

2 Higgs pair production in the four bottom quarks final state

3

Ana Luísa Moreira de Carvalho

4

Thesis to obtain the Master of Science Degree in

5

Physics Engineering

6

Supervisor(s): Dr. José Ricardo Gonçalo
Prof. Pedro Abreu

7

Examination Committee

8

Chairperson: Prof. Full Name

Supervisor: Prof. Full Name 1 (or 2)

Member of the Committee: Prof. Full Name 3

9

Month Year

Dedicated to someone special...

11 Acknowledgments

- 12 A few words about the university, financial support, research advisor, dissertation readers, faculty or
- 13 other professors, lab mates, other friends and family...

¹⁴ **Resumo**

- ¹⁵ Inserir o resumo em Português aqui com o mximo de 250 palavras e acompanhado de 4 a 6 palavras-chave...
- ¹⁶

- ¹⁷ **Palavras-chave:** palavra-chave1, palavra-chave2,...

¹⁸ **Abstract**

¹⁹ Insert your abstract here with a maximum of 250 words, followed by 4 to 6 keywords...

²⁰ **Keywords:** keyword1, keyword2,...

Contents

| | | |
|----|--|------|
| 22 | Acknowledgments | v |
| 23 | Resumo | vii |
| 24 | Abstract | ix |
| 25 | List of Tables | xiii |
| 26 | List of Figures | xv |
| 27 | Nomenclature | 1 |
| 28 | Glossary | 1 |
| 29 | 1 Introduction | 1 |
| 30 | 2 The standard model and beyond | 3 |
| 31 | 2.1 The Standard Model of Particle Physics | 3 |
| 32 | 2.1.1 Higgs pair production | 10 |
| 33 | 2.2 Going beyond | 13 |
| 34 | 3 Collider experiments | 17 |
| 35 | 3.1 Experimental aspects | 18 |
| 36 | 3.2 The Large Hadron Collider | 18 |
| 37 | 3.2.1 The ATLAS detector | 20 |
| 38 | 3.3 Jet reconstruction | 23 |
| 39 | 3.3.1 Jet properties and substructure observables | 25 |
| 40 | 3.3.2 Jet grooming algorithms | 26 |
| 41 | 3.4 Future Colliders | 27 |
| 42 | 3.4.1 The hadronic Future Circular Collider | 27 |
| 43 | 3.4.2 FCC-hh baseline detector | 28 |
| 44 | 3.4.3 FCC-hh physics program | 31 |
| 45 | 4 State of the art | 33 |
| 46 | 4.1 Searches for Higgs pair production at the LHC | 33 |
| 47 | 4.2 Feasibility studies for high-luminosity and future colliders | 34 |
| 48 | 4.3 Hadronic calorimeter granularity studies | 36 |

| | | |
|----|---|----|
| 49 | 5 Sample generation and analysis tools | 39 |
| 50 | 5.1 Fast simulation workflow | 39 |
| 51 | 5.1.1 Particle flow and calorimeter reconstruction in Delphes | 40 |
| 52 | 5.1.2 FCC-hh software | 41 |
| 53 | 5.2 Signal and background samples | 42 |
| 54 | 5.2.1 MadGraph | 42 |
| 55 | 5.2.2 Pythia | 43 |
| 56 | 5.2.3 Delphes | 45 |
| 57 | 6 Analysis | 47 |
| 58 | 6.1 Overview of the $hh \rightarrow b\bar{b}b\bar{b}$ channel | 47 |
| 59 | 6.1.1 Signal event characterization | 48 |
| 60 | 6.1.2 Backgrounds | 48 |
| 61 | 6.2 Analysis strategy | 49 |
| 62 | 6.2.1 Implementation of b-tagging | 51 |
| 63 | 6.2.2 Pre-selection | 51 |
| 64 | 6.2.3 Baseline analysis | 52 |
| 65 | 6.2.4 Optimization | 52 |
| 66 | 7 Results | 53 |
| 67 | 7.1 Di-Higgs discovery potential at the FCC-hh | 53 |
| 68 | 7.1.1 Statistical analysis | 53 |
| 69 | 7.2 Hadronic calorimeter granularity studies for future colliders | 53 |
| 70 | 7.3 Benchmarking against ATLAS analysis at $\sqrt{s} = 13$ TeV | 53 |
| 71 | 8 Conclusions | 59 |
| 72 | Bibliography | 61 |
| 73 | A Jet radius discussion | 69 |
| 74 | B Additional background processes | 70 |
| 75 | C Softdrop mass | 71 |
| 76 | D Samples generation: parameters | 72 |
| 77 | D.1 MadGraph | 72 |
| 78 | D.2 Pythia | 72 |

79 List of Tables

| | | |
|-----|---|----|
| 80 | 2.1 Summary of the particle content of the SM. | 5 |
| 81 | 3.1 ATLAS tile and liquid argon hadronic calorimeters: summary of the pseudorapidity cover- | |
| 82 | ages and transversal segmentation (granularity). | 21 |
| 83 | 3.2 Comparison between the working parameters of the LHC and of the FCC-hh. The val- | |
| 84 | ues of the number bunches and of the number of events per bunch crossing are given | |
| 85 | assuming a bunch spacing of 25 ns. | 29 |
| 86 | 3.3 Hadronic calorimeter layout, granularity and energy resolution. | 30 |
| 87 | 5.1 Summary of the cross sections, k factors and total number of events of the samples used | |
| 88 | in the analysis. | 44 |
| 89 | 5.2 Summary of the benchmark granularity configurations of the HCAL. | 45 |
| 90 | 6.1 b-Tagging (black), c (blue) and light (red) mistag probabilities as a function of η and p_T of | |
| 91 | the (sub)jet. The momentum dependent factor, $(1 - p_T/15000)$, is common to the three | |
| 92 | probabilities. | 51 |
| 93 | 7.1 Cumulative efficiency, in percentage, of each event selection criterion for the signal sam- | |
| 94 | ples (SM and 2HDM) and for the background sample ($4b+j$, $jj+0/1/2\ j$ and $t\bar{t}+0/1/2\ j$). The | |
| 95 | absolute value of expected events after some key selection cuts is shown in curved brack- | |
| 96 | ets. The number of expected events is normalized to $\mathcal{L} = 30\ ab^{-1}$. The double horizontal | |
| 97 | line marks the pre-selection cuts. These results were obtained using the FCC-hh baseline | |
| 98 | detector design, as implemented in Delphes by the FCC-hh study group. | 54 |
| 99 | D.1 Generator (MadGraph5) level cuts for the signal and background samples. | 73 |
| 100 | D.2 Pythia settings for the signal and background samples. | 73 |

List of Figures

| | | | |
|-----|-----|--|----|
| 102 | 2.1 | oi | 4 |
| 103 | 2.2 | Postulated shape of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). | 7 |
| 104 | 2.3 | Feynman diagrams of Higgs pair production from gluon fusion. Triple vertex diagram (a) on the left and box diagram (b) on the right. The minus sign between the diagrams indicates that they interfere destructively. | 10 |
| 105 | 2.4 | Total cross sections at the NLO in QCD for the six largest HH production channels at pp colliders. The thickness of the lines corresponds to the scale and PDF uncertainties added linearly. | 12 |
| 106 | 2.5 | Here, α_1 , α_2 and α_3 represent the coupling constants of $U(1)$, $SU(2)$ and $SU(3)$, respec- tively. What is shown is the variation of the inverse of the coupling with the energy: on the left for the SM and on the right for the MSSM.[REF] | 14 |
| 107 | 2.6 | BSM contributions to Higgs pair production. On the left (figure from [22]) η represent the spin-2 Kaluza-Klein graviton that couples to a pair of SM Higgs bosons, H , at tree level. On the right (figure from [23]) \tilde{t} is the supersymmetric partner of the top quark that gives a contribution to the quantum loop in the Higgs pair production s-channel diagram. | 16 |
| 108 | 3.1 | ATLAS detector. | 22 |
| 109 | 3.2 | FCC detector baseline concept. | 30 |
| 110 | 4.1 | Exclusion limits on the mass of a spin 0 resonance from searches for Higgs pair production with the CMS detector. | 35 |
| 111 | 4.2 | Azimuthal distribution of the energy deposits in the ECAL and HCAL for a pair of K_L^0 with $E = 100$ GeV for an hadronic calorimeter with $20 \text{ cm} \times 20 \text{ cm}$ (left) and $5 \text{ cm} \times 5 \text{ cm}$ (right). Figures from [66] (based on [65]). | 37 |
| 112 | 4.3 | Jet mass | 38 |
| 113 | 4.4 | τ_{32} | 38 |
| 114 | 6.1 | Higgs pair branching ratios. | 48 |
| 115 | 6.2 | Example of diagrams that contribute to the QCD multijet background: five final state jets, four of which are b-jets (left), three final state jets (middle) and four final state jets (right). Here, q stands for a light quark/jet. | 50 |
| 116 | 6.3 | Dominant diagrams of $pp \rightarrow t\bar{t}$ at LO. | 50 |

| | | |
|-----|---|----|
| 131 | 6.4 Event topology targeted by the boosted analysis region. The blob represents the interaction between the gluons and the Higg bosons that is represented by the Feynman | 50 |
| 132 | diagrams shown in figure 2.3. | |
| 133 | | |
| 134 | 6.5 Minimun ΔR between b quarks and subjets of the $R = 0.8$ jets. There is a minimum | 52 |
| 135 | around $\Delta R = 0.3$. It corresponds to the maximum ΔR for which the b quark and the | |
| 136 | subjet are matched. | |
| 137 | 7.1 oi | 54 |
| 138 | 7.2 SUPERIMPOSE ONE BKG? MAYBE NOT | 55 |
| 139 | 7.3 S/\sqrt{B} as a function of the detector configuration for $\mathcal{L} = 30 \text{ ab}^{-1}$ (black) and $\mathcal{L} = 3 \text{ ab}^{-1}$ | |
| 140 | (blue) and for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles). . . | 56 |
| 141 | 7.4 Signal efficiency as a function of the detector configuration for particle flow jets (squares) | |
| 142 | and pure calorimeter (HCAL) jets (triangles). | 57 |
| 143 | A.1 $\Delta R(b, \bar{b})$ distributions for $p_T(h_1) > 200/300 \text{ GeV}$ (solid blue/black) and for $p_T(h_1, h_2) >$ | |
| 144 | $200/300 \text{ GeV}$ (dashed blue/black). | 69 |
| 145 | C.1 Correlation between the maximum ΔR between the Higgs candidate jet and one of the b | |
| 146 | quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$. . | 71 |

¹⁴⁷ Chapter 1

¹⁴⁸ Introduction

¹⁴⁹ NUMBERS? (luminosity, cross sections, BR,...)

¹⁵⁰

¹⁵¹

¹⁵² THESIS GOAL

¹⁵³ It is the ultimate goal of particle physics to discover and study all of Nature's fundamental particles
¹⁵⁴ and to understand their interactions. Through a joint endeavor of theorists and experimentalists, models
¹⁵⁵ that describe particle's dynamics and properties can be precisely probed at collider experiments such
¹⁵⁶ as the Large Hadron Collider (LHC).

¹⁵⁷ We know today that matter particles interact by means of four fundamental forces: electromagnetic,
¹⁵⁸ weak, strong and gravitational, each associated with a mediator particle. We even know that a very
¹⁵⁹ special particle, the Higgs boson, is responsible for generating the mass of all of these particles through
¹⁶⁰ a mechanism called Electroweak Symmetry Breaking (EWSB). All of this knowledge is beautifully sum-
¹⁶¹ marized in the Standard Model of Particle Physics (SM) that was developed in the 1960's, long before
¹⁶² many of the particles it predicts were discovered. The extraordinary precision of the predictions it de-
¹⁶³ livers make it a very successful model. Its most recent prediction, the Higgs boson, was discovered in
¹⁶⁴ 2012 at the LHC which marks an important point in the history of particle physics: we have now found
¹⁶⁵ all the particles predicted by the SM and yet we know that it cannot be the whole story. Mainly because
¹⁶⁶ there are experimental evidences it cannot explain.

¹⁶⁷ From the theoretical point of view, this is enough motivation to construct models that extend the SM
¹⁶⁸ but that can still deliver predictions that are compatible with experimental data. From the experimental
¹⁶⁹ standpoint, this is an indication that we need to keep increasing the precision of our measurements and
¹⁷⁰ probing new kinematic regimes in the hope of finding some discrepancy with the SM or some hint that
¹⁷¹ some new phenomenon might be taking place.

¹⁷² A higher precision requires a larger integrated luminosity and the exploration of new kinematic
¹⁷³ regimes ask for a larger center of mass energy. Very recently, the upgrade of the LHC to its High-
¹⁷⁴ Luminosity (HL) version has began. It is expected to work for a period of ten years between 2026 and
¹⁷⁵ 2036 and it will extend the experimental reach of the LHC. In order to keep extending the physics reach

176 of the LHC and HL-LHC, new colliders with unprecedently high CM energies are currently being de-
177 signed in the hope that they begin to deliver data shortly after the HL-LHC has reached its full discovery
178 potential. One of these projects is the hadronic Future Circular Collider (FCC-hh) that is expected to
179 work at a CM energy of 100 TeV and to deliver a total integrated luminosity ten times larger than what is
180 expected by the end of the HL-LHC operation.

181 The next step for the FCC-hh project is the submission of a Conceptual Design Report by the end
182 of 2018. This document will be used as input for the next meeting of the European Strategy for Particle
183 Physics that will take place in the beginning of 2019. It should present a baseline design for the detector,
184 a first cost estimate and analysis for physics benchmark processes.

185 Both in the HL-LHC and in future colliders, one of the most important benchmark processes is the
186 production of pairs of Higgs bosons. Firstly, this process is predicted by the SM but has not yet been
187 measured which is due to its very small cross section and overwhelming backgrounds. Furthermore,
188 it provides unique insight into the EWSB mechanism because it is sensitive to the shape of the Higgs
189 potential and can even be used to probe physics beyond the SM (BSM).

190 The work presented on this thesis is a Monte Carlo study that accesses the feasibility of the search
191 for pairs of Higgs bosons at the FCC-hh in the final state with four b quarks. We choose this final
192 state because it benefits from the large branching fraction of the Higgs boson to a pair of b quarks.
193 However, in this channel, the SM multijet background is extremely overwhelming. Although challenging,
194 this gives us the chance to explore the boosted kinematic regime and jet substructure observables in
195 order to maximize the rejection of this background. We also evaluate the sensitivity of our analysis to
196 BSM benchmark signal processes. From the detector standpoint, we evaluate how the granularity of the
197 hadronic calorimeter influences the analysis' sensitivity and the power to resolve jet's substructure.

198 Chapter 2 presents an overview of the SM. It summarizes its particle content and interactions and
199 introduces the mathematical formulation of the EWSB breaking mechanism. The successes and short-
200 comings of the SM are also discussed and several BSM models are introduced and their motivations
201 discussed. Finally, a theoretical description of the production of Higgs pairs is provided.

202 The FCC-hh baseline accelerator and detector were highly based on LHC and its current experi-
203 ments, namely ATLAS and CMS. In chapter 3, after a brief discussion of the general features of particle
204 accelerators, we introduce the LHC and the ATLAS experiment. A discussion of jet reconstruction is
205 included. We then introduce the FCC-hh accelerator that is expected to very similar to the LHC except
206 larger in circumference and with more powerful magnets. The current baseline detector design for the
207 FCC-hh is discussed and its features compared to ATLAS.

²⁰⁸ Chapter 2

²⁰⁹ The standard model and beyond

²¹⁰ The Standard Model is the theoretical framework that summarizes our present knowledge of particle
²¹¹ physics. In section 2.1, we provide an overview of this model, focusing on the Higgs mechanism. In
²¹² section 2.2, we motivate the need to explore models beyond the SM and introduce some of the most
²¹³ well known BSM models. In section 2.1.1, we provide a theoretical description of the production of Higgs
²¹⁴ boson pairs.

²¹⁵ 2.1 The Standard Model of Particle Physics

²¹⁶ IDEAS FOR SECTION

- ²¹⁷ - Historic development + biggest breakthroughs
 - ²¹⁸ - Theoretical description of SM: Symmetry group, Lagrangian, ...
 - ²¹⁹ - Biggest successes + experimental confirmations
- ²²⁰
- ²²¹ - Theoretical description of Higgs mechanism: symmetry breaking + mass generation for gauge
 - ²²² bosons
 - ²²³ - Yukawa couplings as a source of mass for leptons

²²⁴

²²⁵ The Standard Model (SM) of particle physics summarizes our present knowledge of fundamental
²²⁶ particles and their interactions. It is formulated in the framework of Quantum Field Theory (QFT) and
²²⁷ describes the subatomic world in terms of fields whose excitations are the particles we can detect. The
²²⁸ particle content of the SM is summarized in figure 2.1. There are two types of fundamental particles:
²²⁹ matter particles and force carriers.

