaking, both H and S acquire a VEV, where the Higgs VEV v_0 is measured to be ∼246 GeV, using Fermi's constant (see Eq. (1.40)), and the new S VEV x_0 is unknown. Following the same procedure as for the Higgs mechanism (Section 1.1.3), we expand around the two fields: $\mathrm{H} \to (v_0 + h)/\sqrt{2}$ and $\mathrm{S} \to x_0 + s$. To obtain physical states, the mass matrix is again diagonalized, leading to[3]:

$$\begin{aligned} \mathrm{H} &= h\cos(\alpha) + s\sin(\alpha) \ , \\ \mathrm{X} &= -h\sin(\alpha) + s\cos(\alpha) \ , \end{aligned} \quad (6.3)$$

where the particle H can be identified with the measured 125 GeV Higgs boson, X is the new scalar arising from the additional field S, and α is the mixing angle. Following [415], we consider only the scenario where S is heavier than the Higgs boson. Notice that $\mathrm{X} \to \mathrm{HH}$ becomes kinematically possible as soon as $m_\mathrm{X} > 2m_\mathrm{H}$. The association of the first particle in Eq. (6.3) to the SM H boson reduces the number of uncorrelated parameters in the singlet model from seven to five. The couplings of the H boson and the X particle to any SM particle k are then simply given by:

$$g_{\mathrm{H}kk} = g_{\mathrm{H}kk}^{\mathrm{SM}}\cos\alpha \quad \text{and} \quad g_{\mathrm{X}kk} = -g_{\mathrm{H}kk}^{\mathrm{SM}}\sin\alpha \ , \quad (6.4)$$

where the angle α is the parameter that enables an indirect connection of the new scalar with the gauge and fermionic sectors of the SM. When $\sin(\alpha) = 0$, $\cos(\alpha) = 1$ and the SM is recovered. The hardest constraints on the singlet model usually come from searches for di-boson final states, single Higgs combinations, and the measurements of the W mass [429].

In the next Sections, we use the real singlet model for a dedicated study of interference effects arising from a non-zero X resonance width. We are particularly interested in

[3] To simplify the notation, we are using the same letter H to refer to both the particle and the field.

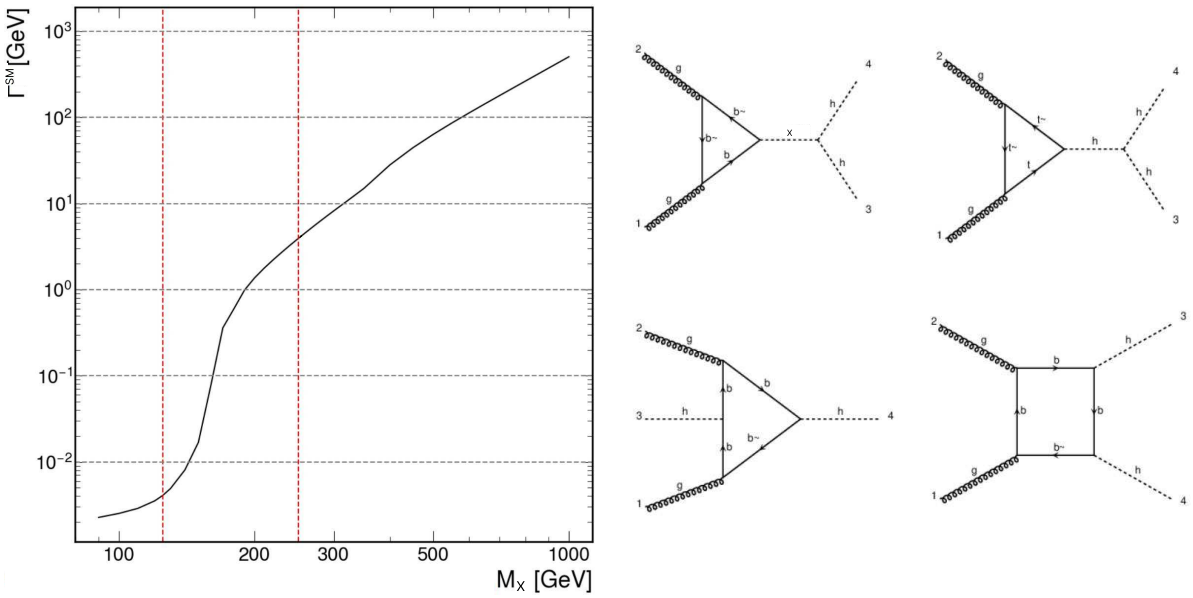

Figure 6.3: (Left) Width of a scalar boson of mass m_X with the same decay modes as the SM H boson, as a function of m_X, as interpolated from the values published in [49]. The $\Gamma^{SM}(m_X)$ curve is used in Eq. (6.5). The red dashed lines are for guidance only, at 125 GeV and 250 GeV. Please note that the 2D scans in this section start at $m_X = 280$ GeV. (Right) Example of 4 of the 20 LO Feynman diagrams considered by the singlet model UFO used for this chapter's study. They include triangle and box diagrams, featuring bottom and top quarks, the SM Higgs and the new scalar X.

knowing whether the NWA holds everywhere and, if not, in which regions of the phase space.

6.3 Methodology

In the real singlet model, for $m_X > 2m_H$, the width of the resonance X is given by:

$$\Gamma_X = \sin^2\alpha\,\Gamma^{SM}(m_X) + \Gamma_{X\to HH} \equiv \sin^2\alpha\,\Gamma^{SM}(m_X) + \frac{\lambda^2_{HHX}\sqrt{1-4m_H^2/m_X^2}}{8\pi m_X}\ , \qquad (6.5)$$

where λ_{HHX} is the trilinear coupling between two H bosons and the new particle X, and $\Gamma^{SM}(m_X)$ represents the width of a scalar boson of mass m_X with the same decay modes of the SM H boson. The width $\Gamma^{SM}(m_X)$ has been calculated by interpolating the values published in [49] and is shown in Fig. 6.3 (left). In addition to α, m_X and λ_{HHX}, this singlet model also depends on the trilinear H coupling modifier $\kappa_\lambda \equiv \lambda_{HHH}/\lambda^{SM}_{HHH}$, and on an additional scalar coupling. The values of the branching ratio of $X \to HH$ with respect to the entire phase-space of the particle X are encoded by $\Gamma_{X\to HH}/\Gamma_X$ and displayed in Fig. 6.4 for two values of m_X, as a function of $\sin\alpha$ and λ_{HHX}. The plots are meant to provide an idea of the range of parameters considered, and how they affect Γ_X. In particular, notice that, as expected, $\mathcal{B}(X \to HH)$ is zero along the $\lambda_{HHX} = 0$ horizontal line, as no coupling exists connecting the H and X sectors. On the contrary, $\mathcal{B}(X \to HH)$ is always one along the $\sin\alpha = 0$ ($\cos\alpha = 1$) vertical line, since in that case X does not couple to any SM particle except the Higgs boson. Also note how the branching ratio

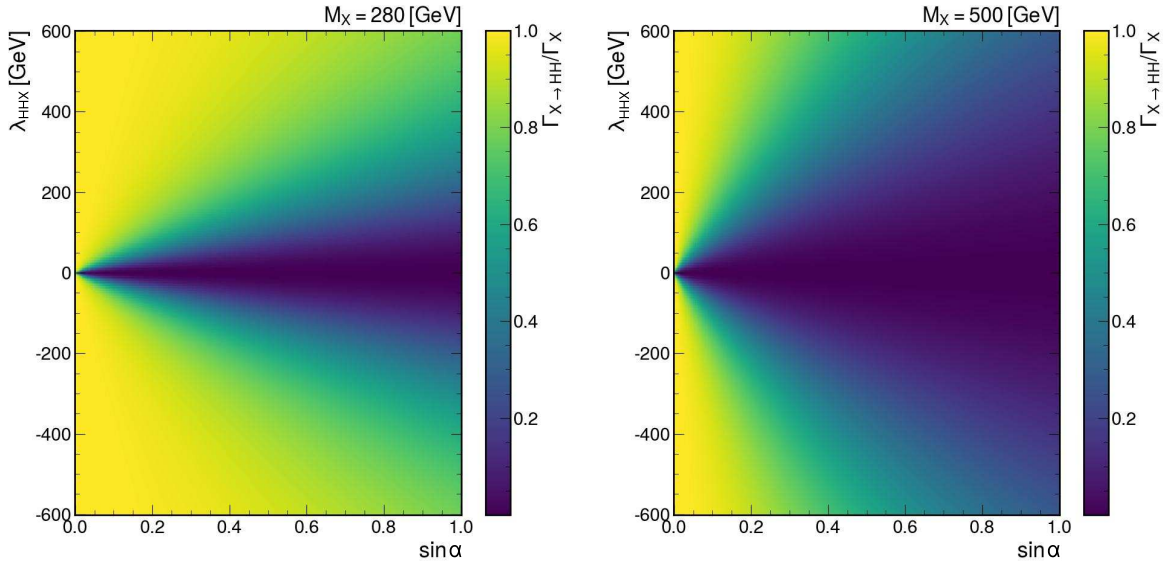

Figure 6.4: Singlet model $\Gamma_{X\to HH}/\Gamma_X$ branching ratio scan, following Eq. (6.5), as a function of $\sin\alpha$ and λ_{HHX}, for $m_X = 280\,\text{GeV}$ (left) and $m_X = 500\,\text{GeV}$ (right). The point at $\sin\alpha = 0$ and $\lambda_{HHX} = 0\,\text{GeV}$ corresponds to the SM scenario, where Eq. (6.5) is not well defined.

changes with increasing m_X, given the extremely significant increase of $\Gamma^{SM}(m_X)$ with the resonance mass.

We use the `MadGraph5_aMC@NLO` generator, version 2.9.7 [241], to simulate inclusive HH events in the singlet model at LO. A custom Universal FeynRules Output (UFO) [430] model based on [415] adds a heavy scalar boson to the SM with couplings to SM particles as defined in Eq. (6.4), following what is described in Section 6.2. In total, the model considers 20 ggF LO Feynman diagrams, 4 of which are shown in Fig. 6.3 (right). The samples are created according to the following parameter grid, using $\kappa_\lambda = 1$:

- m_X [GeV]: 280, 300, 400, 500, 600, 700, 800, 900, 1000;
- $\sin\alpha$: 0.00, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.99;
- λ_{HHX} [GeV]: −600, −500, −400, −300, −200, −100, −50, 0, 50, 100, 200, 300, 400, 500, 600.

The values of m_X are chosen based on the signal samples used in the HH combination presented in Section 5.5. The resonant, nonresonant, and total HH cross sections in the singlet model for each combination of grid points are generated separately, to what amounts to a very computationally intensive task, with the parallel generation of 4860 grid points. The nonresonant cross section is obtained by setting the g_{Xkk} coupling in Eq. (6.4) to zero, and the resonant-only cross section by setting g_{Hkk} to zero. We perform parameter scans in m_X, $\sin\alpha$, and λ_{HHX} for the three cross sections above, and additionally for the *interference ratio*, defined as:

$$R_{int} = \frac{\sigma^{\text{total}} - \left(\sigma^{\text{resonant-only}} + \sigma^{\text{nonresonant}}\right)}{\sigma^{\text{resonant-only}} + \sigma^{\text{nonresonant}}} \,. \tag{6.6}$$

The variable R_{int} provides information concerning the relative strength of the interference between SM and BSM diagrams. When no interference is present, the total cross

section is equal to the sum of its resonant and nonresonant components. We consider the ggF production mode due to its dominant contribution to the cross section. Some total cross section scans are shown in Fig. 6.5, where two-dimensional histograms containing the total cross section as a function of $\sin\alpha$ and λ_{HHX} are displayed for four m_{X} values. The UFO model and procedure are validated using the program `hpair` [431, 432] where κ_λ-dependent results in the nonresonant scenario are found to agree with the NLO predictions of [38].

6.4 Results

Any conclusions to be drawn from this study naturally depend on the thresholds imposed on $\mathrm{R_{int}}$ and $\Gamma_{\mathrm{X}}/m_{\mathrm{X}}$. In the following, we choose as thresholds $\mathrm{R_{int}} = \pm 10$ and $\pm 20\%$, and $\Gamma_{\mathrm{X}}/m_{\mathrm{X}} = 5$, 10 and 20 %. The $\Gamma_{\mathrm{X}}/m_{\mathrm{X}}$ values are defined taking into account jet resolutions achieved by CMS, which amount typically to 15 to 20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV [232]. The contours corresponding to the thresholds above and the exclusion limits derived from the HH combination in the singlet model are shown in Fig. 6.6. The red HH exclusion region uses the total cross section histograms as shown in Fig. 6.5. The HH combination technically did not use the real singlet model, but rather the Randall-Sundrum, NWA spin-0 model, usually considered for spin-0 resonance searches [88]. Given that the two models have the same scalar coupling structure, interpretations in the context of the singlet model remain valid. The $\Gamma_{\mathrm{X}}/m_{\mathrm{X}}$ curves are obtained using Eq. (6.5), and $\mathrm{R_{int}}$ curves use 2D histograms similar to the total cross section ones, after considering Eq. (6.6). Positive and negative interference ratio contours are shown in green ($\mathrm{R_{int}} = +10\%, +20\%$) and blue ($\mathrm{R_{int}} = -10\%, -20\%$). The contours are found to swap positions at $m_{\mathrm{X}} = 400$ GeV, likely because of the peak of the nonresonant HH distribution. For a given m_{X}, the quadratic dependence of Γ_{X} on both $\sin\alpha$ and λ_{HHX} according to Eq. (6.5) leads to the dotted black elliptical isolines of constant $\Gamma_{\mathrm{X}}/m_{\mathrm{X}}$. Those lines denote thresholds beyond which $\Gamma_{\mathrm{X}}/m_{\mathrm{X}}$ exceeds 5%, 10% or 20%, implying that the NWA might no longer be accurate. The experimental bound from the HH combination discussed in Section 1.4.1 is obtained from the 95% CL upper limit on $\sigma(\mathrm{pp}\to\mathrm{X})\mathcal{B}(\mathrm{X}\to\mathrm{HH})$, with the X production cross section growing with increasing $\sin\alpha$, and $\mathcal{B}(\mathrm{X}\to\mathrm{HH})$ growing with increasing λ_{HHX}. We note that large values of $\sin\alpha$, corresponding to regions where the H boson is less SM-like, also tend to be excluded by precision measurements of the Higgs boson, usually at $\sim$0.16 and above [415]. For most of the studied mass points, sizable interference ratios occur only in parameter regions to which the current measurements are not yet sensitive, or at too large values of $\sin\alpha$. In particular, for large resonance masses, with m_{X} above 600 GeV, where interference effects tend to grow, measurements are far below the current sensitivity and might only play a role when the full data set from the HL-LHC becomes available [433]. However, there are regions at intermediate m_{X} where the interpretation of NWA-based limits for HH derived in the singlet model would solicit some care already in the Run 2 combination (for example $m_{\mathrm{X}} = 500$ GeV, $\sin\alpha = 0.2$ and $\lambda_{\mathrm{HHX}} = 400$ GeV). It is important to note that such interpretations are generally model dependent.

The differential cross sections as a function of m_{HH} are shown for a subset of points from the ($\sin\alpha$, λ_{HHX}) parameter space in Fig. 6.7 for $m_{\mathrm{X}} = 280$ GeV, and in Fig. 6.8 for $m_{\mathrm{X}} = 500$ GeV. The parameters are chosen such that $\Gamma_{\mathrm{X}}/m_{\mathrm{X}} = 5\%$, which is well below the detector resolution, and $\mathrm{R_{int}} = \pm 10\%$ or $\pm 20\%$ so that sizable interference

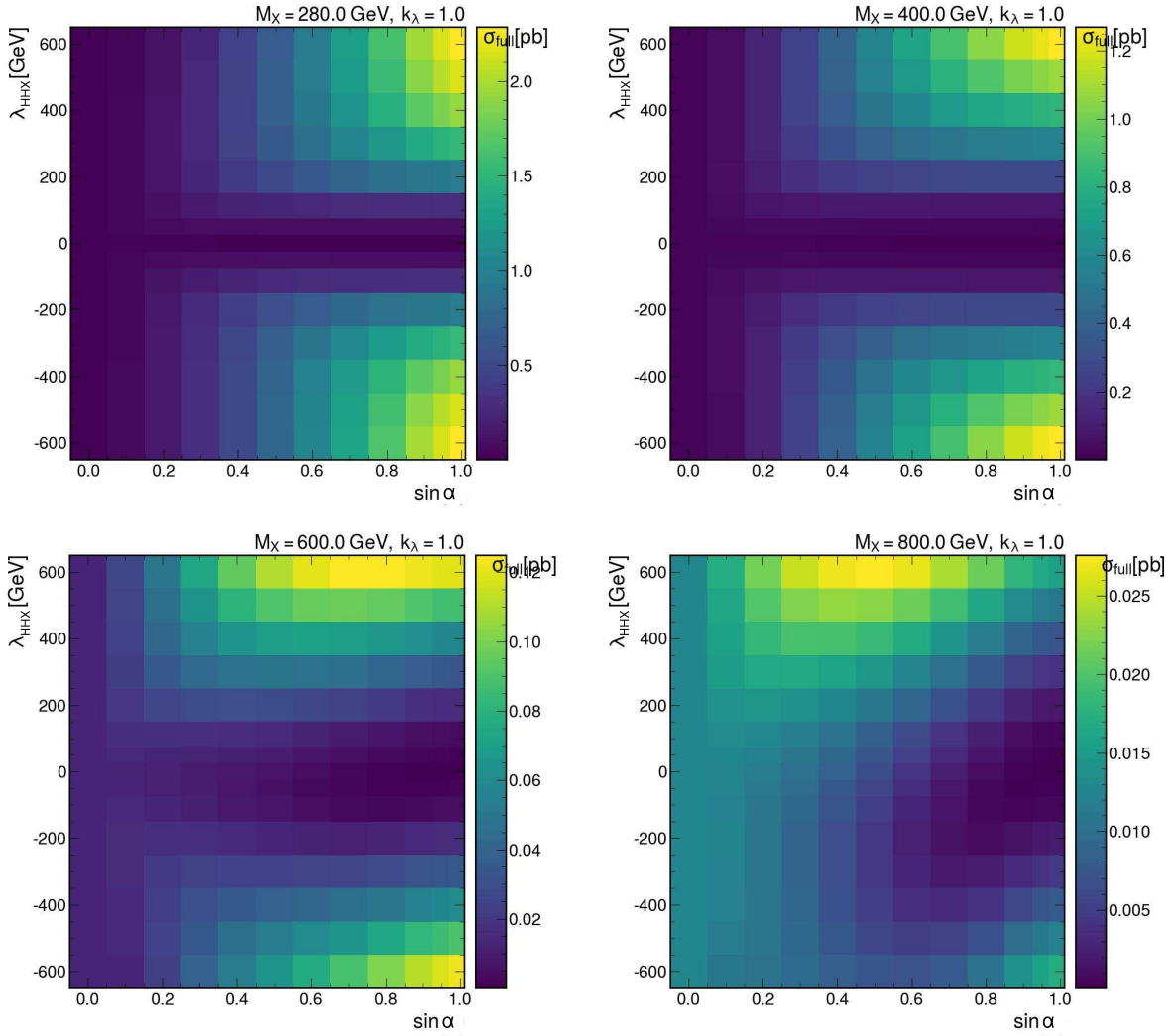

Figure 6.5: Two-dimensional maps for the total cross section of the singlet model described in Section 6.2, as a function of $\sin\alpha$ and λ_{HHX}, for $m_{\mathrm{X}} = 280\,\mathrm{GeV}$ (top left), $m_{\mathrm{X}} = 400\,\mathrm{GeV}$ (top right), $m_{\mathrm{X}} = 600\,\mathrm{GeV}$ (bottom left) and $m_{\mathrm{X}} = 800\,\mathrm{GeV}$ (bottom right). Similar maps were obtained for all other m_{X} values mentioned in the text. Similar maps for the interference ratios (Eq. (6.6)) are later used to draw contour maps for singlet model phase-space regions that are excluded with current HH measurements.

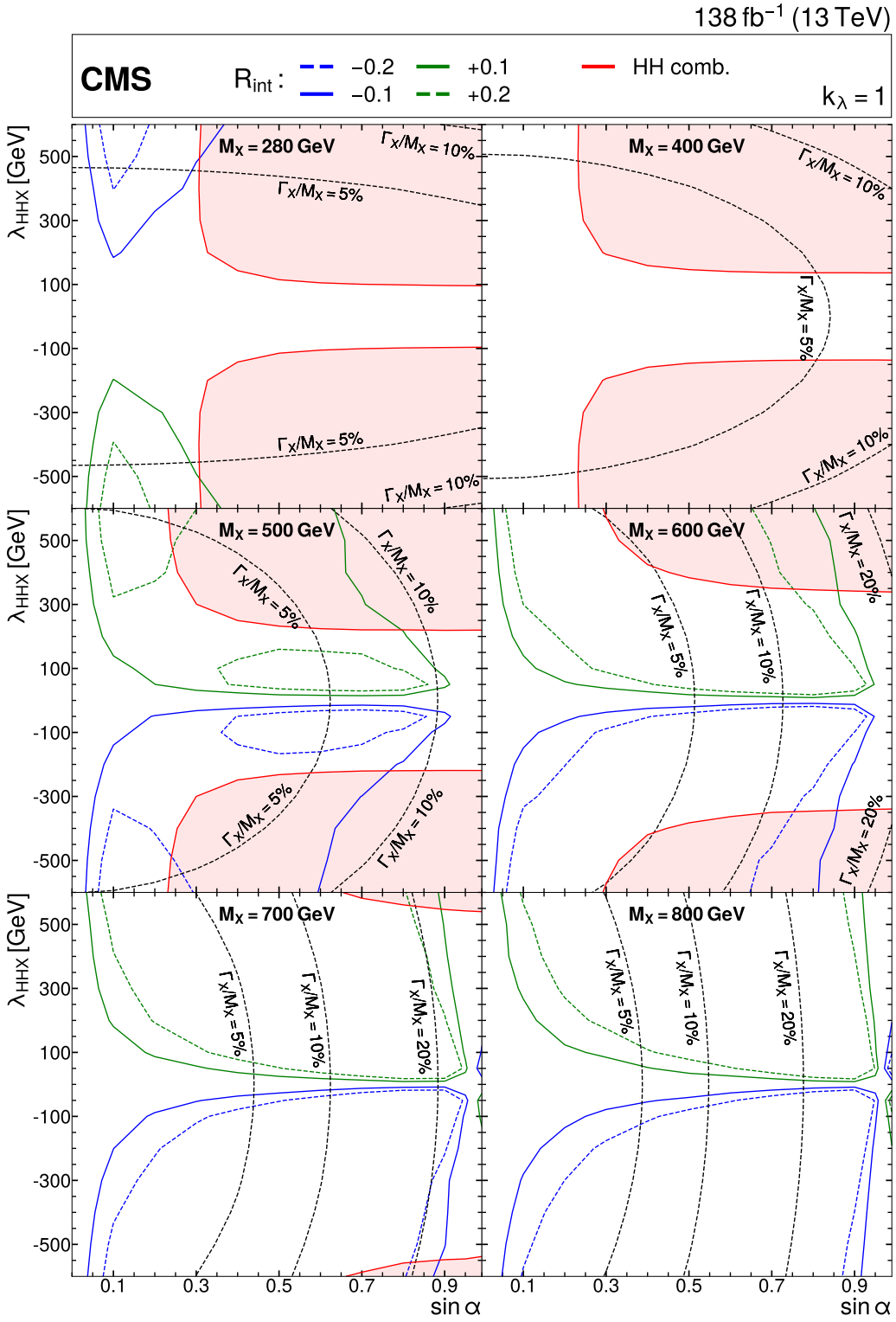

Figure 6.6: Contours of the variable R_{int} as defined in Eq. (6.6) and discussed in the text, in the ($\sin\alpha$, λ_{HHX}) plane for the singlet model with $\kappa_\lambda = 1$ and different resonance masses m_X between (upper left) 280 and (lower right) 800 GeV. Contours are shown for R_{int} values of (dashed blue) −0.2, (solid blue) −0.1, (solid green) 0.1, and (dashed green) 0.2. Regions that are excluded, at 95% CL, from the combined likelihood analysis of the HH analyses presented in this report are indicated by red filled areas. Dashed black lines indicate constant relative widths of 5, 10, and 20%. Published in [7].

effects are expected. The lineshapes show points in parameter space where the R_{int} contours intersect with lines of constant $\Gamma_X/m_X = 5\%$ in Fig. 6.6. The mass points of $m_X = 280\,\text{GeV}$ and $500\,\text{GeV}$ have been chosen because these values are on the left- and right-hand side of the peak in the m_{HH} distribution for nonresonant SM HH production. The signature indicating the presence of the X resonance, including interference effects, can be assessed as the difference between σ^{full} (red graph) and $\sigma^{\text{nonresonant}}$ (green graph). In the $m_X = 280\,\text{GeV}$ case, the resonance peak is at a mass where the nonresonant background is low in comparison. The shape of central part of the peak is thus not much affected, and a classical bump hunt should still work. Still, the signature develops a peak-dip structure for a negative interference ratio, and a shoulder-like enhancement towards large masses for a positive interference ratio. For a precision measurement, which is not yet in our reach, a distortion of the signal shape would have to be taken into account. For $m_X = 500\,\text{GeV}$, in the top panels of Fig. 6.8, the signal shape is found to be strongly modified by the interference effect. Similar distortion effects as seen for $m_X = 280\,\text{GeV}$, albeit on reversed sides of the peak, are visible for $m_X = 500\,\text{GeV}$. The distortion however occurs in a $(\sin\alpha,\, \lambda_{HHX})$ parameter region still relatively far from the regions currently probed, as can be appreciated in Fig. 6.6. The total cross section, for all m_X values, is modified as specified by R_{int}. Although the expected interference effects clearly depend on the underlying model, they can be expected to be of mounting importance in the future as the LHC data set increases.

Recent projections for HH results [7] also enable to study the level at which the parameters of the singlet model might be excluded in the future, namely at the end of the HL-LHC. Using $\kappa_\lambda = 1$, limits are derived in the $(\sin\alpha,\, \lambda_{HHX})$ plane from the combination of resonant HH searches. Resonances between 280 and 1000 GeV are probed using Run 2 data and are projected to integrated luminosities corresponding to 300, 1000, and $3000\,\text{fb}^{-1}$. Projected exclusion regions at 95 % CL are shown in Fig. 6.9, for four different luminosity values, following systematics scenarios mentioned in Section 5.5 ("S1" and "S2"). The HL-LHC dataset of $3000\,\text{fb}^{-1}$ has the potential to considerably expand the present exclusion regions in the $(\sin\alpha,\, \lambda_{HHX})$ plane for all values of m_X. Compared to the present limits, the largest improvement is observed for large masses, $m_X = 600\,\text{GeV}$ and higher, where large regions of the $(\sin\alpha,\, \lambda_{HHX})$ plane can be probed, and part of the plane even up to $m_X = 1\,\text{TeV}$.

We can also inspect the impact the expected limits obtained in Chapter 5 have on singlet model and its interference with the SM. The results are shown in Fig. 6.10, where two things become clear. Firstly, the $bb\tau\tau$ results obtained in this Thesis, as had been discussed before, clearly drive the sensitivity of the CMS combination for intermediate masses. Secondly, the validity of the NWA becomes less evident for $m_X = 500/600\,\text{GeV}$, still using Run 2 data. Our result is also overlayed with the combined HH projections, as shown in Fig. 6.11. We perform better than the full combination at the end of Run 3, using the $bb\tau\tau$ channel only.

We finalize this Chapter by noting that all conclusions drawn above are model dependent, and may well be different for other BSM scenarios. This means that distortions could be stronger and affected parameter regions could be larger. We also recall that a (reasonable) choice of parameter thresholds had to be made, and which could have been different. This represents however the first dedicated interference study within CMS, hinting already at the necessity of avoiding the NWA in future HH analyses.

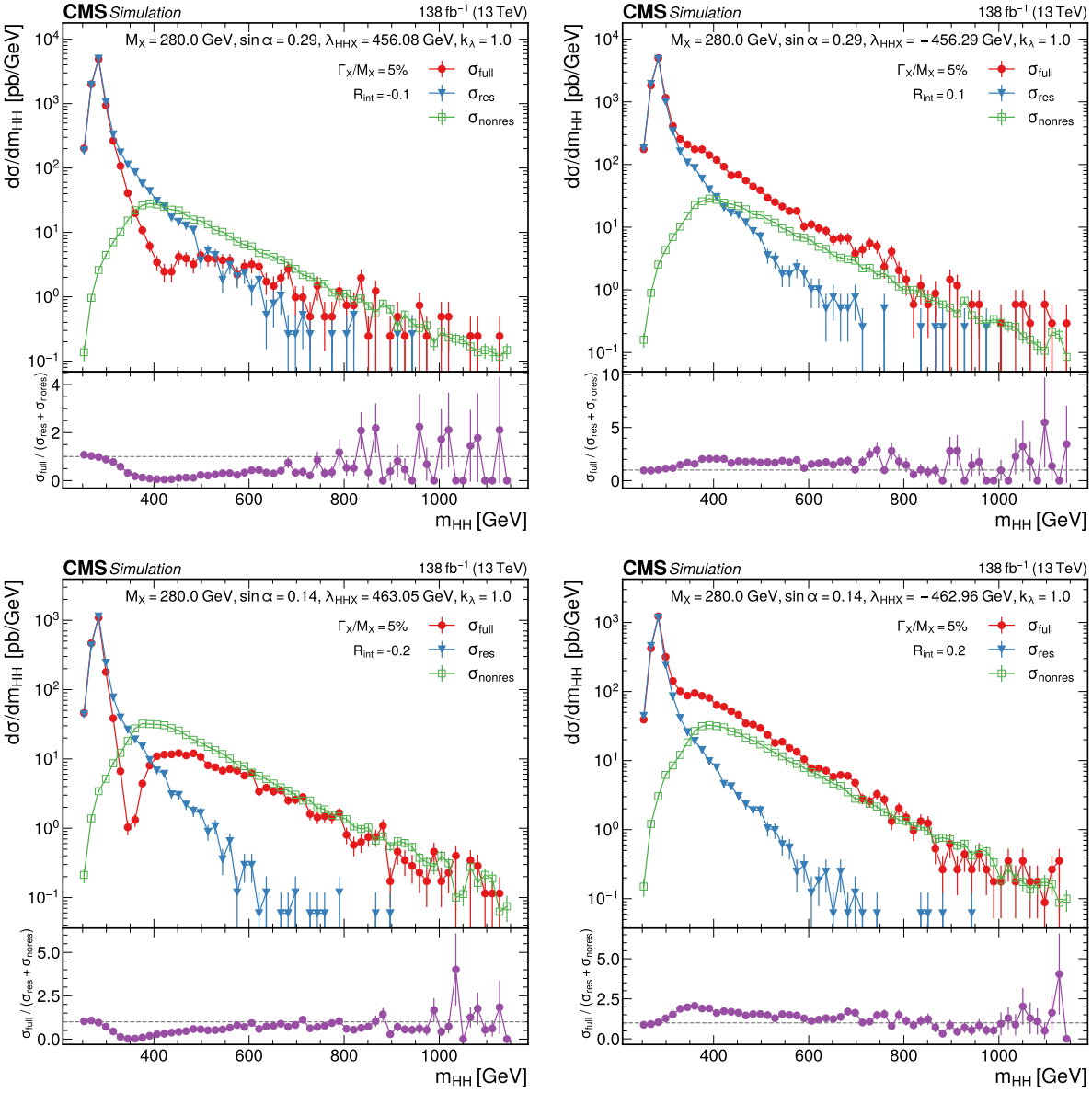

Figure 6.7: Expected differential cross sections for HH production, as a function of m_{HH}, for the real-singlet model with $m_{\mathrm{X}} = 280\,\mathrm{GeV}$ and $\Gamma_{\mathrm{X}}/m_{\mathrm{X}} = 5\%$. The parameters $\sin\alpha$ and λ_{HHX} have been chosen such that (upper row) $\mathrm{R_{int}} = \pm 10\%$ and (lower row) $\mathrm{R_{int}} = \pm 20\%$, (left) negative and (right) positive values of $\mathrm{R_{int}}$. The total cross section for HH production σ^{full} (red line, labeled as σ_{full}) is compared to the cross sections $\sigma^{\text{resonant-only}}$ (blue line, labeled as σ_{res}) and $\sigma^{\mathrm{nonresonant}}$ (green line, labeled as σ_{nores}) considering only resonant and nonresonant production. In the lower panels the ratio of σ^{full} over $(\sigma^{\text{resonant-only}} + \sigma^{\mathrm{nonresonant}})$ is shown. Published in [7].

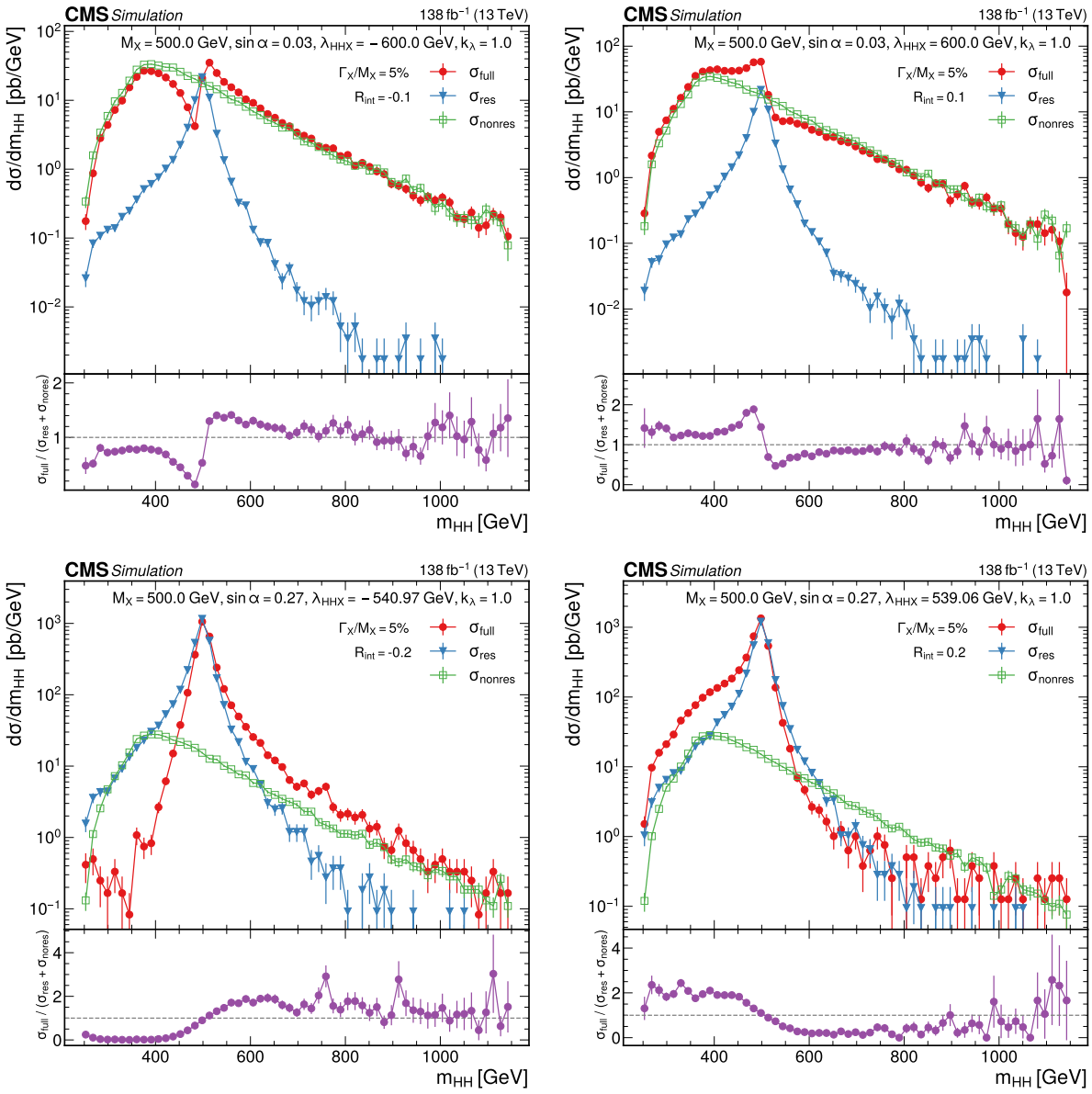

Figure 6.8: Expected differential cross sections for HH production, as a function of m_{HH} for the real-singlet model with $m_{\mathrm{X}} = 500\,\mathrm{GeV}$ and $\Gamma_{\mathrm{X}}/m_{\mathrm{X}} = 5\%$. The parameters $\sin\alpha$ and λ_{HHX} have been such that (upper row) $\mathrm{R_{int}} = \pm 10\%$ and (lower row) $\mathrm{R_{int}} = \pm 20\%$, for (left) negative and (right) positive values of $\mathrm{R_{int}}$. The total section for HH production σ^{full} (red line, labeled as σ_{full}) is compared to the cross sections $\sigma^{\text{resonant-only}}$ (blue line, labeled as σ_{res}) and $\sigma^{\mathrm{nonresonant}}$ (green line, labeled as σ_{nores}) considering only resonant and nonresonant production. In the lower panels the ratio of σ^{full} over $(\sigma^{\text{resonant-only}} + \sigma^{\mathrm{nonresonant}})$ is shown. Published in [7].

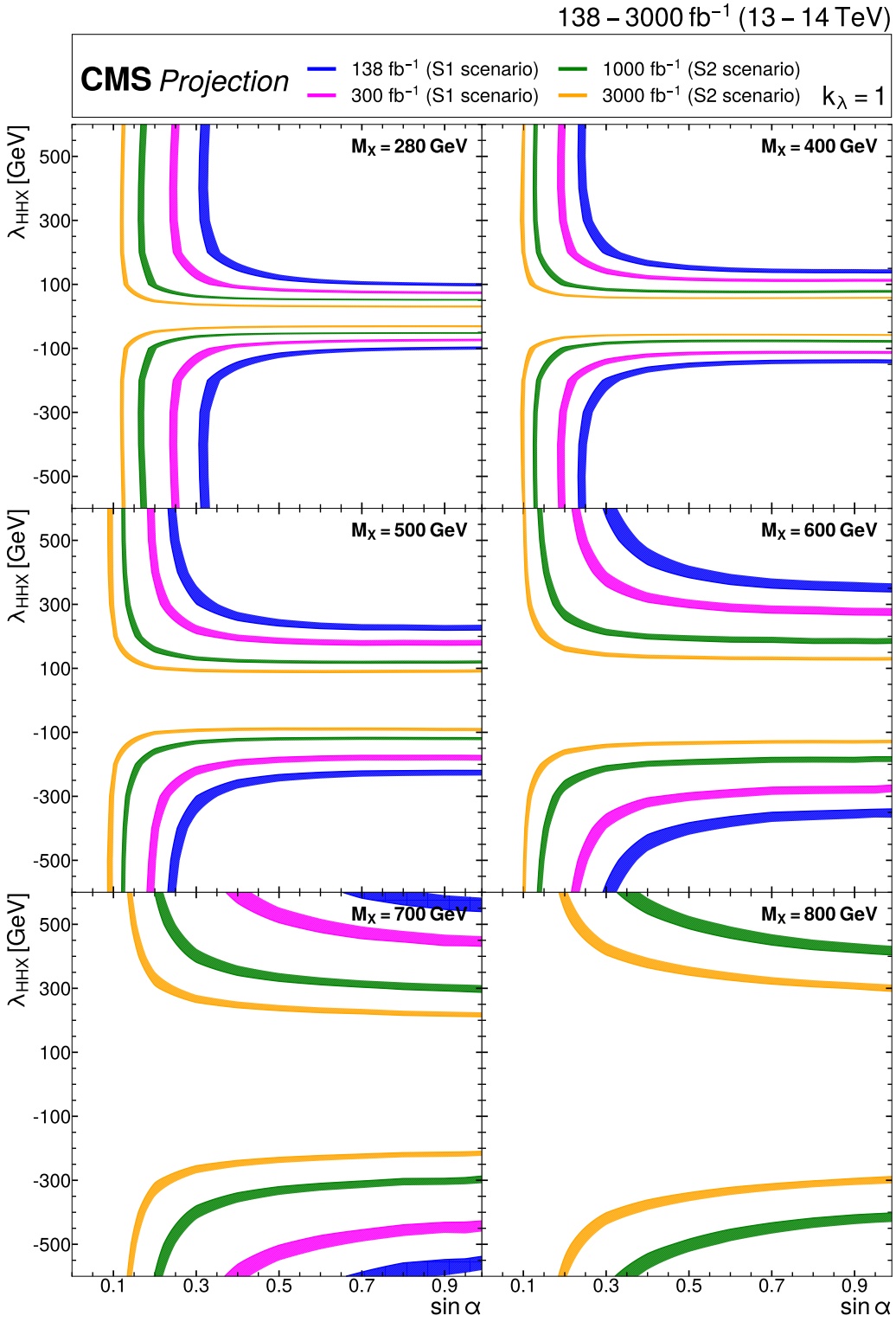

Figure 6.9: Exclusion contours at 95% CL, in the ($\sin\alpha$, λ_{HHX}) plane for $\kappa_\lambda = 1$ in the real-singlet model. These contours are obtained from the combined likelihood analysis of the HH searches discussed in Section 5.5 for (upper left to lower right) $m_{\mathrm{X}} = 280$, 400, 500, 600, 700, and 1000 GeV. The expected limits from the Run 2 dataset have been projected to integrated luminosities of 300, 1000, and 3000 fb^{-1}. Excluded areas are indicated by the direction of the hatching along the exclusion contours. Published in [7].

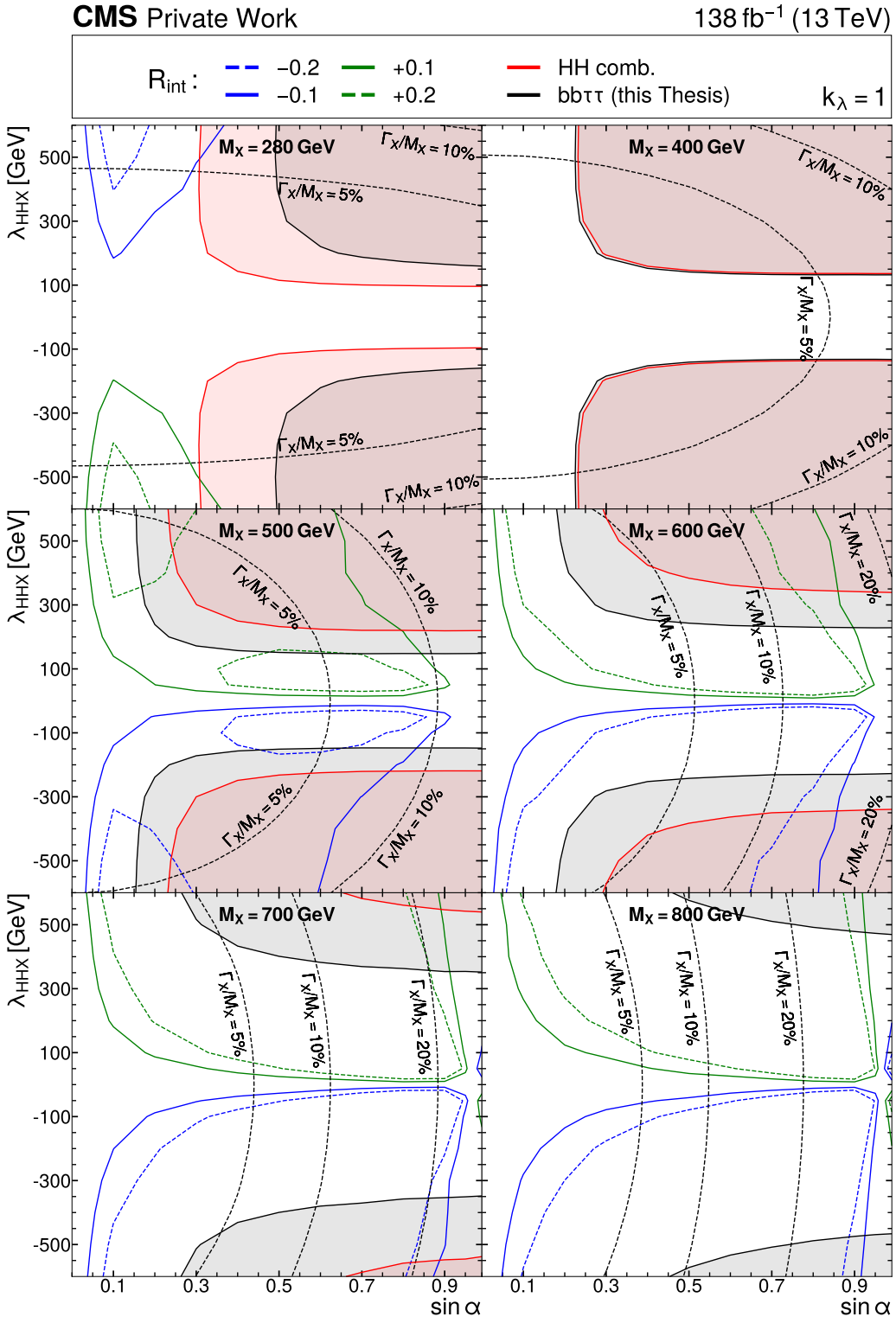

Figure 6.10: This Figure is identical to Fig. 6.6, except for the addition of the black contours, referring to the results obtained in Chapter 5. The validity of the NWA is put into question more strongly, especially at $m_X = 500/600\,\text{GeV}$.

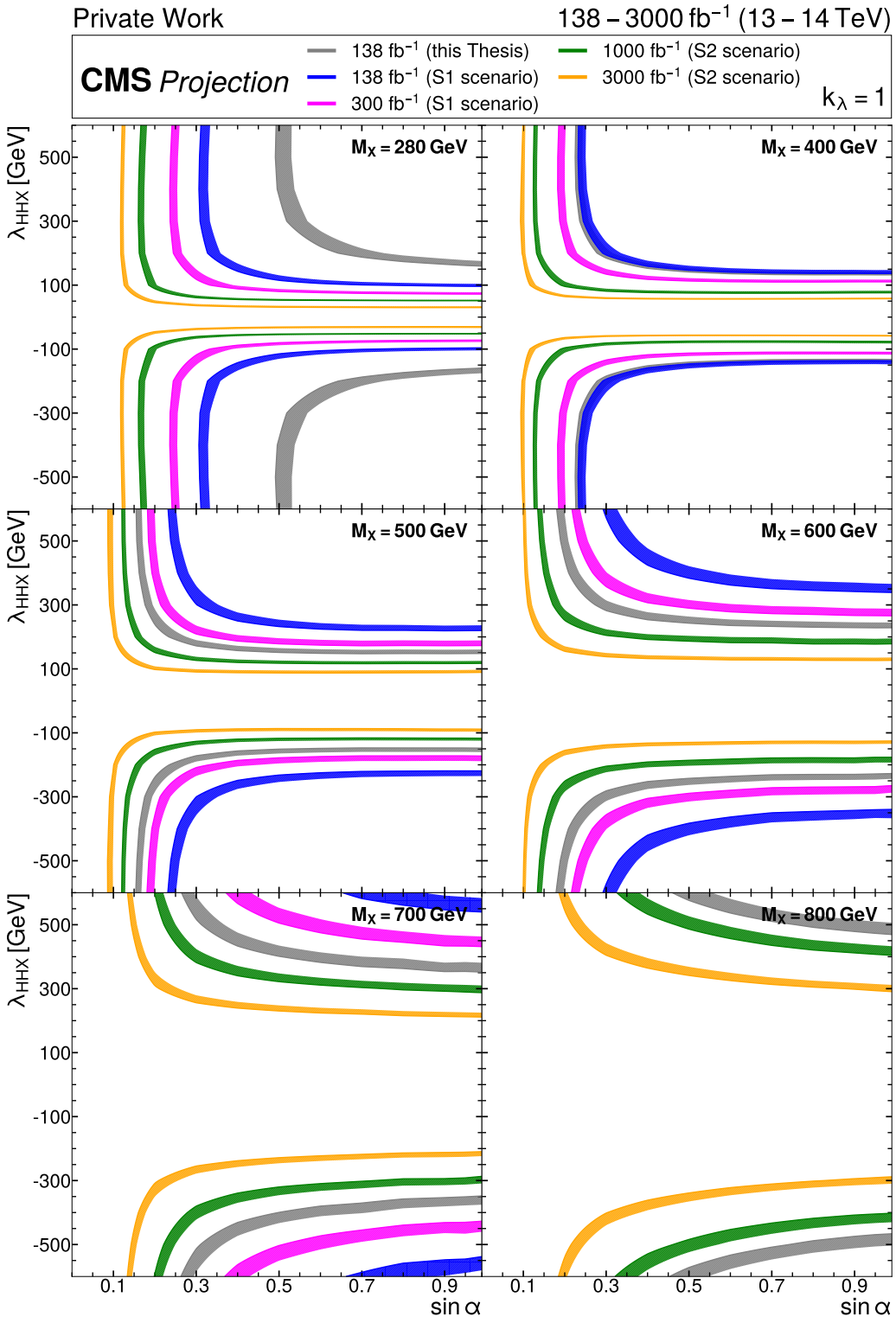

Figure 6.11: This Figure is identical to Fig. 6.9, except for the addition of the black contours, referring to the results obtained in Chapter 5. Our Run 2 result outperforms the combined projections for the end of Run 3, starting at $m_X = 400\,\text{GeV}$ mass points.

Conclusion

The work presented in this Thesis covers some of the central axes of Experimental Particle Physics. By exploiting the current and future capabilities of the CMS detector at CERN's LHC, original findings have been provided to the field, across a range of topics. The contributions span developments for the reconstruction of Trigger Primitives in the future High Granularity Calorimeter (HGCAL), and studies related to the production of Higgs boson pairs. The latter is approached from the perspective of a physics analysis, through the resonant production of the bb$\tau\tau$ final state, and from a phenomenological angle, by investigating the impact of removing the Narrow Width Approximation from HH interpretations, namely on double Higgs production cross section upper limits.

The upcoming High-Luminosity LHC (HL-LHC) phase promises to bring a ten-fold increase in the integrated luminosity delivered to the experiments, with a center-of-mass energy of 14 TeV. This additional data will strongly benefit the sensitivity of physics measurements, especially for rare processes, which in turn leads to more precise theoretical calculations. However, the HL-LHC also brings major challenges. The expected huge increase of radiation levels requires the complete hardware refurbishment of many CMS detectors, especially when located close to the beamline. Large upgrades are furthermore required in the triggering and reconstruction of physics objects, in order to sustain the tremendous data rates to be processed, and to disentangle the $\sim$200 pile-up proton interactions at each bunch-crossing. In particular, the HGCAL is bound to become a flagship detector for CMS. It will bring impressive spatial and timing resolutions, within a lateral and longitudinal granularity never before seen in a High Energy Physics experiment. The successful reconstruction in such a crowded environment starts from an effective trigger strategy, in which Trigger Primitives play a major role. This Thesis proposed new algorithms for the reconstruction of Trigger Primitives, focusing on the mitigation of cluster splits during the clustering reconstruction stage. Such a task was achieved through a software framework which was developed from scratch, includes a simplified version of HGCAL's geometry, and serves as a basis for further studies. The algorithms have the potential to be implemented in the firmware, and improve HGCAL's capabilities. Importantly, cluster splitting is shown to be almost fully removed, using techniques that will not require additional firmware resources.

On the physics analysis side, CMS is strongly engaged on studying the production of Higgs boson pairs, exploiting both nonresonant and resonant processes as a gateway to New Physics. These studies are deeply motivated by the intrinsic structure of the Standard Model (SM) of Particle Physics, which contains a yet unmeasured coupling of the Higgs boson with itself, and by a wide range of Beyond the SM (BSM) frameworks, where new particles couple to the Higgs sector and can profoundly alter the understanding of fundamental particle interactions. In this work, a search for the resonant production of a pair of Higgs bosons decaying into a pair of b quarks and a pair of τ leptons is performed, using CMS Run 2 data collected at 13 TeV. 95% CL upper limits on the production cross section of spin-0 and spin-2 BSM particles are computed, for masses from 250 GeV to 3 TeV. When comparing the expected limits obtained to the first 2016-only

CMS results, an improvement much larger than a simple luminosity scaling is observed. The performance increase comes from updated samples, an extended trigger strategy and analysis phase-space, the optimization of many selection stages, and the development and implementation of powerful discriminants. This work provides the tightest limits ever obtained in the $\mathrm{bb}\tau\tau$ channel for most of the scanned mass range, and demonstrates that this channel will drive CMS' sensitivity at intermediate mass values. The results should thus be included in a future $\mathrm{X} \to \mathrm{HH}$ combination. Moreover, for some mass points, the Run 2 results reach and surpass the most recent projection of $\mathrm{X} \to \mathrm{HH} \to \mathrm{bb}\tau\tau$ results for the HL-LHC. The techniques employed in the analysis are already serving as a starting point for future $\mathrm{bb}\tau\tau$ studies, which will further test the consistency of the SM and beyond.

The above resonant analysis was complemented by a phenomenological study, which probed the impact that resonance effects have on HH cross section upper limits. For the first time in CMS, the Narrow Width Approximation was not used, in order to assess interference effects between BSM particles decaying to a Higgs boson pair, and the continuous SM HH background. The study is performed using the singlet model. For most of the studied mass points, the effects are sizable only in parameter regions to which current measurements are not yet sensitive. However, there are regions at intermediate m_{X} values where the interpretation of NWA-based limits requires some care already in the Run 2 combination. Such concerns will definitely play a role when the full HL-LHC dataset becomes available. Additionally, clear distortions in the m_{HH} lineshapes are observed, which will have to be taken into account as soon as HH processes can be precisely measured. It is important to note that such interpretations are generally model dependent, and that a reasonable choice of parameter thresholds had to be made. Notwithstanding such limitations, the study hints at the necessity of avoiding the NWA in future HH analyses. Finally, this Thesis provides the most recent projections of singlet model upper limits, for a variety of resonance masses, to HL-LHC luminosity values.

I conclude this Thesis by expressing my sincere wish that the contributions here discussed will, first and foremost, be useful, and that might hopefully provide a modest basis for the next groundbreaking discovery we all wish to witness.

A Appendices

A.1 Summary of ATLAS Results on Higgs Boson Pair Production

In this Section, we provide a very brief overview on the most up-to-date results obtained by the ATLAS Collaboration on Higgs boson pair production. A comparison with CMS results can be made by looking at Sections 1.4.1 and 1.4.2. The channels covered by ATLAS are virtually the same as covered by CMS, but occasionally structured a bit differently: $HH \to bbbb$ resolved [434] and boosted [435], $HH \to bb\gamma\gamma$ [436], $HH \to bb\tau\tau$ [351], $HH \to bb\ell\ell + E_T^{miss}$ [437] and combined multi-leptonic decays [438].