²³⁰ Matter particles are the building blocks of all the matter in our world. They come in two groups,
²³¹ leptons and quarks. Quarks make up atomic nuclei and leptons, namely electrons and muons can orbit
²³² atomic nuclei forming atoms. Both quarks and leptons are fermions which means they have half-integer
²³³ spin. There are six quarks: three of the 'up type' (up, charm and top represented by u , d and t) with
²³⁴ electric charge of $+2/3$ and three of 'down type' (down, strange and bottom represented by d , s and b)

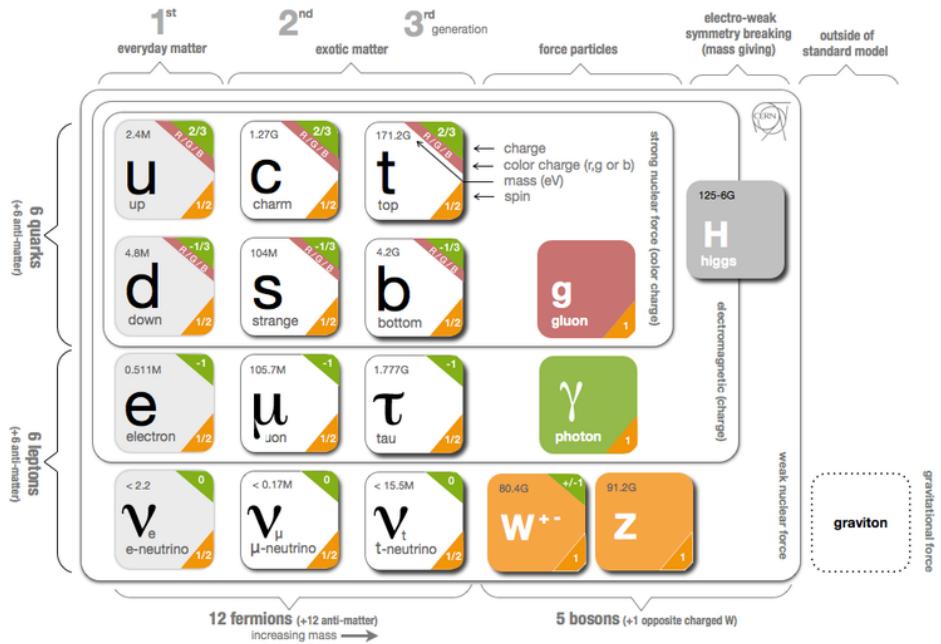


Figure 2.1: oi

with electric charge of $-1/3$. Similarly, we have three leptons with charge -1 (electron, muon and tau represented by e , μ and τ) and three neutral leptons (electron, muon and tau neutrinos represented by ν with the symbol of the corresponding charged lepton as subscript) that are, within the SM, massless. We can classify quarks and leptons in three generations, each composed of an up type and down type quark or of a charged lepton and the corresponding neutrino.

The force carriers, technically called gauge bosons, are particles associated with the fundamental interactions: strong, electromagnetic, weak and gravitational ¹. Each interaction can be interpreted as the result of the exchange of the corresponding gauge boson. Gluons (g) and photons (γ) are the mediators of the strong and electromagnetic interactions, respectively. They are massless, electrically neutral and have spin 1. The W^+ , W^- and Z bosons are the mediators of the weak interaction and have a mass of 82 and 91 GeV, respectively. The W^+ and W^- bosons have electric charges of $+1$ and -1 , respectively and spin 1. The Z boson is electrically neutral and also has spin 1. The gauge bosons can also be referred to as vector bosons because they have spin equal to one.

In addition to matter particles and gauge bosons, the theoretical formulation of the SM rests on the existence of the Higgs boson that is an electrically neutral and spin 0 particle. It has a mass of 125 GeV and it interacts with every particle that has mass.

The fermions, gauge bosons and Higgs boson properties, namely, electric charge, spin and mass are summarized in Table 2.1.

Historically, an empirically successful quantum theory of electromagnetism, Quantum Electrodynamics (QED), was developed in the late 1940's. In the early 1950's there were high hopes that quantum

¹ The gauge boson that corresponds to the gravitational force has not yet been found. In addition we still do not have a theory that successfully describes gravitation in the framework of QFT so we will not include the gravitational force or its gauge boson in any of the following discussions.

| Type | Particles | Electric charge | Spin | Mass |
|--------------|----------------------------|-----------------|------|---------------------------|
| Quarks | u, c, t | 2/3 | 1/2 | 0.0022, 1.27, 173.21 GeV |
| | d, s, b | -1/3 | | 0.0047, 0.096, 4.18 GeV |
| Leptons | e, μ, τ | -1 | 0 | 0.51, 105.66, 1776.86 MeV |
| | ν_e, ν_μ, ν_τ | 0 | | < 2 eV |
| Gauge bosons | g | 0 | 1 | 0 |
| | γ | 0 | | $< 10^{-18}$ eV |
| | W^\pm | ± 1 | | 80.385 GeV |
| | Z | 0 | | 91.1876 GeV |
| Higgs boson | H | 0 | 0 | 125.09 GeV |

Table 2.1: Summary of the particle content of the SM.

255 theories could also be formulated for the weak and strong interactions. This is the context in which
 256 Yang-Mills theories emerged. They extend the concept of gauge theory from abelian groups, that lead
 257 to the development of QED, to non-abelian gauge groups. However, the quanta of the fields predicted
 258 by these theories must be massless in order to maintain gauge invariance. Therefore, they were set
 259 aside until the 1960's when the idea of particles acquiring mass through symmetry breaking in massless
 260 theories was put forward by Goldstone [1], Nambu and Jona-Lasinio [2]. In the following paragraphs we
 261 discuss in more detail the caveats of Yang-Mills theories and the phenomenon of Spontaneous Symme-
 262 try Breaking (SSB) as the basis of the modern Higgs mechanism. We then describe this mechanism in
 263 the framework of the SM.

264 On the one hand, if one takes a Yang-Mills theory, it becomes clear that it is not possible to include
 265 in the Lagrangian a mass term for the gauge bosons because it is not invariant under a gauge transfor-
 266 mation. This would not be a problem if we just wanted to describe electromagnetic or strong interactions
 267 because the gauge bosons associated with these interactions, the photon and the gluon, are indeed
 268 massless. However, for the weak interactions this is not the case. Even before the discovery of the
 269 Z and W^\pm bosons [3, 4] there were experimental evidence of the short range character of the weak
 270 interactions which indicated that the corresponding gauge bosons should be massive.

271 On the other hand, spontaneous symmetry breaking (SSB) is a phenomenon through which the
 272 invariance of a system under a certain symmetry group is destroyed [5]. The system may then be
 273 invariant under a subgroup of the initial symmetry but the invariance under the original symmetry group
 274 is no longer present. In particle physics, this happens because the vacuum of the system (lowest energy
 275 states) does not share the symmetry of the Lagrangian. The SSB mechanism predicts the existence
 276 of scalar massless particles, the Nambu-Goldstone boson, (the number depends on the number of
 277 generators of the original and final symmetry groups) as a consequence of the Goldstone theorem [1].
 278 Though, when considering this mechanism we get once again massless particles which does not seem
 279 to be a step in the right direction if we wish to describe weak interactions.

280 However, the real breakthrough occurs when we combine a theory with local gauge invariance with
 281 the mechanism of SSB. In this case the Nambu-Goldstone bosons do not appear and it is possible to
 282 give mass to the gauge bosons. This is the Higgs mechanism, proposed independently by P.W. Higgs
 283 [6], F. Englert and R. Brout [7] and by G. Guralnik, C. R. Hagen and T. Kibble [8] in 1964.

²⁸⁴ The SM is a non-abelian gauge theory with spontaneous symmetry breaking. It is locally invariant
²⁸⁵ under the following symmetry group:

$$SU_{color}(3) \times SU_L(2) \times U_Y(1), \quad (2.1)$$

²⁸⁶ where the $SU_{color}(3)$ group describes the strong interactions (QCD) and the $SU_L(2) \times U_Y(1)$ group
²⁸⁷ describes the electroweak interactions. Here, L stands for left and Y stands for hypercharge. In the SM
²⁸⁸ the Higgs mechanism, which we now describe, is realized in the latter. The Lagrangian corresponding
²⁸⁹ to the Higgs and gauge sectors of this theory is given by:

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) - \frac{1}{4} W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \quad (2.2)$$

where the Higgs potential, $V(\phi^\dagger \phi)$, is given by:

$$V(\phi^\dagger \phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \quad (2.3)$$

²⁹⁰ $W_{\mu\nu}^a$ and $B_{\mu\nu}$ are the field tensors, defined as a function of the gauge fields of $SU(2)$ and $U(1)$, respectively,
²⁹¹ W_μ^a ($a = 1, 2, 3$) and B_μ :

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g \epsilon^{abc} W_\mu^b W_\nu^c \quad (2.4)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.5)$$

²⁹² where g is the coupling constant associated with the $SU(2)$ group and ϵ^{abc} is the completely anti-symmetric tensor in 3 dimensions. The covariant derivative, D_μ , is introduced to preserve local gauge invariance and is given by:

$$D_\mu \phi = \left(\partial_\mu + ig W_\mu^a T^a + i \frac{g'}{2} B_\mu \right) \phi. \quad (2.6)$$

²⁹⁵ $T^a = \frac{\tau^a}{2}$ (where τ^a are the Pauli matrices) are the $SU(2)$ group generators in the fundamental representation and g' is the coupling constant associated with the $U(1)$ group.

Due to the requirement of Lorentz invariance, only the scalar field, ϕ , can have a vacuum expectation value (VEV), v , different from zero ². The values of v are determined by the minima of the potential:

$$v = 0 \quad \text{or} \quad v = \sqrt{\frac{-\mu^2}{2\lambda}}. \quad (2.7)$$

²⁹⁷ For the equation on the right (for which we get $v \neq 0$) we only obtain a real value for v (which is a
²⁹⁸ requirement for the VEV of a theory) if $\mu^2 < 0$. Therefore we conclude that the equation on the right
²⁹⁹ corresponds to $\mu^2 \leq 0$ while the equation on the left corresponds to $\mu^2 \geq 0$. In both cases λ has to

²The other fields that appear in Eq. 2.2 are vector fields. If they were to acquire a VEV different from zero that would break Lorentz invariance.

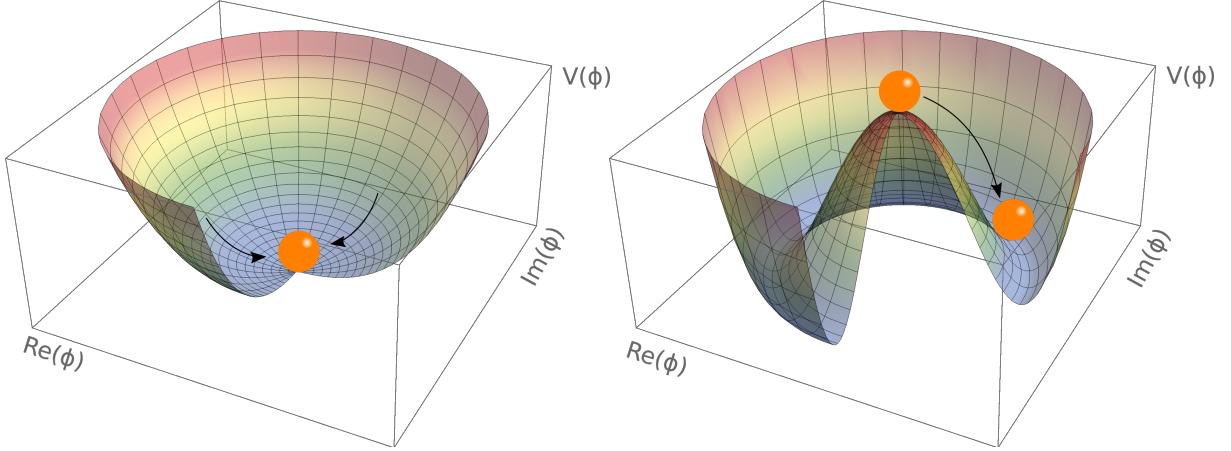


Figure 2.2: Postulated shape of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

be larger than zero to guarantee that the energy is bounded from below³ because in Eq. 2.3 λ is the coefficient of the term with the highest power in ϕ and therefore determines the concavity of the potential.

The shapes of the Higgs potential for $\mu^2 > 0$ and $\mu^2 < 0$ are shown in Figure 2.2 on the left and right, respectively. For $\mu^2 > 0$ (left) we have a single minimum located at $\langle\phi\rangle = 0$. For $\mu^2 < 0$ (right) the potential has the shape of a 'Mexican hat'. There is an infinite number of minima located in a circumference centered at zero. In this case the minima occur for $\langle\phi\rangle, \langle\phi^\dagger\rangle \neq 0$. Therefore the fields acquire a VEV different than zero and this what leads to the SSB.

We can now write the scalar field in terms of its minimum value, v , and of oscillations around that minimum, h (which corresponds to the Higgs field):

$$\phi = \begin{bmatrix} 0 \\ v + \frac{h}{\sqrt{2}} \end{bmatrix} \quad (\text{unitary gauge}). \quad (2.8)$$

If we expand the first term of the Lagrangian shown in Eq. 2.2 using Eq. 2.6 and Eq. 2.8 and taking into consideration that $W_\mu^a T^a$ represents a sum over all values of a we get

$$\mathcal{L} = \frac{1}{4} \left(v^2 + \frac{h^2}{2} + \frac{2}{\sqrt{2}} vh \right) [g^2 (W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu} + W_\mu^3 W^{3\mu}) - 2gg' B^\mu W_\mu^3 + g'^2 B_\mu B^\mu] + \dots . \quad (2.9)$$

We see that for the W_μ^1 and W_μ^2 fields we have only terms that are quadratic in these fields. These correspond to mass terms. However, for the W_μ^3 and B_μ fields there is a term that mixes the two fields. To obtain the physical states of the theory we need to transform these fields in order to get rid of the mixing term which is not physical. We can start by writing the last three terms of Eq. 2.9 in a matrix form and diagonalize the corresponding matrix:

$$\begin{bmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{bmatrix} \begin{bmatrix} W_\mu^3 \\ B_\mu \end{bmatrix} \xrightarrow{\text{Diagonalization}} \begin{bmatrix} 0 & 0 \\ 0 & g^2 + g'^2 \end{bmatrix} \begin{bmatrix} A_\mu \\ Z_\mu \end{bmatrix} . \quad (2.10)$$

³In a purely mathematical formulation this means that the function that represents the Higgs potential is concave upwards.

A_μ and Z_μ are the physical fields that are related with W_μ^3 and B_μ by means of a rotation matrix:

$$\begin{bmatrix} A_\mu \\ Z_\mu \end{bmatrix} = \begin{bmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{bmatrix} \begin{bmatrix} W_\mu^3 \\ B_\mu \end{bmatrix} \quad (2.11)$$

where θ_W is the Weinberg angle. By inverting this relation we can write W_μ^3 and B_μ as a function of A_μ and Z_μ . Replacing in 2.9 and imposing that the A_μ field has zero mass we can determine θ_W : $\tan \theta_W = \frac{g'}{g}$. The Lagrangian of 2.9 takes then the form

$$\mathcal{L} = \frac{1}{2} (v^2 g^2) (W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu}) + \frac{1}{2} (v^2 [g^2 + g'^2]) Z_\mu Z^\mu + \dots \quad (2.12)$$

309 where we show only the mass terms for the gauge bosons. Note that, by construction, there is no mass
 310 term for A_μ which allows us to identify this field with the photon. W_μ^1 and W_μ^2 are related to the W^\pm
 311 boson and Z_μ corresponds to the Z boson. We have shown that it is the fact that $v \neq 0$ that allows for
 312 the existence of non-zero mass terms for the W^\pm and Z bosons.

If we now expand the second term of the Higgs potential (2.3) using 2.8 we get, among other terms,

$$\mathcal{L} = -h^3 \sqrt{-\mu^2 \lambda} - h^4 \lambda + \dots . \quad (2.13)$$

313 These terms encode the Higgs self interactions and represent, respectively, the three and four point
 314 interactions. We see that the coupling constants of these interactions depend on the parameters of the
 315 Higgs potential, μ^2 and λ .