Recent combined nonresonant results from the most sensitive searches with full Run 2 data are obtained. In Fig. A.1, observed and expected 95% CL upper limits on the ggF (left) and VBF (right) HH signal strength are depicted. The three most sensitive channels are the silver bullets, bbbb, $bb\gamma\gamma$ and $bb\tau\tau$, as expected. The ggF combination leads to an observed (expected) limit on the HH cross section at 2.9 (2.4) the SM expectation. The VBF combination is naturally much worse, the analysis being mostly used to constrain κ_{2V}. In Fig. A.2, the Higgs boson self-coupling modifier is constrained to the observed (expected) values of $-1.2 < \kappa_\lambda < 7.2$ ($-1.6 < \kappa_\lambda < 7.2$), while the coupling modifier between two Higgs bosons and two vector bosons is constrained to the observed (expected) values of $0.57 < \kappa_{2V} < 1.48$ ($0.41 < \kappa_{2V} < 1.65$). The limits refer to a 95% CL. A two dimensional fit is shown in Fig. A.3 (left). It is done in order to reduce model dependence, since the uni-dimensional fits are run by fixing all other coupling modifiers to their SM values. Finally, a ggF HH EFT interpretation is performed for the three silver bullets, ignoring VBF production, as shown in Fig. A.3 (right). Constrains on various BSM couplings can be inspected in Ref. [439].

Resonant results are shown in Fig. A.4, were we can emphasize the small excess observed at $m_X = 1\,\text{TeV}$ for $bb\tau\tau$, with a local (global) significance of $3.1\,\sigma$ ($2.0\,\sigma$). A statistical combination of the three silver bullets is also shown. The other channels do not observe any tension.

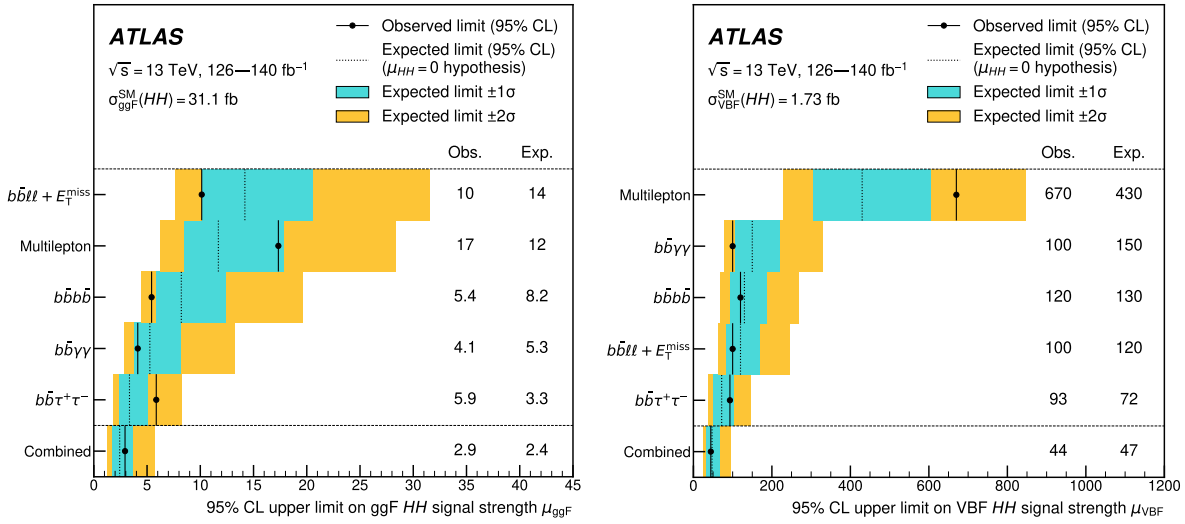

Figure A.1: Observed and expected 95% CL upper limits on the signal strength for the inclusive ggF HH (left) and VBF HH production (right), in the bb$\tau\tau$, bb$\gamma\gamma$, bbbb, multi-lepton and bb$\ell\ell$+$E_{\mathrm{T}}^{\mathrm{miss}}$ decay channels, and their statistical combination. The ggF or VBF HH production cross section is fixed to the SM predicted value for $m_{\mathrm{H}} = 125\,\mathrm{GeV}$ when deriving limits on the corresponding signal strength. The expected limits, along with the $\pm1\sigma$ and $\pm2\sigma$ bands, are calculated under the assumption of no HH process, and with all nuisances profiled to the observed data. Taken from [439].

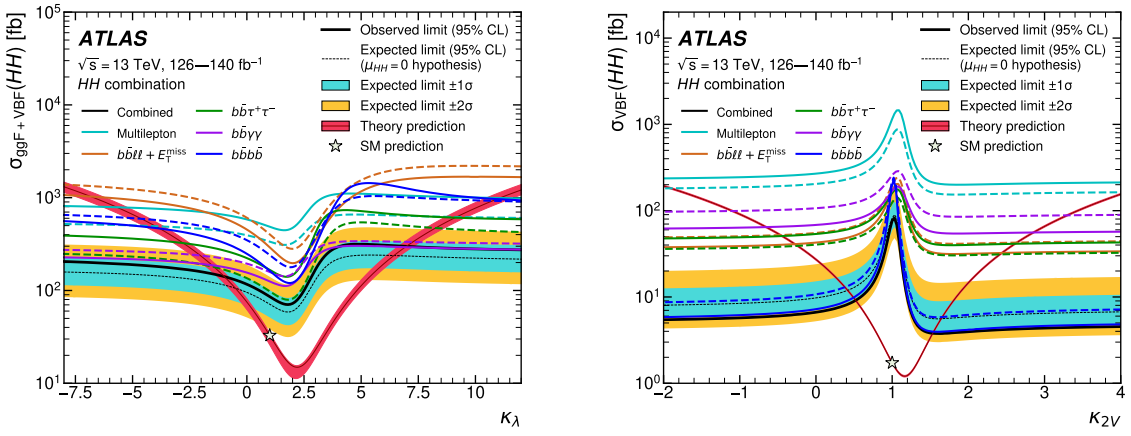

Figure A.2: Observed (solid lines) and expected (dashed lines) 95% CL exclusion limits on the HH production cross sections of the inclusive ggF and VBF processes as a function of κ_λ (left) and the VBF process as a function of $\kappa_{2\mathrm{V}}$ (right), for the bb$\gamma\gamma$ (purple), bb$\tau\tau$ (green), multi-lepton (cyan), bbbb (blue) and bb$\ell\ell$+$E_{\mathrm{T}}^{\mathrm{miss}}$ (brown) decay channels and their combination (black). The expected limits assume no HH production or no VBF HH production, respectively, for the left and right plots. In the right plot, the ggF HH production cross section is assumed to be as predicted by the SM. The red line shows the theory prediction for the ggF and VBF HH production cross section as a function of κ_λ (left), and the predicted VBF HH cross section as a function of $\kappa_{2\mathrm{V}}$ (right). The bands surrounding the red cross section lines indicate the theoretical uncertainties on the predicted cross sections. Taken from [439].

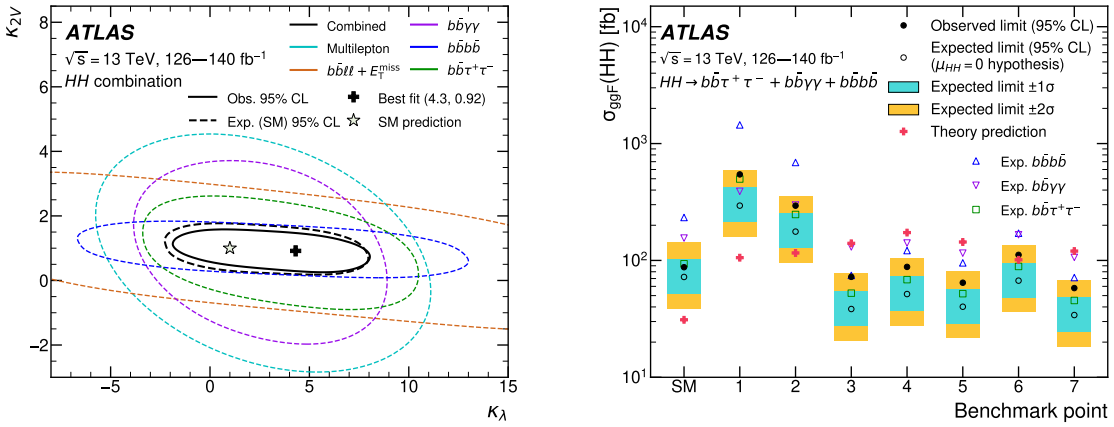

Figure A.3: (Left) Expected 95% CL contours in the κ_{2V}–κ_λ plane, corresponding to the individual decay channels and their combination, are illustrated using dashed lines. The observed contour from the combined results is depicted by a solid black line. The SM prediction is marked by a star, and the combined best-fit value is indicated by a cross. (Right) Observed and expected 95% CL combined upper limits on the cross section for the SM and seven BSM HH EFT benchmarks in the ggF process, describing representative signal kinematics and m_{HH} shape features obtained by varying multiple EFT coefficients. The expected limits from the bb$\tau\tau$, bb$\gamma\gamma$ and bbbb decay channels are presented as well. Theoretical predictions, estimated using specific sets of coefficient values defined in the benchmarks, are shown as red cross dots. Taken from [439].

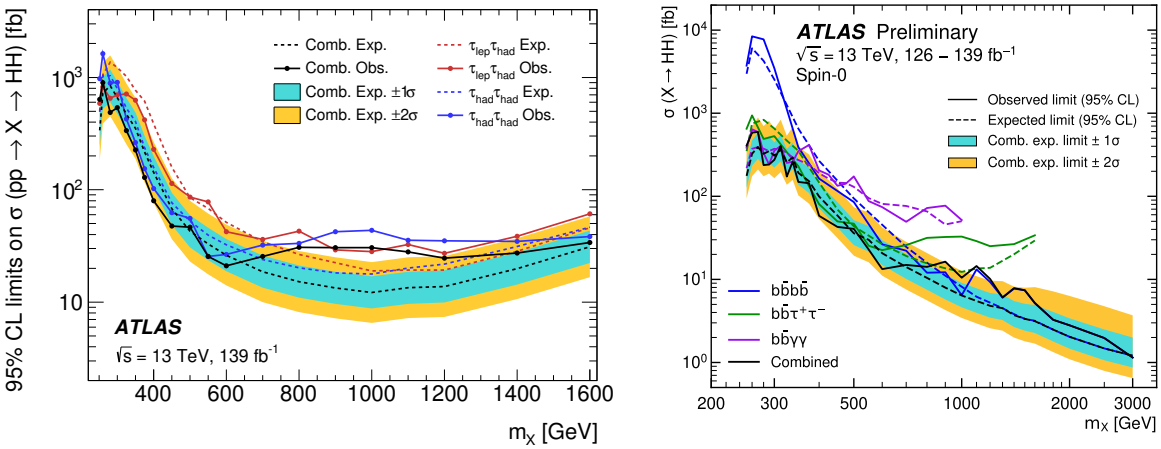

Figure A.4: Expected and observed 95% CL upper limits for the cross section of a scalar resonance decaying to two Higgs bosons, as a function of its mass m_{X}. (Left) bb$\tau\tau$ decay channel, exhibiting a moderate excess at 1 TeV with a local (global) significance of $3.1\,\sigma$ ($2.0\,\sigma$). (Right) The statistical combination of the bb$\gamma\gamma$, bb$\tau\tau$ and bbbb channels. Taken from [439] and [440].

A.2 Final Results in Table Format

m_X [GeV]	spin-0 ($\mathcal{X}$)	spin-2	ATLAS ($\mathcal{Y}$)	CMS Comb. ($\mathcal{Z}$)	$\mathcal{Y}/\mathcal{X}$	$\mathcal{Z}/\mathcal{X}$
250/251	$223.63^{+322.58}_{-155.93}$	$229.49^{+331.95}_{-160.89}$	$338.34^{+132.53}_{-94.55}$	–	**1.51**	–
260	$619.14^{+878.27}_{-438.53}$	$591.8^{+841.84}_{-419.94}$	$723.67^{+283.47}_{-202.23}$	$253.25^{+125.17}_{-81.00}$	**1.16**	0.41
270	$685.55^{+977.94}_{-483.91}$	$632.81^{+902.71}_{-446.68}$	–	$245.73^{+126.36}_{-81.59}$	–	0.36
280	$689.45^{+983.51}_{-486.67}$	$572.27^{+825.46}_{-401.19}$	$840.89^{+329.39}_{-234.98}$	$251.92^{+127.53}_{-82.42}$	**1.22**	0.37
300	$570.31^{+822.65}_{-399.82}$	$447.27^{+646.94}_{-310.77}$	$660.05^{+258.55}_{-184.45}$	$263.78^{+130.38}_{-84.11}$	**1.16**	0.46
320	$455.08^{+661.87}_{-316.84}$	$337.89^{+495.47}_{-233.62}$	–	$235.49^{+109.82}_{-73.60}$	–	0.52
325	–	–	$471.96^{+184.87}_{-131.89}$	–	–	–
350	$297.85^{+436.76}_{-205.94}$	$221.68^{+325.06}_{-152.4}$	$351.08^{+137.52}_{-98.11}$	$206.10^{+101.05}_{-64.91}$	**1.18**	0.69
375	–	–	$216.74^{+84.90}_{-60.57}$	–	–	–
400	$143.07^{+212.07}_{-97.66}$	$98.14^{+145.48}_{-67.23}$	$142.98^{+56.01}_{-39.96}$	$131.64^{+61.39}_{-40.82}$	1.00	0.92
450	$71.78^{+106.97}_{-48.65}$	$47.85^{+71.88}_{-32.2}$	$67.65^{+26.50}_{-18.91}$	$85.32^{+42.51}_{-27.50}$	0.94	**1.19**
500	$37.60^{+57.68}_{-24.99}$	$26.86^{+40.99}_{-17.77}$	$42.71^{+16.73}_{-11.93}$	$64.22^{+30.97}_{-20.54}$	**1.14**	**1.71**
550	$28.81^{+43.97}_{-19.06}$	$18.07^{+27.57}_{-11.95}$	$32.89^{+12.88}_{-9.19}$	$51.84^{+24.59}_{-16.20}$	**1.12**	**1.80**
600	$21.00^{+32.71}_{-13.89}$	$15.62^{+24.1}_{-10.44}$	$26.18^{+10.25}_{-7.32}$	$37.46^{+19.11}_{-12.35}$	**1.25**	**1.78**
650	$20.02^{+29.91}_{-13.25}$	$14.16^{+20.93}_{-8.99}$	–	–	–	–
700	$15.14^{+23.34}_{-10.36}$	$11.23^{+16.6}_{-7.87}$	$18.67^{+7.31}_{-5.22}$	$28.52^{+13.75}_{-9.05}$	**1.23**	**1.88**
750	$12.21^{+18.82}_{-8.56}$	$9.28^{+14.31}_{-6.51}$	–	–	–	–
800	$10.25^{+16.47}_{-7.19}$	$8.3^{+12.27}_{-5.27}$	$15.23^{+5.97}_{-4.26}$	$15.94^{+7.94}_{-5.20}$	–	**1.56**
850	$10.25^{+15.16}_{-6.51}$	$7.32^{+10.59}_{-4.49}$	–	–	**1.49**	–
900	$8.30^{+13.33}_{-5.82}$	$6.35^{+9.99}_{-4.12}$	$13.40^{+5.25}_{-3.75}$	$7.55^{+4.00}_{-2.48}$	**1.61**	0.91
1000	$6.35^{+10.8}_{-4.12}$	$5.37^{+8.45}_{-3.48}$	$12.19^{+4.78}_{-3.41}$	$4.76^{+2.47}_{-1.53}$	**1.92**	0.75
1100	–	–	$13.41^{+5.25}_{-3.75}$	$2.92^{+1.53}_{-0.94}$	–	–
1200	–	–	$13.75^{+5.44}_{-3.84}$	$1.96^{+1.02}_{-0.63}$	–	–
1250	$6.35^{+9.99}_{-4.12}$	$5.37^{+8.45}_{-3.48}$	–	–	–	–
1300	–	–	–	$1.51^{+0.79}_{-0.49}$	–	–
1400	–	–	$19.84^{+7.77}_{-5.54}$	$1.25^{+0.66}_{-0.41}$	–	–
1500	$8.30^{+12.27}_{-4.91}$	$8.3^{+12.8}_{-5.22}$	–	$1.02^{+0.56}_{-0.34}$	–	0.12
1600	–	–	$31.05^{+12.16}_{-8.68}$	$0.89^{+0.49}_{-0.30}$	–	–
1700	–	–	–	$0.79^{+0.45}_{-0.27}$	–	–
1750	$9.28^{+15.49}_{-5.89}$	$11.23^{+18.03}_{-7.48}$	–	–	–	–
1800	–	–	–	$0.68^{+0.40}_{-0.24}$	–	–
1900	–	–	–	$0.63^{+0.37}_{-0.22}$	–	–
2000	$12.21^{+20.38}_{-7.75}$	$16.11^{+26.65}_{-10.0}$	–	$0.60^{+0.35}_{-0.21}$	–	0.05
2200	–	–	–	$0.46^{+0.30}_{-0.17}$	–	–
2400	–	–	–	$0.38^{+0.26}_{-0.14}$	–	–
2500	$21.00^{+35.39}_{-13.41}$	$37.6^{+61.88}_{-23.48}$	–	–	–	–
2600	–	–	–	$0.35^{+0.24}_{-0.13}$	–	–
2800	–	–	–	$0.35^{+0.23}_{-0.13}$	–	–
3000	$33.69^{+55.45}_{-21.38}$	$61.52^{+102.23}_{-36.95}$	–	$0.32^{+0.23}_{-0.12}$	–	0.01

Table A.1: Nominal values of the cross section 95% CL upper limits shown in Section 5.4.1, in fb. The errors correspond to $1\,\sigma$ intervals. The combination (ATLAS result) was taken from Ref. [7] ([345]). Both refer to spin-0 hypotheses. The bb$\tau\tau$ result in the combination comes from Ref. [101]. The first row refers to 250 GeV for all measurements except for ATLAS, which considered 251 GeV. All values in bold show regions where our analysis demonstrates an improvement with respect to other analyses.

A.3 Fraction of Muons with Large Transverse Momentum

In Table A.2 we present the fractions of $\tau_\mu\tau_h$ events with $p_T(\mu) > 120\,\text{GeV}$. We include all radion signal samples and the two most important backgrounds, DY and $t\bar{t}$. The fraction is small for most samples, and its maximum value is $\sim$20 %.

Sample	**baseline**	**res1b**	**res2b**	**boosted**
250 GeV	0.001 ± 0.014	0.001 ± 0.023	0.001 ± 0.020	0.000 ± 1.000
260 GeV	0.002 ± 0.014	0.001 ± 0.023	0.002 ± 0.021	0.000 ± 0.715
270 GeV	0.004 ± 0.014	0.002 ± 0.023	0.001 ± 0.021	0.000 ± 1.000
280 GeV	0.005 ± 0.014	0.003 ± 0.023	0.004 ± 0.020	0.000 ± 1.000
300 GeV	0.007 ± 0.015	0.005 ± 0.024	0.004 ± 0.022	0.000 ± 0.503
320 GeV	0.010 ± 0.015	0.006 ± 0.024	0.007 ± 0.021	0.000 ± 0.318
350 GeV	0.017 ± 0.013	0.012 ± 0.022	0.016 ± 0.020	0.040 ± 0.192
400 GeV	0.025 ± 0.012	0.019 ± 0.020	0.019 ± 0.017	0.000 ± 0.157
450 GeV	0.043 ± 0.011	0.035 ± 0.018	0.037 ± 0.016	0.065 ± 0.087
500 GeV	0.061 ± 0.010	0.056 ± 0.018	0.054 ± 0.015	0.083 ± 0.064
550 GeV	0.070 ± 0.012	0.072 ± 0.020	0.064 ± 0.017	0.049 ± 0.056
600 GeV	0.091 ± 0.011	0.082 ± 0.020	0.086 ± 0.016	0.087 ± 0.037
650 GeV	0.106 ± 0.010	0.103 ± 0.021	0.103 ± 0.016	0.093 ± 0.027
700 GeV	0.118 ± 0.010	0.110 ± 0.021	0.111 ± 0.017	0.127 ± 0.020
750 GeV	0.133 ± 0.010	0.119 ± 0.022	0.125 ± 0.018	0.141 ± 0.017
800 GeV	0.155 ± 0.009	0.149 ± 0.022	0.140 ± 0.020	0.163 ± 0.015
850 GeV	0.157 ± 0.009	0.126 ± 0.024	0.148 ± 0.021	0.166 ± 0.013
900 GeV	0.169 ± 0.009	0.145 ± 0.026	0.166 ± 0.022	0.177 ± 0.013
1000 GeV	0.182 ± 0.012	0.150 ± 0.038	0.165 ± 0.034	0.193 ± 0.016
1250 GeV	0.219 ± 0.013	0.185 ± 0.046	0.195 ± 0.041	0.235 ± 0.015
1500 GeV	0.205 ± 0.016	0.188 ± 0.062	0.202 ± 0.052	0.205 ± 0.019
1750 GeV	0.190 ± 0.019	0.170 ± 0.067	0.180 ± 0.060	0.195 ± 0.022
2000 GeV	0.189 ± 0.022	0.159 ± 0.084	0.218 ± 0.087	0.188 ± 0.026
2500 GeV	0.178 ± 0.030	0.201 ± 0.120	0.194 ± 0.120	0.188 ± 0.035
3000 GeV	0.149 ± 0.040	0.093 ± 0.175	0.065 ± 0.186	0.165 ± 0.048
DY	0.013 ± 0.005	0.011 ± 0.017	0.012 ± 0.045	0.063 ± 0.174
TT	0.068 ± 0.001	0.023 ± 0.001	0.018 ± 0.001	0.157 ± 0.014

Table A.2: Fraction of events in the $\tau_\mu\tau_h$ channel where the first lepton, a muon, has a p_T above 120 GeV, for the three analysis categories and for the **baseline** selection. The selections and categories are detailed in Section 4.4. The signal samples refer to the radion spin-0 hypothesis.

A.4 Additional Plots on MET Scale Factors

In Figs. A.5 and A.6, we show the data/MC curves coming from the $\mathrm{MET}_{\text{no-}\mu}$ trigger efficiency fits, for the four Run 2 data-taking periods. The methods to derive the $\mathrm{MET}_{\text{no-}\mu}$ SFs are explained in Section 4.7.4. Fig. A.5 shows the result of testing five different fits in order to find the ideal fitting range. We tried to reach a balance between domain size and fit quality. We chose the fit starting at 150 GeV. Fig. A.6 shows a comparison of the same quantity between the $\tau_\mu\tau_\mathrm{h}$ and $\tau_\mu\tau_\mu$ channels, for all Run 2 data-taking years. The pairs of distributions are always found to be compatible, validating the $\tau_\mu\tau_\mu$ channel in this context.

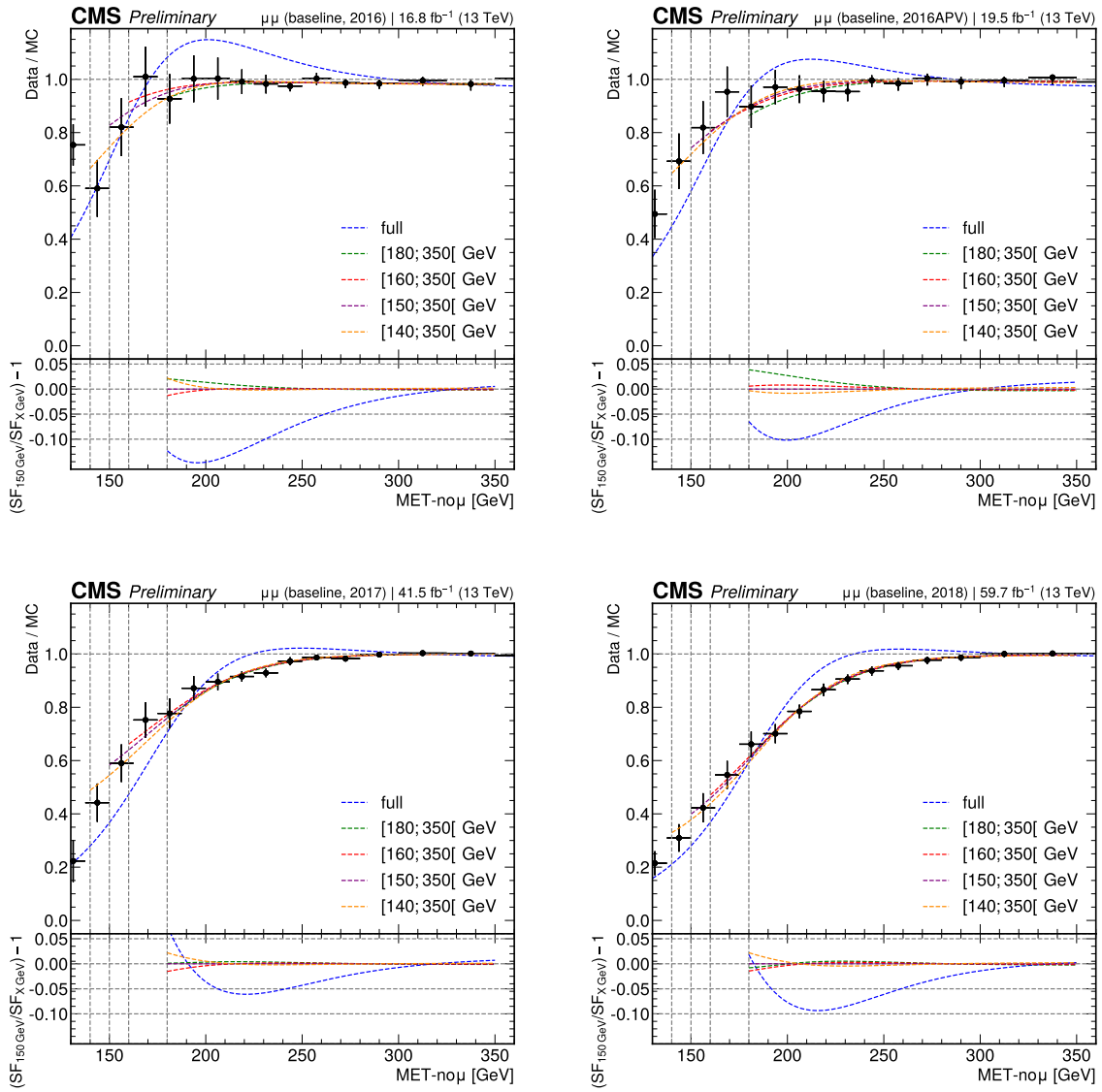

Figure A.5: Data/MC SFs of $\text{MET}_{\text{no-}\mu}$ trigger efficiencies, for the four Run 2 data-taking periods. SFs are extracted from the ratio of the sigmoid fits of data and MC efficiency curves, implemented to smooth out the SFs' distributions. Five different fit ranges were tested, and zoomed in the turn-on region to better display differences. All fits are reasonably compatible except for the full range fit, which cannot describe the data. We decided to use the fit starting at 150 GeV for all data periods.

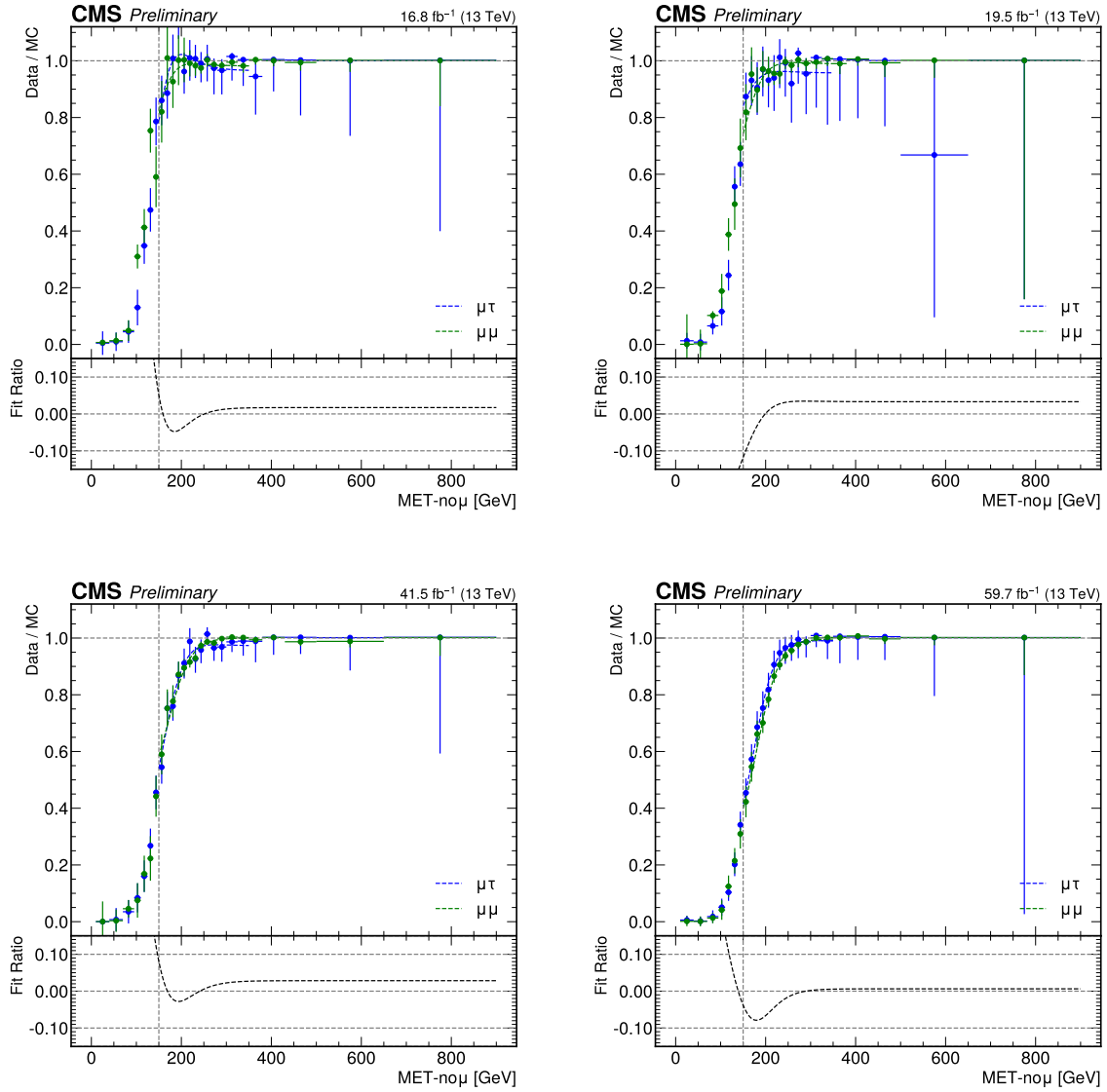

Figure A.6: Data/MC SFs of $\mathrm{MET}_{\text{no-}\mu}$ trigger efficiencies, for the four Run 2 data-taking periods. SFs are extracted from the ratio of the sigmoid fits of data and MC efficiency curves, implemented to smooth out the SFs' distributions. The SFs are observed to be compatible between the $\tau_\mu\tau_\mathrm{h}$ and $\tau_\mu\tau_\mu$ channels, within statistical uncertainties.

A.5 B, C and D distributions for the ABCD Method

In Section 4.6.1 the data-driven ABCD method to estimate the multijet background was presented. For completeness, we here provide some examples of data and MC distributions in the B, C and D regions. Those quantities are the p_{T} and η of the first tau lepton and of the first b jet. The distributions are shown in Figs. A.7 and A.8, for the **res1b** and **res2b** categories, respectively. We display the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel, where the multijet background is more significant. The difference between the data and the total MC distribution corresponds to the multijet background in each of those regions. The statistical uncertainties of the C region (middle) are smaller than for the B region (left), indicating that using the C region for the shape in the ABCD method leads to overall smaller statistical uncertainties.

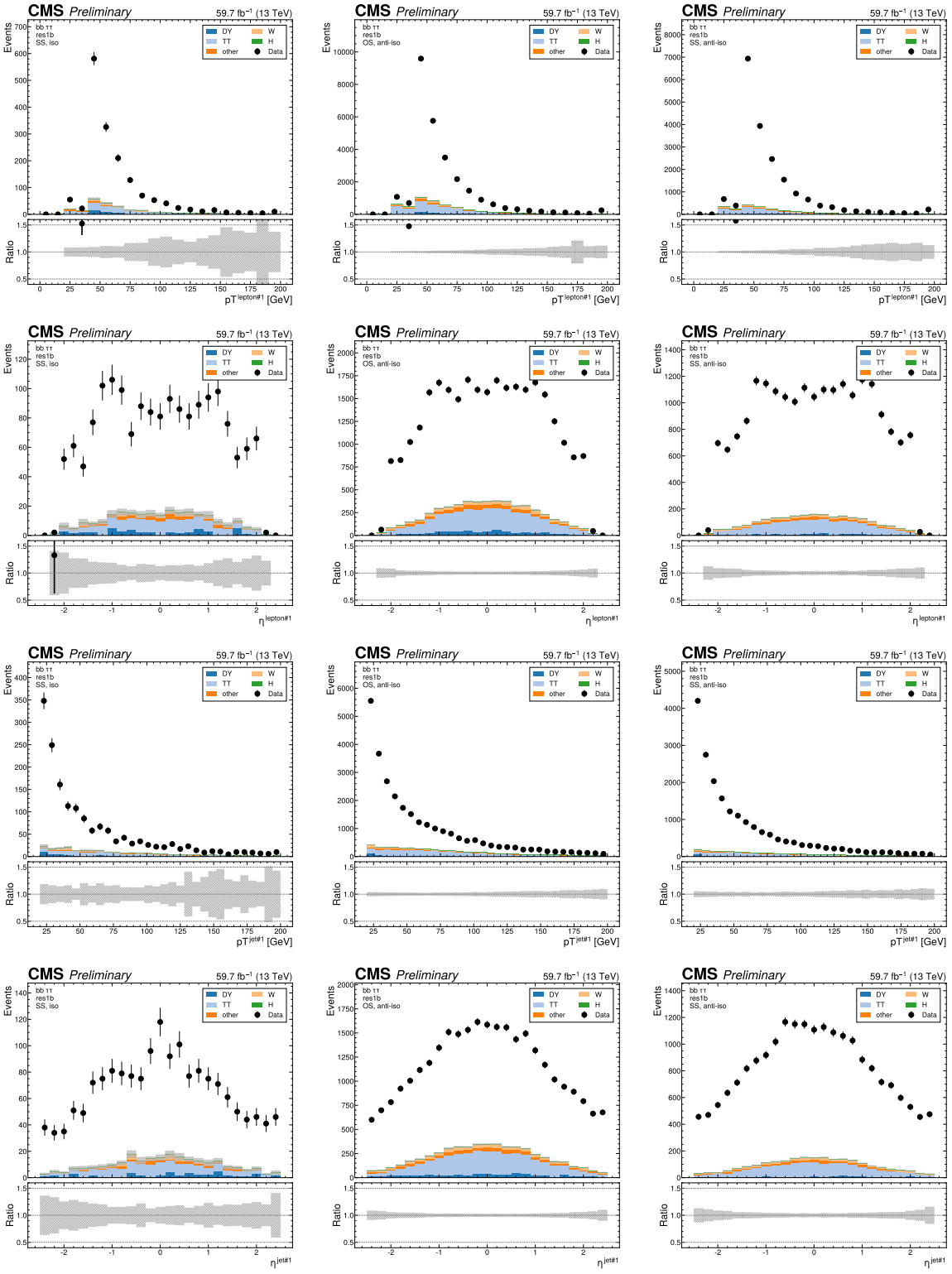

Figure A.7: Control plots for the multijet background, in the **res1b** category, for the three regions of the ABCD method used to estimate the multijet normalization and shape in the SR. The B, C, and D regions are shown in the left, middle and right columns, respectively, corresponding to SS/iso, OS/anti-iso and SS/anti-iso. Each row shows the distributions of a different variable, from top to bottom: the p_{T} of the first tau lepton, the η of the same lepton, the p_{T} of the first b-jet, and the η of the same lepton.

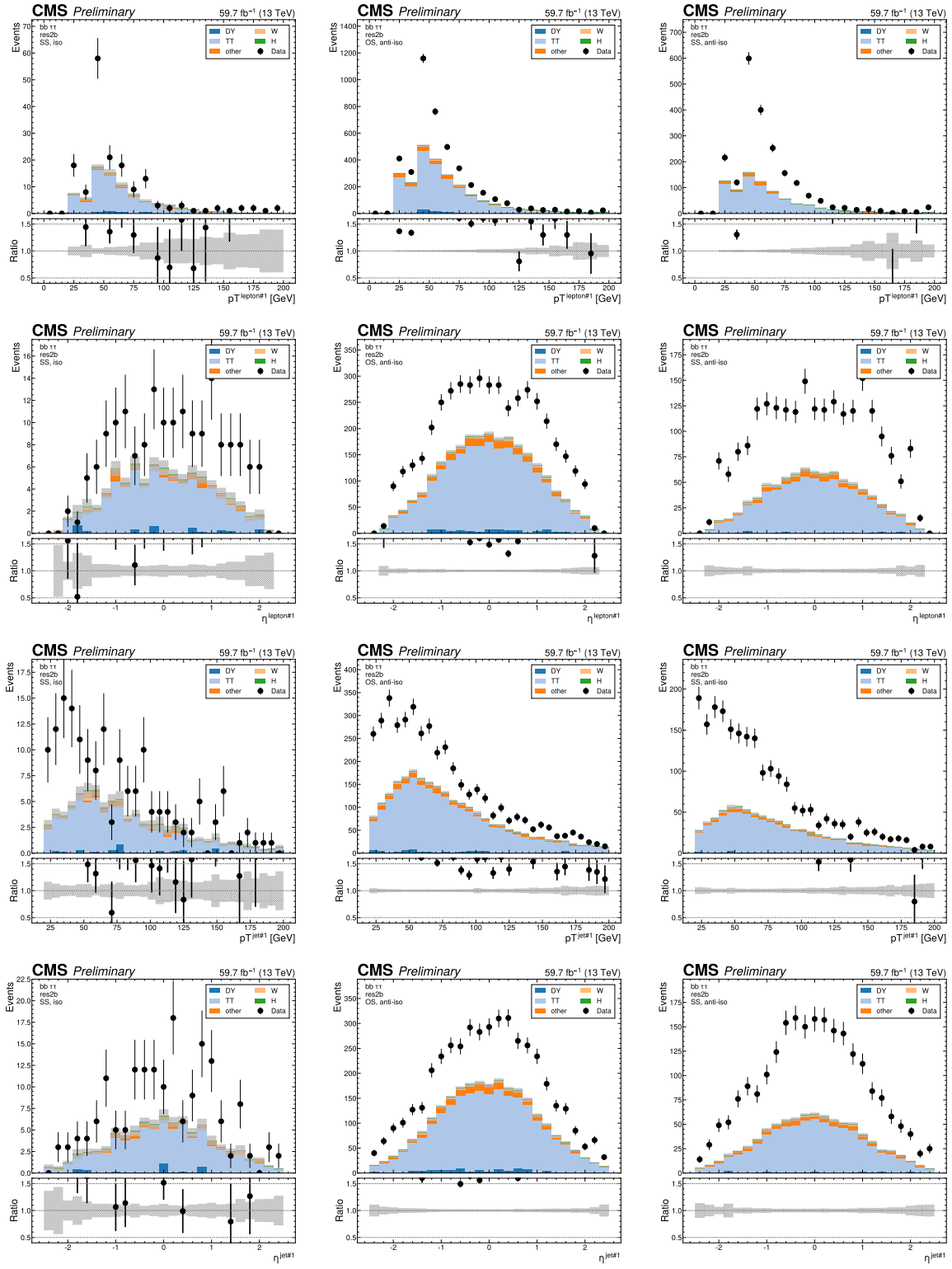

Figure A.8: Control plots for the multijet background, in the **res2b** category, for the three regions of the ABCD method used to estimate the multijet normalization and shape in the SR. The B, C, and D regions are shown in the left, middle and right columns, respectively, corresponding to SS/iso, OS/anti-iso and SS/anti-iso. Each row shows the distributions of a different variable, from top to bottom: the p_{T} of the first tau lepton, the η of the same lepton, the p_{T} of the first b-jet, and the η of the same lepton.

B List of acronyms

2HDM Two-Higgs Doublet Model . . . 28
ADC Analog-to-digital converter . . . 129
ELU Exponential Linear Unit . . . 188
AFP ATLAS Forward Proton Project . . . 104
SVFit Secondary Vertex Fit . . . 187
AHCAL Analogue Hadron Calorimeter . . . 123
ALICE A Large Ion Collider Experiment . . . 59
APD Avalanche Photodiode . . . 70
ASIC application-specific integrated circuit . . . 128
ATCA Advanced Telecommunications Computing Architecture . . . 127
ATCA Advanced Telecommunications Computing Architecture . . . 127
ATLAS A Toroidal LHC Apparatus . . . 5
BBR Beam-Beam Remnant . . . 97
BC Best-Choice . . . 131
BDT Boosted Decision Tree . . . 91
BE back-end . . . 110
BMTF Barrel Muon Track Finder . . . 80
BPIX barrel pixel . . . 69
BRIL Beam Radiation, Instrumentation, and Luminosity . . . 111
BR branching ratio . . . 31
BSM Beyond the Standard Model . . . 5
BTL Barrel Timing Layer . . . 108
BU Builder Unit . . . 81
BX bunch-crossing . . . 70
CDF Collider Detector at Femilab . . . 97
CE-E silicon electromagnetic calorimeter . . . 118
CE-H silicon hadronic calorimeter . . . 121
CEP Central Exclusive Production . . . 105
CKM Cabibbo-Kobayashi-Maskawa . . . 16
CL confidence level . . . 19
CMB Cosmic Microwave Background . . . 29
CMSSW CMS Software . . . 81

PRL profile likelihood ratio . . . 232
CMS Compact Muon Solenoid . . . 5
CNN Convolutional Neural Network . . . 177
DL deep learning . . . 217
LBN Lorentz Boost Network . . . 219
CPU Central Processing Unit . . . 81
CP charge-parity . . . 28
CR control region . . . 153
CSC Cathode Strip Chamber . . . 74
CTPPS CMS TOTEM Precision Proton Spectrometer . . . 104
CT Correlator Trigger . . . 116
D2S data-to-surface . . . 129
DAQ data acquisition . . . 79
DCS detector control system . . . 127
DM dark matter . . . 61
DNN Deep Neural Network . . . 156
DPS Double Parton Scattering . . . 104
DSS detector safety system . . . 127
DTH400 DAQ and TCDS Hub . . . 129
DT Drift Tube . . . 74
DY Drell-Yan . . . 94
EB ECAL Barrel . . . 70
ECAL electromagnetic calorimeter . . . 69
ECON-D DAQ concentrator chip . . . 130
ECON-T TPG concentrator chip . . . 130
ECON concentrator chip . . . 131
ED extra dimension . . . 37
EE ECAL Endcap . . . 70
EFT Effective Field Theory . . . 33
EMF endcap muon trigger finder . . . 128
EMTF Endcap Muon Track Finder . . . 80
EM electromagnetic . . . 6
ETL Endcap Timing Layer . . . 108
ETL Endcap Timing Layer . . . 108
EWSB electroweak symmetry breaking . . . 5
EW electroweak . . . 5
FASER ForwArd Search ExpeRiment . . . 62

HP High Purity . . . 178
HSCP heavy stable charged particles . . . 116
I2C inter-integrated circuit . . . 129
ID Inner Detector . . . 59
IP Interaction Point . . . 53
ISR initial state radiation . . . 98
IT Inner Tracker . . . 107
JEC jet energy correction . . . 178
JER jet energy resolution . . . 179
JES jet energy scale . . . 225
KF Kalman Filter . . . 86
L1A L1-Accept . . . 79
L1 Level-1 . . . 77
LEP Large Electron–Positron Collider . . . 34
LFV Lepton Flavour Violation . . . 21
LHCb LHC beauty . . . 59
LHCf LHC forward . . . 63
LHC Large Hadron Collider . . . 16
LLP long-lived particle . . . 70
LLR Leprince Ringuet Laboratory . . . 136
LO leading order . . . 31
LP Low Purity . . . 178
LS Long Shutdown . . . 57
LUCID Luminosity Cherenkov Integrating Detector . . . 125
LUT Look-Up Table . . . 144
Linac4 Linear Accelerator . . . 54
MAPP MoEDAL Apparatus for Penetrating Particles . . . 64
MB Minimum Bias . . . 98
MC Monte Carlo . . . 96
ME0 Muon Endcap 0 . . . 111
MET missing transverse energy . . . 22
MFT Muon Forward Tracker . . . 126
MIP minimum ionizing particle . . . 75
ML Machine Learning . . . 36
MOND modified newtonian dynamics . . . 29
MPI multiple-parton interaction . . . 97
MP Medium Purity . . . 178

MSSM Minimal Supersymmetric Standard Model . . . 38
MTD MIP Timing Detector . . . 107
MVA multivariate analysis . . . 174
MoEDAL Monopole & Exotics Detector At the LHC . . . 64
N2HDM next-to-minimal 2HDM . . . 41
NLO next-to-leading order . . . 31
NMSSM next-to-minimal MSSM . . . 41
NNLO next-to-next-to-leading order . . . 24
NN Neural Network . . . 116
NP New Physics . . . 5
NWA Narrow Width Approximation . . . 37
NZS non zero-suppression . . . 128
NbTi Niobium-Titanium . . . 55
OMTF Overlap Muon Track Finder . . . 80
OT Outer Tracker . . . 107
PCB printed circuit board . . . 118
PDF parton distribution function . . . 31
PD Primary Dataset . . . 81
PF Particle Flow . . . 84
PID particle identification . . . 62
PLL phase-locked loop . . . 129
PMNS Pontecorvo-Maki-Nakagawa-Sakata . . . 7
PMT photomultiplier tube . . . 60
PM Polyethylene Moderator . . . 122
PNet Particle Net . . . 83
POG Physics Object Group . . . 173
POI Parameter of Interest . . . 221
PS Proton Synchrotron . . . 54
PUPPI Pileup Per Particle Identification . . . 115
PU pile-up . . . 56
PV primary vertex . . . 56
PaaS platform-as-a-service . . . 136
ParT Particle Transformer . . . 177
QCD Quantum Chromodynamics . . . 6
QED Quantum Electrodynamics . . . 10
QGP Quark Gluon Plasma . . . 62
RICH Ring-imaging Cherenkov detector . . . 61

RNN Recurrent Neural Network . . . 177
RPC Resistive Plate Chamber . . . 60
RP Roman Pot . . . 64
RTL Register-transfer level . . . 152
RU Readout Unit . . . 108
S1 Stage 1 . . . 132
S2I source-to-image . . . 136
S2 Stage 2 . . . 132
SCT Semiconductor Tracker . . . 59
SEE single event effect . . . 107
SFO-EWPT strong first-order electroweak phase transition . . . 244
SF scale factor . . . 154
SL superlayer . . . 75
SMEFT Standard Model EFT . . . 34
SMOG System for Measuring the Overlap with Gas . . . 61
SM Standard Model . . . v
SND@LHC Scattering and Neutrino Detector at the LHC . . . 63
SPS Super Proton Synchrotron . . . 54
SR signal region . . . 185
SS stainless steel . . . 120
STC super trigger cell . . . 131
SUSY Supersymmetry . . . 5
Sci scintillator . . . 60
SiPM silicon photomultiplier . . . 73
Si silicon . . . 59
SoA Structure of Arrays . . . 118
TCDS time and control distribution system . . . 129
TC trigger cell . . . 121
TDAQ trigger and data acquisition . . . 127
TDC time-to-digital converter . . . 129
TEC Tracker EndCap . . . 68
TEC Tracker EndCap . . . 68
TES Tau Energy Scale . . . 198
TIB Tracker Inner Barrel . . . 68
TICL The Iterative CLustering algorithm . . . 117
TID Tracker Inner Disk . . . 68
TMT Time Multiplexed Trigger . . . 80

Bibliography

[1] The 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: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020 (cited on pages 1, 5, 17, 57).

[2] The 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 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021 (cited on pages 1, 5, 17, 57).

[3] The CMS Collaboration. "Observation of a new boson with mass near 125 GeV in *pp* collisions at $\sqrt{s} = 7$ and 8 TeV". In: *Journal of High Energy Physics* 2013.6 (June 2013). ISSN: 1029-8479. DOI: 10.1007/jhep06(2013)081. URL: http://dx.doi.org/10.1007/JHEP06(2013)081 (cited on pages 1, 5, 17, 57).

[4] Bruno Alves. *HH searches: CMS HH experimental overview*. Accessed: 21 September 2024. Nov. 2023. URL: https://indico.cern.ch/event/1276727/contributions/5643184/ (cited on page 1).

[5] Bruno Alves. *Di-Higgs searches at CMS*. Tech. rep. Accepted for publication by PoS. Geneva: CERN, Apr. 2024. URL: https://cds.cern.ch/record/2901955 (cited on pages 1, 27).

[6] Bruno Alves. *Techniques for SMEFT fit*. Talk at QCD@LHC2022, IJCLab Orsay, France. Accessed: 20 August 2024. Dec. 2022. URL: https://indico.cern.ch/event/1150707/contributions/5150450/ (cited on pages 1, 34).

[7] The CMS Collaboration. *Searches for Higgs boson production through decays of heavy resonances*. Accepted for publication at Phys. Rept. 2024. arXiv: 2403.16926 [hep-ex]. URL: https://arxiv.org/abs/2403.16926 (cited on pages 2, 40, 42, 241, 244, 251–255, 264).

[8] Bruno Alves. "Cluster reconstruction in the HGCAL at the Level 1 trigger". In: *EPJ Web of Conferences* 295 (May 2024), S02006. DOI: 10.1051/epjconf/202429502006. URL: https://doi.org/10.1051/epjconf/202429502006 (cited on pages 2, 127, 130, 133, 144).

[9] Bruno Alves and Marco Chiusi. "Standalone framework for the emulation of HGCAL firmware trigger primitives in the CMS online trigger system". In: *Zenodo* (Oct. 2023). DOI: 10.5281/zenodo.10001296. URL: https://doi.org/10.5281/zenodo.10001296 (cited on pages 2, 136, 140–142).

[10] The CDF Collaboration. "Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab". In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2626 (cited on page 5).

[11] The D0 Collaboration. "Observation of the Top Quark". In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2632 (cited on page 5).

[12] Eite Tiesinga et al. "CODATA recommended values of the fundamental physical constants: 2018*". In: *Rev. Mod. Phys.* 93.2 (2021), p. 025010. DOI: 10.1103/RevModPhys.93.025010 (cited on page 7).

[13] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097 (cited on pages 7, 16, 75, 92, 93, 134, 153, 243).

[14] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. eprint: https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf. URL: https://doi.org/10.1143/PTP.28.870 (cited on page 7).

[15] K.S. Hirata et al. "Experimental study of the atmospheric neutrino flux". In: *Physics Letters B* 205.2 (1988), pp. 416–420. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(88)91690-5. URL: https://www.sciencedirect.com/science/article/pii/0370269388916905 (cited on page 7).

[16] A. Quadt. "Top quark physics at hadron colliders". In: *The European Physical Journal C - Particles and Fields* 48.3 (Dec. 2006), pp. 835–1000. DOI: 10.1140/epjc/s2006-02631-6. URL: https://dx.doi.org/10.1140/epjc/s2006-02631-6 (cited on page 7).

[17] Neutelings, Izaak. *CMS coordinate system – TikZ.net*. Accessed: 18 May 2024. 2021. URL: https://tikz.net/axis3d_cms/ (cited on pages 8, 65, 67).

[18] Emmy Noether. "Invariant variation problems". In: *Transport Theory and Statistical Physics* 1.3 (Jan. 1971), pp. 186–207. ISSN: 1532-2424. DOI: 10.1080/00411457108231446. URL: http://dx.doi.org/10.1080/00411457108231446 (cited on page 7).

[19] Francis Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., 1984. ISBN: 0-471-88741-2 (cited on pages 8, 9).

[20] David Griffiths. *Introduction to Elementary Particles*. Second, Revised Edition. WILEY-VCH Verlag GmbH & Co. KGaA, 2008. ISBN: 978-3-527-40601-2 (cited on pages 8, 13, 29).

[21] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413 (cited on page 10).

[22] David E. Morrissey and Michael J. Ramsey-Musolf. “Electroweak baryogenesis”. In: *New Journal of Physics* 14.12 (Dec. 2012), p. 125003. DOI: 10.1088/1367-2630/14/12/125003. URL: https://dx.doi.org/10.1088/1367-2630/14/12/125003 (cited on page 13).

[23] Fedor Bezrukov and Mikhail Shaposhnikov. “The Standard Model Higgs boson as the inflaton”. In: *Physics Letters B* 659.3 (2008), pp. 703–706. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2007.11.072. URL: https://www.sciencedirect.com/science/article/pii/S0370269307014128 (cited on page 13).

[24] Min Deng, Jinhan Yu, and Donna G. Blackmond. “Symmetry breaking and chiral amplification in prebiotic ligation reactions”. en. In: *Nature* 626.1019 (Feb. 2024). ISSN: 1476-4687. DOI: 10.1038/s41586-024-07059-y. URL: https://doi.org/10.1038/s41586-024-07059-y (cited on page 13).

[25] William D. Piñeros and Tsvi Tlusty. “Spontaneous chiral symmetry breaking in a random driven chemical system”. en. In: *Nature Communications* 13.2244 (Apr. 2022). ISSN: 2041-1723. DOI: 10.1038/s41467-022-29952-8. URL: https://doi.org/10.1038/s41467-022-29952-8 (cited on page 13).

[26] Aron Beekman, Louk Rademaker, and Jasper van Wezel. “An introduction to spontaneous symmetry breaking”. In: *SciPost Physics Lecture Notes* (Dec. 2019). ISSN: 2590-1990. DOI: 10.21468/scipostphyslectnotes.11. URL: http://dx.doi.org/10.21468/SciPostPhysLectNotes.11 (cited on page 13).

[27] Gabriel Raya and Luca Ambrogioni. *Spontaneous Symmetry Breaking in Generative Diffusion Models*. 2023. arXiv: 2305.19693 [cs.LG]. URL: https://arxiv.org/abs/2305.19693 (cited on page 13).

[28] J. Goldstone. “Field Theories with Superconductor Solutions”. In: *Nuovo Cim.* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722 (cited on page 14).

[29] Simone Amoroso et al. “Compatibility and combination of world W-boson mass measurements”. In: *Eur. Phys. J. C* 84.5 (2024), p. 451. DOI: 10.1140/epjc/s10052-024-12532-z. arXiv: 2308.09417 [hep-ex] (cited on page 15).

[30] The ALEPH Collaboration. “Precision electroweak measurements on the Z resonance”. In: *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008 (cited on page 15).

[31] The CDF Collaboration. “High-precision measurement of the W boson mass with the CDF II detector”. In: *Science* 376.6589 (2022), pp. 170–176. DOI: 10.1126/science.abk1781 (cited on page 15).

[32] The CMS Collaboration. *A measurement of the W boson mass in proton-proton collisions at* $\sqrt{s} = 13\,\text{TeV}$. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2910372 (cited on page 15).