316 In addition to being responsible for giving mass to the gauge bosons the Higgs field is also respon-
 317 sible for the mass of the fermions (leptons and quarks). However, the mechanism through which this
 318 occurs is fundamentally different. In the case of the leptons the mass terms are placed explicitly in the
 319 Lagrangian:

$$\mathcal{L}_{\text{fermions}} = G_1 \bar{L} \phi R + G_2 \bar{L} \phi_c R + \text{hermitian conjugate} \quad (2.14)$$

320 where L denotes a left-handed fermion doublet and R denotes a right-handed fermion singlet. Here, left
 321 and right refer to helicity states. G_1 and G_2 are arbitrary coupling constants that can be written in terms
 322 of the fermion's mass and the VEV. ϕ is given by 2.8 and ϕ_c is given by (after the spontaneous symmetry
 323 breaking and in the unitary gauge):

$$\phi_c = \begin{bmatrix} v + \frac{h}{\sqrt{2}} \\ 0 \end{bmatrix}. \quad (2.15)$$

We now take a quick detour to motivate why fermions are represented as chiral states (left and right) of the $SU(2)$ symmetry. We base this discussion on [9]. In the context of the unification of the electromagnetic and weak forces, formalized by Weinberg, Glashow and Salam in 1960, both interactions are interpreted as manifestations of the electroweak force. Weak charged currents are axial vector currents which means they couple only to left handed fermions while weak neutral currents, as well as QED, cou-

ple to both helicity states. This suggested that fermions were better represented as left-handed doublets and right-handed singlets of the $SU(2)$ symmetry group. The left handed doublets, L , are defined as:

$$L : \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L \quad (2.16)$$

where l represents an electron, muon or tau, u is any quark of the up type and d' is a quark of the down type. The right-handed states are singlets, define as:

$$R : l_R, u_R, d'_R. \quad (2.17)$$

324 In 1956, C. S. Wu *et al.* showed that the weak interaction violates parity conservation [10]. In 1958, M.
325 Goldhaber *et al.* conducted an experiment that showed that neutrinos are left-handed and anti-neutrinos
326 are right-handed [11] which is why the SM does not include a right-handed state for neutrinos. We can
327 now continue the discussion of the mass generation mechanism for fermions.

328 The first term in equation 2.14 gives mass to down type fermions (electron, muon, tau, down, strange
329 and bottom quarks) and the second to up type fermions (up, charm and top quarks). In addition, these
330 terms give rise to the interaction terms between the Higgs field and the fermions. Take, as an example,
331 $\bar{L} = (\bar{t}, \bar{b})_L$ and $R = b_R$. For the first term of 2.14 we get:

$$G_1 \bar{L} \phi R = G_1 (\bar{t}, \bar{b})_L \begin{bmatrix} 0 \\ v + \frac{h}{\sqrt{2}} \end{bmatrix} b_R = G_1 v \bar{b}_L b_R + \frac{G_1}{\sqrt{2}} \bar{b}_L b_R h. \quad (2.18)$$

The first term is the mass term for b quarks. Therefore we can redefine $G_1 v = m_b$ and obtain:

$$G_1 \bar{L} \phi R = m_b \bar{b}_L b_R + \frac{m_b}{v \sqrt{2}} \bar{b}_L b_R h. \quad (2.19)$$

332 The second term gives the interaction between the Higgs boson and the fermions, in this case, the b
333 quarks. The strength of this interaction is directly proportional to the mass of the corresponding fermion.

334 In the SM formalism, neutrinos as massless particles. However, there is no reason why they cannot
335 acquire mass through the mechanism that we just described. Nonetheless, the usual argument is that it
336 would be unnatural for the same mechanism to produce the mass of very heavy particles, such as the
337 top quark, and the mass of very light particle, such as the neutrinos. Therefore, BSM models that try to
338 explain the mass generation for neutrinos usually resort to a different mechanism.

339 The SM has delivered extremely accurate predictions about the existence and properties of new
340 particles which makes it a very successful theory. It predicted the existence of the W and Z bosons [12],
341 the gluon, the charm and top quarks and the Higgs boson. [REFS] In addition, the SM prediction for the
342 value of the anomalous magnetic dipole moment of the electron (calculated up to order α^5) agrees with
343 the measured value up to the 11th decimal place, making it the most precise measurement in science.

344 2.1.1 Higgs pair production

345 IDEAS FOR SECTION

- 346 - Main production process and Feynman diagrams
- 347 - Computation of cross section + dependency with COM energy and triple coupling
- 348 - Sensitivity to shape of Higgs potential
- 349 - Conclude and motivate why this process should be studied (within and beyond SM)

350

351 Within the SM there are still some processes that have not been measured. One of these is the
 352 production of pairs of Higgs bosons. The experimental challenges and efforts related to this process are
 353 discussed in section 4.1. Here we provide a theoretical description of the process.

354 At the LHC, the main production process of Higgs pairs is gluon-gluon fusion (ggF). Higgs pairs can
 355 also be produced through vector boson (V) fusion (VBF), namely of W and Z bosons, in association with
 356 a pair of top quarks ($t\bar{t}h$) or through Higgs strahlung (Vh)⁴. The ggF production process has a cross
 357 section ten times larger than the next most common production process which is VBF (approximately
 358 50 pb *versus* 4 pb) and therefore it is the dominant contribution when we study inclusive production.
 359 For this reason we focus the following discussion on this production mode. The leading order Feynman
 360 diagrams for Higgs pair production via ggF are shown in figure 2.3.

361 The diagram on the left has an off-shell (virtual) Higgs boson, h^* , that couples to gluons by the usual
 362 heavy quark triangle (same mechanism as in single Higgs production). h^* then decays to two on-shell
 363 Higgs bosons. This diagram contains the three point interaction between Higgs bosons and therefore it
 364 is the one that allows us to probe this coupling. In the diagram on the right, the two Higgs bosons couple
 365 to the gluons by a box of heavy quarks and are directly radiated from a quark. The largest contributions
 366 for these quantum loops come from heavy quarks, such as the top and bottom, because the coupling
 367 constant of the Higgs boson to fermions is directly proportional to the fermions mass (see section 2.1).

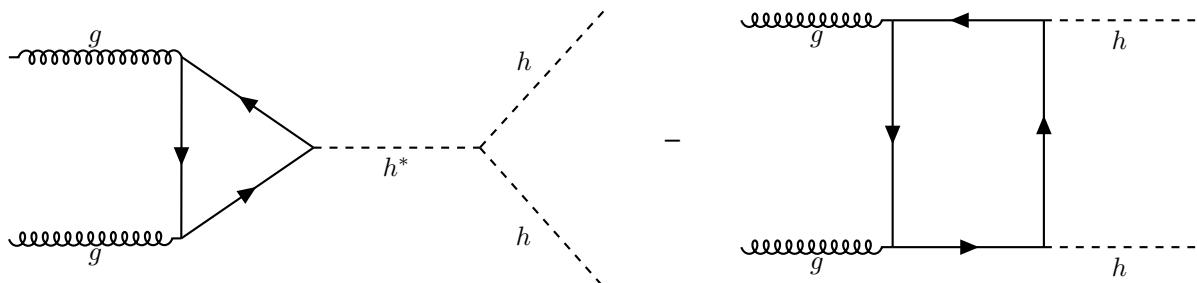


Figure 2.3: Feynman diagrams of Higgs pair production from gluon fusion. Triple vertex diagram (a) on the left and box diagram (b) on the right. The minus sign between the diagrams indicates that they interfere destructively.

The amplitudes for the box, \mathcal{M}_\square , and triangle, \mathcal{M}_Δ , diagrams scale as [13]:

$$\mathcal{M}_\square \sim \frac{\alpha_s}{4\pi} y_t^2, \quad \mathcal{M}_\Delta \sim \lambda_{hhh} \frac{\alpha_s}{4\pi} y_t^2 \frac{m_h^2}{\hat{s}} \left(\log \frac{m_t^2}{\hat{s}} + i\pi \right)^2 \quad (2.20)$$

368 where \hat{s} is the CM energy and y_t is the Yukawa coupling of the top quark.

⁴In this process, at LO, a Higgs boson is radiated from a vector boson

At a CM energy of 13 TeV the cross section for Higgs pair production, as predicted by NLO calculations, is very small, approximately 30 fb. It is suppressed due to the destructive interference between the LO diagrams that leads to a $\sim 50\%$ suppression of the total cross section [13]. Furthermore, the cross section of the triangle diagram (the one that gives us access to the value of the Higgs triple coupling) is smaller than the one of the box diagram, approximately 4 fb compared to 30 fb⁵, and is strongly suppressed for larger values of the CM energy which can be seen directly from the expression for the amplitude in 2.20. This means that the Higgs trilinear coupling mostly affects the Higgs pair production at threshold, in particular, the m_{hh} distribution. The tail of this distribution (high invariant mass of the Higgs pair), however, is mostly determined by the box diagram contribution [13].

The LO calculation for the cross section of Higgs pair production has been performed, for example, in [14]. A value of the order of 10 fb is reported. The NLO calculation is a theoretical challenge: several two-loop diagrams that take into account virtual and real radiation have to be considered. In addition, top quark mass effects can be included in various approximations. This leads to corrections with different signs which suggests that the uncertainty on the cross section due to top quark mass effects are of the order of $\pm 10\%$ at NLO. Therefore, a calculation including the full top mass dependence was of the utmost importance. This result became available recently [15]:

$$\sigma_{gg \rightarrow hh}^{\text{NLO}} = 27.80^{+13.8\%}_{-12.8\%}(\text{scale}) \pm 0.3\%(\text{stat.}) \pm 0.1\%(\text{int.}) \text{ fb} \quad (2.21)$$

where the dependence of the result on the variation of the scales by a factor of two around the central scale, the statistical error coming from the limited number of phase space points evaluated and the error coming from the numerical integration of the amplitude are shown.

This result shows that the introduction of NLO contributions produces a significantly different result. Therefore the inclusion of such effects is necessary if we wish to obtain an accurate result that can be compared to experimental values. The analytical expressions for the NLO cross section are long and complex so we abstain from reproducing them here. Nonetheless, the LO cross section can be written in a compact form and it allows us to discuss some key features of the process. Therefore, we will present it here, based on [14] and [16].

The partonic LO cross section for $gg \rightarrow hh$ can be written

$$\hat{\sigma}_{gg \rightarrow hh}^{\text{LO}}(\hat{s}) \sim \int_{\hat{t}_-}^{\hat{t}_+} d\hat{t} \ (|C_\Delta F_\Delta + C_\square F_\square|^2 + |C_\square G_\square|^2) \quad (2.22)$$

where \hat{s} and \hat{t} are the Mandelstam variables and, in addition, \hat{s} can be identified with the square of the partonic CM energy of the process. The integration limits, \hat{t}_\pm , are derived from a momentum parametrization in the CM frame, leading to $\hat{t}_\pm = m_h^2 - \frac{\hat{s}}{2}(1 \mp \beta_h)$, where $\beta_h^2 = 1 - 4\frac{m_h^2}{\hat{s}}$ [16]. F_Δ , F_\square and G_\square are form factors whose full expressions can be found, for example, in [14]. C_Δ and C_\square can

⁵These values are obtained using MadGraph5. They are shown here to give a rough estimate of the difference between the values of the cross sections of both diagrams.

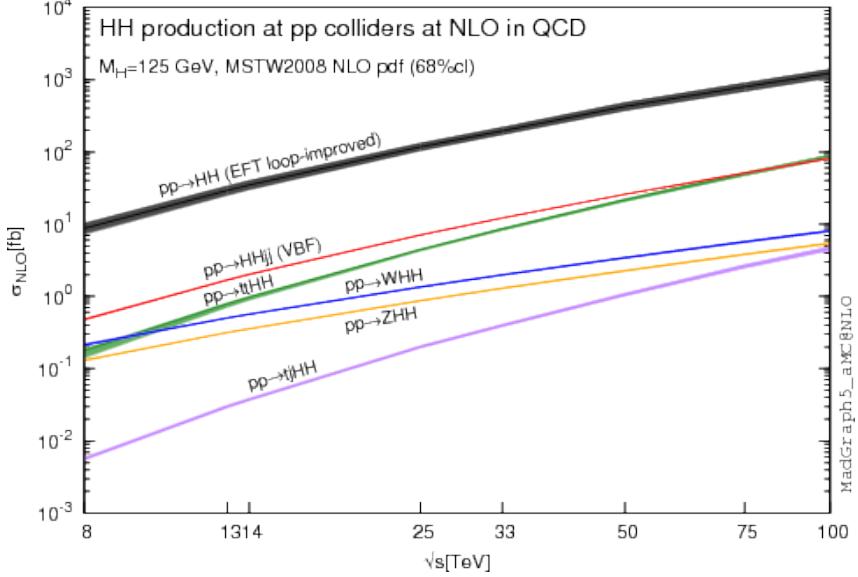


Figure 2.4: Total cross sections at the NLO in QCD for the six largest HH production channels at pp colliders. The thickness of the lines corresponds to the scale and PDF uncertainties added linearly.

be interpreted as generalized couplings and are given by

$$C_{\Delta} = \lambda_{hhh} \frac{m_Z^2}{\hat{s} - m_h^2}, \quad C_{\square} = 1 \quad (2.23)$$

where \hat{s} is the CM energy of the process. If we take the limit $m_Q^2 \gg \hat{s} \sim M_h^2$ (where m_Q is the mass of the quarks that contribute to the quantum loops) we can get simple expressions for the remaining form factors:

$$F_{\Delta} = \frac{2}{3} + \mathcal{O}(\hat{s}/m_Q^2), \quad F_{\square} = -\frac{2}{3} + \mathcal{O}(\hat{s}/m_Q^2), \quad G_{\square} = \mathcal{O}(\hat{s}/m_Q^2). \quad (2.24)$$

In this limit, the partonic cross section is simply given by

$$\hat{\sigma}_{gg \rightarrow hh}^{\text{LO}}(\hat{s}) \sim \int_{\hat{t}_-}^{\hat{t}_+} d\hat{t} \left| \lambda_{hhh} \frac{m_Z^2}{\hat{s} - m_h^2} - 1 \right|^2. \quad (2.25)$$

There are two important points that are worth discussing. Firstly, the total cross section has terms that are proportional to the Higgs triple coupling, λ_{hhh} , which can be read directly from Eq. 2.25. On the one hand, this means that measuring this process gives us access to the value of λ_{hhh} and therefore provides valuable insight into the shape of the Higgs potential and ultimately into the EWSB mechanism in the SM. On the other hand, if λ_{hhh} has a value that is different from the one predicted by the SM, that will affect the measured value of the cross section and can lead to hints of new physics.

Secondly, although this is not evident from 2.25, the cross section for di-Higgs production increases with s . This can be seen in figure 2.4 that shows the variation of the total (integrated) NLO cross section with the CM energy for the six largest production channels. Note that increasing the CM energy from 13 to 100 TeV increases the inclusive cross section by approximately two orders of magnitude which is a consequence of the increased phase space that becomes available.

Therefore, the increase in the cross section of rare processes, such as Higgs pairs production, as

399 the CM energy of collision experiments increases supports the claim that future colliders, with higher
400 CM energies, might be our chance of discovering and precisely studying these processes.

401 2.2 Going beyond

402 IDEAS FOR SECTION

- 403 - Why we need to study physics BSM (motivation: theoretical+experimental)
 - 404 - Brief description of most well known and well studied BSM models that include changes in the Higgs
 - 405 sector, predict heavier Higgs
 - 406 - How these can be probed using Higgs pair production
-

407

408 Despite the success of the SM, there is evidence, both experimental and theoretical in nature, that
409 indicate that it cannot be the final theory of particle physics. This led to the development of alternative
410 models that extend the SM but that can still reproduce its successful predictions. These are referred to
411 as Beyond the Standard Model (BSM) models.

412 On the one hand, there are several pieces of experimental evidence that the SM cannot explain.
413 These include the nature of dark matter, postulated to explain the experimental observations of the
414 velocity of far away galaxies [17], the asymmetry between matter and anti-matter in the present Universe
415⁶ and the fact the neutrinos oscillate between flavors which implies that they have a non-zero mass. This
416 phenomenon was measured independently by two collaborations, the Sudbury Neutrino Observatory
417 (SNO) and the Super-Kamiokande, in 1998 and 2001-2002, respectively, [18–20].

418 On the other hand, its theoretical formulation also has some weaknesses: it accurately describes
419 particles interactions at the electroweak scale (~ 246 GeV) but it does not include gravity which means it
420 cannot be valid at the Planck scale ($\sim 10^{19}$ GeV) where gravity cannot be overlooked; it has a lot (over 20)
421 of free parameters whose values have to be tuned to fit experimental observations, and there is a large
422 discrepancy between the mass scales associated with the electroweak and gravitational interactions
423 (this is one of the simplest formulations of what is known as the hierarchy problem).

424 When faced with these weaknesses, or rather hints of incompleteness, the theoretical community
425 put a great effort into the development of models that add new ingredients to the SM but that simultane-
426 ous lead to predictions that are compatible with existent experimental measurements. In the following
427 paragraphs we will introduce and briefly describe some of the most well studied (both theoretically and
428 experimentally) BSM models. We follow the discussion presented in [21] as a starting point.