[33] G. ’t Hooft. “Renormalizable Lagrangians for massive Yang-Mills fields”. In: *Nuclear Physics B* 35.1 (1971), pp. 167–188. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(71)90139-8. URL: https://www.sciencedirect.com/science/article/pii/0550321371901398 (cited on page 15).

[34] Benjamin W. Lee, C. Quigg, and H. B. Thacker. "Weak interactions at very high energies: The role of the Higgs-boson mass". In: *Phys. Rev. D* 16 (5 Sept. 1977), pp. 1519–1531. DOI: 10.1103/PhysRevD.16.1519. URL: https://link.aps.org/doi/10.1103/PhysRevD.16.1519 (cited on page 17).

[35] M. J. G. Veltman. "Second Threshold in Weak Interactions". In: *Acta Phys. Pol. B* 8.6-7 (1977), pp. 475–492. URL: https://cds.cern.ch/record/2792151 (cited on page 17).

[36] N. Cabibbo et al. "Bounds on the fermions and Higgs boson masses in grand unified theories". In: *Nuclear Physics B* 158.2 (1979), pp. 295–305. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(79)90167-6. URL: https://www.sciencedirect.com/science/article/pii/0550321379901676 (cited on page 17).

[37] David d'Enterria. *On the Gaussian maximum at a mass m_H 125 GeV of the product of decay probabilities of the Standard Model Higgs boson*. 2014. arXiv: 1208.1993 [hep-ph]. URL: https://arxiv.org/abs/1208.1993 (cited on pages 17, 20).

[38] The CMS Collaboration. "A portrait of the Higgs boson by the CMS experiment ten years after the discovery". en. In: *Nature* 607.7917 (July 2022), pp. 60–68. ISSN: 0028-0836. DOI: 10.1038/s41586-022-04892-x. URL: https://www.nature.com/articles/s41586-022-04892-x (cited on pages 17, 22, 47, 48, 50, 105, 106, 241, 249).

[39] The ATLAS Collaboration. "A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery". en. In: *Nature* 607.7917 (July 2022), pp. 52–59. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-022-04893-w. URL: https://www.nature.com/articles/s41586-022-04893-w (cited on page 17).

[40] The CMS Collaboration. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV". In: *The European Physical Journal C* 75.5 (May 2015). DOI: 10.1140/epjc/s10052-015-3351-7. URL: https://doi.org/10.1140%5C%2Fepjc%5C%2Fs10052-015-3351-7 (cited on page 17).

[41] The ATLAS Collaboration. "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment". In: *The European Physical Journal C* 76.1 (Jan. 2016). DOI: 10.1140/epjc/s10052-015-3769-y. URL: https://doi.org/10.1140%5C%2Fepjc%5C%2Fs10052-015-3769-y (cited on page 17).

[42] The CMS Collaboration. *Measurement of the Higgs boson mass and width using the four leptons final state*. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2871702 (cited on pages 17, 18, 243).

[43] The ATLAS Collaboration. "Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4l$ decay channel using 139 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions recorded by the ATLAS detector at the LHC". In: *Physics Letters B* 843 (Aug. 2023), p. 137880. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2023.137880. URL: http://dx.doi.org/10.1016/j.physletb.2023.137880 (cited on page 17).

[44] The ATLAS Collaboration. "Combined Measurement of the Higgs Boson Mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ Decay Channels with the ATLAS Detector Using $\sqrt{s} = 7$, 8, and 13 TeV pp Collision Data". In: *Phys. Rev. Lett.* 131.25 (2023), p. 251802. DOI: 10.1103/PhysRevLett.131.251802. arXiv: 2308.04775 [hep-ex] (cited on pages 18, 243).

[45] Fabrizio Caola and Kirill Melnikov. "Constraining the Higgs boson width with ZZ production at the LHC". In: *Phys. Rev. D* 88 (5 Sept. 2013), p. 054024. DOI: 10.1103/PhysRevD.88.054024. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.054024 (cited on page 18).

[46] John M. Campbell, R. Keith Ellis, and Ciaran Williams. "Bounding the Higgs width at the LHC using full analytic results for $gg \to e^-e^+\mu^-\mu^+$". In: *Journal of High Energy Physics* 2014.4 (2014). ISSN: 1029-8479. DOI: 10.1007/JHEP04(2014)060. URL: https://doi.org/10.1007/JHEP04(2014)060 (cited on page 18).

[47] Nikolas Kauer and Giampiero Passarino. "Inadequacy of zero-width approximation for a light Higgs boson signal". In: *Journal of High Energy Physics* 2012.8 (Aug. 2012). ISSN: 1029-8479. DOI: 10.1007/jhep08(2012)116. URL: http://dx.doi.org/10.1007/JHEP08(2012)116 (cited on page 18).

[48] The ATLAS Collaboration. "Evidence of off-shell Higgs boson production from ZZ leptonic decay channels and constraints on its total width with the ATLAS detector". In: *Phys. Lett. B* 846 (2023), p. 138223. DOI: 10.1016/j.physletb.2023.138223. arXiv: 2304.01532 [hep-ex] (cited on page 18).

[49] D. de Florian et al. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. Geneva: CERN, 2017. DOI: 10.23731/CYRM-2017-002. URL: https://cds.cern.ch/record/2227475 (cited on pages 19, 20, 22, 225, 243, 247).

[50] The CMS Collaboration. *Search for the SM Higgs boson produced in association with bottom quarks in final states with leptons*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2892701 (cited on page 19).

[51] Kathrin Becker et al. *Recommended predictions for the boosted-Higgs cross section*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2669113 (cited on page 20).

[52] The ATLAS Collaboration. "Direct constraint on the Higgs–charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector". In: *The European Physical Journal C* 82.8 (2022). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-022-10588-3. URL: https://doi.org/10.1140/epjc/s10052-022-10588-3 (cited on page 20).

[53] The ATLAS and CMS Collaborations. "Evidence for the Higgs Boson Decay to a Z Boson and a Photon at the LHC". In: *Physical Review Letters* 132.2 (Jan. 2024). ISSN: 1079-7114. DOI: 10.1103/physrevlett.132.021803. URL: http://dx.doi.org/10.1103/PhysRevLett.132.021803 (cited on page 21).

[54] The CMS Collaboration. "Search for the Higgs boson decay to a pair of electrons in proton-proton collisions at s=13TeV". In: *Phys. Lett. B* 846 (2023), p. 137783. DOI: 10.1016/j.physletb.2023.137783. arXiv: 2208.00265 [hep-ex] (cited on page 21).

[55] The ATLAS Collaboration. "Search for the Higgs boson decays H $\to ee$ and H $\to e\mu$ in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Physics Letters B* 801 (Feb. 2020), p. 135148. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2019.135148. URL: http://dx.doi.org/10.1016/j.physletb.2019.135148 (cited on page 21).

[56] David d'Enterria, Andres Poldaru, and George Wojcik. "Measuring the electron Yukawa coupling via resonant s-channel Higgs production at FCC-ee". In: *The European Physical Journal Plus* 137.2 (Feb. 2022). ISSN: 2190-5444. DOI: 10.1140/epjp/s13360-021-02204-2. URL: http://dx.doi.org/10.1140/epjp/s13360-021-02204-2 (cited on page 21).

[57] The CMS Collaboration. "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector". In: *Physics Letters B* 779 (Apr. 2018), pp. 283–316. DOI: 10.1016/j.physletb.2018.02.004. URL: https://doi.org/10.1016%2Fj.physletb.2018.02.004 (cited on pages 21, 173, 174).

[58] The CMS Collaboration. "Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 83.7 (2023), p. 562. DOI: 10.1140/epjc/s10052-023-11452-8. arXiv: 2204.12957. URL: https://cds.cern.ch/record/2807752 (cited on pages 21, 173, 174).

[59] The CMS Collaboration. "Analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of High Energy Physics* 2022.6 (June 2022). DOI: 10.1007/jhep06(2022)012. URL: https://doi.org/10.1007%2Fjhep06%282022%29012 (cited on pages 21, 173, 174).

[60] The CMS Collaboration. "Evidence for Higgs boson decay to a pair of muons". In: *Journal of High Energy Physics* 2021.1 (Jan. 2021). DOI: 10.1007/jhep01(2021)148 (cited on page 21).

[61] R. Frederix et al. "Higgs pair production at the LHC with NLO and parton-shower effects". In: *Physics Letters B* 732 (2014), pp. 142–149. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2014.03.026. URL: https://www.sciencedirect.com/science/article/pii/S0370269314001828 (cited on pages 23, 31).

[62] The CMS Collaboration. *LHC Higgs WG4 group*. Twiki. Accessed: 2024-04-17. 2024. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGHH (cited on pages 24, 160).

[63] Andrea Dainese et al. *Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC*. Tech. rep. Geneva, Switzerland, 2019. DOI: 10.23731/CYRM-2019-007. URL: https://cds.cern.ch/record/2703572 (cited on pages 23, 99, 245).

[64] F. Maltoni, E. Vryonidou, and M. Zaro. "Top-quark mass effects in double and triple Higgs production in gluon-gluon fusion at NLO". In: *Journal of High Energy Physics* 2014.11 (Nov. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep11(2014)079. URL: http://dx.doi.org/10.1007/JHEP11(2014)079 (cited on page 24).

[65] D. Rainwater and D. Zeppenfeld. "Observing $\vec{H}\ W^{(*)}W^{(*)} \to e^{\pm}\mu^{\mp}p_T$ in weak boson fusion with dual forward jet tagging at the CERN LHC". In: *Phys. Rev. D* 60 (11 Nov. 1999), p. 113004. DOI: 10.1103/PhysRevD.60.113004. URL: https://link.aps.org/doi/10.1103/PhysRevD.60.113004 (cited on page 24).

[66] The CMS Collaboration. *Search for Higgs boson pair production with one associated vector boson in proton-proton collisions at* $\sqrt{s} = 13\,\text{TeV}$. Tech. rep. 2024. arXiv: 2404.08462 [hep-ex]. URL: https://arxiv.org/abs/2404.08462 (cited on pages 24, 27, 43).

[67] The ATLAS Collaboration. "Search for Higgs boson pair production in association with a vector boson in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *The European Physical Journal C* 83.6 (June 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11559-y. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11559-y (cited on page 24).

[68] The CMS Collaboration. "Measurement of the $B_s^0 \to \mu^+\mu^-$ decay properties and search for the $B^0 \to \mu^+\mu^-$ decay in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Physics Letters B* 842 (2023), p. 137955. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2023.137955. URL: https://www.sciencedirect.com/science/article/pii/S0370269323002897 (cited on page 27).

[69] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, Sept. 2013. ISBN: 9781139525367. DOI: 10.1017/cbo9781139525367. URL: http://dx.doi.org/10.1017/cbo9781139525367 (cited on pages 29, 38, 104).

[70] C. L. Bennett et al. "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps And Results". In: *The Astrophysical Journal Supplement Series* 208.2 (Sept. 2013), p. 20. ISSN: 1538-4365. DOI: 10.1088/0067-0049/208/2/20. URL: http://dx.doi.org/10.1088/0067-0049/208/2/20 (cited on page 29).

[71] N. Aghanim et al. "Planck 2018 results: VI. Cosmological parameters". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201833910. URL: http://dx.doi.org/10.1051/0004-6361/201833910 (cited on pages 29, 30).

[72] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. "Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657−558: Direct Evidence for the Existence of Dark Matter". In: *The Astrophysical Journal* 604.2 (Apr. 2004), pp. 596–603. ISSN: 1538-4357. DOI: 10.1086/381970. URL: http://dx.doi.org/10.1086/381970 (cited on page 29).

[73] M. Markevitch et al. "Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657-56". In: *The Astrophysical Journal* 606.2 (May 2004), pp. 819–824. ISSN: 1538-4357. DOI: 10.1086/383178. URL: http://dx.doi.org/10.1086/383178 (cited on page 29).

[74] Mordehai Milgrom. "MOND theory". In: *Canadian Journal of Physics* 93.2 (Feb. 2015), pp. 107–118. ISSN: 1208-6045. DOI: 10.1139/cjp-2014-0211. URL: http://dx.doi.org/10.1139/cjp-2014-0211 (cited on page 29).

[75] Sebastian Baum et al. *The Final Frontier for Proton Decay*. 2024. arXiv: 2405.15845 [hep-ph]. URL: https://arxiv.org/abs/2405.15845 (cited on page 29).

[76] The Super-Kamiokande Collaboration. "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ with an enlarged fiducial volume in Super-Kamiokande I-IV". In: *Physical Review D* 102.11 (Dec. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.112011. URL: http://dx.doi.org/10.1103/PhysRevD.102.112011 (cited on page 29).

[77] The Super-Kamiokande Collaboration. "Search for proton decay via $p \to \mu^+K^0$ in 0.37 megaton-years exposure of Super-Kamiokande". In: *Phys. Rev. D* 106 (7 Oct. 2022), p. 072003. DOI: 10.1103/PhysRevD.106.072003. URL: https://link.aps.org/doi/10.1103/PhysRevD.106.072003 (cited on page 29).

[78] E. Kh. Akhmedov, G. C. Branco, and M. N. Rebelo. "Seesaw mechanism and structure of neutrino mass matrix". In: *Physics Letters B* 478.1–3 (Apr. 2000), pp. 215–223. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(00)00282-3. URL: http://dx.doi.org/10.1016/S0370-2693(00)00282-3 (cited on page 30).

[79] G. Degrassi et al. "Probing the Higgs self coupling via single Higgs production at the LHC". In: *Journal of High Energy Physics* 2016.12 (Dec. 2016). ISSN: 1029-8479. DOI: 10.1007/jhep12(2016)080. URL: http://dx.doi.org/10.1007/JHEP12(2016)080 (cited on pages 31, 32, 48, 49).

[80] The CMS Collaboration. "Sensitivity to new physics in final states with multiple gauge and Higgs bosons". In: *Journal of High Energy Physics* 2022.9 (2018). ISSN: 1029-8479. DOI: 10.1007/JHEP09(2022)038 (cited on page 35).

[81] O. J. P. Éboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi. "$pp \to jje^\pm\mu^\pm\nu\nu$ and $jje^\pm\mu^\mp\nu\nu$ at $\mathcal{O}(\alpha_{\rm em}^6)$ and $\mathcal{O}(\alpha_{\rm em}^4\alpha_S^2)$ for the study of the quartic electroweak gauge boson vertex at CERN LHC". In: *Phys. Rev. D* 74 (7 Oct. 2006), p. 073005. DOI: 10.1103/PhysRevD.74.073005. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.073005 (cited on page 35).

[82] Eduardo da Silva Almeida, O. J. P. Éboli, and M. C. Gonzalez–Garcia. "Unitarity constraints on anomalous quartic couplings". In: *Phys. Rev. D* 101 (11 June 2020), p. 113003. DOI: 10.1103/PhysRevD.101.113003. URL: https://link.aps.org/doi/10.1103/PhysRevD.101.113003 (cited on page 35).

[83] E. Brost and L. Cadamuro. "Searching for Pairs of Higgs Bosons in the LHC Run 2 Dataset". In: *Symmetry* 14.7 (2022). ISSN: 2073-8994. DOI: 10.3390/sym14071467 (cited on pages 35, 245).

[84] Florian Goertz et al. "Higgs boson pair production in the D=6 extension of the SM". In: *Journal of High Energy Physics* 2015.4 (2018). ISSN: 1029-8479. DOI: 10.1007/JHEP04(2015)167. URL: https://doi.org/10.1007/JHEP04(2015)167 (cited on page 35).

[85] Alexandra Carvalho and Martino Dall'Osso and Pablo de Castro Manzano and Tommaso Dorigo and Florian Goertz and Maxime Gouzevich and Mia Tosi. *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*. 2016. arXiv: 1608.06578 [hep-ph]. URL: https://arxiv.org/abs/1608.06578 (cited on page 35).

[86] Alexandra Carvalho et al. "Higgs boson pair production in the D=6 extension of the SM". In: *Journal of High Energy Physics* 2016.4 (2016). ISSN: 1029-8479. DOI: 10.1007/JHEP04(2016)126. URL: https://doi.org/10.1007/JHEP04(2016)126 (cited on page 36).

[87] Lisa Randall and Raman Sundrum. "Large Mass Hierarchy from a Small Extra Dimension". In: *Physical Review Letters* 83.17 (Oct. 1999), pp. 3370–3373. ISSN: 1079-7114. DOI: 10.1103/physrevlett.83.3370. URL: http://dx.doi.org/10.1103/PhysRevLett.83.3370 (cited on page 37).

[88] Alexandra Carvalho. "Gravity particles from Warped Extra Dimensions, Predictions for LHC". In: (2018). arXiv: 1404.0102 [hep-ph]. URL: https://arxiv.org/abs/1404.0102 (cited on pages 37, 160, 249).

[89] G. C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Physics Reports* 516.1 (2012). Theory and phenomenology of two-Higgs-doublet models, pp. 1–102. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2012.02.002. URL: https://www.sciencedirect.com/science/article/pii/S0370157312000695 (cited on page 38).

[90] H.E. Haber and G.L. Kane. "The search for supersymmetry: Probing physics beyond the standard model". In: *Physics Reports* 117.2 (1985), pp. 75–263. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(85)90051-1. URL: https://www.sciencedirect.com/science/article/pii/0370157385900511 (cited on page 38).

[91] Jihn E. Kim. "Light pseudoscalars, particle physics and cosmology". In: *Physics Reports* 150.1 (1987), pp. 1–177. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(87)90017-2. URL: https://www.sciencedirect.com/science/article/pii/0370157387900172 (cited on page 38).

[92] James M. Cline, Kimmo Kainulainen, and Michael Trott. "Electroweak baryogenesis in two Higgs doublet models and B meson anomalies". In: *Journal of High Energy Physics* 2011.11 (Nov. 2011). ISSN: 1029-8479. DOI: 10.1007/jhep11(2011)089. URL: http://dx.doi.org/10.1007/JHEP11(2011)089 (cited on page 38).

[93] Abdelhak Djouadi and Jérémie Quevillon. "The MSSM Higgs sector at a high MSUSY: reopening the low $\tan\beta$ regime and heavy Higgs searches". In: *Journal of High Energy Physics* 2013.10 (Oct. 2013). ISSN: 1029-8479. DOI: 10.1007/jhep10(2013)028. URL: http://dx.doi.org/10.1007/JHEP10(2013)028 (cited on page 39).

[94] The CMS Collaboration. "Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of High Energy Physics* 2023.7 (2023). ISSN: 1029-8479. DOI: 10.1007/JHEP07(2023)073 (cited on page 39).

[95] David Eriksson, Johan Rathsman, and Oscar Stål. "2HDMC – two-Higgs-doublet model calculator". In: *Computer Physics Communications* 181.1 (Jan. 2010), pp. 189–205. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2009.09.011. URL: http://dx.doi.org/10.1016/j.cpc.2009.09.011 (cited on page 40).

[96] R. Harlander et al. *Interim recommendations for the evaluation of Higgs production cross sections and branching ratios at the LHC in the Two-Higgs-Doublet Model*. 2014. arXiv: 1312.5571 [hep-ph]. URL: https://arxiv.org/abs/1312.5571 (cited on page 40).

[97] G. Chalons and F. Domingo. "Analysis of the Higgs potentials for two doublets and a singlet". In: *Physical Review D* 86.11 (Dec. 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.115024. URL: http://dx.doi.org/10.1103/PhysRevD.86.115024 (cited on page 41).

[98] Chien-Yi Chen, Michael Freid, and Marc Sher. "Next-to-minimal two Higgs doublet model". In: *Physical Review D* 89.7 (Apr. 2014). ISSN: 1550-2368. DOI: 10.1103/physrevd.89.075009. URL: http://dx.doi.org/10.1103/PhysRevD.89.075009 (cited on page 41).

[99] Margarete Mühlleitner et al. "The N2HDM under theoretical and experimental scrutiny". In: *Journal of High Energy Physics* 2017.3 (Mar. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep03(2017)094. URL: http://dx.doi.org/10.1007/JHEP03(2017)094 (cited on page 41).

[100] Margarete Mühlleitner et al. "The N2HDM under theoretical and experimental scrutiny". In: *Journal of High Energy Physics* 2017.3 (Mar. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep03(2017)094. URL: http://dx.doi.org/10.1007/JHEP03(2017)094 (cited on page 41).

[101] The CMS Collaboration. "Search for a heavy Higgs boson decaying into two lighter Higgs bosons in the $bb\tau\tau$ final state at 13 TeV". In: *Journal of High Energy Physics* 2021.57 (2021). DOI: 10.1007/JHEP11(2021)057. URL: https://doi.org/10.1007/JHEP11%282021%29057 (cited on pages 41, 42, 201, 236, 241, 264).

[102] Bartosz Dziewit, Joris Vergeest, and Marek Zrałek. "Flavor invariance of leptonic Yukawa terms in the 3HDM". In: *Physics Letters B* 853 (2024), p. 138667. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2024.138667. URL: https://www.sciencedirect.com/science/article/pii/S0370269324002259 (cited on page 41).

[103] Panagiotis Stylianou and Georg Weiglein. "Constraints on the trilinear and quartic Higgs couplings from triple Higgs production at the LHC and beyond". In: *The European Physical Journal C* 84.4 (Apr. 2024). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-024-12722-9. URL: http://dx.doi.org/10.1140/epjc/s10052-024-12722-9 (cited on page 41).

[104] The CMS Collaboration. "Search for Higgs boson pair production in the $bb\mathrm{W}^{+}\mathrm{W}^{-}$ decay mode in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Journal of High Energy Physics* 2024.7 (July 2024). ISSN: 1029-8479. DOI: 10.1007/jhep07(2024)293. URL: http://dx.doi.org/10.1007/JHEP07(2024)293 (cited on pages 42, 43).

[105] The CMS Collaboration. "Search for heavy resonances decaying to a pair of Lorentz-boosted Higgs bosons in final states with leptons and a bottom quark pair at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Journal of High Energy Physics* 2022.5 (May 2022). ISSN: 1029-8479. DOI: 10.1007/jhep05(2022)005. URL: http://dx.doi.org/10.1007/JHEP05(2022)005 (cited on page 42).

[106] The CMS Collaboration. "Search for Higgs boson pairs decaying to $\mathrm{WW}^{*}\mathrm{WW}^{*}$, $\mathrm{WW}^{*}\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Journal of High Energy Physics* 2023.7 (2018). ISSN: 1029-8479. DOI: 10.1007/JHEP07(2023)095 (cited on pages 42, 43).

[107] The CMS Collaboration. *Search for a new resonance decaying into two spin-0 bosons in a final state with two photons and two bottom quarks in proton-proton collisions at* $\sqrt{s} = 13\,\mathrm{TeV}$. May 2024. DOI: 10.1007/jhep05(2024)316. URL: http://dx.doi.org/10.1007/JHEP05(2024)316 (cited on pages 42, 241).

[108] The CMS Collaboration. "Search for a massive scalar resonance decaying to a light scalar and a Higgs boson in the four b quarks final state with boosted topology". In: *Phys. Lett. B* 842 (2023), p. 137392. DOI: 10.1016/j.physletb.2022.137392. arXiv: 2204.12413 [hep-ex] (cited on pages 42, 241).

[109] The ATLAS Collaboration. "Combination of searches for resonant Higgs boson pair production using pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ with the ATLAS detector". In: *Physical Review Letters* 132.23 (). ISSN: 1079-7114. DOI: 10.1103/physrevlett.132.231801. URL: http://dx.doi.org/10.1103/PhysRevLett.132.231801 (cited on page 43).

[110] The CMS Collaboration. "Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Physics Letters B* 778 (Mar. 2018), pp. 101–127. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2018.01.001. URL: http://dx.doi.org/10.1016/j.physletb.2018.01.001 (cited on pages 43, 154, 162, 182, 236, 239).

[111] The CMS Collaboration. "Search for nonresonant Higgs boson pair production in final state with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: (2022) (cited on pages 43, 154, 159, 162, 185, 187, 196, 218).

[112] The CMS Collaboration. "Search for Nonresonant Pair Production of Highly Energetic Higgs Bosons Decaying to Bottom Quarks". In: *Phys. Rev. Lett.* 131 (4 July 2023), p. 041803. DOI: 10.1103/PhysRevLett.131.041803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.041803 (cited on page 43).

[113] The CMS Collaboration. "Search for Higgs boson pair production in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Phys. Rev. Lett.* 129 (8 Aug. 2022), p. 081802. DOI: 10.1103/PhysRevLett.129.081802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.129.081802 (cited on pages 43, 194).

[114] The CMS Collaboration. "Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of High Energy Physics* 2021.257 (2021). ISSN: 1029-8479. DOI: 10.1007/JHEP03(2021)257 (cited on pages 43, 47).

[115] The CMS Collaboration. "Search for nonresonant Higgs boson pair production in the four leptons plus two b jets final state in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of High Energy Physics* 2023.6 (2023). ISSN: 1029-8479. DOI: 10.1007/JHEP06(2023)130 (cited on page 43).

[116] The CMS Collaboration. *Search for the nonresonant and resonant production of a Higgs boson in association with an additional scalar boson in the $\gamma\gamma\tau\tau$ final state*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2893031 (cited on page 43).

[117] The CMS Collaboration. *Search for nonresonant Higgs boson pair production in the $WW\gamma\gamma$ channel in pp collisions at $\sqrt{s} = 13\,\text{TeV}$*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2840773 (cited on page 43).

[118] The CMS Collaboration. *Higgs PAG Summary Plots*. Twiki. Accessed: 2024-04-14. 2023. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryResultsHIG (cited on page 46).

[119] The CMS Collaboration. *Search for highly energetic double Higgs boson production in the two bottom quark and two vector boson all-hadronic final state*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2904879 (cited on page 47).

[120] Huilin Qu, Congqiao Li, and Sitian Qian. *Particle Transformer for Jet Tagging*. 2024. arXiv: 2202.03772 [hep-ph]. URL: https://arxiv.org/abs/2202.03772 (cited on pages 47, 106, 177, 240).

[121] Christoph Englert et al. "VBS $W^{\pm}W^{\pm}H$ production at the HL-LHC and a 100 TeV pp-collider". In: *International Journal of Modern Physics A* 32.18 (2017), p. 1750106. DOI: 10.1142/S0217751X17501068. eprint: https://doi.org/10.1142/S0217751X17501068. URL: https://doi.org/10.1142/S0217751X17501068 (cited on page 47).

[122] The CMS Collaboration. *Search for HHWW couplings in the VBS production of $W^{\pm}W^{\pm}H$, with $H \to bb$ decays*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2905615 (cited on page 47).

[123] Fabio Maltoni et al. "Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC". In: *The European Physical Journal C* 77.12 (Dec. 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5410-8. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5410-8 (cited on pages 48, 50).

[124] G. Degrassi, M. Fedele, and P.P. Giardino. "Constraints on the trilinear Higgs self coupling from precision observables". In: *Journal of High Energy Physics* 2017.155 (Apr. 2017). ISSN: 1029-8479. DOI: 10.1007/JHEP04(2017)155 (cited on page 48).

[125] Fabio Monti et al. *Modelling of the single-Higgs simplified template cross-sections (STXS 1.2) for the determination of the Higgs boson trilinear self-coupling*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2803606 (cited on page 48).

[126] Wojciech Bizoń et al. “Constraints on the trilinear Higgs coupling from vector boson fusion and associated Higgs production at the LHC”. In: *Journal of High Energy Physics* 2017.7 (July 2017). ISSN: 1029-8479. DOI: 10.1007/jhep07(2017)083. URL: http://dx.doi.org/10.1007/JHEP07(2017)083 (cited on page 48).

[127] S. Di Vita et al. “A global view on the Higgs self-coupling”. In: *Journal of High Energy Physics* 2017.9 (Sept. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep09(2017)069. URL: http://dx.doi.org/10.1007/JHEP09(2017)069 (cited on page 50).

[128] The CMS Collaboration. *Constraints on the Higgs boson self-coupling with combination of single and double Higgs boson production*. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2882424 (cited on pages 50, 51).

[129] The ATLAS Collaboration. “Constraints on the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: *Physics Letters B* 843 (2023), p. 137745. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2023.137745. URL: https://www.sciencedirect.com/science/article/pii/S0370269323000795 (cited on page 50).

[130] European Organization for Nuclear Research (CERN). “Convention for the Establishment of a European Organization for Nuclear Research”. Accessed: 29 August 2024. 1954. URL: https://council.web.cern.ch/en/content/convention-establishment-european-organization-nuclear-research (cited on page 53).

[131] The ALEPH, DELPHI, L3 and OPAL Collaborations. “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP”. In: *Physics Reports* 532.4 (Nov. 2013), pp. 119–244. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2013.07.004. URL: http://dx.doi.org/10.1016/j.physrep.2013.07.004 (cited on page 53).

[132] O. Aberle et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2020. DOI: 10.23731/CYRM-2020-0010. URL: https://cds.cern.ch/record/2749422 (cited on pages 53, 102, 240).

[133] Ewa Lopienska. “The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022”. In: *CERN-GRAPHICS-2022-001* (2022). General Photo. URL: https://cds.cern.ch/record/2800984 (cited on page 54).

[134] Maurizio Vretenar et al. *Linac4 design report*. Vol. 6. CERN Yellow Reports: Monographs. Geneva: CERN, 2020. DOI: 10.23731/CYRM-2020-006. URL: https://cds.cern.ch/record/2736208 (cited on page 54).

[135] Dan Faircloth and Scott Lawrie. “An overview of negative hydrogen ion sources for accelerators”. In: *New Journal of Physics* 20.2 (Feb. 2018), p. 025007. DOI: 10.1088/1367-2630/aaa39e. URL: https://dx.doi.org/10.1088/1367-2630/aaa39e (cited on page 54).

[136] Valery Lebedev and Vladimir Shiltsev, eds. *Accelerator Physics at the Tevatron Collider*. Particle Acceleration and Detection. Springer New York, 2014, nil. DOI: 10.1007/978-1-4939-0885-1. URL: http://dx.doi.org/10.1007/978-1-4939-0885-1 (cited on page 56).

[137] Sören Möller. *Accelerator Technology*. Particle Acceleration and Detection. Springer International Publishing, 2020, nil. DOI: 10.1007/978-3-030-62308-1. URL: http://dx.doi.org/10.1007/978-3-030-62308-1 (cited on page 56).

[138] The CMS Collaboration. "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of High Energy Physics* 2018.7 (July 2018). ISSN: 1029-8479. DOI: 10.1007/JHEP07(2018)161. URL: https://doi.org/10.1007/JHEP07(2018)161 (cited on pages 56, 78).

[139] M. Lamont. "Chapter 4: Operational Experience from LHC Run 1 & 2 and Consolidation in View of Run 3 and the HL-LHC". In: *Adv. Ser. Direct. High Energy Phys.* 31 (2024), pp. 101–120. DOI: 10.1142/9789811278952_0004. URL: http://cds.cern.ch/record/2906147 (cited on pages 57, 59).

[140] The CMS Collaboration. *Public CMS Luminosity Information*. Twiki. Accessed: 30 August 2024. 2023. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults (cited on pages 58, 102).

[141] J. Wenninger. "Machine Protection and Operation for LHC". In: (Aug. 2016). DOI: 10.5170/CERN-2016-002.377 (cited on page 59).

[142] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (August 2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08003 (cited on page 59).

[143] Bartosz Mindur. "ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 845 (2017). Proceedings of the Vienna Conference on Instrumentation 2016, pp. 257–261. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2016.04.026. URL: https://www.sciencedirect.com/science/article/pii/S0168900216301905 (cited on page 59).

[144] C. W. Fabjan and D. Fournier. "Calorimetry". In: *Particle Physics Reference Library: Volume 2: Detectors for Particles and Radiation*. Ed. by Christian W. Fabjan and Herwig Schopper. Cham: Springer International Publishing, 2020, pp. 201–280. ISBN: 978-3-030-35318-6. DOI: 10.1007/978-3-030-35318-6_6. URL: https://doi.org/10.1007/978-3-030-35318-6_6 (cited on pages 60, 122).

[145] Peter W. Krieger. "The ATLAS Liquid Argon Calorimeter: One year of LHC operation and future upgrade plans for HL-LHC". In: (May 2011). DOI: 10.1109/ANIMMA.2011.6172835 (cited on page 60).

[146] Christian W. Fabjan and Fabiola Gianotti. "Calorimetry for particle physics". In: *Reviews of Modern Physics* 75 (4 2003), pp. 1243–1246. ISSN: 00346861. DOI: 10.1103/RevModPhys.75.1243 (cited on pages 60, 70, 118).

[147] The ATLAS Collaboration. "The ATLAS trigger system for LHC Run 3 and trigger performance in 2022". In: *Journal of Instrumentation* 19.06 (June 2024), P06029. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/06/p06029. URL: http://dx.doi.org/10.1088/1748-0221/19/06/P06029 (cited on page 60).

[148] The ATLAS Collaboration. "Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Phys. Rev. Lett.* 121 (8 Aug. 2018), p. 081801. DOI: 10.1103/PhysRevLett.121.081801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.081801 (cited on page 60).

[149] The LHCb Collaboration. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08005 (cited on page 60).

[150] The LHCb Collaboration. *Framework TDR for the LHCb Upgrade II: Opportunities in flavour physics, and beyond, in the HL-LHC era*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2776420 (cited on pages 61, 126).

[151] S. Erhan, M. Medinnis, and Peter E. Schlein. "Collider beauty physics at LHC". In: (1990). DOI: 10.5170/CERN-1990-010-V-2.239. URL: https://cds.cern.ch/record/215733 (cited on page 61).

[152] Daniel Denegri. "Standard Model physics at the LHC (pp collisions)". In: (1990). DOI: 10.5170/CERN-1990-010-V-1.56. URL: https://cds.cern.ch/record/214585 (cited on page 61).

[153] The LHCb Collaboration. *LHCb inner tracker: Technical Design Report*. Technical design report. LHCb. revised version number 1 submitted on 2002-11-13 14:14:34. Geneva: CERN, 2002. URL: https://cds.cern.ch/record/582793 (cited on page 61).

[154] Sean Benson et al. "The LHCb Turbo Stream". In: *Journal of Physics: Conference Series* 664.8 (Dec. 2015), p. 082004. DOI: 10.1088/1742-6596/664/8/082004. URL: https://dx.doi.org/10.1088/1742-6596/664/8/082004 (cited on page 61).

[155] Luciano Libero Pappalardo. "LHCb Fixed Target Results and Prospects". In: *The XVIII International Conference on Strangeness in Quark Matter (SQM 2019)*. Ed. by Domenico Elia et al. Cham: Springer International Publishing, 2020, pp. 489–493. ISBN: 978-3-030-53448-6 (cited on page 61).

[156] Atri Bhattacharya et al. "Forward neutrinos from charm at the Large Hadron Collider". In: *Physical Review D* 109.1 (Jan. 2024). ISSN: 2470-0029. DOI: 10.1103/physrevd.109.014040. URL: http://dx.doi.org/10.1103/PhysRevD.109.014040 (cited on page 61).

[157] The LHCb Collaboration. "Measurement of Antiproton Production in $p-\text{He}$ collisions at $\sqrt{s_{\text{NN}}} = 110\,\text{GeV}$". In: *Physical Review Letters* 121.22 (Nov. 2018). ISSN: 1079-7114. DOI: 10.1103/physrevlett.121.222001. URL: http://dx.doi.org/10.1103/PhysRevLett.121.222001 (cited on page 61).

[158] The ALICE Collaboration. In: 3.08 (Aug. 2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08002 (cited on page 62).

[159] The ALICE Collaboration. *Upgrade of the ALICE Readout & Trigger System*. Tech. rep. 2013. URL: https://cds.cern.ch/record/1603472 (cited on page 62).

[160] Jakub Kvapil et al. "ALICE Central Trigger System for LHC Run 3". In: *EPJ Web of Conferences* 251 (2021). Ed. by C. Biscarat et al., p. 04022. ISSN: 2100-014X. DOI: 10.1051/epjconf/202125104022. URL: http://dx.doi.org/10.1051/epjconf/202125104022 (cited on page 62).

[161] FASER Collaboration. *The FASER Detector*. 2022. arXiv: 2207.11427 [physics.ins-det] (cited on page 62).

[162] FASER Collaboration. "Detecting and studying high-energy collider neutrinos with FASER at the LHC". In: *The European Physical Journal C* 80.1 (Jan. 2020). DOI: 10.1140/epjc/s10052-020-7631-5. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-020-7631-5 (cited on page 62).

[163] The FASER Collaboration. "First Direct Observation of Collider Neutrinos with FASER at the LHC". In: *Phys. Rev. Lett.* 131 (3 July 2023), p. 031801. DOI: 10.1103/PhysRevLett.131.031801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.031801 (cited on page 63).

[164] The FASER Collaboration. "First Measurement of ν_e and ν_μ Interaction Cross Sections at the LHC with FASER's Emulsion Detector". In: *Phys. Rev. Lett.* 133 (2 July 2024), p. 021802. DOI: 10.1103/PhysRevLett.133.021802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.133.021802 (cited on page 63).

[165] The SND@LHC Collaboration. *SND@LHC: the scattering and neutrino detector at the LHC*. May 2024. DOI: 10.1088/1748-0221/19/05/p05067. URL: http://dx.doi.org/10.1088/1748-0221/19/05/P05067 (cited on page 63).

[166] The SND@LHC Collaboration. "Observation of Collider Muon Neutrinos with the SND@LHC Experiment". In: *Phys. Rev. Lett.* 131 (3 July 2023), p. 031802. DOI: 10.1103/PhysRevLett.131.031802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.031802 (cited on page 63).

[167] The LHCf Collaboration. "The LHCf detector at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08006–S08006. DOI: 10.1088/1748-0221/3/08/s08006. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08006 (cited on page 63).

[168] The Pierre Auger Collaboration. "The Pierre Auger Cosmic Ray Observatory". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 798 (2015), pp. 172–213. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2015.06.058. URL: https://www.sciencedirect.com/science/article/pii/S0168900215008086 (cited on page 63).

[169] H. Kawai et al. "Telescope Array Experiment". In: *Nuclear Physics B - Proceedings Supplements* 175-176 (2008). Proceedings of the XIV International Symposium on Very High Energy Cosmic Ray Interactions, pp. 221–226. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2007.11.002. URL: https://www.sciencedirect.com/science/article/pii/S0920563207007992 (cited on page 63).

[170] The TOTEM Collaboration. "The TOTEM Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08007–S08007. DOI: 10.1088/1748-0221/3/08/s08007. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08007 (cited on page 63).

[171] James Pinfold et al. *Technical Design Report of the MoEDAL Experiment*. Tech. rep. 2009. URL: https://cds.cern.ch/record/1181486 (cited on page 64).

[172] Andrew Haas et al. "Looking for milli-charged particles with a new experiment at the LHC". In: *Physics Letters B* 746 (June 2015), pp. 117–120. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2015.04.062. URL: http://dx.doi.org/10.1016/j.physletb.2015.04.062 (cited on page 64).

[173] A. Ball et al. "Search for millicharged particles in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Physical Review D* 102.3 (Aug. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.032002. URL: http://dx.doi.org/10.1103/PhysRevD.102.032002 (cited on page 64).

[174] Alessandro De Angelis and Mário Pimenta. *Introduction to Particle and Astroparticle Physics*. Undergraduate Lecture Notes in Physics. Springer Cham, 2018. ISBN: 978-3-319-78181-5. DOI: 10.1007/978-3-319-78181-5. URL: https://link.springer.com/book/10.1007/978-3-319-78181-5 (cited on page 66).

[175] Tai Sakuma. "Cutaway diagrams of CMS detector". In: (2019). URL: https://cds.cern.ch/record/2665537 (cited on page 68).

[176] The CMS Collaboration. *The CMS tracker system project: Technical Design Report*. Tech. rep. Geneva, 1997. URL: https://cds.cern.ch/record/368412 (cited on page 67).

[177] The CMS Collaboration. *The CMS tracker: addendum to the Technical Design Report*. Tech. rep. Geneva, 2000. URL: https://cds.cern.ch/record/490194 (cited on page 67).

[178] Philip Allport. "Applications of silicon strip and pixel-based particle tracking detectors". In: *Nature Reviews Physics* 1.9 (Sept. 2019). DOI: 10.1038/s42254-019-0081-z. URL: https://doi.org/10.1038/s42254-019-0081-z (cited on page 67).

[179] The CMS Collaboration. "Displaced tracking and vertexing calibration using neutral K mesons". In: (2024). URL: https://cds.cern.ch/record/2890104 (cited on page 67).

[180] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.QZ28.FLHW. URL: https://cds.cern.ch/record/2272264 (cited on pages 68, 107, 108).

[181] Paolo Azzurri. "The CMS Silicon Strip Tracker". In: *J. Phys.: Conf. Ser.* 41 (2006), pp. 127–134. DOI: 10.1088/1742-6596/41/1/011. URL: https://cds.cern.ch/record/914891 (cited on page 69).

[182] The CMS Collaboration. "The CMS Phase-1 pixel detector upgrade". In: *Journal of Instrumentation* 16.02 (Feb. 2021), P02027. DOI: 10.1088/1748-0221/16/02/P02027. URL: https://dx.doi.org/10.1088/1748-0221/16/02/P02027 (cited on pages 70, 106).

[183] The CMS Collaboration. *The CMS electromagnetic calorimeter project: Technical Design Report*. Tech. rep. Geneva, 1997. URL: https://cds.cern.ch/record/349375 (cited on pages 69, 71, 72).

[184] The CMS Collaboration. "Performance of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: (Mar. 2024). arXiv: 2403.15518 [physics.ins-det] (cited on pages 70, 81).

[185] The CMS Collaboration. "Energy resolution of the barrel of the CMS Electromagnetic Calorimeter". In: *Journal of Instrumentation* 2.04 (Apr. 2007), P04004. DOI: 10.1088/1748-0221/2/04/P04004. URL: https://dx.doi.org/10.1088/1748-0221/2/04/P04004 (cited on page 71).

[186] The CMS Collaboration. In: *Journal of Instrumentation* 8.09 (Sept. 2013), P09009–P09009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/09/p09009. URL: http://dx.doi.org/10.1088/1748-0221/8/09/P09009 (cited on page 71).

[187] The CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report*. Tech. rep. Geneva, 1997. URL: https://cds.cern.ch/record/357153 (cited on page 71).

[188] Richard Wigmans. *Calorimetry: Energy Measurement in Particle Physics*. Oxford University Press, Sept. 2017. ISBN: 9780198786351. DOI: 10.1093/oso/9780198786351.001.0001. URL: https://doi.org/10.1093/oso/9780198786351.001.0001 (cited on pages 71, 88, 118).

[189] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (August 2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08004 (cited on pages 73, 74).

[190] The CMS Collaboration. *The CMS magnet project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. DOI: 10.17181/CERN.6ZU0.V4T9. URL: https://cds.cern.ch/record/331056 (cited on page 73).

[191] The CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: *Journal of Instrumentation* 12.10 (Oct. 2017), P10003–P10003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/10/p10003. URL: http://dx.doi.org/10.1088/1748-0221/12/10/P10003 (cited on pages 74, 85, 86, 91–93).

[192] The CMS Collaboration. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical design report. CMS. URL: https://cds.cern.ch/record/706847 (cited on pages 75, 79).

[193] Amrutha Samalan. "Muon Detector Development for the CMS Phase-2 Upgrade and Muon Radiography Applications". Presented 26 March 2024. Ghent University, 2024. URL: https://cds.cern.ch/record/2894453 (cited on pages 76, 111).

[194] A. Colaleo et al. *CMS Technical Design Report for the Muon Endcap GEM Upgrade*. Tech. rep. 2015. URL: https://cds.cern.ch/record/2021453 (cited on pages 77, 111, 112).

[195] The CMS Collaboration. *Trigger data formats for the CSC and GEM systems for Run-3*. Tech. rep. Geneva: CERN, 2021 (cited on pages 77, 111).

[196] The CMS Collaboration. *Stairway to discovery: a report on the CMS programme of cross section measurements from millibarns to femtobarns*. 2024. arXiv: 2405.18661 [hep-ex]. URL: https://arxiv.org/abs/2405.18661 (cited on page 78).

[197] The CMS Collaboration. "Measurement of the inelastic proton–proton cross section at $\sqrt{s} = 7\,\text{TeV}$". In: *Physics Letters B* 722.1–3 (May 2013), pp. 5–27. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2013.03.024. URL: http://dx.doi.org/10.1016/j.physletb.2013.03.024 (cited on page 78).

[198] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Level-1 Trigger*. Tech. rep. CMS-TDR-021. CERN, 2020. URL: https://cds.cern.ch/record/2714892 (cited on pages 79, 113–115, 126).

[199] The CMS Collaboration. *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 (cited on page 79).

[200] The CMS Collaboration. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10017–P10017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/10/p10017. URL: http://dx.doi.org/10.1088/1748-0221/15/10/P10017 (cited on page 79).

[201] Somnath Choudhury. "Performance of the High-Level Trigger System at CMS in LHC Run-2". In: *IEEE Transactions on Nuclear Science* 68.8 (2021), pp. 2035–2042. DOI: 10.1109/TNS.2021.3087618 (cited on page 79).

[202] Alexandre Zabi. "Opérations et optimisations du système de déclenchement calorimétrique de l'expérience CMS au LHC". Habilitation à diriger des recherches. Laboratoire Leprince Ringuet Ecole Polytechnique, Jan. 2016. URL: https://hal.science/tel-03030251 (cited on pages 79, 80, 126).

[203] M. Baber et al. "Development and testing of an upgrade to the CMS Level-1 calorimeter trigger". In: *Journal of Instrumentation* 9.01 (Jan. 2014), p. C01006. DOI: 10.1088/1748-0221/9/01/C01006. URL: https://dx.doi.org/10.1088/1748-0221/9/01/C01006 (cited on page 80).

[204] A. Tapper and D. Acosta. *CMS Technical Design Report for the Level-1 Trigger Upgrade*. Tech. rep. 2013. URL: https://cds.cern.ch/record/1556311 (cited on page 80).

[205] I Bird et al. *Update of the Computing Models of the WLCG and the LHC Experiments*. Tech. rep. 2014. URL: https://cds.cern.ch/record/1695401 (cited on page 81).

[206] Andrea Bocci. "CMS High Level Trigger performance comparison on CPUs and GPUs". In: *J. Phys. Conf. Ser.* 2438.1 (2023), p. 012016. DOI: 10.1088/1742-6596/2438/1/012016 (cited on page 81).

[207] Thomas Reis. "Developing GPU-compliant algorithms for CMS ECAL local reconstruction during LHC Run 3 and Phase 2". In: *J. Phys. Conf. Ser.* 2438.1 (2023), p. 012027. DOI: 10.1088/1742-6596/2438/1/012027 (cited on page 81).

[208] The CMS Collaboration. "Particle Flow Reconstruction on Heterogeneous Architecture for CMS". In: (2022). URL: https://cds.cern.ch/record/2842374 (cited on page 81).

[209] The CMS Collaboration. "Performance of the CMS muon trigger system in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of Instrumentation* 16.07 (July 2021), P07001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/07/p07001. URL: http://dx.doi.org/10.1088/1748-0221/16/07/P07001 (cited on page 81).

[210] The CMS Collaboration. *Enriching the physics program of the CMS experiment via data scouting and data parking*. 2024. arXiv: 2403.16134 [hep-ex]. URL: https://arxiv.org/pdf/2403.16134 (cited on pages 81, 82).

[211] The CMS Collaboration. "Data Parking and Data Scouting at the CMS Experiment". In: (2012). URL: https://cds.cern.ch/record/1480607 (cited on page 81).

[212] The CMS Collaboration. "Novel strategy targeting HH and HHH production at High Level Trigger in Run 3". In: *CERN Document Server: 2868787* (2023). URL: https://cds.cern.ch/record/2868787 (cited on pages 81, 83, 106, 240).

[213] The CMS Collaboration. *Novel strategy targeting HH and HHH production at High Level Trigger in Run 3 (CMS DP-2023/050)*. Twiki. Accessed: 2024-07-8. 2023. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/Run3HH (cited on page 83).

[214] The CMS Collaboration. *Search for pair-produced multijet resonances using data scouting*. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2866497 (cited on page 83).

[215] Andrea Bocci et al. *Heterogeneous reconstruction of tracks and primary vertices with the CMS pixel tracker*. 2020. arXiv: 2008.13461 [physics.ins-det]. URL: https://arxiv.org/abs/2008.13461 (cited on page 83).

[216] The CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical design report. CMS. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757 (cited on page 84).

[217] The ALEPH Collaboration. *ALEPH: Technical Report 1983*. Geneva: CERN, 1983. URL: https://cds.cern.ch/record/300680 (cited on page 84).

[218] Milos Dordevic. "The CMS Particle Flow Algorithm". In: *EPJ Web Conf.* 191 (2018), p. 02016. DOI: 10.1051/epjconf/201819102016. URL: https://cds.cern.ch/record/2678077 (cited on page 86).

[219] Matthew Nguyen. *Particle flow in CMS*. Accessed: 12 July 2024. July 2023. URL: https://indico.bnl.gov/event/19985/#2-particle-flow-in-cms (cited on page 86).

[220] Wolfgang Adam et al. *Track Reconstruction in the CMS tracker*. Tech. rep. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/934067 (cited on page 86).

[221] The CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of Instrumentation* 13.06 (June 2018), P06015–P06015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/06/p06015. URL: http://dx.doi.org/10.1088/1748-0221/13/06/P06015 (cited on page 90).

[222] 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. URL: https://dx.doi.org/10.1088/0954-3899/31/9/N01 (cited on page 91).

[223] The 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. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/05/p05014. URL: http://dx.doi.org/10.1088/1748-0221/16/05/P05014 (cited on pages 91, 174).

[224] The CMS Collaboration. "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of Instrumentation* 13.10 (Oct. 2018), P10005–P10005. DOI: 10.1088/1748-0221/13/10/p10005. URL: https://doi.org/10.1088%2F1748-0221%2F13%2F10%2Fp10005 (cited on page 93).

[225] The CMS Collaboration. "Performance of τ-lepton reconstruction and identification in CMS". In: *Journal of Instrumentation* 7.01 (Jan. 2012), P01001–P01001. DOI: 10.1088/1748-0221/7/01/p01001. URL: https://doi.org/10.1088%2F1748-0221%2F7%2F01%2Fp01001 (cited on page 93).

[226] The CMS Collaboration. "Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS". In: *Journal of Instrumentation* 11.01 (Jan. 2016), P01019–P01019. DOI: 10.1088/1748-0221/11/01/p01019. URL: https://doi.org/10.1088%2F1748-0221%2F11%2F01%2Fp01019 (cited on page 93).

[227] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. DOI: 10.1088/1126-6708/2008/04/063. URL: https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063 (cited on page 94).

[228] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Physical Review D* 48.7 (Oct. 1993), pp. 3160–3166. ISSN: 0556-2821. DOI: 10.1103/physrevd.48.3160. URL: http://dx.doi.org/10.1103/PhysRevD.48.3160 (cited on page 94).

[229] M. Wobisch and T. Wengler. *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*. 1999. arXiv: hep-ph/9907280 [hep-ph]. URL: https://arxiv.org/abs/hep-ph/9907280 (cited on page 94).

[230] Yu. L. Dokshitzer et al. "Better jet clustering algorithms". In: *Journal of High Energy Physics* 1997.08 (Aug. 1997), pp. 001–001. ISSN: 1029-8479. DOI: 10.1088/1126-6708/1997/08/001. URL: http://dx.doi.org/10.1088/1126-6708/1997/08/001 (cited on page 94).

[231] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *The European Physical Journal C* 72.3 (Mar. 2012). DOI: 10.1140/epjc/s10052-012-1896-2. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-012-1896-2 (cited on page 94).

[232] Garvita Agarwal. "Jet Energy Scale and Resolution Measurements in CMS". In: *Proceedings of 41st International Conference on High Energy physics — PoS(ICHEP2022)*. ICHEP2022. Sissa Medialab, Dec. 2022, p. 652. DOI: 10.22323/1.414.0652. URL: http://dx.doi.org/10.22323/1.414.0652 (cited on pages 95, 249).

[233] The CMS Collaboration, Offline Software and Computing. *Evolution of the CMS Computing Model towards Phase-2*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2751565 (cited on page 97).

[234] Gennaro Corcella et al. "HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)". In: *Journal of High Energy Physics* 2001.01 (Jan. 2001), pp. 010–010. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2001/01/010. URL: http://dx.doi.org/10.1088/1126-6708/2001/01/010 (cited on page 97).

[235] Manuel Bähr et al. "Herwig++ physics and manual". In: *The European Physical Journal C* 58.4 (Nov. 2008), pp. 639–707. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-008-0798-9. URL: http://dx.doi.org/10.1140/epjc/s10052-008-0798-9 (cited on page 97).

[236] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *The European Physical Journal C* 76.4 (Apr. 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4018-8. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4018-8 (cited on page 97).

[237] T Gleisberg et al. "Event generation with SHERPA 1.1". In: *Journal of High Energy Physics* 2009.02 (Feb. 2009), pp. 007–007. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2009/02/007. URL: http://dx.doi.org/10.1088/1126-6708/2009/02/007 (cited on page 97).

[238] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 2006.05 (May 2006), pp. 026–026. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2006/05/026. URL: http://dx.doi.org/10.1088/1126-6708/2006/05/026 (cited on page 97).

[239] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024 (cited on page 97).

[240] C. Oleari. "The POWHEG BOX". In: *Nuclear Physics B - Proceedings Supplements* 205-206 (2010). Loops and Legs in Quantum Field Theory, pp. 36–41. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2010.08.016. URL: https://www.sciencedirect.com/science/article/pii/S0920563210001994 (cited on page 97).

[241] 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: *JHEP* 07 (2014). DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph] (cited on pages 97, 248).

[242] The CMS Collaboration. “Extraction and validation of a new set of CMS Pythia8 tunes from underlying-event measurements”. In: *The European Physical Journal C* 80.1 (Jan. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7499-4. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7499-4 (cited on page 97).

[243] Christian Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph]. URL: https://arxiv.org/abs/2203.11601 (cited on page 98).

[244] S. Agostinelli et al. “Geant4 – a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688 (cited on page 97).

[245] J. Allison et al. “Geant4 developments and applications”. In: *IEEE Transactions on Nuclear Science* 53.1 (2006), pp. 270–278. DOI: 10.1109/TNS.2006.869826 (cited on page 97).

[246] Sezen Sekmen. *Recent Developments in CMS Fast Simulation*. 2017. arXiv: 1701.03850 [physics.ins-det]. URL: https://arxiv.org/abs/1701.03850 (cited on pages 97–99).

[247] The CMS Collaboration. *Monte Carlo production tools – What is the difference between premix and classical mixing?* Accessed: 21 June 2024. URL: https://cms-pdmv.gitbook.io/project/untitled-4#what-is-the-difference-between-premix-and-classical-mixing (cited on page 98).

[248] G. Petrucciani, A. Rizzi, and C. Vuosalo. “Mini-AOD: A New Analysis Data Format for CMS”. In: *Journal of Physics: Conference Series* 664.7 (Dec. 2015), p. 072052. DOI: 10.1088/1742-6596/664/7/072052. URL: https://dx.doi.org/10.1088/1742-6596/664/7/072052 (cited on page 99).

[249] Andrea Rizzi, Giovanni Petrucciani and Marco Peruzzi. “A further reduction in CMS event data for analysis: the NANOAOD format”. In: *EPJ Web of Conferences* 214 (September 2019), S06021. DOI: 10.1051/epjconf/201921406021. URL: https://doi.org/10.1051/epjconf/201921406021 (cited on page 99).