429 It is a well known consequence of renormalization in QFT that the coupling constants become de-
430 pendent on the energy scale at which the theory is probed. As the energy scale increases the $U(1)$
431 coupling constant gets larger while the $SU(2)$ and $SU(3)$ coupling constants get smaller. If one extrap-
432 olates far enough these become nearly equal at an energy scale of approximately 10^{15} GeV. Although
433 this matching is far from perfect, it sparked the idea that these three forces could be unified at an energy
434 scale of 10^{15} GeV. Grand Unification Theories (GUT) try to combine $SU(3) \times SU(2) \times U(1)$ into a larger

6Or why do we live in an Universe made entirely out of matter?

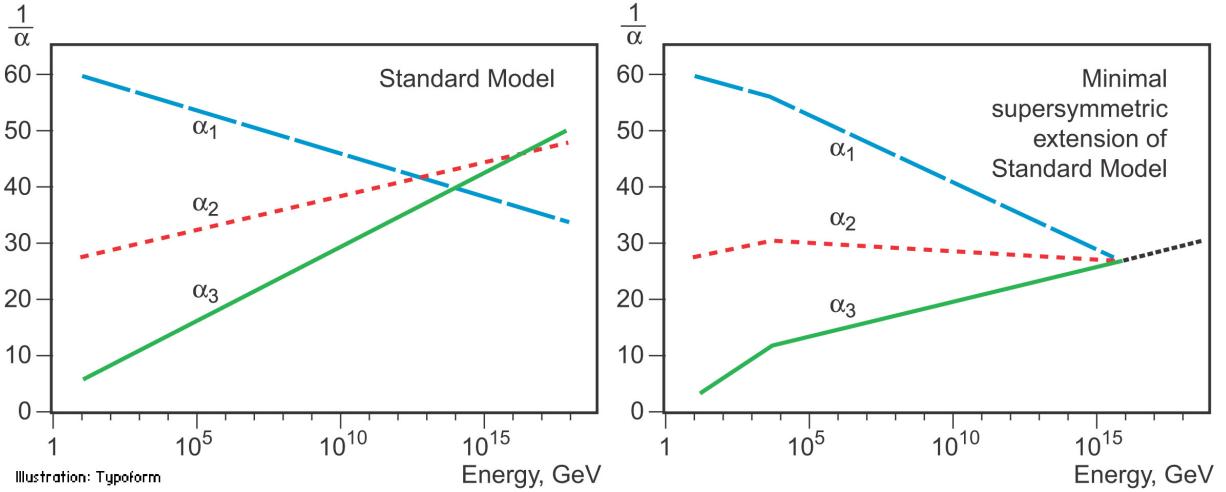


Figure 2.5: Here, α_1 , α_2 and α_3 represent the coupling constants of $U(1)$, $SU(2)$ and $SU(3)$, respectively. What is shown is the variation of the inverse of the coupling with the energy: on the left for the SM and on the right for the MSSM.[REF]

435 symmetry group.

436 Supersymmetric (SUSY) models introduce a new symmetry that links fermions and bosons. For
 437 each boson(fermion) of the SM it introduces a fermionic(bosonic) partner. Apart from spin, the super-
 438 symmetric partners would share the same mass and quantum numbers. Since we have not found any
 439 supersymmetric particles in the LHC this means that supersymmetry is necessarily a broken symmetry
 440 and, if they exist, new particles should have a larger mass (outside of the present reach of the LHC)
 441 than their SM partners. SUSY models were introduced because they offer a natural fix for the hierarchy
 442 problem. In addition, the Minimal Supersymmetric extension of the SM (MSSM) also leads to a better
 443 convergence of the coupling constants as can be seen in figure 2.5. From the standpoint of GUT this is
 444 extremely appealing.

445 An early proposal of a theory that could unify gravity with electromagnetism was given by Theodore
 446 Kaluza in 1921. In particular, he showed that these two forces could stem from a single tensor with
 447 the introduction of an extra space dimension. In 1926, Oscar Klein offered an explanation for this extra
 448 dimension; he proposed that it had a circular topology such that at each point of the four dimensional
 449 space-time we would have a circle with a small radius. This theory has more degrees of freedom
 450 (because it is formulated in a higher dimensional space-time) and therefore it predicts new particles that
 451 are usually known as Kaluza-Klein gravitons (and their excited states). Nonetheless, the Kaluza-Klein
 452 does not provide a satisfactory explanation for the hierarchy problem. Therefore, in 1999, Lisa Randall
 453 and Raman Sundrum introduced a new model that does. This model introduces only two new particles:
 454 a spin 2 graviton (and its Kaluza-Klein excitations) and a radion, that is a spin 0 neutral particle.

455 Models with two Higgs doublets (2HDM) are one of the simplest possible extensions of the Higgs
 456 sector of the SM. They are appealing because while the fermionic sector is rather complex, having three
 457 families, the scalar sector is quite simple, having a single particle, which seems unnatural. This type of
 458 structure is realized in various new physics models including SUSY models. In addition, they provide
 459 an additional source for CP violation which could help explain the matter-anti-matter asymmetry in the

460 Universe.

461 When studying the 2HDM there are several restrictions that can be placed on the parameters of
462 the model so as to reduce its complexity. A generally well accepted assumption is that there is no CP
463 violation coming, at tree-level, from the Higgs sector. This is the so-called CP-conserving 2HDM. In this
464 case, we obtain five Higgs bosons that are CP eigenstates. Three of them are neutral, h_1 , h_2 and h_3 ,
465 and the other two are charged, h^\pm . h_1 is usually taken to be the SM Higgs boson and its mass is set to
466 125 GeV.

467 Simplified dark matter (DM) models are based on the exchange of a single particle between DM and
468 SM particles and try to explain the nature of DM and how it interacts with the SM. The particle exchanged
469 is called a mediator and, depending on the specific model, it can be neutral or electrically charged and
470 have spin 0, 1 or 2. The simplest possible scenario is a neutral scalar mediator.

471 The 2HDM and dark matter model with a spin 0 mediator are explored in this work as sources of
472 alternative di-Higgs production processes. For these models, the main alteration to the SM occurs
473 because new heavy particles, namely, h_2, h_3 and the DM mediator can couple to the Higgs bosons
474 through the s-channel diagram. This corresponds to replacing the off-shell Higgs boson by one of these
475 particles in the Feynman diagram on the left in figure 2.3.

476 The spin-2 graviton predicted by Kaluza-Klein and Randal-Sundrum models can couple directly to
477 gluons and then decay producing a Higgs pair, as it illustrated in figure 2.6 on the left. SUSY particles
478 can contribute to the quantum loops in the Higgs production Feynmann diagrams. An example is shown
479 in figure 2.6 on the right where the top quark is replaced by its supersymmetric partner is the s-channel
480 diagram. These are two examples of how BSM models can change the Higgs pair production process
481 but there are many other. The crucial point is that some BSM contributions can lead to an enhancement
482 of the cross section for Higgs pair production with respect to the SM. Experimentally, this would mean
483 that we would not need as much sensitivity and therefore the process could be measured with less
484 data and therefore sooner. This is the reason why a lot of the searches performed at the LHC focus on
485 this type of scenario. In addtion, if a new heavy particle couples to the Higgs boson via an s-channel
486 diagram, it could lead to the existence of a peak in the Higgs pair invariant mass spectrum (assuming
487 that we have enough experimental resolution and that there are not other processes coming into play).

488 Moreover, BSM models introduce new free parameters (in addition to the SM ones) that can be
489 constrained using the experimental results obtained at the LHC (and other experiments). This reduces
490 the available parameter space of the models and may even reject some of them.

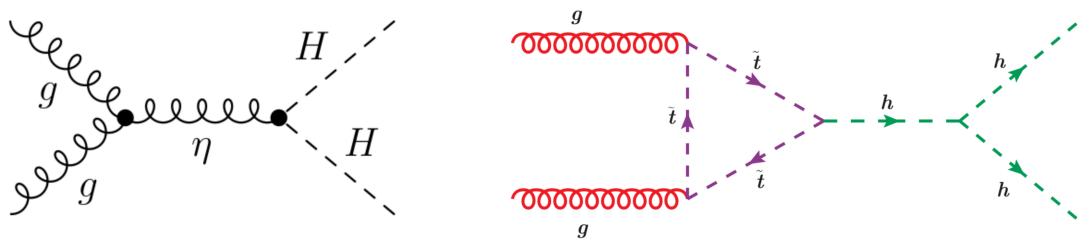


Figure 2.6: BSM contributions to Higgs pair production. On the left (figure from [22]) η represent the spin-2 Kaluza-Klein graviton that couples to a pair of SM Higgs bosons, H , at tree level. On the right (figure from [23]) \tilde{t} is the supersymmetric partner of the top quark that gives a contribution to the quantum loop in the Higgs pair production s-channel diagram.

491 **Chapter 3**

492 **Collider experiments**

493 In this chapter we start by providing an overview of the goals and main challenges of modern collider
494 experiments. The definitions of some key quantities that describe an accelerator are introduced and a
495 brief discussion on how they influence the discovery potential of an accelerator is presented.

496 In section 3.2 we introduce the LHC and in section 3.2.1 we describe the ATLAS experiment, includ-
497 ing brief discussions on b-tagging, trigger and data acquisition algorithms and systems. In section 3.3
498 we introduce the concept of a hadronic jet and describe how these objects are reconstructed in a general
499 collider experiment. Jet properties and substructure observables, as well as jet grooming algorithms,
500 are introduced in sections 3.3.1 and 3.3.2, respectively.

501 In section 3.4 we shift the focus to future collider experiments and accelerators and motivate their
502 need. In sections 3.4.1 and 3.4.2 we introduce the concept of the hadronic Future Circular Collider and
503 describe the baseline detector design, respectively.

504 Collider experiments are the best tool we have to explore matter's most fundamental structure. When
505 we accelerate a particle we increase its momentum. If we take into account the wave-particle duality and
506 the De Broglie expression, $\lambda = h/p$, where λ is the wavelength and p is the particle's momentum, we
507 can see that a particle with a large p will have a small λ . The wavelength gives us the dimension scale
508 of the objects we can probe with a given wave. If we want to probe very small particles (subatomic and
509 smaller) we need very small λ and therefore very large p . Conceptually, this is the basic idea behind
510 modern particle accelerators.

511 In practice, charged particles can be accelerated and their trajectories controlled by means of electro-
512 magnetic fields. However, this is not without numerous technical challenges. When a charged particle is
513 subject to an acceleration perpendicular to its velocity (which is exactly what happens in circular acceler-
514 ators) it emits electromagnetic radiation, called synchrotron radiation. The power emitted is proportional
515 to the fourth power of the particle's energy and inversely proportional to the radius squared and to the
516 fourth power of the particle's mass. This radiation limits the maximum energy that can be achieved in
517 electron-positron colliders. In proton-proton colliders, however, the energy is limited by the maximum
518 magnetic field that can be achieved. Therefore, there is also the need for extremely powerful magnets
519 which are usually implemented using technology based in superconductivity. Using superconducting

520 magnets raises another challenge: they can only operate at very low temperatures, close to the ab-
521 solute zero. In addition, in order to sustain a stable beam it is necessary that the beam pipe has an
522 environment very close to absolute vacuum.

523 3.1 Experimental aspects

One of the most important parameters of a particle accelerator is the time integrated luminosity, $\int \mathcal{L}(t)dt$. For a given process with cross section σ , it determines the number of event that will be produced, N :

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

where $\mathcal{L}(t)$ is the instantaneous luminosity that is a measure of the number of collisions per bunch crossing. The instantaneous luminosity is given by:

$$\mathcal{L} = f_{coll} \frac{n_1 n_2}{4\pi \sigma_x \sigma_y}, \quad (3.2)$$

524 where f_{coll} is the collision frequency, n_1 and n_2 are the number of protons in each bunch and σ_x and σ_y
525 characterize the transverse beam size in the horizontal and vertical directions.

526 To increase the chances of measuring a rare process, or to increase the statistical significance of the
527 measurement of an already discovered process, we want to increase N as much as possible. To do so
528 we can either increase the cross section of the process or the integrated luminosity.

529 While the cross sections of most physics processes increase when the CM energy goes from 13 to
530 100 TeV, many BSM models predict new processes, or new contributions to existing processes, whose
531 cross sections increase more rapidly than the SM backgrounds. In addition, by conservation of energy,
532 a larger CM energy implies that particles with larger mass can be created. Based on Eq. 3.2, we can
533 tune its parameters to obtain the highest possible luminosity. Nonetheless, because we are dealing with
534 charged particles, there is a limit on how close the bunches can be and on how many protons we can
535 pack in a bunch. Moreover, the beam's transverse dimensions cannot be infinitely reduced. A smarter
536 way to increase the number of collisions that an accelerator can produce is to run for a longer time,
537 therefore increasing the integration time in Eq. 3.1.

538 In conclusion, the CM energy and the integrated luminosity are two of the main parameters that drive
539 the discovery potential of an accelerator.

540 3.2 The Large Hadron Collider

541 IDEAS FOR SECTION

- 542 - The LHC: goals and physics programs
- 543 - Successes, discoveries
- 544 - Upgrades and outlook

545

546 The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. It
547 is housed by the European Organization for Nuclear Research (CERN) which focuses on fundamental
548 particle physics with the goal of probing matter's most elementary structure. Ever since its creation, in
549 1954, CERN has housed many accelerators and experiments and played a key role in the development
550 of fundamental and applied science.

551 The LHC consists of a 27-kilometer ring located beneath the Franco-Swiss border, near Geneva.
552 Most of its running time is dedicated to accelerating protons up to a maximum center of mass energy
553 (\sqrt{s}) of 13 TeV and colliding them at the center of the two general purpose experiments, ATLAS (which
554 is described in section 3.2.1) and CMS. The LHCb experiment also records data from proton-proton
555 collisions but it is dedicated to the study of beauty particles. The ALICE experiment is optimized to study
556 heavy-ion collisions at a CM energy of 2.76 TeV.

557 The acceleration of charged particles at the LHC is based on radio frequency (RF) cavities. These
558 cavities are shaped to sustain a resonant electromagnetic field that oscillates at a frequency of 400 MHz.
559 During the acceleration stage, charged particles passing through the cavities feel an overall force that
560 propels them forward. When the LHC is running at full energy, a perfectly timed proton with exactly the
561 right energy feels a zero net force when passing the cavities. Protons with a slightly different energies
562 arriving slightly earlier or later are decelerated or accelerated in order to keep the beam sorted in dis-
563 crete packages with the same energy. These are called bunches. There are 2808 bunches circulating at
564 the same time, each containing approximately 10^{11} protons. The bunches are spaced by 25 ns. Further-
565 more, the successful operation of the LHC also relies on superconducting magnets made of Niobium-
566 Titanium filaments chilled to -271.3°C and on an ultra high vacuum (of the order of $10^{-10} - 10^{-11}\text{ mbar}$)
567 inside the beam pipes. The magnets are placed along the LHC ring and produce dipole and quadrupole
568 electromagnetic fields. The dipole magnets create a nominal field of 8.3 T and bend the beam along the
569 tunnel. The quadrupole magnets focus the beam at the interaction points. The ultra high vacuum greatly
570 reduces the probability that the beam interacts with any particle. It is crucial to keep a stable beam to
571 continuously maintain collisions during long runs.

572 One of the main research goals of the LHC was to discover the Higgs boson. This was achieved
573 in 2012 when ATLAS and CMS reported the discovery of a particle consistent with the boson predicted
574 by the Higgs mechanism, with a mass of 125 GeV [24],[25]. Ever since, efforts have been directed
575 to measuring its mass, couplings, spin-parity properties with increasing precision using different decay
576 channels and production modes.

577 ATLAS and CMS reported evidence (measurement with a significance greater than three sigma)
578 for the Higgs decaying to $b\bar{b}$ [26, 27] [UPDATE WITH DISCOVERY]. The searches targeted the VH
579 production mode. It offers the best sensitivity to the $h b\bar{b}$ Yukawa coupling because requiring a vector
580 boson helps reduce the SM backgrounds, namely the ones from QCD interactions. CMS reported the
581 first observation (measurement with a significance greater than five sigma) of the Higgs boson decaying
582 to a pair of tau leptons [28]. In addition, the observation of Higgs boson production in association with a
583 $t\bar{t}$ pair was very recently reported by both collaborations [29, 30]. Moreover, precision measurements of

584 the masses of the Higgs [31, 32] and W [33] bosons and of the top quark [34, 35] were also performed.
585 So far, no conclusive signs of new physics were seen at the LHC.

586 Future prospects for the LHC include its upgrade to the High Luminosity-LHC (HL-LHC) after the
587 scheduled long shutdown of 2024-2026. This upgrade will increase the size of the dataset to 3000 fb^{-1}
588 over the course of ten years [36]. During the shutdown, the ATLAS detector will be upgraded.

589 In the Higgs sector, the high value of the integrated luminosity will improve the statistical precision of
590 already measured channels and the discovery potential of rare processes [37].

591 3.2.1 The ATLAS detector

592 The ATLAS detector has a cylindrical geometry and a multi layered structure. Its dimensions are 25
593 meters in height (diameter) and 44 meters in length and it weights approximately 7000 tonnes. In the fol-
594 lowing paragraphs we describe the detector's layers and their functionalities. A schematic representation
595 of the detector as well as the appropriate coordinate system can be found in figure 3.1.

596 A combination of cartesian and cylindrical coordinates is used to describe the detector. In both cases,
597 the origin is defined to coincide with the interaction point. The Cartesian system is right-handed and the
598 z axis is defined to be the direction of the beam. The x-axis point from the interaction point to the center
599 of the LHC ring and the y-axis point upwards. The azimuthal angle, ϕ , is measured around the beam axis
600 and the polar angle, θ , from the beam line. The pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. Another
601 commonly used quantity is the rapidity, y , defined as a function of a particle's energy, E , and longitudinal
602 momentum, p_L : $y = \frac{1}{2} \ln \left(\frac{E+p_L}{E-p_L} \right)$. In the limit where a particle's mass is negligible with respect to its
603 momentum the pseudorapidity converges to the definition of rapidity. In addition, the angular distance
604 between two points, ΔR , is defined as $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where rapidity can also be used instead
605 of the pseudorapidity.

606 The detector consists of an inner detector (ID) or tracker, electromagnetic (EM) and hadronic calorime-
607 ters and a muon spectrometer (MS). The magnet configuration consists of a thin superconducting
608 solenoid that surrounds the ID cavity and three superconducting toroids (one barrel and two end-caps)
609 arranged with an eight-fold azimuthal symmetry around the calorimeters.

610 The ID covers the pseudorapidity range $|\eta| < 2.5$ and it makes up the innermost layer of the detector.
611 It consists of silicon pixel, silicon micro-strip, and straw tube transition radiation tracking detectors. The
612 tracker provides precision measurements of the positions and momenta of charged particles. As a
613 charged particle transverses the several layers of the ID it ionizes the medium creating electrical signals
614 that can be read out. These individual electrical signals are then combined to reconstruct the trajectory
615 of the particle.

616 Lead/Liquid-Argon (LAr) sampling EM calorimeters cover the pseudorapidity range $|\eta| < 3.2$. The
617 EM calorimeter has an accordion like structure with layers of showering material (lead) interleaved with
618 layers of active material (liquid argon). These calorimeters provide measurements of the energy of
619 electrons and photons. The interaction of these particles with the lead layers induces the production
620 of an EM shower whose energy is measured in the liquid argon layers. The granularity of the EM

| TileCal | Barrel | Extended barrel |
|------------------------|--|------------------------|
| Coverage | $ \eta < 1.0$ | $0.8 < \eta < 1.7$ |
| Granularity | 0.1×0.1 | 0.1×0.1 |
| LAr calorimeter | End-cap | Forward |
| Coverage | $1.5 < \eta < 3.2$ | $3.2 < \eta < 4.9$ |
| Granularity | 0.1×0.1 for $1.5 < \eta < 2.5$ 0.2×0.2 for $2.5 < \eta < 3.2$ | 0.2×0.2 |

Table 3.1: ATLAS tile and liquid argon hadronic calorimeters: summary of the pseudorapidity coverages and transversal segmentation (granularity).

621 calorimeter strongly depends on the longitudinal layer and on the pseudorapidity region.

622 The hadronic calorimetry in the pseudorapidity range $|\eta| < 1.7$ is provided by a scintillator-tile
623 calorimeter (TileCal) which is divided in a central barrel and two smaller end-cap barrels, one on each
624 side of the central barrel. The active components are scintillator tiles made of polystyrene that are in-
625 terleaved with steel plates as the passive material. The scintillation light emitted by the tiles when an
626 ionising particle crosses the calorimeter is collected on both ends of the tiles by wavelength-shifting op-
627 tical fibers. The light signal emitted is proportional to the particle's energy. The TileCal is composed of
628 several cells with transverse segmentation $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ¹ [38].

629 For $|\eta| > 1.5$ LAr calorimeters extend the pseudorapidity range to $|\eta| = 4.9$. The LAr calorimeter is
630 divided in end-cap and forward. These cover the pseudorapidity ranges $1.5 < |\eta| < 3.2$ and $3.2 < |\eta| <$
631 4.9 , respectively. The active material is liquid-argon and the absorbers are copper and tungsten for the
632 end-cap and forward calorimeters, respectively. In the end-cap LAr calorimeters the segmentation is
633 $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $1.5 < |\eta| < 2.5$ and 0.2×0.2 for $2.5 < |\eta| < 3.2$. In the forward LAr calorimeter
634 the segmentation is $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$.

635 The granularity of the ATLAS hadronic calorimeters is summarized in table 3.1.

636 The hadronic calorimeters provide measurements of the energy of hadrons, jets, τ leptons and miss-
637 ing transverse energy (E_T^{miss}). Approximately one third of the energy of jets is deposited in this layer.
638 In the TileCal, the jet energy resolution is given by $\sigma/E \sim 50\%/\sqrt{E} + 3\%$ [38], where the first term is the
639 stochastic term that derives from sampling fluctuations and follows a Poisson distribution and the second
640 term is a constant that depends on the characteristics of the calorimeter. For the LAr calorimeter, the jet
641 energy resolution is given by $\sigma/E \sim 60\%/\sqrt{E} + 2\%$ [39].

642 The MS is the outermost layer of the detector and it is dedicated to detecting muons that travel
643 through the previous layers almost without interacting. This layer provides measurements of the muons
644 transverse momenta. It is composed of Monitored Drift Tubes (MDT) and Cathode Strip Chambers
645 (CSC) that provide high precision measurements of the muons' momentum in the pseudorapidity range
646 $|\eta| < 2.7$ and of Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) dedicated to triggering
647 purposes for $|\eta| < 2.4$.

¹The TileCal is composed of three longitudinal layers. Only the first two have a segmentation equal to 0.1×0.1 . In the third layer the segmentation is 0.2×0.1 . However, most of the energy of hadronic showers is deposited in the first layers and therefore this detail is not very relevant for this work.

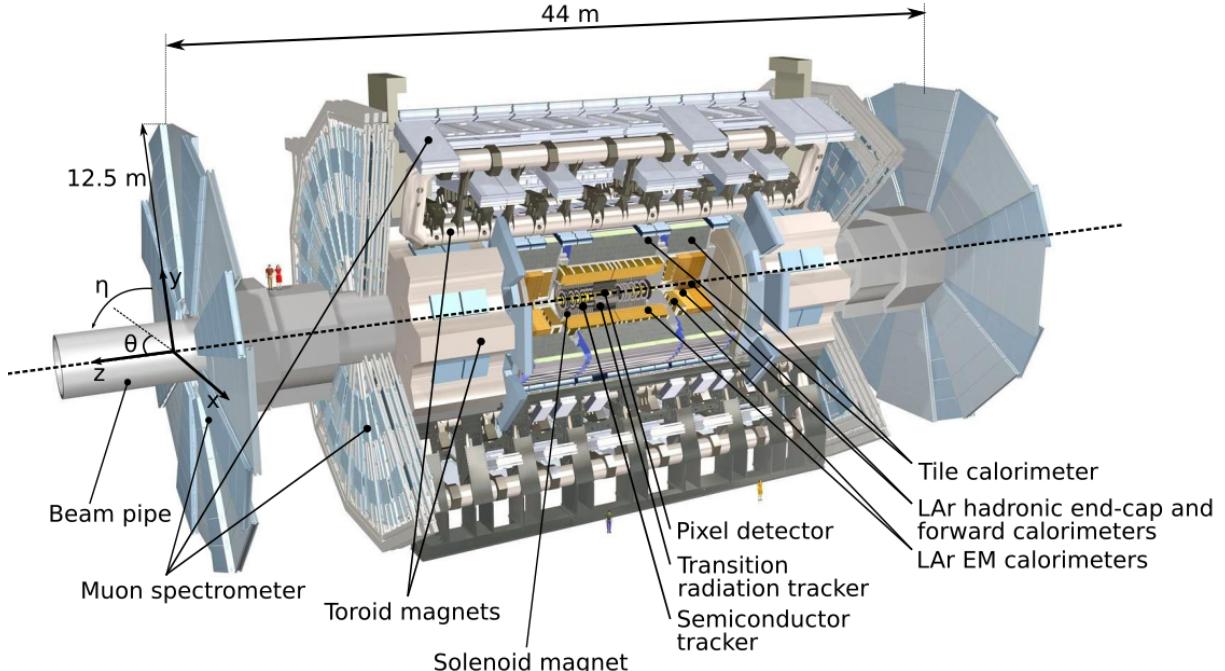


Figure 3.1: ATLAS detector.

648 **b-Tagging**

649 Each collision produces a large number of hadronic jets (we refer to section 3.3 for a detailed description
 650 of jets and how they are reconstructed). For this work, jets initiated by a b quark (b jets) are particularly
 651 important: we are looking for a Higgs pair decaying to four b quarks which leads to an experimental
 652 signature that consists of four b-jets. b-Tagging algorithms determine, with a given probability, if a jet
 653 was originated by a b quark.

654 When a b quark is produced it hadronizes almost instantly, producing a B hadron. B hadrons have a
 655 life time of ~ 1 ps and can be highly relativistic meaning that they can travel a few millimeters to a few
 656 centimeters inside the inner detector before decaying. When they decay there is often a reconstructible
 657 secondary vertex that is slightly displaced from the primary vertex where the b-quark was produced.
 658 The existence of a secondary vertex is used by b-Tagging algorithms to identify, or tag, a jet as coming
 659 from a b-quark. It is important to note that a complete b-tagging algorithm relies on the reconstruction of
 660 a secondary vertex which can only be done using the information from the inner detector. This implies
 661 that, in ATLAS, we can only b-tag jets that are produced in the region $|\eta| < 2.5$.

662 In ATLAS, b tagging algorithms are applied to the sub-set of tracks that are associated with that
 663 jet. The matching between tracks and calorimeter-based jets is performed using the ghost association
 664 technique [40]². The identification of b-jets in ATLAS is based on distinct strategies encoded on three
 665 b-tagging algorithms: impact parameter-based algorithms, an inclusive secondary vertex reconstruction
 666 algorithm and a decay chain multi-vertex reconstruction algorithm. The output of these algorithms are

²This procedure works by introducing ghost versions of the measured tracks that have the same direction but infinitesimally small p_T such that they do not modify the properties of the calorimeter jets. The jets are then reclustered and a track is considered to be associated with a given jet if its ghost version is contained in the jet after reclustering.

667 combined in a multivariate discriminant based on a Boosted Decision Tree (BDT) which provides the
668 best discrimination between the different jet flavors [41].

669 The impact parameter-based algorithms, IP2D and IP3D, use as discriminant variables the trans-
670 verse impact parameter significance and the transverse and longitudinal impact parameter significance,
671 respectively. The secondary vertex finding algorithm, SV, explicitly reconstructs a displaced secondary
672 vertex inside the jet by trying to find pairs of tracks with a common origin. The decay chain multi-vertex
673 reconstruction algorithm, JetFitter, tries to reconstruct the full b-hadron decay chain. This approach
674 allows to resolve b- and c-hadrons vertices even if there are not two tracks associated with them.

675 **Trigger and data acquisition**

676 The LHC delivers approximately 1000 million proton-proton collisions per second, which corresponds to
677 an event rate of 1 GHz. On the one hand, only a small fraction of these events result in interesting physics
678 processes. On the other hand, the detector does not have enough storage and read out capabilities to
679 record all the collisions. The triggering and data acquisition systems are responsible for selecting a
680 manageable rate of events for permanent storage and further analysis.

681 The trigger is responsible for selecting events with interesting experimental signatures. The trigger
682 system in Run-2 consists of a hardware-based first level trigger (Level-1) and a software-based high
683 level trigger (HLT). The Level-1 trigger takes as input coarse granularity calorimeter and muon detector
684 information and reduces the event rate to 100 kHz. The HLT uses full granularity detector information
685 and reduces the rate to approximately 1 kHz [42].

686 [USE ANOTHER EX?] Take, as an example, the inclusive cross section for the production of jets
687 which is of the order of 10^6 fb. If we consider a luminosity of 0.8 fb^{-1} per day, this corresponds to a rate
688 of approximately 200 Hz, just for the production of jets. Note that this value is the same as the maximum
689 allowed rate for all physics processes after all the triggers. Therefore, it is clear that we cannot simply
690 take all the jet events. We need to place tight cuts that allow us to reduce this rate to an acceptable
691 (within the available quotas) rate. This is of particular importance for processes with very large cross
692 sections such as QCD multijet production.

693 **3.3 Jet reconstruction**

694 A jet is a collimated spray of hadrons that is interpreted as coming from a single initial parton such
695 that it approximately retains information about its physical properties, namely 3-momentum, mass and
696 charge. The existence of such objects is a direct consequence of the confinement property of Quan-
697 tum Chromodynamics (QCD). Quarks and gluons, the fundamental degrees of freedom of QCD, are
698 not asymptotically free. They are confined inside hadrons. Therefore, when one of these particles is
699 produced it undergoes showering and hadronization processes that lead to the formation of hadrons.

700 At a particle detector we are interested in reconstructing jets. These are the objects that are used in
701 the physics analysis. Working with jets instead of hadrons is an advantage because it greatly reduces the

702 number of objects we need to analyze per event. In addition, they work as a *proxy* for the fundamental
703 partons produced in the event.

704 Jets are obtained through jet finding algorithms. These are clustering algorithms that group together
705 experimental quantities (energy deposits in the calorimeters, tracks or particle flow objects) using a
706 sequential recombination scheme. A crucial property of these algorithms is that they should be infrared
707 and collinear safe, meaning that the outcome of the algorithm (namely the number of jets and their
708 properties) should not be significantly modified by the emission of soft radiation or by a collinear splitting.

709 The most widely used jet finding algorithms are the k_T [43], the Cambridge/Aachen (C/A) [44] and
710 the anti- k_T [45]. They are based on distance measurements:

$$d_{ij} = \min(k_{Ti}^{2p}, k_{Tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad d_{iB} = k_{Ti}^{2p}, \quad (3.3)$$

711 where d_{ij} is the distance between entities i and j and d_{iB} is the distance between entity i and the beam
712 (B). $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{Ti} , y_i and ϕ_i are the transverse momentum, rapidity and azimuth
713 of entity i . Here, k_T stands for the transverse momentum. For $p = 1, 0, -1$ we get the k_T , C/A and
714 anti- k_T algorithms, respectively.

715 The clustering algorithm starts by computing all d_{ij} and d_{iB} distances. If the smallest distance is a
716 d_{ij} , the four momenta of particle i and j are summed and the distances are updated. If the smallest
717 distance is a d_{iB} , particle i is removed and called a jet. This procedure is iterated until all particles are
718 clustered in jets.

719 It is worth discussing further the anti- k_T algorithm since it is the default algorithm used in most
720 analyses, including this one. In this algorithm, the Δ_{ij} distance between constituents i and j is weighted
721 by the inverse of the transverse momentum of the constituent with a largest k_T . This feature implies that
722 particles with larger momenta will have a smaller d_{ij} distance and therefore will be clustered first. This
723 prevents soft particles from being clustered among themselves before clustering the hardest particles.

724 In ATLAS, the transverse and longitudinal segmentation of the calorimeters allow for a three dimensional
725 reconstruction of particle showers which is based in a topological clustering algorithm. Topo-
726 clusters of calorimeters cells are seeded by cells whose absolute energy exceeds the electronic and
727 pile-up noise by four standard deviations. The topo-clusters are then expanded by adding all adjacent
728 cells with absolute energy two standard deviations above noise. Finally, all cells neighbouring the pre-
729 vious set are also added. After energy calibration, the topo-clusters are fed as input to a jet finding
730 algorithm.

731 **Boosted kinematic regime**

732 With the increase of the CM energy of particle colliders the production of particles with a transverse
733 momentum much larger than their mass became a reality. In this kinematic regime (referred to as boosted
734 due to the high Lorentz boost of the particles) traditional jet reconstruction algorithms, that rely on a one-
735 to-one correspondence between jets and partons, begin to fail [46].

Due to the high Lorentz boosts, decay products of heavy resonances get more collimated. The

angular separation of the decay products is approximately [47]:

$$\Delta R \sim \frac{2m}{p_T} \quad (3.4)$$

where p_T and m are the transverse momentum and mass of the decaying particle. In addition to decaying particles with a larger p_T , another event topology that can produce highly collimated particles is the decay of a particle with a very large mass (directly seen from 3.4). This scenario is of particular interest for new physics searches because BSM models often predict the existence of heavy particles.

Take, as an example, the decay of a Higgs boson to two b quarks. Considering that the p_T of the Higgs is approximately 200 GeV (which is a reasonable and commonly used value in boosted Higgs bosons searches) we get $\Delta R \sim 1$. For resolved jets, the default jet radius parameter used in ATLAS is 0.4. We see that for a $p_T \sim 200$ GeV the angular separation between the decay products of a Higgs boson is already similar to the default jet diameter, which can jeopardize the ability to resolve the individual decay products.