[250] R. Tomás et al. “Operational scenario of first high luminosity LHC run”. In: *Journal of Physics: Conference Series* 2420.1 (Jan. 2023), p. 012003. DOI: 10.1088/1742-6596/2420/1/012003. URL: https://dx.doi.org/10.1088/1742-6596/2420/1/012003 (cited on page 102).

[251] Markus Zerlauth and Oliver Bruning. "Status and prospects of the HL-LHC project". In: *PoS* EPS-HEP2023 (2024), p. 615. DOI: 10.22323/1.449.0615 (cited on pages 102, 103).

[252] Jory Sonneveld. *Silicon radiation damage at the LHC experiments*. Radiation effects in the LHC experiments and impact on operation and performance. URL: https://indico.cern.ch/event/769192/contributions/3288030/attachments/1793552/2925431/20190211_silicon_raddam_sonneveld.pdf Accessed: 24 June 2024. Feb. 2019 (cited on page 102).

[253] HiLumi HL-LHC Project. *The HL-LHC project*. https://hilumilhc.web.cern.ch/content/hl-lhc-project. Accessed: 10 August 2024. 2024 (cited on page 102).

[254] The CMS Collaboration. *Measurement of the Drell-Yan forward-backward asymmetry and of the effective leptonic weak mixing angle using proton-proton collisions at 13 TeV*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2893842 (cited on page 103).

[255] Andrea Wulzer. "An Equivalent Gauge and the Equivalence Theorem". In: *Nuclear Physics B* 885 (2014), pp. 97–126. ISSN: 0550-3213. DOI: https://doi.org/10.1016/j.nuclphysb.2014.05.021. URL: https://www.sciencedirect.com/science/article/pii/S0550321314001631 (cited on page 103).

[256] Michael S. Chanowitz and Mary K. Gaillard. "The TeV physics of strongly interacting W's and Z's". In: *Nuclear Physics B* 261 (1985), pp. 379–431. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(85)90580-2. URL: https://www.sciencedirect.com/science/article/pii/0550321385905802 (cited on page 104).

[257] The CDF Collaboration. "Measurement of the cross section for prompt isolated diphoton production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\,\text{TeV}$". In: *Physical Review D* 84.5 (Sept. 2011). ISSN: 1550-2368. DOI: 10.1103/physrevd.84.052006. URL: http://dx.doi.org/10.1103/PhysRevD.84.052006 (cited on page 104).

[258] The CMS Collaboration. *Observation of four top quark production in proton-proton collisions at* $\sqrt{s} = 13\,\text{TeV}$. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2853304 (cited on page 104).

[259] The ATLAS Collaboration. "Observation of four-top-quark production in the multilepton final state with the ATLAS detector". In: *The European Physical Journal C* 83.6 (June 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11573-0. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11573-0 (cited on page 104).

[260] Qing-Hong Cao, Shao-Long Chen, and Yandong Liu. "Probing Higgs width and top quark Yukawa coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ productions". In: *Physical Review D* 95.5 (Mar. 2017). ISSN: 2470-0029. DOI: 10.1103/physrevd.95.053004. URL: http://dx.doi.org/10.1103/PhysRevD.95.053004 (cited on page 104).

[261] M Albrow et al. *CMS-TOTEM Precision Proton Spectrometer*. Tech. rep. 2014. URL: https://cds.cern.ch/record/1753795 (cited on page 104).

[262] L. Adamczyk et al. *Technical Design Report for the ATLAS Forward Proton Detector*. Tech. rep. 2015. URL: https://cds.cern.ch/record/2017378 (cited on page 104).

[263] João Varela. *The CT-PPS project and timing detectors*. Accessed: 26 June 2024. June 2015. URL: https://indico.cern.ch/event/396441/contributions/1836636/ (cited on page 105).

[264] Michael Pitt. *Physics perspectives of a CMS near-beam proton spectrometer at HL-LHC*. Accessed: 26 June 2024. Mar. 2021. URL: https://indico.cern.ch/event/955960/contributions/4251256/ (cited on page 105).

[265] Andre David. "Higgs Boson: 10 Years Turning the Possible into the Known". In: *LHEP* 2023 (2023), p. 447. DOI: 10.31526/lhep.2023.447. URL: https://cds.cern.ch/record/2882582 (cited on page 106).

[266] Huilin Qu and Loukas Gouskos. "Jet tagging via particle clouds". In: *Phys. Rev. D* 101 (5 Mar. 2020), p. 056019. DOI: 10.1103/PhysRevD.101.056019. URL: https://link.aps.org/doi/10.1103/PhysRevD.101.056019 (cited on pages 106, 176, 177, 183, 239).

[267] The CMS Collaboration. "Search for ZZ and ZH production in the bbbb final state using proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *preprint arXiv:2403.20241* (2024). URL: https://arxiv.org/abs/2403.20241 (cited on pages 106, 240).

[268] The CMS Collaboration. "Development of the CMS detector for the CERN LHC Run 3". In: *Journal of Instrumentation* 19.05 (May 2024), P05064. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/05/p05064. URL: http://dx.doi.org/10.1088/1748-0221/19/05/P05064 (cited on page 106).

[269] The CMS Collaboration. *CMS Technical Design Report for the Pixel Detector Upgrade*. Tech. rep. 2012. URL: https://cds.cern.ch/record/1481838 (cited on page 106).

[270] J. Christiansen and M. Garcia-Sciveres. *RD Collaboration Proposal: Development of pixel readout integrated circuits for extreme rate and radiation*. Tech. rep. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1553467 (cited on page 107).

[271] Alkiviadis Papadopoulos. "Design and construction of the CMS Inner Tracker for the HL-LHC Upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1064 (2024), p. 169409. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2024.169409. URL: https://www.sciencedirect.com/science/article/pii/S0168900224003358 (cited on page 107).

[272] D. Abbaneo. "The CMS tracker for HL-LHC". In: *CMS Workshop: Perspectives on Physics and on CMS at Very High Luminosity, HL-LHC*. 2013, pp. 179–191 (cited on page 107).

[273] The CMS Collaboration. *A MIP Timing Detector for the CMS Phase-2 Upgrade*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2667167 (cited on pages 107, 109).

[274] Olmo Cerri. *CMS precision timing physics impact for the HL-LHC upgrade*. Tech. rep. 2018. arXiv: 1810.00860. URL: https://cds.cern.ch/record/2641474 (cited on page 107).

[275] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Tech. rep. This is the final version, approved by the LHCC. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283187 (cited on pages 110, 128).

[276] R. Jiménez Estupiñán et al. "Phase-II Upgrade of the CMS Electromagnetic Calorimeter Detector Control and Safety Systems for the High Luminosity Large Hadron Collider". In: *Proc. 19th Int. Conf. Accel. Large Exp. Phys. Control Syst. (ICALEPCS'23)* (Cape Town, South Africa, Oct. 9–13, 2023). International Conference on Accelerator and Large Experimental Physics Control Systems 19. JACoW Publishing, Geneva, Switzerland, Feb. 2024, pp. 1516–1521. ISBN: 978-3-95450-238-7. DOI: 10.18429/JACoW-ICALEPCS2023-THPDP074. URL: https://jacow.org/icalepcs2023/papers/thpdp074.pdf (cited on page 110).

[277] A. Bornheim. "Design studies for the Phase II upgrade of the CMS Barrel Electromagnetic Calorimeter". In: *Journal of Instrumentation* 12.03 (Mar. 2017), p. C03018. DOI: 10.1088/1748-0221/12/03/C03018. URL: https://dx.doi.org/10.1088/1748-0221/12/03/C03018 (cited on page 110).

[278] Gabriella Pásztor. "The Phase-2 Upgrade of the CMS Detector". In: *PoS* LHCP2022 (2023), p. 045. DOI: 10.22323/1.422.0045. URL: https://cds.cern.ch/record/2880161 (cited on page 111).

[279] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger*. Tech. rep. This is the final version of the document, approved by the LHCC. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2759072 (cited on pages 113, 117, 129).

[280] J. Duarte et al. "Fast inference of deep neural networks in FPGAs for particle physics". In: *Journal of Instrumentation* 13.07 (July 2018), P07027–P07027. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/07/p07027. URL: http://dx.doi.org/10.1088/1748-0221/13/07/P07027 (cited on page 116).

[281] Thiago R. F. P. Tomei. *The High-Level Trigger for the CMS Phase-2 Upgrade*. 2022. arXiv: 2211.03684 [hep-ex]. URL: https://arxiv.org/abs/2211.03684 (cited on page 117).

[282] Bruno Alves, Felice Pantaleo, and Marco Rovere. "Clustering in the Heterogeneous Reconstruction Chain of the CMS HGCAL Detector". In: *Journal of Physics: Conference Series* 2438.1 (Feb. 2023), p. 012015. DOI: 10.1088/1742-6596/2438/1/012015. URL: https://dx.doi.org/10.1088/1742-6596/2438/1/012015 (cited on page 117).

[283] Bruno Alves et al. "Heterogeneous techniques for rescaling energy deposits in the CMS Phase-2 endcap calorimeter". In: *EPJ Web Conf.* 251 (2021), p. 04017. DOI: 10.1051/epjconf/202125104017. URL: https://doi.org/10.1051/epjconf/202125104017 (cited on page 117).

[284] A. Matthes et al. “Tuning and optimization for a variety of many-core architectures without changing a single line of implementation code using the Alpaka library”. In: June 2017. arXiv: 1706.10086. URL: https://arxiv.org/abs/1706.10086 (cited on page 118).

[285] Erik Zenker et al. “Alpaka - An Abstraction Library for Parallel Kernel Acceleration”. In: IEEE Computer Society, May 2016. arXiv: 1602.08477. URL: http://arxiv.org/abs/1602.08477 (cited on page 118).

[286] Benjamin Worpitz. *Investigating performance portability of a highly scalable particle-in-cell simulation code on various multi-core architectures*. Master Thesis. Sept. 2015. DOI: 10.5281/zenodo.49768. URL: http://dx.doi.org/10.5281/zenodo.49768 (cited on page 118).

[287] Christian R. Trott et al. “Kokkos 3: Programming Model Extensions for the Exascale Era”. In: *IEEE Transactions on Parallel and Distributed Systems* 33.4 (2022), pp. 805–817. DOI: 10.1109/TPDS.2021.3097283 (cited on page 118).

[288] H. Carter Edwards, Christian R. Trott, and Daniel Sunderland. “Kokkos: Enabling many-core performance portability through polymorphic memory access patterns”. In: *Journal of Parallel and Distributed Computing* 74.12 (2014). Domain-Specific Languages and High-Level Frameworks for High-Performance Computing, pp. 3202–3216. ISSN: 0743-7315. DOI: https://doi.org/10.1016/j.jpdc.2014.07.003. URL: http://www.sciencedirect.com/science/article/pii/S0743731514001257 (cited on page 118).

[289] Ganesh Parida. “Run-3 Commissioning of CMS Online HLT reconstruction using GPUs”. In: *EPJ Web of Conferences* 295 (May 2024), S11020. DOI: 10.1051/epjconf/202429511020. URL: https://doi.org/10.1051/epjconf/202429511020 (cited on page 118).

[290] Andrea Bocci et al. “Performance portability for the CMS Reconstruction with Alpaka”. In: *J. Phys. Conf. Ser.* 2438.1 (2023), p. 012058. DOI: 10.1088/1742-6596/2438/1/012058 (cited on page 118).

[291] Richard Wigmans. “New developments in calorimetric particle detection”. In: *Progress in Particle and Nuclear Physics* 103 (Nov. 2018), pp. 109–161. ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2018.07.003. URL: http://dx.doi.org/10.1016/j.ppnp.2018.07.003 (cited on pages 118, 125).

[292] The CMS Electromagnetic Calorimeter Group. “Radiation hardness qualification of PbWO4 scintillation crystals for the CMS Electromagnetic Calorimeter”. In: *Journal of Instrumentation* 5.03 (Mar. 2010), P03010. DOI: 10.1088/1748-0221/5/03/P03010. URL: https://dx.doi.org/10.1088/1748-0221/5/03/P03010 (cited on page 118).

[293] Richard Wigmans and Michele Livan. *Calorimetry for Collider Physics, an Introduction*. UNITEXT for Physics. Springer Cham, July 2019. ISBN: 978-3-030-23652-6. DOI: 10.1007/978-3-030-23653-3. URL: https://doi-org.ezproxy.cern.ch/10.1007/978-3-030-23653-3 (cited on pages 118, 123).

[294] D. Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. Upgrade Project Leader Deputies: Lucia Silvestris (INFN-Bari), Jeremy Mans (University of Minnesota) Additional contacts: Lucia.Silvestris@cern.ch, Jeremy.Mans@cern.ch. Geneva, 2015. DOI: 10.17181/CERN.VU8I.D59J. URL: https://cds.cern.ch/record/2020886 (cited on page 118).

[295] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. CMS-TDR-019. CERN, 2018. DOI: 10.17181/CERN.IV8M.1JY2 (cited on pages 118, 119, 121–123, 126, 127, 129).

[296] Brad Cox. "The Shashlik Calorimeter, a LYSO/W plate Calorimeter for Precision EM Calorimeter in the High Luminosity LHC". In: *PoS* ICHEP2016 (2016), p. 239. DOI: 10.22323/1.282.0239. URL: https://cds.cern.ch/record/2310100 (cited on page 118).

[297] The CMS Collaboration. *HGCAL Website: Material budget from vertex up to in front of Muon Stations*. https://hgcal.web.cern.ch/MaterialBudget/MaterialBudgetFromVertexUpToInfrontOfMuonStations/. Accessed: 21 June 2023 (cited on pages 119, 139).

[298] Alessandro Tarabini. "Measurement of Higgs boson fiducial cross sections with the CMS detector and electromagnetic reconstruction with the high-granularity endcap calorimeter". Ec. Polytech., Palaiseau, 2024. URL: https://cds.cern.ch/record/2891919 (cited on page 121).

[299] Helmuth Spieler. *Silicon Detectors - presentation*. http://indico.cern.ch/event/124392/contributions/1339904/attachments/74582/106976/IntroSilicon.pdf. 2012 (cited on page 121).

[300] The CMS Collaboration. *Lateral layout of Si sensors in the endcap calorimeter for Phase 2*. Tech. rep. Geneva: CERN, 2024 (cited on page 122).

[301] E. Fretwurst et al. *Radiation Damage in Silicon Detectors Caused by Hadronic and Electromagnetic Irradiation*. 2002. arXiv: physics/0211118 [physics.ins-det]. URL: https://arxiv.org/abs/physics/0211118 (cited on page 122).

[302] Cor Claeys and Eddy Simoen. "Basic Radiation Damage Mechanisms in Semiconductor Materials and Devices". In: *Radiation Effects in Advanced Semiconductor Materials and Devices*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 9–52. ISBN: 978-3-662-04974-7. DOI: 10.1007/978-3-662-04974-7_2. URL: https://doi.org/10.1007/978-3-662-04974-7_2 (cited on page 122).

[303] S. Tadeja and M. Guthoff and N. Smirnov and P. Kicsiny and A. Lokhovitksiy. *Radiation Simulation Plotting tool v.2.4*. https://radiationtool.web.cern.ch/rpt/. Accessed: 2 July 2024. 2024 (cited on page 124).

[304] T.T. Böhlen et al. "The FLUKA Code: Developments and Challenges for High Energy and Medical Applications". In: *Nuclear Data Sheets* 120 (2014), pp. 211–214. ISSN: 0090-3752. DOI: https://doi.org/10.1016/j.nds.2014.07.049. URL: https://www.sciencedirect.com/science/article/pii/S0090375214005018 (cited on page 124).

[305] Alfredo Ferrari et al. “FLUKA: a multi-particle transport code”. In: *CERN Yellow report* 2005-10 (Jan. 2005). DOI: 10.2172/877507 (cited on page 124).

[306] Felix Sefkow and Frank Simon. “A highly granular SiPM-on-tile calorimeter prototype”. In: *Journal of Physics: Conference Series* 1162 (Jan. 2019), p. 012012. ISSN: 1742-6596. DOI: 10.1088/1742-6596/1162/1/012012. URL: http://dx.doi.org/10.1088/1742-6596/1162/1/012012 (cited on page 123).

[307] The CMS HGCAL Collaboration. “Response of a CMS HGCAL silicon-pad electromagnetic calorimeter prototype to 20–300 GeV positrons”. In: *Journal of Instrumentation* 17.05 (May 2022), P05022. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/05/p05022. URL: http://dx.doi.org/10.1088/1748-0221/17/05/P05022 (cited on page 125).

[308] The CMS HGCAL Collaboration. *Performance of the CMS High Granularity Calorimeter prototype to charged pion beams of 20−300 GeV/c*. 2023. arXiv: 2211.04740 [physics.ins-det]. URL: https://arxiv.org/abs/2211.04740 (cited on page 125).

[309] The CMS HGCAL Collaboration. “Using graph neural networks to reconstruct charged pion showers in the CMS High Granularity Calorimeter”. In: (June 2024). arXiv: 2406.11937 [physics.ins-det] (cited on page 125).

[310] The CMS HGCAL Collaboration. “Timing performance of the CMS High Granularity Calorimeter prototype”. In: *Journal of Instrumentation* 19.04 (Apr. 2024), P04015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/04/p04015. URL: http://dx.doi.org/10.1088/1748-0221/19/04/P04015 (cited on page 125).

[311] J. Borg et al. “SKIROC2_CMS an ASIC for testing CMS HGCAL”. In: *Journal of Instrumentation* 12.02 (Feb. 2017), p. C02019. DOI: 10.1088/1748-0221/12/02/C02019. URL: https://dx.doi.org/10.1088/1748-0221/12/02/C02019 (cited on page 125).

[312] Terzo S. et al. “Novel 3D Pixel Sensors for the Upgrade of the ATLAS Inner Tracker”. In: 9 (2021), p. 624668. DOI: 10.3389/fphy.2021.624668. URL: https://dx.doi.org/10.3389/fphy.2021.624668 (cited on page 125).

[313] Federico Lasagni Manghi. *ATLAS Upgrade for HL-LHC*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2902993 (cited on page 126).

[314] The LHCb Collaboration. *Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2244311 (cited on page 126).

[315] The ALICE Collaboration. *Technical Design report for the ALICE Inner Tracking System 3 - ITS3 ; A bent wafer-scale monolithic pixel detector*. Tech. rep. Co-project Manager: Magnus Mager, magnus.mager@cern.chds. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2890181 (cited on page 126).

[316] The ALICE Collaboration. *Technical Design Report of the ALICE Forward Calorimeter (FoCal)*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2890281 (cited on page 126).

[317] The ALICE Collaboration. *Letter of intent for ALICE 3: A next generation heavy-ion experiment at the LHC*. Tech. rep. 202 pages, 103 captioned figures, 19 tables. Geneva: CERN, 2022. arXiv: 2211.02491. URL: https://cds.cern.ch/record/2803563 (cited on page 126).

[318] Jean-Baptiste Sauvan. *HDR: Algorithms for the Level 1 Trigger Primitives Generation in the High-Granularity Calorimeter at the High-Luminosity LHC*. Tech. rep. LLR, 2022 (cited on page 126).

[319] Erich Frahm, Jeremiah Mans, and Nadja Strobbe. *CMS HGC Scintillator Motherboard Design*. Tech. rep. CMS internal document (restricted). Geneva: CERN, CMS-CE-ES-0230 (cited on page 127).

[320] PICMG. *R3.0 AdvancedTCA Base Specification*. https://www.picmg.org/openstandards/advancedtca. Accessed: 14 July 2024 (cited on page 128).

[321] Andrew Rose et al. "Serenity: An ATCA prototyping platform for CMS Phase-2". In: *Proceedings of Topical Workshop on Electronics for Particle Physics — PoS(TWEPP2018)* (May 2019). DOI: 10.22323/1.343.0115. URL: http://dx.doi.org/10.22323/1.343.0115 (cited on pages 128, 132).

[322] lbGBT team. *The lpGBTv1 manual* (cited on pages 128, 131).

[323] Jan Troska et al. "The VTRx+, an optical link module for data transmission at HL-LHC". In: *PoS* TWEPP-17 (2017), p. 048. DOI: 10.22323/1.313.0048 (cited on page 129).

[324] François Vasey. *Versatile Link+ Status and Plans*. https://espace.cern.ch/project-Versatile-Link-Plus/Shared%20Documents/Presentations/2020_VLplus_ACES_Vasey_26May2020.pdf. Accessed: 15 July 2024 (cited on page 129).

[325] Samtec. *Firefly Micro Flyover System*. https://www.samtec.com/optics/systems/firefly/. Accessed: 15 July 2024 (cited on page 129).

[326] Paul Aspell et al. *HGCROC3: Working document on specification*. https://edms.cern.ch/file/2324379/1/HGCROC3_Spec_Working_Document_v2.0.pdf. CMS internal document (restricted). 2021 (cited on pages 129, 130).

[327] F. Bouyjou et al. "HGCROC3: the front-end readout ASIC for the CMS High Granularity Calorimeter". In: *Journal of Instrumentation* 17.03 (Mar. 2022), p. C03015. DOI: 10.1088/1748-0221/17/03/C03015. URL: https://dx.doi.org/10.1088/1748-0221/17/03/C03015 (cited on pages 129, 130).

[328] The CMS Collaboration. *The HGCAL Backend TDAQ Systems for the Pre-series Design Review*. Tech. rep. Geneva: CERN, 2022 (cited on pages 130, 132, 133).

[329] Iouliia Skliarova. "A Survey of Network-Based Hardware Accelerators". In: *Electronics* 11.7 (2022), p. 1029. DOI: https://doi.org/10.3390/electronics11071029. URL: https://www.mdpi.com/2079-9292/11/7/1029 (cited on page 131).

[330] Louis Portalès. "L1 Triggering on High-Granularity Information at the HL-LHC". In: *Instruments* 6.4 (2022). ISSN: 2410-390X. DOI: 10.3390/instruments6040071. URL: https://www.mdpi.com/2410-390X/6/4/71 (cited on page 131).

[331] K. E. Batcher. “Sorting Networks and Their Applications”. In: *Proceedings of the April 30–May 2, 1968, Spring Joint Computer Conference*. AFIPS ’68 (Spring). Atlantic City, New Jersey, 1968, pp. 307–314. ISBN: 9781450378970. DOI: 10.1145/1468075.1468121 (cited on page 131).

[332] Cristina Martin Perez, Jean-Baptiste Sauvan, and S. Webb. *Updates on FE data reduction studies*. Accessed: 17 July 2024. Jan. 2020. URL: https://indico.cern.ch/event/879499/#17-summary-of-fe-studies (cited on page 131).

[333] Cristina Martin Perez. *Search for the Higgs Boson Produced in Association with Top Quarks with the CMS Detector at the LHC*. Series Title: Springer Theses. Springer Cham, 2022. ISBN: 978-3-030-90208-7. DOI: 10.1007/978-3-030-90206-3. URL: https://link.springer.com/book/10.1007/978-3-030-90206-3 (cited on page 131).

[334] James Hirschauer and Ralph Wickwire. *Trigger Concentrator ASIC (ECON-T) for the CMS HL-LHC endcap calorimeter*. https://edms.cern.ch/file/2206779/1/ECON-T_specification_working_doc_v9_2mar2022.pdf. CMS internal document (restricted). 2022 (cited on page 131).

[335] G. van Rossum. *Python tutorial*. Tech. rep. CS-R9526. Amsterdam: Centrum voor Wiskunde en Informatica (CWI), May 1995 (cited on page 136).

[336] Charles R. Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2 (cited on page 136).

[337] Jim Pivarski et al. *Uproot*. Sept. 2017. DOI: 10.5281/zenodo.4340632. URL: https://github.com/scikit-hep/uproot5 (cited on page 136).

[338] Enrico Guiraud, Axel Naumann, and Danilo Piparo. *TDataFrame: functional chains for ROOT data analyses*. Version v1.0. Jan. 2017. DOI: 10.5281/zenodo.260230. URL: https://doi.org/10.5281/zenodo.260230 (cited on page 136).

[339] Miguel Grinberg. *Flask web development: developing web applications with python*. O’Reilly Media, Inc., 2018 (cited on page 137).

[340] Amit Patel. *Hexagonal Grids*. URL: https://www.redblobgames.com/grids/hexagons/. Accessed: 19 July 2024. Oct. 2021 (cited on page 137).

[341] Apache. *Parquet*. URL: https://www.redblobgames.com/grids/hexagons/. Accessed: 19 July 2024. 2013 (cited on page 140).

[342] Sudah Ahujda. *Optimization studies for egamma*. Accessed: 19 July 2024. Oct. 2019. URL: https://indico.cern.ch/event/857805/#4-optimization-studies-for-ega (cited on page 143).

[343] AMD. *Vitis HLS*. URL: https://www.xilinx.com/products/design-tools/vitis/vitis-hls.html. Accessed: 21 July 2024. Oct. 2024 (cited on page 152).

[344] Atilim Gunes Baydin et al. “Automatic differentiation in machine learning: a survey”. In: *ArXiv* abs/1502.05767 (2015). URL: https://api.semanticscholar.org/CorpusID:3766791 (cited on page 152).

[345] The ATLAS Collaboration. *Search for resonant and nonresonant Higgs boson pair production in the $bb\tau\tau$ decay channel using 13 TeV pp collision data from the ATLAS detector*. 2022. DOI: 10.48550/ARXIV.2209.10910. URL: https://arxiv.org/abs/2209.10910 (cited on pages 153, 236, 264).

[346] HIG MC group. *Summary table of samples produced for the 1 Billion campaign, with 25ns bunch-crossing*. Twiki. Accessed: 2024-09-19 (cited on pages 159, 160).

[347] The CMS Collaboration, GEN group. *Standard Model Cross Sections for CMS at $\sqrt{s} = 13\,\text{TeV}$*. Twiki. Accessed: 2024-09-19 (cited on page 159).

[348] The CMS Collaboration, GEN group. *How to Compute Cross Sections with the GenXSecAnalyzer*. Twiki. Accessed: 2024-09-19 (cited on pages 159, 160).

[349] The CMS Collaboration. *Measurement of differential $\text{t}\bar{\text{t}}$ production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$*. 2021. DOI: 10.1103/PhysRevD.104.092013. arXiv: 2108.02803 [hep-ex] (cited on page 159).

[350] LHC Higgs cross-section Working Group. *SM Higgs production cross sections at $\sqrt{s} = 13$, 14 and 27 TeV TeV (update in CERN HL-LHC YR 2019)* (cited on page 160).

[351] The ATLAS Collaboration. "Search for the nonresonant production of Higgs boson pairs via gluon fusion and vector-boson fusion in the $bb\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Physical Review D* 110.3 (Aug. 2024). ISSN: 2470-0029. DOI: 10.1103/physrevd.110.032012. URL: http://dx.doi.org/10.1103/PhysRevD.110.032012 (cited on pages 167, 261).

[352] Victor Lendermann et al. "Combining Triggers in HEP data analysis". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 604.3 (2009), pp. 707–718. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2009.03.173. URL: https://www.sciencedirect.com/science/article/pii/S0168900209006627 (cited on page 171).

[353] A. Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*. 2009. arXiv: physics/0703039 [physics.data-an]. URL: https://arxiv.org/abs/physics/0703039 (cited on page 174).

[354] The CMS Collaboration. "Identification of hadronic tau lepton decays using a deep neural network". In: *Journal of Instrumentation* 17.07 (July 2022), P07023. DOI: 10.1088/1748-0221/17/07/p07023. URL: https://doi.org/10.1088%2F1748-0221%2F17%2F07%2Fp07023 (cited on pages 175, 176, 180).