A possible workaround is to use a single jet with a larger R parameter to reconstruct both decay products. The problem with doing so is that we no longer have information about each individual decay product which means we are loosing some information about the event. We cannot, for example, compute angular variables between the decay products. This led to the development of techniques and observables that allow for the exploration of the intrinsic structure (or substructure) of these large- R jets. Some of these techniques are introduced and discussed in the following section.

3.3.1 Jet properties and substructure observables

TO INCLUDE (it will depend on the ones that are most important to the analysis):

- tau_N
- energy correlation functions and ratios
- Fox Wolfram
- ...
- Jet mass definition!

In general, jet substructure variables aim to quantify the existence of energy clusters insider a jet. Each cluster is interpreted as corresponding to an individual jet. These are called subjets because they are contained inside the large- R jet that was actually reconstructed. Once they are identified, the subjets can be handled and used for the analysis like normal jets. However, a lot of the substructure techniques do not focus on reconstructing the subjets but rather on determining whether or not they exist, how many they are and how are they distributed inside the large- R jet.

Heavy resonances decaying to two(three) particles will produce large- R jets that are consistent with the existence of two(three) energy clusters. Examples of such topologies are the Higgs and top decays: the Higgs boson always decays to pairs of particles (leptons, quarks or bosons) and the top quark

770 decays, with a probability close to 100%, to a b quark and a W boson that then decays to a pair of
 771 leptons or quarks. These topologies are usually referred to as two or three prong. In contrast, jets
 772 initiated by a gluon or quark splitting are not expected to have a meaningful substructure. The energy
 773 is expected to be concentrated around the jet axis following an isotropic distribution and to become less
 774 dense as we approach the jet's border. This is the signature of a one-prong topology. The previous
 775 discussion is valid at LO and captures the generic features of jets that are targeted by jet substructure
 776 techniques. Nonetheless, other effects may come into play. Take, for example, a highly virtual gluon with
 777 a high p_T . If it splits into two quarks the resulting jet may have two subjets and thus mimic the topology
 778 of a heavy resonance decay.

The N-subjetiness variable [48], τ_N , may be used to identify jets compatible with N subjets. It is given
 by

$$\tau_N = \frac{1}{d_0} \sum_k p_T^k \min(\Delta R_1^k, \dots, \Delta R_N^k), \quad d_0 = \sum_k p_T^k R_0 \quad (3.5)$$

779 where the index k runs over all particles in a jet, the indexes 1 to N identify the number of axis inside the
 780 jet and R_0 is the radius of the jet. This variable will have a small value if the particles with the highest p_T
 781 (the ones we are most interested in) are clustered around the axis (because ΔR will be small) and will
 782 have a larger value otherwise. A jet with small τ_N is considered to be consistent with having N or fewer
 783 subjets because all its constituents are aligned with the axis.

784 We usually use ratios of τ_N variables ($\tau_{MN} = \tau_M / \tau_N$). Of particular interest for this work are the τ_{21} ,
 785 τ_{31} and τ_{32} ratios. These observables can take values between zero and one. For τ_{21} , for example, a
 786 small value (close to zero) indicates that the jet is more compatible with two subjets than with one and
 787 therefore it can help discriminate between two-prong and one-prong jets. The same applies to every
 788 other ratio.

In addition to jet substructure observables, more standard variables, from which we highlight the jet
 mass, can also be used to discriminate between jets coming from QCD background and jets resulting
 from heavy resonance decays. In the latter, the jet's invariant mass should roughly correspond to the
 mass of the resonance. The invariant mass of a jet, M , is calculated from the energies and momenta of
 its constituents as follows:

$$M = \left(\sum_i E_i \right)^2 - \left(\sum_i \vec{p}_i \right)^2 \quad (3.6)$$

789 where E_i and \vec{p}_i are the energy and three-momentum of the i^{th} constituent. However, the mass resolu-
 790 tion is not expected to be very good for large-R jets because of all the extra QCD radiation that may be
 791 caught inside the jet. A possible workaround is the use of jet grooming algorithms which we describe in
 792 the following section.

793 3.3.2 Jet grooming algorithms

794 The main goal of jet grooming algorithms is to remove contamination of softer jet constituents from pileup
 795 or underlying event and to leave behind the hard substructure. The main advantage of such algorithms is
 796 that they improve the mass resolution of jets. These features are of particular interest in high luminosity

797 environments such as the HL-LHC and future high energy colliders.

798 There are three main jet grooming algorithms: trimming, pruning and mass drop filtering. In this
799 work we do not use pruning or trimming, although these techniques might be worth exploring in future,
800 more comprehensive, studies. In particular, they can be useful to help reject pileup contributions. The
801 mass drop filtering procedure isolates relatively symmetric subjets within a jet, each with a significantly
802 smaller mass than the original jet [47]. This technique was developed and optimized using C/A jets for
803 the search of Higgs decaying to $b\bar{b}$ pairs [49]. It works as follows:

- The last step of the C/A clustering is undone such that the jet is split into subjets, j_1 and j_2 with $m_{j_1} > m_{j_2}$. We require that there is a significant difference between the mass of the original jet, m_{jet} , and m_{j_1} : $m_{j_1}/m_{\text{jet}} < \mu_{\text{frac}}$, where μ_{frac} is a parameter of the algorithm. In addition the splitting is required to be relatively symmetric:

$$\frac{\min[(p_T^{j_1})^2, (p_T^{j_2})^2]}{(m_{\text{jet}})^2} \times \Delta R_{j_1, j_2}^2 > y_{\text{cut}} \quad (3.7)$$

804 where y_{cut} is a parameter that defines the energy sharing between the subjets. It is usually taken
805 to be ~ 0.09 . If these two criteria are not met the jet is discarded.

- 806 • The subjets are clustered with the C/A algorithm with radius parameter $R_{\text{filt}} = \min[0.3, \Delta R_{j_1, j_2}/2]$.
807 All jet's constituents that are outside of the three hardest subjets are discarded and we obtain
808 the filtered jet and its subjets.

809 The subjets that are identified within a jet are interpreted as corresponding to the decay products of
810 the particle that produced the original jet. There are usually only two decay products. A third subjet is
811 allowed in order to account for extra QCD radiation.

812 3.4 Future Colliders

813 As we already argued in the beginning of this chapter, a larger CM energy is one of the factors driving
814 the discovery potential of an accelerator. As far as we know today, proton-proton colliders are the main,
815 and possibly only, man-made experimental tool available to explore particle physics in the energy range
816 on tens of TeV. With this in mind, new hadronic colliders with CM energies of the order of tens of TeV
817 have been proposed. The main projects are the hadronic Future Circular Collider (FCC-hh) lead by
818 CERN and the Super Proton-Proton Collider (SPPC) proposed by China.

819 In this work we focus on FCC-hh and use the established detector's baseline design as our starting
820 point for this study. In the following section we describe in detail the FCC-hh accelerator and detector.

821 3.4.1 The hadronic Future Circular Collider

822 IDEAS FOR SECTION

- 823 - Center of mass energy ✓
824 - Basic design: circular, circumference, location, ... ✓
825 - CERN FCC study group: creation, mission, main tasks ✓

- 826 - The role of the hadronic calorimeter
827 - Basic structure: tracker, calorimeters, muon chambers ✓
828 - Technical specifications: nb. bunches, magnets, ... ✓
829 - Technical dimensions and materials (predicts) of each layer ✓
830 - Focus on hadronic calorimeter: material being studied, resolution, granularity ✓
831 - What processes we want to study at the FCC
832 - Focus on Higgs pair production

833

834

835 CERN's FCC study group was launched as a result of a recommendation made in the 2013 update
836 of the European Strategy for Particle Physics that '*Europe needs to be in a position to propose an*
837 *ambitious post-LHC accelerator project at CERN by the time of the next Strategy update*'. It investigates
838 the technological challenges and physics opportunities of a future circular collider. The design and
839 infrastructure are driven by a proton-proton collider (FCC-hh) requirements. Electron-positron (FCC-ee)
840 and electron-proton (FCC-eh) colliders are also being analyzed. The main goal of this effort is to deliver
841 a Conceptual Design Report (CDR) by the end of 2018. This document will include a first cost estimate
842 to be submitted for the next update of the European Strategy for Particle Physics, foreseen by 2018.

843 The FCC-hh baseline design consists of a proton-proton circular collider with a maximum CM energy
844 of 100 TeV housed by a 100 km tunnel in the area of Geneva. This machine will extend the research
845 program of the LHC (and of the HL-LHC) after these have reached their full discovery potential, by
846 around 2040. In addition, it will allow for the exploration of an entirely new kinematic regime, probing
847 energy scales where new physics may come into play. A possible way of defining the target luminosity of
848 this machine is to require that within the first year of operation it surpasses the exploration potential of the
849 LHC [50]. Comprehensive studies [50, 51] indicate that this can be achieved with an integrated luminosity
850 of the order of 10 ab^{-1} per experiment. Considering a reasonable operation period of 10 years this leads
851 to integrated luminosity per experiment of the order of 1 ab^{-1} per year.

852 The FCC-hh is expected to work with a dipole field of 16 T and to provide a peak instantaneous lumi-
853 nosity thirty times larger than the LHC. The number of bunches is expected to be almost a factor of four
854 larger than for the LHC and the number of events per bunch crossing is expected to be approximately
855 1000. The latter brings a lot of technical challenges because it means that the mean pileup expected
856 for the FCC-hh is almost 40 times larger than for the LHC. This requires the development (or improve-
857 ment of already existing techniques) of techniques and algorithms that allow us to further reject pileup
858 contributions.

859 Some relevant parameters of the LHC and FCC-hh are summarized in table 3.2.

860 3.4.2 FCC-hh baseline detector

861 The design of the FCC-hh baseline detector, which we describe in detail in this section, has been greatly
862 based on that of the ATLAS and CMS experiments, in particular the central barrel. The layers and sub

| Parameter | LHC | FCC-hh |
|---|------|--------|
| Circumference [km] | 27 | 100 |
| CM energy [TeV] | 13 | 100 |
| Luminosity (peak) [$10^{34} \text{cm}^{-2}\text{s}^{-1}$] | 1 | 30 |
| Dipole field [T] | 8.33 | 16 |
| Nb. of bunches | 2808 | 10600 |
| Nb. of events per bunch crossing | 27 | 1026 |

Table 3.2: Comparison between the working parameters of the LHC and of the FCC-hh. The values of the number bunches and of the number of events per bunch crossing are given assuming a bunch spacing of 25 ns.

detectors are arranged in the same order and perform very similar roles. The geometry is cylindrical and therefore we can use exactly the same coordinate system that was introduced in section 3.2.1. The dimensions are very close to the ones of ATLAS: 25 meters in height (diameter) and 48 m in length. A schematic representation of the FCC detector is shown in figure 3.2.

The detector consists of trackers, EM and hadronic calorimeters and MS. The magnet configuration consists of three solenoid magnets (one central-barrel and two forward) that surround the central barrel calorimeters and the forward trackers.

The tracker covers the pseudorapidity range $|\eta| < 6$ and is divided in three sub systems: inner, outer and forward. The inner and outer trackers and the forward tracker are expected to cover the pseudorapidity ranges $|\eta| < 2.5$ and $2.5 < |\eta| < 6.0$, respectively. The inner tracker will be instrumented with pixel detectors while the outer and forward tracker will have layers of both pixel and strip detectors.

The EM calorimeter covers the pseudorapidity range $|\eta| < 6$. It is divided in barrel, end-cap and forward. These cover the pseudorapidity ranges $|\eta| < 1.5$, $1.4 < |\eta| < 2.5$ and $2.3 < |\eta| < 6$, respectively. The proposed layout for the EM calorimeter is a LAr sampling configuration with lead, glue and steel plates as absorbers. The granularity is expected to be two to four times better than for the ATLAS ECAL. For the barrel calorimeter, the goal energy resolution is $10\%/\sqrt{E} \oplus 1\%$.

The hadronic calorimeter covers the pseudorapidity range $|\eta| < 6$. It is also divided in barrel, end-cap and forward that cover the pseudorapidity ranges $|\eta| < 1.3$, $1.0 < |\eta| < 1.8$ and $2.3 < |\eta| < 6.0$, respectively. The proposed layout for the barrel calorimeter consists of scintillator tiles interleaved with lead and stainless steel plates as absorbers. The end-cap and forward calorimeters are expected to be based on liquid argon with copper plates as absorbers. For the barrel and end-cap calorimeters, the expected segmentation $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ while for the forward calorimeter it is $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$. Overall, this corresponds to approximately four times the ATLAS HCAL granularity. The energy resolution is expected to be $40\%/\sqrt{E} \oplus 2.5\%$, $50\%/\sqrt{E} \oplus 3\%$ and $100\%/\sqrt{E} \oplus 5\%$, for the barrel, end-cap and forward calorimeters, respectively. The pseudorapidity coverage, layout, granularity and energy resolution of the hadronic calorimeters are summarized in table 3.3.

The muon spectrometer is divided in barrel, end-cap and forward regions that cover the pseudorapidity ranges $|\eta| < 1.0$, $1.0 < |\eta| < 2.5$ and $2.5 < |\eta| < 6.0$, respectively. The muon's system layout is a layered structure of gas chambers.

| Parameter | Barrel | End-cap | Forward |
|--|------------------------------|----------------------------|-----------------------------|
| η coverage | $ \eta < 1.3$ | $1.0 < \eta < 1.8$ | $2.3 < \eta < 6.0$ |
| Layout | Sci-Pb-Steel (1 : 1.3 : 3.3) | LAr-Cu (1 : 5) | LAr-Cu (1 : 200) |
| Granularity ($\Delta\eta \times \Delta\phi$) | 0.025×0.025 | 0.025×0.025 | 0.05×0.05 |
| Energy resolution (σ_E/E) | $40\%/\sqrt{E} \oplus 2.5\%$ | $50\%/\sqrt{E} \oplus 3\%$ | $100\%/\sqrt{E} \oplus 5\%$ |

Table 3.3: Hadronic calorimeter layout, granularity and energy resolution.

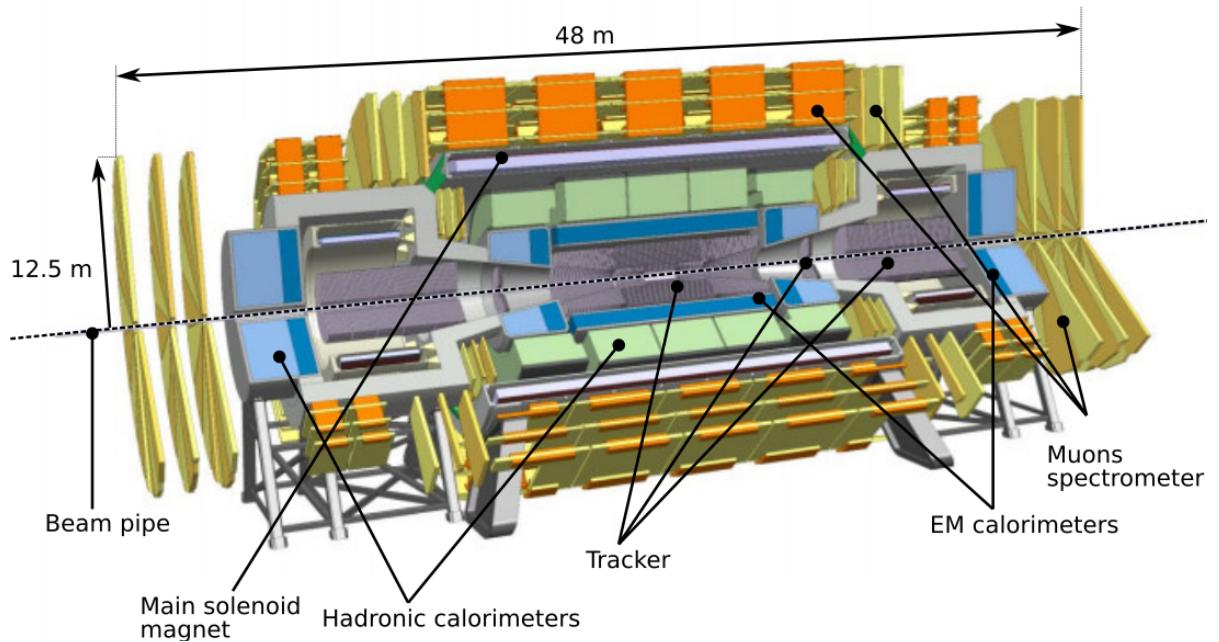


Figure 3.2: FCC detector baseline concept.

892 3.4.3 FCC-hh physics program

893 The FCC-hh physics program is vast and diverse. Within the SM, it includes the study of gauge bosons
894 pair production and heavy flavor production, the measurement of the top quark's properties and the
895 the study of the EWSB mechanism via multi-Higgs production. In addition, heavy ions collisions will
896 allow for a deeper understanding of the Quark-Gluon Plasma. From a BSM standpoint, searches for
897 supersymmetric and dark matter particles in new kinematic regimes can be pursued. A comprehensive
898 review of the physics potential of the FCC-hh can be found here [13]. This document collects the results
899 of the many studies that have been carried out since the beginning of the FCC initiative, in 2014.