[355] The CMS Collaboration. "Performance of the CNN-based tau identification algorithm with Domain Adaptation using Adversarial Machine Learning for Run 2". In: (2024). URL: https://cds.cern.ch/record/2904699 (cited on page 175).

[356] B. Vormwald. "The CMS Phase-1 pixel detector – experience and lessons learned from two years of operation". In: *Journal of Instrumentation* 14.07 (July 2019), p. C07008. DOI: 10.1088/1748-0221/14/07/C07008. URL: https://dx.doi.org/10.1088/1748-0221/14/07/C07008 (cited on page 176).

[357] Schunk L. Marzani S. and Soyez G. "The jet mass distribution after Soft Drop". In: *Eur Phys J C Part Fields* 78.2 (Feb. 2018). DOI: 10.1140/epjc/s10052-018-5579-5. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6560847/ (cited on page 176).

[358] The CMS Collaboration. *Identification of b quark jets at the CMS Experiment in the LHC Run 2*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2138504 (cited on page 177).

[359] The CMS Collaboration. In: *Journal of Instrumentation* 13.05 (May 2018), P05011. DOI: 10.1088/1748-0221/13/05/P05011. URL: https://dx.doi.org/10.1088/1748-0221/13/05/P05011 (cited on page 177).

[360] 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. URL: https://dx.doi.org/10.1088/1748-0221/15/12/P12012 (cited on page 177).

[361] The CMS Collaboration. "Performance summary of AK4 jet b tagging with data from proton-proton collisions at 13 TeV with the CMS detector". In: (2023). URL: https://cds.cern.ch/record/2854609 (cited on page 177).

[362] Ashish Vaswani et al. *Attention Is All You Need*. 2023. arXiv: 1706.03762 [cs.CL]. URL: https://arxiv.org/abs/1706.03762 (cited on page 177).

[363] The CMS Collaboration. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *Journal of Instrumentation* 6.11 (Nov. 2011), P11002–P11002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/6/11/p11002. URL: http://dx.doi.org/10.1088/1748-0221/6/11/P11002 (cited on page 178).

[364] The CMS Collaboration. "Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV". In: (2021). URL: https://cds.cern.ch/record/2792322 (cited on pages 178, 179).

[365] The CMS Collaboration. "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018". In: (2020). URL: https://cds.cern.ch/record/2715872 (cited on page 179).

[366] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735. eprint: https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf. URL: https://doi.org/10.1162/neco.1997.9.8.1735 (cited on page 182).

[367] François Chollet et al. *Keras*. https://keras.io. 2015 (cited on pages 183, 190).

[368] Zhilu Zhang and Mert Sabuncu. "Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio et al. Vol. 31. Curran Associates, Inc., 2018, pp. 8778–8788. URL: https://proceedings.neurips.cc/paper/2018/file/f2925f97bc13ad2852a7a551802feea0-Paper.pdf (cited on page 183).

[369] Ilya Loshchilov and Frank Hutter. *Decoupled Weight Decay Regularization*. 2019. arXiv: 1711.05101 [cs.LG]. URL: https://arxiv.org/abs/1711.05101 (cited on pages 183, 189).

[370] Lorenzo Bianchini et al. "Reconstruction of the Higgs mass in $H \to \tau\tau$ Events by Dynamical Likelihood techniques". In: *Journal of Physics: Conference Series* 513.2 (June 2014), p. 022035. DOI: 10.1088/1742-6596/513/2/022035. URL: https://dx.doi.org/10.1088/1742-6596/513/2/022035 (cited on page 187).

[371] Luca Lista. *Statistical Methods for Data Analysis, With Applications in Particle Physics*. Lecture Notes in Physics. Springer International Publishing, 2023, nil. DOI: 10.1007/978-3-031-19934-9. URL: http://dx.doi.org/10.1007/978-3-031-19934-9 (cited on pages 188, 221, 233).

[372] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. "Learning Representations by Back-Propagating Errors". In: *Cognitive Modeling*. Cognitive Modeling. The MIT Press, 2002, pp. 213–222. DOI: 10.7551/mitpress/1888.003.0013. URL: http://dx.doi.org/10.7551/mitpress/1888.003.0013 (cited on pages 188, 217).

[373] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. *Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)*. 2016. arXiv: 1511.07289 [cs.LG]. URL: https://arxiv.org/abs/1511.07289 (cited on page 188).

[374] Sergey Ioffe and Christian Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. 2015. arXiv: 1502.03167 [cs.LG]. URL: https://arxiv.org/abs/1502.03167 (cited on pages 188, 219).

[375] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/ (cited on page 190).

[376] Lucas Corcodilos. "The 2D Alphabet background modeling method and its use in the search for an excited bottom quark". PhD thesis. Johns Hopkins University, 2021. URL: https://inspirehep.net/literature/2047785 (cited on page 194).

[377] K. Rose. "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems". In: *Proceedings of the IEEE* 86.11 (1998), pp. 2210–2239. DOI: 10.1109/5.726788 (cited on page 198).

[378] The CMS Collaboration. "Observation of $\gamma\gamma \to \tau\tau$ in proton–proton collisions and limits on the anomalous electromagnetic moments of the τ lepton". In: *Reports on Progress in Physics* 87.10 (Sept. 2024), p. 107801. DOI: 10.1088/1361-6633/ad6fcb. URL: https://dx.doi.org/10.1088/1361-6633/ad6fcb (cited on page 201).

[379] The CMS Collaboration. "Calibration of the mass-decorrelated ParticleNet tagger for boosted $b\bar{b}$ and $c\bar{c}$ jets using LHC Run 2 data". In: (2022). URL: https://cds.cern.ch/record/2805611 (cited on page 211).

[380] The CMS Collaboration. "Search for Higgs Boson Decay to a Charm Quark-Antiquark Pair in Proton-Proton Collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Physical Review Letters* 131.6 (Aug. 2023). ISSN: 1079-7114. DOI: 10.1103/physrevlett.131.061801. URL: http://dx.doi.org/10.1103/PhysRevLett.131.061801 (cited on page 211).

[381] Yanyan Gao et al. "Spin determination of single-produced resonances at hadron colliders". In: *Physical Review D* 81.7 (Apr. 2010). ISSN: 1550-2368. DOI: 10.1103/physrevd.81.075022. URL: http://dx.doi.org/10.1103/PhysRevD.81.075022 (cited on page 217).

[382] Sara Bolognesi et al. "Spin and parity of a single-produced resonance at the LHC". In: *Physical Review D* 86.9 (Nov. 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.095031. URL: http://dx.doi.org/10.1103/PhysRevD.86.095031 (cited on page 217).

[383] Ian Anderson et al. "Constraining anomalous HVV interactions at proton and lepton colliders". In: *Physical Review D* 89.3 (Feb. 2014). ISSN: 1550-2368. DOI: 10.1103/physrevd.89.035007. URL: http://dx.doi.org/10.1103/PhysRevD.89.035007 (cited on page 217).

[384] Andrei V. Gritsan et al. "Constraining anomalous Higgs boson couplings to the heavy-flavor fermions using matrix element techniques". In: *Physical Review D* 94.5 (Sept. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.055023. URL: http://dx.doi.org/10.1103/PhysRevD.94.055023 (cited on page 217).

[385] Andrei V. Gritsan et al. "New features in the JHU generator framework: Constraining Higgs boson properties from on-shell and off-shell production". In: *Physical Review D* 102.5 (Sept. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.056022. URL: http://dx.doi.org/10.1103/PhysRevD.102.056022 (cited on page 217).

[386] F. Rosenblatt. "The Perceptron: A probabilistic model for information storage and organization in the brain". In: *Psychological Review* 65.6 (1957), pp. 386–408 (cited on page 217).

[387] Seppo Linnainmaa. "Taylor Expansion of the Accumulated Rounding Error". In: *BIT* 16.2 (1976), pp. 146–160. DOI: 10.1007/bf01931367. URL: http://dx.doi.org/10.1007/bf01931367 (cited on page 217).

[388] The CMS Collaboration. *Prospects for HH measurements at the HL-LHC*. Tech. rep. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2652549 (cited on page 218).

[389] Baldi, Pierre and Cranmer, Kyle and Faucett, Taylor and Sadowski, Peter and Whiteson, Daniel. "Parameterized neural networks for high-energy physics". In: *Eur. Phys. J.* C76.5 (2016), p. 235. DOI: 10.1140/epjc/s10052-016-4099-4. arXiv: 1601.07913 [hep-ex] (cited on page 218).

[390] Angela Giraldi. *Optimisation of a multivariate analysis technique for the* $\mathbf{t\bar{t}}$ *background rejection in the search for Higgs boson pair production in* $\mathbf{b\bar{b}}\tau^+\tau^-$ *decay channel with the CMS experiment at the LHC*. Presented 20 Apr 2018. Mar. 2018. URL: https://cds.cern.ch/record/2623854 (cited on page 218).

[391] Gao Huang et al. *Densely Connected Convolutional Networks*. 2018. arXiv: 1608.06993 [cs.CV] (cited on page 219).

[392] M. Erdmann et al. "Lorentz Boost Networks: autonomous physics-inspired feature engineering". In: *Journal of Instrumentation* 14.06 (June 2019), P06006–P06006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/06/p06006. URL: http://dx.doi.org/10.1088/1748-0221/14/06/P06006 (cited on page 219).

[393] Hao Li et al. *Visualizing the Loss Landscape of Neural Nets*. 2018. arXiv: 1712.09913 [cs.LG] (cited on page 219).

[394] Kaiming He et al. *Deep Residual Learning for Image Recognition*. 2015. arXiv: 1512.03385 [cs.CV] (cited on page 219).

[395] Geoffrey E. Hinton et al. *Improving neural networks by preventing co-adaptation of feature detectors*. 2012. arXiv: 1207.0580 [cs.NE]. URL: https://arxiv.org/abs/1207.0580 (cited on page 219).

[396] Prajit Ramachandran and Barret Zoph and Quoc V. Le. "Searching for Activation Functions". In: *CoRR* abs/1710.05941 (2017). arXiv: 1710.05941. URL: http://arxiv.org/abs/1710.05941 (cited on page 219).

[397] Cheng Guo and Felix Berkhahn. "Entity Embeddings of Categorical Variables". In: *CoRR* abs/1604.06737 (2016). arXiv: 1604.06737. URL: http://arxiv.org/abs/1604.06737 (cited on page 219).

[398] Kaiming He et al. *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. 2015. arXiv: 1502.01852 [cs.CV]. URL: https://arxiv.org/abs/1502.01852 (cited on page 219).

[399] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. Ed. by Yoshua Bengio and Yann LeCun. 2015. URL: http://arxiv.org/abs/1412.6980 (cited on page 220).

[400] R. Barlow and C. Beeston. "Fitting using finite Monte Carlo samples". In: *Computer Physics Communications* 77.2 (1993), pp. 219–228. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(93)90005-W (cited on page 228).

[401] A. L. Read. "Presentation of search results: the CLs technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002). (From the workshop 'Advanced Statistical Techniques in Particle Physics', 18–22 March 2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313. URL: https://dx.doi.org/10.1088/0954-3899/28/10/313 (cited on pages 228, 233).

[402] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). DOI: 10.1140/epjc/s10052-011-1554-0. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-011-1554-0 (cited on pages 228, 232, 234).

[403] The LHC Higgs Combination Group. *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1379837 (cited on pages 228, 233).

[404] Glen Cowan. *Statistical Data Analysis*. Vol. 1998. Oxford Science Publications. Great Clarendon Street, Oxfort OX2 6DP: Oxfort University Press, 1998. ISBN: 0-19-850156-0 (cited on page 229).

[405] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360 (cited on page 232).

[406] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2 (1999), pp. 435–443. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(99)00498-2. URL: https://www.sciencedirect.com/science/article/pii/S0168900299004982 (cited on page 233).

[407] The CMS Collaboration. "Search for heavy resonances decaying into two Higgs bosons or into a Higgs boson and a W or Z boson in proton-proton collisions at 13 TeV". In: *Journal of High Energy Physics* 2019.1 (Jan. 2019). ISSN: 1029-8479. DOI: 10.1007/jhep01(2019)051. URL: http://dx.doi.org/10.1007/JHEP01(2019)051 (cited on page 236).

[408] Alberto Belvedere et al. "Dispelling the $\sqrt{\mathcal{L}}$ myth for the High-Luminosity LHC". In: *Eur. Phys. J. C* 84.7 (2024), p. 715. DOI: 10.1140/epjc/s10052-024-13032-w. arXiv: 2402.07985 [hep-ph] (cited on page 239).

[409] The CMS Collaboration. "Jet energy scale and resolution of jets with ParticleNet p_{T} regression using Run3 data collected by the CMS experiment in 2022 and 2023 at 13.6 TeV". In: (2024). URL: https://cds.cern.ch/record/2904704 (cited on page 240).

[410] Pablo de Castro Manzano et al. "Hemisphere Mixing: a fully data-driven model of QCD multijet backgrounds for LHC searches". In: *PoS* EPS-HEP2017 (2017), p. 370. DOI: 10.22323/1.314.0370. URL: https://pos.sissa.it/314/370/pdf (cited on page 240).

[411] D. Berdine, N. Kauer, and D. Rainwater. "Breakdown of the Narrow Width Approximation for New Physics". In: *Phys. Rev. Lett.* 99 (11 Sept. 2007), p. 111601. DOI: 10.1103/PhysRevLett.99.111601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.99.111601 (cited on page 243).

[412] The CMS Collaboration. "A measurement of the Higgs boson mass in the diphoton decay channel". In: *Phys. Lett. B* 805 (2020), p. 135425. DOI: 10.1016/j.physletb.2020.135425. arXiv: 2002.06398 [hep-ex] (cited on page 243).

[413] The ATLAS and CMS Collaborations. "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex] (cited on page 243).

[414] Marcela Carena, Zhen Liu, and Marc Riembau. "Probing the electroweak phase transition via enhanced di-Higgs boson production". In: *Phys. Rev. D* 97 (9 May 2018), p. 095032. DOI: 10.1103/PhysRevD.97.095032. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.095032 (cited on page 243).

[415] Papaefstathiou, Andreas and White, Graham. "The electro-weak phase transition at colliders: confronting theoretical uncertainties and complementary channels". In: *JHEP* 05 (2021), p. 099. DOI: 10.1007/JHEP05(2021)099. arXiv: 2010.00597 [hep-ph] (cited on pages 244, 246, 248, 249).

[416] Papaefstathiou, Andreas and Robens, Tania and Tetlalmatzi-Xolocotzi, Gilberto. "Triple Higgs boson production at the Large Hadron Collider with Two Real Singlet scalars". In: *Journal of High Energy Physics* 2021.5 (May 2021). ISSN: 1029-8479. DOI: 10.1007/jhep05(2021)193. URL: http://dx.doi.org/10.1007/jhep05(2021)193 (cited on page 244).

[417] Steven D. Bass, Albert De Roeck, and Marumi Kado. "The Higgs boson implications and prospects for future discoveries". en. In: *Nature Reviews Physics* 3.9 (Sept. 2021), pp. 608–624. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00341-2. URL: https://doi.org/10.1038/s42254-021-00341-2 (cited on page 244).

[418] K. Kajantie et al. "Is There a Hot Electroweak Phase Transition at $m_H \gtrsim m_W$?" In: *Phys. Rev. Lett.* 77 (14 Sept. 1996), pp. 2887–2890. DOI: 10.1103/PhysRevLett.77.2887. URL: https://link.aps.org/doi/10.1103/PhysRevLett.77.2887 (cited on page 244).

[419] Chiara Caprini et al. "Detecting gravitational waves from cosmological phase transitions with LISA: an update". In: *Journal of Cosmology and Astroparticle Physics* 2020.03 (Mar. 2020), pp. 024–024. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/03/024. URL: http://dx.doi.org/10.1088/1475-7516/2020/03/024 (cited on page 244).

[420] Michele Armano et al. *LISA Pathfinder*. 2019. arXiv: 1903.08924 [astro-ph.IM]. URL: https://arxiv.org/abs/1903.08924 (cited on page 244).

[421] Shuichi Sato et al. "The status of DECIGO". In: *Journal of Physics: Conference Series* 840.1 (May 2017), p. 012010. DOI: 10.1088/1742-6596/840/1/012010. URL: https://dx.doi.org/10.1088/1742-6596/840/1/012010 (cited on page 244).

[422] Sidney Coleman. "Fate of the false vacuum: Semiclassical theory". In: *Phys. Rev. D* 15 (10 May 1977), pp. 2929–2936. DOI: 10.1103/PhysRevD.15.2929. URL: https://link.aps.org/doi/10.1103/PhysRevD.15.2929 (cited on page 245).

[423] Sidney Coleman and Frank De Luccia. "Gravitational effects on and of vacuum decay". In: *Phys. Rev. D* 21 (12 June 1980), pp. 3305–3315. DOI: 10.1103/PhysRevD.21.3305. URL: https://link.aps.org/doi/10.1103/PhysRevD.21.3305 (cited on page 245).

[424] Gudrun Hiller et al. *Vacuum Stability in the Standard Model and Beyond*. 2024. arXiv: 2401.08811 [hep-ph]. URL: https://arxiv.org/abs/2401.08811 (cited on page 245).

[425] Joan Elias-Miró et al. “Higgs mass implications on the stability of the electroweak vacuum”. In: *Physics Letters B* 709.3 (2012), pp. 222–228. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.02.013. URL: https://www.sciencedirect.com/science/article/pii/S037026931200144X (cited on page 245).

[426] Tania Robens. “More Doublets and Singlets”. In: (2022). arXiv: 2205.06295 [hep-ph]. URL: https://arxiv.org/abs/2205.06295 (cited on page 245).

[427] Marcela Carena et al. *Probing the Electroweak Phase Transition with Exotic Higgs Decays*. 2022. arXiv: 2203.08206 [hep-ph]. URL: https://arxiv.org/abs/2203.08206 (cited on page 246).

[428] Papaefstathiou, Andreas and White, Graham. “The Electro-Weak Phase Transition at Colliders: Discovery Post-Mortem”. In: *JHEP* 02 (2022), p. 185. DOI: 10.1007/JHEP02(2022)185. arXiv: 2108.11394 [hep-ph] (cited on page 246).

[429] Tania Robens. *Constraining extended scalar sectors at current and future colliders – an update*. 2022. arXiv: 2209.15544 [hep-ph]. URL: https://arxiv.org/abs/2209.15544 (cited on page 246).

[430] Adam Alloul et al. “FeynRules 2.0 – A complete toolbox for tree-level phenomenology”. In: *Comput. Phys. Commun.* 185 (2014), p. 2250. DOI: 10.1016/j.cpc.2014.04.012. arXiv: 1310.1921 [hep-ph] (cited on page 248).

[431] Ramona Gröber et al. “NLO QCD corrections to Higgs pair production including dimension-6 operators”. In: *JHEP* 09 (2015), p. 092. DOI: 10.1007/JHEP09(2015)092. arXiv: 1504.06577 [hep-ph] (cited on page 249).

[432] R. Gröber, M. Mühlleitner, and M. Spira. “Higgs pair production at NLO QCD for CP-violating Higgs sectors”. In: *Nuclear Physics B* 925 (2017), pp. 1–27. ISSN: 0550-3213. DOI: https://doi.org/10.1016/j.nuclphysb.2017.10.002. URL: https://www.sciencedirect.com/science/article/pii/S0550321317303231 (cited on page 249).

[433] Andrea Dainese et al., eds. *Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC*. Vol. 7/2019. CERN Yellow Reports: Monographs. CERN, 2019. ISBN: 978-92-9083-549-3. DOI: 10.23731/CYRM-2019-007 (cited on page 249).

[434] The ATLAS Collaboration. “Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector”. In: *Phys. Rev. D* 108 (5 Sept. 2023), p. 052003. DOI: 10.1103/PhysRevD.108.052003. URL: https://link.aps.org/doi/10.1103/PhysRevD.108.052003 (cited on page 261).

[435] The ATLAS Collaboration. “Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector”. In: *Phys. Rev. D* 108 (5 Sept. 2023), p. 052003. DOI: 10.1103/PhysRevD.108.052003. URL: https://link.aps.org/doi/10.1103/PhysRevD.108.052003 (cited on page 261).

[436] The ATLAS Collaboration. "Studies of new Higgs boson interactions through nonresonant HH production in the $b\bar{b}\gamma\gamma$ final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Journal of High Energy Physics* 2024.1 (Jan. 2024). ISSN: 1029-8479. DOI: 10.1007/jhep01(2024)066. URL: http://dx.doi.org/10.1007/JHEP01(2024)066 (cited on page 261).

[437] The ATLAS Collaboration. "Search for non-resonant Higgs boson pair production in the bb$\ell\ell$+$E_{\text{T}}^{\text{miss}}$ final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Journal of High Energy Physics* 2024.2 (Feb. 2024). ISSN: 1029-8479. DOI: 10.1007/jhep02(2024)037. URL: http://dx.doi.org/10.1007/JHEP02(2024)037 (cited on page 261).

[438] The ATLAS Collaboration. *Search for non-resonant Higgs boson pair production in final states with leptons, taus, and photons in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector*. 2024. arXiv: 2405.20040 [hep-ex]. URL: https://arxiv.org/abs/2405.20040 (cited on page 261).

[439] The ATLAS Collaboration. "Combination of searches for Higgs boson pair production in *pp* collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector." In: *Physical Review Letters* 133.10 (Sept. 2024). ISSN: 1079-7114. DOI: 10.1103/physrevlett.133.101801. URL: http://dx.doi.org/10.1103/PhysRevLett.133.101801 (cited on pages 261–263).

[440] The ATLAS Collaboration. *Combination of searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ decay channels using pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2786865 (cited on page 263).

ECOLE DOCTORALE

Titre : Recherche de production résonante de paires de bosons de Higgs dans le canal de désintégration bb$\tau\tau$ et développements dans la reconstruction des primitives de déclenchement du Calorimètre à Haute Granularité avec le détecteur CMS au LHC

Mots clés : Higgs, HGCAL, CMS, LHC, Primitives de Déclenchement

Résumé : Nous presentons une analyse de la production résonante de paires de bosons de Higgs (HH), se désintégrant en quarks b et en leptons τ, avec l'expérience CMS au Grand Collisionneur de Hadrons (LHC) du CERN. L'analyse exploite les 138 fb^{-1} de données de collisions proton-proton collectés entre 2016 et 2018 à une énergie au centre de masse de 13 TeV. Ce processus résonant est fortement motivé par un grand nombre de théories capables de répondre aux lacunes actuelles du Modèle Standard (SM). Le canal de désintégration étudié présente plusieurs avantages expérimentaux, à savoir une signature de l'état final relativement pure, équilibrée par un rapport d'embranchement modérée de 7.3%. Les résultats d'une étude similaire ont été récemment publiés par la collaboration ATLAS, rapportant une tension avec le SM pour une masse invariante du système HH d'environ 1 TeV. L'analyse effectuée ici vise donc à confirmer ou à rejeter un tel excès. Les limites supérieures attendues, avec un niveau de confiance de 95%, sont fixées sur la production de signatures de nouvelle physique, et représentent une amélioration considérable par rapport aux résultats antérieurs de CMS et d'ATLAS. Par ailleurs, ces travaux s'attaquent à une simplification majeure exploitée par les recherches résonantes en physique des hautes énergies, à savoir l'Approximation de Faible Largeur (NWA), qui suppose que la largeur des nouvelles résonances est négligeable par rapport à la résolution du détecteur, ignorant les effets d'interférence potentiels. Nous montrons que le niveau de sensibilité actuel des analyses HH est tel qu'il remet en question la validité de la NWA, ce qui laisse présager la nécessité d'éviter complètement cette approximation dans des analyses futures.
Les travaux décrits portent également sur l'amélioration de la sensibilité des détecteurs. Le futur LHC à haute luminosité (HL-LHC) apportera un grand nombre de collisions par croisement de paquets de protons et des niveaux de radiation extrêmement élevés, qui ne pourront être soutenus que par un programme très important de mise à niveau des détecteurs au sein de CMS. L'une des sections modernisées sera celle des bouchons, où le nouveau calorimètre à haute granularité (HGCAL) sera installé. Le HGCAL offre de nombreuses possibilités d'études et d'optimisations, et deviendra certainement une pierre angulaire de la prochaine phase HL-LHC de CMS, en fournissant des résolutions spatiales et temporelles élevées pour améliorer la reconstruction en ligne et hors ligne des données de physique. Le système de déclenchement de CMS, qui devra supporter les flux de données importants attendus du HL-LHC, sera critique pour le HGCAL. Nous avons spécifiquement développé de nouveaux algorithmes pour permettre la reconstruction robuste des primitives de déclenchement, les éléments constitutifs du premier niveau du système de déclenchement en ligne de CMS. Ces algorithmes comprennent des techniques permettant d'atténuer la création erronée de plusieurs groupes d'énergie à partir de particules individuelles, et le calcul de quantités calorimétriques dans un système de coordonnées modifié. Ces développements font partie d'outils de reconstruction, mis en œuvre à partir de zéro, qui fournissent également une version simplifiée de la géométrie de HGCAL. Les efforts futurs bénéficieront de ces outils.

Title : Search for resonant Higgs boson pair production in the bb$\tau\tau$ decay channel and developments in the reconstruction of High Granularity Calorimeter trigger primitives with the CMS detector at the LHC

Keywords : Higgs, HGCAL, CMS, LHC, Trigger Primitives

Abstract : We perform a search for the resonant production of a pair of Higgs bosons (HH), decaying into a pair of b quarks and a pair of τ leptons, with the CMS experiment at the CERN Large Hadron Collider (LHC). The analysis exploits proton-proton collisions at a center-of-mass energy of 13 TeV, for a total of 138 fb^{-1} collected during the 2016, 2017 and 2018 data-taking years. The gluon-fusion production mode is considered, together with spin-0 and spin-2 hypotheses. This resonant process is strongly motivated by a series of theories able to address current shortcomings of the SM. The decay channel is instead known for its experimental benefits, namely a relatively clean final state signature, balanced by a moderate branching fraction of 7.3%. Additionally, the results of a similar search have been recently reported by the ATLAS Collaboration, where a small tension with the SM was recorded at a resonance mass of 1 TeV. The physics analysis here performed thus aims at confirming or rejecting such an excess. Expected upper limits at a 95% confidence level are set on the production of New Physics signatures, showcasing a compelling improvement over past CMS and ATLAS results. Furthermore, this work tackles a major simplification exploited by resonant searches in High Energy Physics, namely the Narrow Width Approximation (NWA), which assumes that the width of new resonances is negligible when compared to the detector's resolution, ignoring potential interference effects. We show that the current sensitivity level of double Higgs boson analyses is such as to put into question the correctness of the NWA, hinting at the necessity of altogether avoiding such approximation in future HH analyses.
This work is also concerned with sensitivity improvements from a detector perspective. The upcoming High-Luminosity LHC (HL-LHC) will bring large numbers of collisions per proton bunch crossing and extremely high radiation levels, which can only be sustained by a very significant detector upgrade programme within CMS. One of the upgraded sections will be the endcaps, where the novel High Granularity Calorimeter (HGCAL) will be installed. The HGCAL provides ample opportunities for studies and optimizations, and will certainly become a cornerstone of the upcoming CMS HL-LHC phase, providing high spatial and timing resolutions to improve the online and offline reconstruction of physics data. Central to the HGCAL will be the CMS trigger system, which will have to withstand the large rates expected from the HL-LHC. We have specifically developed new algorithms to enable the robust reconstruction of Trigger Primitives, the building blocks of the first level of the online trigger system in CMS. The algorithms include techniques to mitigate the wrongful creation of several energy clusters from single particles, and the computation of calorimetric quantities within a modified coordinate system. These developments are part of a reconstruction framework, implemented from scratch, which also provides a simplified version of HGCAL's geometry. Future efforts will benefit from such tools.

Institut Polytechnique de Paris
91120 Palaiseau, France