900 Regarding Higgs pair production, a maximum precision on the SM cross section of 3% is expected
901 to be achieve using the $b\bar{b}\gamma\gamma$ final state. This would allow to constraint the Higgs triple coupling to be
902 $\lambda_{hhh} \in [0.97, 1.03]$. The $hh \rightarrow b\bar{b}b\bar{b}$ would allow for a 5% precision on the SM cross section and to
903 constraint the triple coupling to be $\lambda_{hhh} \in [0.9, 1.5]$. In spite of the larger background yield, the $b\bar{b}b\bar{b}$
904 channel provides a reasonable number of events in the tail of the m_{hh} distribution. As we discussed in
905 section 2.1.1, the tail of this distribution does not have a large contribution from the triangle diagram.
906 Therefore, the sensitivity to the Higgs triple coupling is expected to be smaller. However, the high
907 energy regime can be more sensitive to new physics contributions which makes the $b\bar{b}b\bar{b}$ channel a very
908 interesting one.

909 The Higgs quartic coupling could be probed through triple Higgs production. In this case the most
910 promising final state seems to be $b\bar{b}b\bar{b}\gamma\gamma$. This channel could constraint the Higgs quartic coupling to be
911 $\lambda_{hhhh} \in [-4, 16]$.

912 Chapter 4

913 State of the art

914 In this chapter we present the state of the art of the searches for Higgs pairs production both at the
915 LHC and in future colliders. In section 4.1, we review the searches that have been conducted at the
916 LHC by the ATLAS and CMS experiments. We include discussions on the different final states that were
917 target and report on the constraints that were derived for the cross section and trilinear coupling. A brief
918 overview of the current constraints on some BSM models is also presented. In section 4.2, we present
919 the results obtained from feasibility studies that access the discovery potential for this process at the
920 HL-LHC and at the FCC-hh.

921 4.1 Searches for Higgs pair production at the LHC

922 IDEAS FOR SECTION

- 923 - Previous searches by ATLAS and CMS at the LHC ✓
 - 924 - Different decay channels used: pros and cons + compare achieved sensitivities ✓
 - 925 Channels: bbbb, bb tau tau, bb gamma gamma, gamma gamma WW, bb ZZ ✓
 - 926 - Models investigated and motivation
 - 927 - Exclusion limits
-

928 The discovery of the Higgs boson is a strong evidence that the Higgs mechanism operates as predicted.
929 However, by itself, it does not guarantee that the shape of the Higgs potential is the one depicted in
930 figure 2.2, on the right. In order to reconstruct the Higgs potential and gain a deeper understanding of
931 the mechanism that leads to the breaking of the electroweak symmetry one must measure the Higgs
932 boson self-couplings, namely its three and four point interactions, whose strengths depend on the values
933 of the parameters of the Higgs potential, as shown in 2.13.

935 However, in the SM, the cross section for the production of Higgs pairs through ggF is extremely
936 small: ~ 30 fb at the current CM energy achieved at the LHC (value computed at Next to Leading
937 Order (NLO) accuracy) [52]. This value is approximately three orders of magnitude smaller than the
938 production cross section of a single Higgs boson. In addition, this value has to be multiplied by the

branching fraction of the chosen decay channel which further reduces the effective cross section of the full process. Nonetheless, ATLAS and CMS have conducted searches for this process whose results are summarized in this section.

The searches performed so far covered different decay channels and targeted not only the SM process but also some BSM scenarios where di-Higgs production is enhanced. Neither could achieve enough statistical significance to declare the measurement of this process in the SM nor have found any significant deviation from the expected values. These searches resulted in upper limits for the cross section of di-higgs production in the SM and for the values of the parameters of BSM benchmark theories. From the limits on the cross section it is also possible to constraint the values of the Higgs self coupling, $k_\lambda = \lambda_{hhh}/\lambda_{SM}$.

The $h \rightarrow bbbb$ channel [53, 54] benefits from the large branching fraction of $h \rightarrow bb$ ($\sim 58\%$). In addition, ATLAS showed that this is the most sensitive channel to resonance masses over 500 GeV [55]. However, this channel suffers with an overwhelming multijet background which drives the need for very tight trigger level cuts in order to bring the event rates down to manageable values.

The $h \rightarrow b\bar{b}\tau\tau$ analysis [55, 56] benefits from a sizable branching fraction ($\sim 7.3\%$) and from a relatively small background contribution from other SM processes. This searches target the semi-leptonic decay of the $\tau\tau$ pair.

The $hh \rightarrow bb\gamma\gamma$ [57, 58], $WW^*\gamma\gamma$ [55, 59], $ZZ\gamma\gamma$ [59] analysis can make use of very efficient diphoton triggers and isolation criteria that greatly reduce multijet background. In addition the excellent mass resolution of $h \rightarrow \gamma\gamma$ can be exploited. The $hh \rightarrow bbWW^*$ channel also benefits from the large branching fraction of $h \rightarrow WW$ ($\sim 21\%$).

The most stringent upper limit on the cross section of Higgs pair production in the SM comes from the $hh \rightarrow bb\gamma\gamma$ channel. The limit is 0.73 pb at 95% confidence level (CL) which is equivalent to ~ 20 times the predicted SM cross-section. The Higgs boson self coupling is constrained at 95% CL to be $-8.2 < k_\lambda < 13.3$. The analyzed data corresponds to an integrated luminosity of 19.7 fb^{-1} of proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$ collected with the CMS detector. [REF]

The upper limits on the cross section (at 95% CL) of Higgs pair production at 8 TeV as a function of the mass of a spin 0 resonance are summarized in figure 4.1. The limits were obtained using data collected with the CMS detector and come from searches using different final states, namely, $b\bar{b}\gamma\gamma$ (blue), $bb\bar{b}\bar{b}$ (red and pink) and $b\bar{b}\tau\tau$ (green). Depending on the analysis, the corresponding integrated luminosity varies between 17.9 fb^{-1} and 19.7 fb^{-1} . The results are usually interpreted in the framework of the Randall-Sundrum models such that the spin 0 resonance corresponds to the radion that decays to a pair of Higgs bosons. For a mass of 1 TeV the upper limit on the cross section is approximately 12 pb.

4.2 Feasibility studies for high-luminosity and future colliders

IDEAS FOR SECTION

- HL-LHC feasibility study [60],[61],[62]
- FCC hh to 4b (Selvaggi et al.) [63], [64]

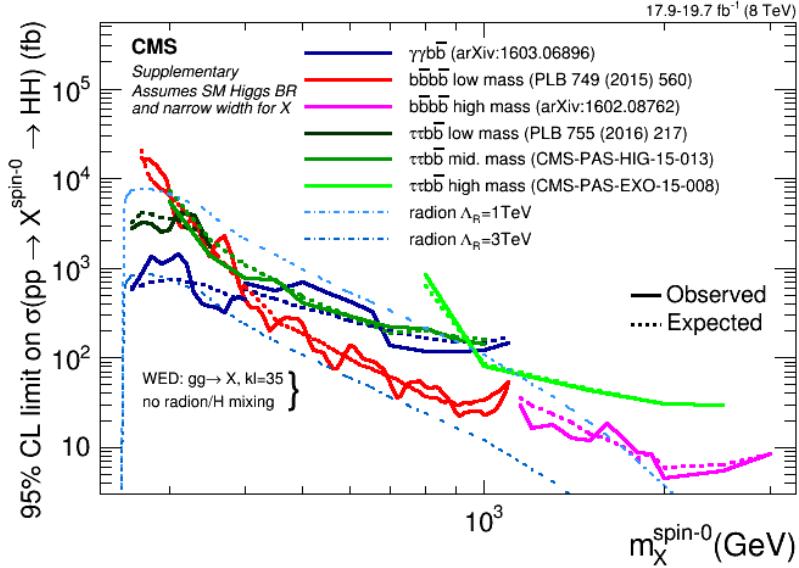


Figure 4.1: Exclusion limits on the mass of a spin 0 resonance from searches for Higgs pair production with the CMS detector.

976

977

978 Without any BSM contribution the discovery (or even evidence) of Higgs pairs production in the
979 four b quarks final state at the LHC is highly unlikely even considering the total expected integrated
980 luminosity of 300 fb^{-1} . Nonetheless, the HL-LHC as well as future colliders pose a good opportunity
981 for the discovery and precision studies of this process. Therefore, Monte Carlo studies assessing the
982 feasibility of searches for $hh \rightarrow b\bar{b}b\bar{b}$ at the HL-LHC and at the FCC-hh have been performed.

983 For the HL-LHC, a study including the $pp \rightarrow b\bar{b}b\bar{b}$, $pp \rightarrow b\bar{b}jj$, $pp \rightarrow jjjj$ and $pp \rightarrow t\bar{t} \rightarrow b\bar{b}jjjj$
984 backgrounds reports a significance (S/\sqrt{B}) of 4 (1.3) for an integrated luminosity of 3000 (300) fb^{-1} [60].
985 The analysis is performed in three orthogonal regions (boosted, intermediate and resolved)¹ and the
986 reported significance is obtained from the combination of these three regions. The highest significance
987 ($S/\sqrt{B} = 2.9$) is achieved in the boosted category. In addition, the impact of pile up is evaluated.
988 Considering a mean pile up of 80 and making use of jet grooming techniques a significance of 3.1 (1.0)
989 for an integrated luminosity of 3000 (300) fb^{-1} is reported (also considering the three analysis regions).
990 This work makes use of artificial neural networks (ANN's) as well as of jet substructure observables to
991 further increases the signal-background separation. Furthermore, single Higgs backgrounds, namely
992 $Z(\rightarrow b\bar{b})h(\rightarrow b\bar{b})$, $t\bar{t}h(\rightarrow b\bar{b})$ and $b\bar{b}h(\rightarrow b\bar{b})$, are shown to be negligible for the analysis, when compared
993 to the dominant QCD multijet background.

994 ATLAS and CMS also carried out preliminary studies for the sensitivity on the trilinear coupling at
995 the HL-LHC. Several channels have been investigated: $hh \rightarrow b\bar{b}\gamma\gamma$, $hh \rightarrow b\bar{b}\tau^+\tau^-$, $hh \rightarrow b\bar{b}W^+W^-$
996 and $hh \rightarrow b\bar{b}b\bar{b}$. The $b\bar{b}\gamma\gamma$ final state is the most sensitive one. In this channel, ATLAS and CMS
997 reported significances of 1.05σ and 1.6σ , respectively, for an integrated luminosity of 3000 fb^{-1} . Taking,
998 as an example, the ATLAS result, the achieved significance translates to an upper limit on the total

¹The event topologies associated with each of these categories are the same that are used in this analysis. They are described in chapter 6 and therefore we abstain from repeating them here.

999 di-Higgs cross section of approximately twice the SM value. This corresponds to an exclusion limit of
1000 $-0.8 < \lambda_{hhh}/\lambda_{SM} < 7.7$. The analysis conducted by ATLAS in this channel is done at generator level.
1001 The energy and momenta of the particles are smeared to simulate the detector's response. A mean pile
1002 up of 200 is considered. No MVA techniques are employed in the analysis.

1003 For the $hh \rightarrow b\bar{b}b\bar{b}$ channel, ATLAS states that a cross section 5.2 times larger than the SM value
1004 can be excluded at 95% CL, with systematics uncertainties being taken into consideration. The Higgs
1005 trilinear coupling is expected to be constrained to $-3.5 < \lambda_{hhh}/\lambda_{SM} < 11$. These results are based
1006 on the extrapolation of the current results obtained with the 2016 dataset, comprising an integrated
1007 luminosity of 10.1 fb^{-1} . This study is based only on the resolved analysis documented in [53].

1008 For the FCC-hh, a recent study that uses as signal sample $pp \rightarrow hhj \rightarrow b\bar{b}b\bar{b}j$ and that includes
1009 only the irreducible $pp \rightarrow b\bar{b}b\bar{b}j$ reports a significance of 6.61 for an integrated luminosity of 30 ab^{-1} ,
1010 considering an analysis that targets the boosted region [63]. The extra jet in the signal sample has $p_T >$
1011 200 GeV which provides the Higgs pair with a large Lorentz boost. EXPLAIN WHY THIS INCREASES
1012 SENSITIVITY TO HIGGS TRIPLE COUPLING. The analysis relies on the jet substructure observable
1013 τ_{21} and on a tight mass cut around the Higgs mass. No multivariate techniques are employed.

1014 A study comparing the feasibility of the search for di-Higgs production in the $b\bar{b}b\bar{b}$ in the HL-LHC and
1015 at the FCC-hh was presented in [64]. Only the irreducible background is considered. For a boosted cut
1016 based analysis, a significance of 1.1 is reported for an integrated luminosity of 3 ab^{-1} at $\sqrt{s} = 14 \text{ TeV}$
1017 (HL-LHC). For an integrated luminosity of 10 ab^{-1} at $\sqrt{s} = 100 \text{ TeV}$, which corresponds to the FCC-hh,
1018 this number is 5.7. While the significance is large for the FCC-hh, the reported signal to background ratio
1019 (S/B) is approximately one order of magnitude smaller which means the measurement might be more
1020 sensitive to systematic uncertainties on the backgrounds.

1021 4.3 Hadronic calorimeter granularity studies

1022 Even before the baseline design for the FCC-hh was established, there were studies regarding the
1023 impact of the granularity of the calorimeters in the spatial resolving power of hadronic showers and on
1024 the resolution of jet mass and substructure variables. These studies targeted the development of future
1025 colliders and greatly influenced the baseline design of the FCC-hh.

1026 The granularity of hadronic calorimeters is a key parameter for future collider detectors because it
1027 determines how well we can resolve energy deposits from pileup vertices and highly-boosted jet topolo-
1028 gies. In [65], the use of smaller calorimeter cells (smaller than the ones that are used in currently
1029 operating hadronic calorimeters) to resolve individual hadrons is investigated. For two Kaons (K_L^0) with
1030 an energy of 100 GeV each and with a truth level separation (ΔR) equal to 0.035 the energy deposited in
1031 the ECAL (blue) and HCAL (red) is shown as a function of ΔR for HCAL cells with sizes $20 \text{ cm} \times 20 \text{ cm}$
1032 ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$) and $5 \text{ cm} \times 5 \text{ cm}$ ($\Delta\eta \times \Delta\phi = 0.022 \times 0.022$) is shown in figure 4.2 on the
1033 left and right, respectively. The ECAL segmentation is equal to $2 \text{ cm} \times 2 \text{ cm}$. For a granularity of
1034 $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$ both hadrons can be resolved in the HCAL.

1035 Additional studies focusing on the jet mass resolution and jet substructure observables resolution

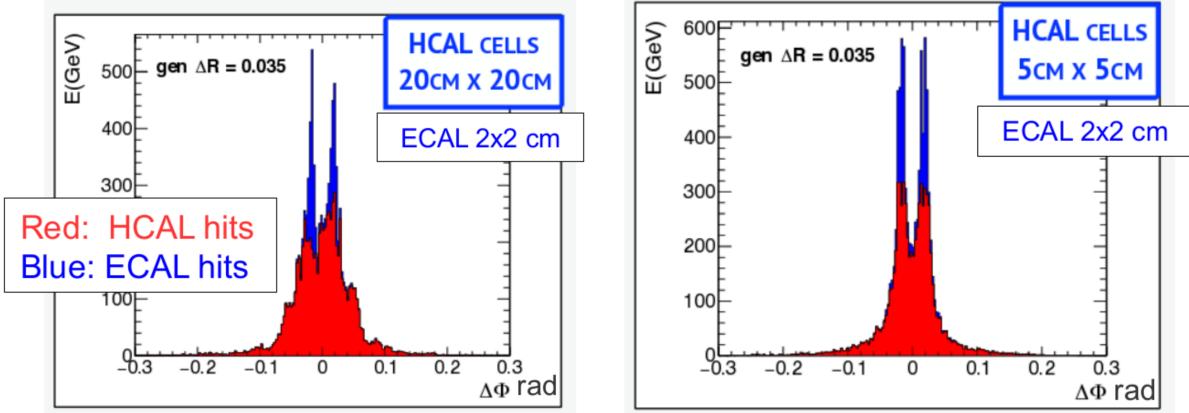


Figure 4.2: Azimuthal distribution of the energy deposits in the ECAL and HCAL for a pair of K_L^0 with $E = 100$ GeV for an hadronic calorimeter with $20\text{ cm} \times 20\text{ cm}$ (left) and $5\text{ cm} \times 5\text{ cm}$ (right). Figures from [66] (based on [65]).

were performed for three calorimeter (HCAL and ECAL) configurations. Some of these studies were presented in major conferences focused on future colliders, namely FCC week 2015 and 2016 and BOOST 2017. Here, we show results presented in [66–68]. The calorimeter configurations tested are: HCAL(ECAL) $0.1(0.025)$ $\eta \times 5.6(1.4)$ deg ϕ , HCAL(ECAL) $0.05(0.012)$ $\eta \times 2.8(0.7)$ deg ϕ and HCAL(ECAL) $0.025(0.006)$ $\eta \times 1.4(0.35)$ deg ϕ .

For $t\bar{t}$ events (generated at NLO with MadGraph5 and passed through Delphes 3.2 to simulate detector response) the (energy flow)jet mass distribution for $p_T(\text{jet} > 3)$ TeV is shown in figure 4.3, on the left. The jet mass resolution² is shown in the same figure, on the left. Compared to $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ cells, the mass resolution improves by 80 % and 120 % for $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$ and $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ cells, respectively ?? [OR GIVE NUMERS IN TERMS OF RMS].

Regarding jet substructure, the resolution of the τ_{32} variable in QCD dijet events generated with Pythia8 is shown in figure 4.4 for particle flow (or energy flow) jets, on the left, and for tower jets, on the right. [MAYBE USE OTHER PICTURE WITHOUT COMPARING EFLOW AND CALO].

In summary, previous studies focus on the impact of granularity in the resolution of jet mass and jet substructure observables.

²The resolution of a given observable, X , is given by $(X^{\text{reco}} - X^{\text{true}})/X^{\% \text{ true}}$.

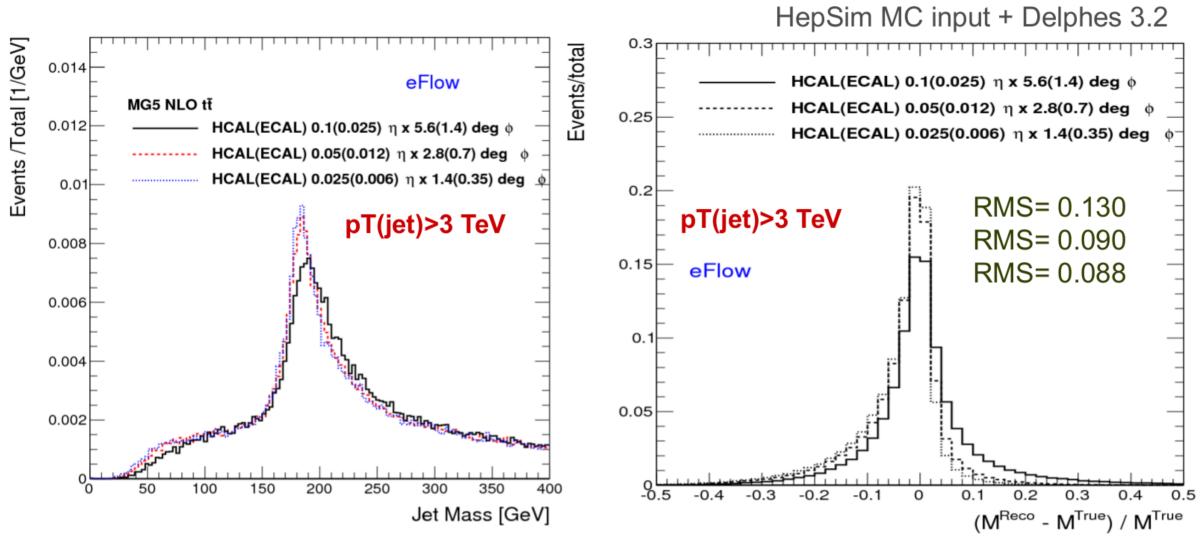


Figure 4.3: Jet mass

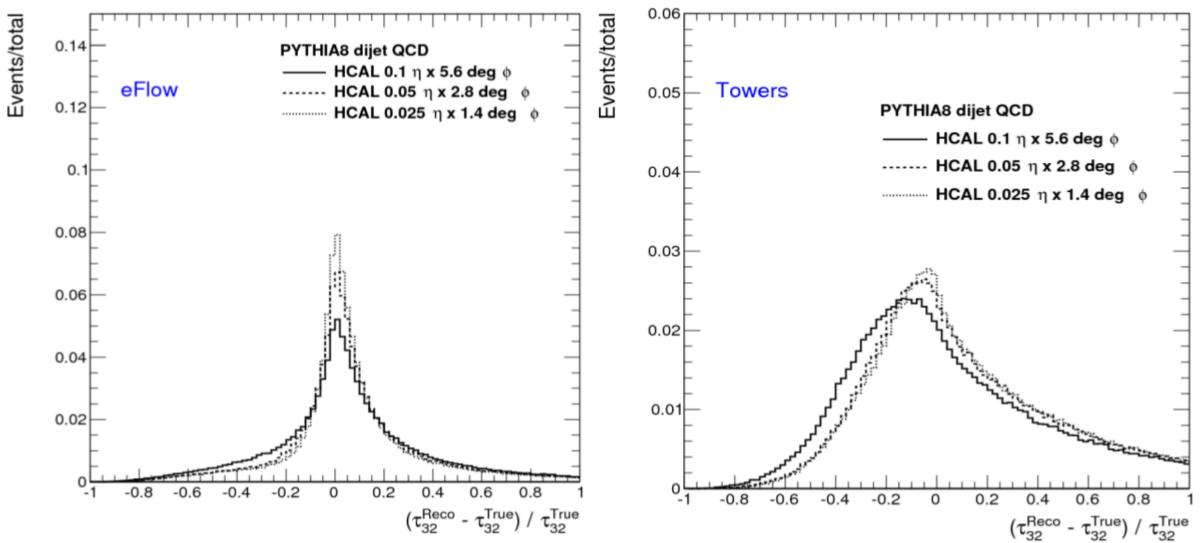


Figure 4.4: τ_{32}

1051

Chapter 5

1052

Sample generation and analysis tools

1053 A crucial component of this work is the generation of the Monte Carlo samples that are used in the anal-
1054 ysis. They are produced using fast simulation. We use the software and machinery that was developed
1055 by the FCC study group at CERN. The simulation work flow and Monte Carlo samples are described in
1056 this chapter.

1057 In section 5.1 we introduce the generators used to produce the Monte Carlo samples and briefly
1058 describe their working principles and functionalities. We then focus on the FCC-hh software, in sub-
1059 section 5.1.2, and explain how the implementation of the previously described simulation work flow is
1060 implemented. A detailed description of the technical settings used to produce the samples is provided
1061 in sections 5.2 and 5.2.3.

1062

5.1 Fast simulation workflow

- 1063 - Samples used for signal and backgrounds
1064 - Number of events, cross section, how they were generated (gen level cuts)
-

1065
1066 The Monte Carlo samples used in this study were generated using MadGraph5_aMC@NLO [69]. The
1067 showering and hadronization are simulated using Pythia 8 [70] and the detector response is parametrized
1068 using Delphes 3 [71] with the FCC-hh card. These samples are available at the CERN EOS storage.

1069 MadGraph5 is a matrix element generator for high energy physics processes, such as decays and
1070 scatterings. The user specifies the desired process in terms of the initial and final states and can impose
1071 additional constraints such as allowing for a number of refined criteria, including forced or forbidden s-
1072 channel resonances, excluded internal particles, and forced decay chains of final state particles [69].
1073 As a result, MadGraph automatically generates the corresponding Feynman diagrams and creates the
1074 necessary code to compute the matrix element at a given point of the phase space.

1075 The Pythia8 program is frequently used for event generation in high energy physics. In this work,
1076 however, we use it only to simulate the showering and hadronization process and not the parton level
1077 hard process which is simulated in MadGraph5. The Les Houches file that is produced by MadGraph5

1078 is used as input to a Pythia8 program that can decay unstable particles, simulate initial and final state
1079 showers as well as the hadronization of coloured particles, such as quarks and gluons. The desired
1080 settings can be specified in an additional file, a card, that is used as input for a Pythia8 run.

1081 Delphes3 allows for a quick and simple simulation of the detector's response [71]. Its goal is to allow
1082 the fast simulation of a multipurpose detector for phenomenological studies. The simulation includes a
1083 track propagation system, electromagnetic and hadronic calorimeters and a muon identification system.
1084 Low level physics objects, such as tracks and energy deposits, and high level physics objects, such as
1085 leptons and jets, are reconstructed from the detector's response and can be used to perform physics
1086 analysis. In the following paragraphs we briefly describe how the detector response is simulated and
1087 how jet reconstruction is performed.

1088 The magnetic field is uniform, axial and parallel to the beam direction. Charged particles follow a
1089 helicoidal trajectory from the interaction point to the calorimeter while neutral particles have a straight
1090 line trajectory. The probability of a charged particle being reconstructed as a track is defined by the
1091 user. Only a smear in the modulus of the transverse momentum is applied (not to the direction). The
1092 tracking efficiency as well as energy and momentum resolutions are specified by the user and may
1093 include a dependency on the particle type, momentum and pseudorapidity. The calorimeters have a
1094 finite segmentation in the (η, ϕ) plane and the cell size can be defined in the configuration file. For
1095 simplicity, the segmentation is the same for the ECAL and in the HCAL and they are perfectly overlaid.
1096 The amount of energy deposited in the calorimeters by each particle type can be defined by the user.
1097 By default, stable hadrons deposit all their energy in the HCAL although in a real detector a significant
1098 fraction of their energy is deposited in the ECAL. The energy resolution of the ECAL and HCAL are
1099 parameterized as a function of η and include stochastic, noise and constant terms. The electromagnetic
1100 and hadronic energy deposits are independently smeared by a log-normal distribution.

1101 Jets can be produced using generator level long-lived particles after showering and hadronization,
1102 tracks, calorimeter towers or particle-flow tracks and towers. These are referred to as generator, track,
1103 calorimeter or particle-flow jets, respectively. For generator jets no detector simulation nor reconstruction
1104 are taken into account. In spite of the type of jet the user can choose the jet clustering algorithm and
1105 the values of its parameters as well as the minimum transverse momentum of the jets that are stored in
1106 the final collection. Delphes integrates the FastJet package [72] and therefore allows jet reconstruction
1107 with the most popular jet reconstruction algorithms, namely, anti- k_T , k_T and C/A. Jets resulting from the
1108 hadronization of a b quark, b jets, are identified if a b quark is found within a ΔR distance from the jet's
1109 axis. The tagging efficiency and mis-tagging probabilities can be defined by the user.

1110 5.1.1 Particle flow and calorimeter reconstruction in Delphes

1111 In Delphes, hadronic jets can be reconstructed using only the information from the HCAL towers or using
1112 a particle flow algorithm that combines information from the tracking system and from the HCAL towers.
1113 These two approaches create jets that are referred to as calorimeter and particle flow jets, respectively.
1114 The latter can also be referred to as energy flow jets (eflow jets in short). In this work we performed the

1115 analysis using both sets of jets and compare the results. Therefore we briefly describe them here.

1116 Calorimeter jets are very simple. They are reconstructed using as input for the jet clustering algo-
1117 rithm the 4-vectors associated with the calorimeter towers, after an energy smearing has been applied.
1118 Therefore the energy resolution is limited by the transverse segmentation of the calorimeters.

1119 The goal of the particle flow approach is to make use of all the available information provided by the
1120 various sub-detectors for reconstructing an event [71]. This approach is used by some experimental
1121 collaborations [73, 74] but the exact implementation depends on the specificities of the experiment. If
1122 the momentum resolution of the tracking system is better than the energy resolution of the calorimeters
1123 it might be convenient to use the tracking information to estimate the momentum of charged particles.
1124 In real experiments, the tracking resolution is only better than the calorimeter's energy resolution up to
1125 some energy threshold. However, in Delphes, it is assumed that it is always convenient to estimate the
1126 momentum of charged particles via the tracker.

1127 The particle flow algorithm works as follows [71]. For each calorimeter tower it counts:

- 1128
- the total energy deposited in ECAL and HCAL, E_{ECAL} and E_{HCAL} , respectively;
 - the total energy deposited in ECAL and HCAL originating from charged particles for which a track
1129 has been reconstructed, $E_{\text{ECAL,trk}}$ and $E_{\text{HCAL,trk}}$, respectively.
1130

Then it defines $\Delta_{\text{ECAL}} = E_{\text{ECAL}} - E_{\text{ECAL,trk}}$, $\Delta_{\text{HCAL}} = E_{\text{HCAL}} - E_{\text{HCAL,trk}}$ and computes $E_{\text{Tower}}^{\text{eflow}}$ given by

$$E_{\text{Tower}}^{\text{eflow}} = \max(0, \Delta_{\text{ECAL}}) + \max(0, \Delta_{\text{HCAL}}). \quad (5.1)$$

1131 All reconstructed tracks result in a particle flow track. If $E_{\text{Tower}}^{\text{eflow}} > 0$ a particle flow tower is created with
1132 energy $E_{\text{Tower}}^{\text{eflow}}$. The particle flow tracks and jets are then used as input for the jet clustering algorithms.

1133 5.1.2 FCC-hh software

1134 FCC software (FCCSW) [75], common to all FCC experiments (electron-electron, electron-hadron and
1135 hadron-hadron) has been developed and is maintained by the FCC study group. The software is based
1136 on Gaudi [76]. An FCC Event Data Model based on podio [77] was also developed. It consists in specific
1137 classes that encode the information about the events.

1138 The FCC-hh study group is responsible for the generation of Monte Carlo quick simulation samples
1139 for the main benchmark processes for the FCC-hh. The samples are generated using the workflow
1140 described in the previous section. CERN users can request rights to run the EventProducer package
1141 [78] and produce samples for any desired process using the machinery that is already implemented. In
1142 addition, the FCC Event Data Model classes are directly accessible and can be used to read the ROOT
1143 files that are produced after the events are pass through Delphes.

1144 The machinery to submit jobs to CERN's batch system is also implemented for both the generation
1145 (MadGraph5) and reconstruction (Pythia8 plus Delphes3) levels.

1146 In this work we make use of this software in order to produce the necessary samples.

1147 **5.2 Signal and background samples**

1148 **5.2.1 MadGraph**

1149 The 4b+j samples are generated with an extra jet with a high p_T at generator level in order to force the
1150 four b quarks to have a high Lorentz boost and therefore increases the probability of the events being
1151 reconstructed with two large-R jets with a two-prong substructure. The 4b+j QCD sample is constituted
1152 by two independent samples that have a different generator level cut in the minimum p_T of the light jets,
1153 namely, $200 < p_{T,j}^{\min} < 500$ GeV and $p_{T,j}^{\min} > 500$ GeV. The jj+0/1/2 j is divided into several individual
1154 samples that are produced in different H_T regions, where H_T is the scalar sum of the p_T of all partons
1155 at generator level. The minimum allowed H_T is 500 GeV. Since these backgrounds are QCD processes,
1156 the p_T distribution of the final state jets falls very steeply as the p_T increases. Therefore, if one were
1157 to generate events for these processes without restricting the phase space, most events would consist
1158 of jets with a very low p_T which are exactly the type of events that are rejected the most by a boosted
1159 analysis (see chapter 6 for more details on the event topology that is targeted and on the analysis
1160 strategy). Also note that for the 4b+j (QCD) and jj+0/1/2 j samples we do not take into account the
1161 regions of the phase space with $p_{T,j} < 200$ GeV and with $0 < H_T < 500$ GeV. We assume that we can
1162 reject most of the events (if not all) events with these kinematic characteristics by going to a sufficiently
1163 boosted region of the phase space.

1164 The 4b+j is generated using the four flavor scheme meaning that the extra parton can be a gluon
1165 or a quark from the first he second geerations. For the jj+0/1/2 j and $t\bar{t}$ +0/1/2 j samples the five flavor
1166 scheme is used and therefore the extra partons can also be b quarks. Notice that after the showering
1167 procedure there could be some overlapp between the 4b+j and jj+0/1/2 j. This is taken care of in our
1168 analysis code: if an event from the jj+0/1/2 j has four b quarks at truth level [GIVE PERCENTAGE] then
1169 we do not consider it because it will certainly overlapp with an event from the 4b+j sample.

1170 For the hh SM sample, one of the Higgs is decayed to $b\bar{b}$ in MadGraph. The reasoning behind this
1171 choice is that most searches for Higgs pair production make use of final channel that include at least two
1172 b quarks in order to keep the cross section times BR of the process large enough an this way the same
1173 generator level sample can be used to perform different analysis. The decay of the remaining Higgs
1174 boson can then be implemented in Pythia, in the case of this work, to $b\bar{b}$. For the hh BSM samples both
1175 Higgs are decayed only in Pythia.

1176 **Signal samples - BSM**

1177 In addtion to the SM di-Higgs signal, we also explore the signature and analysis sensitivity for di-Higgs
1178 signals produced by two benchmark BSM models: CP-conserving 2HDM and a simplified dark matter
1179 model (DM) with a spin 0 mediator. These models were described in section 2.2.

1180 Both models are readily available in FeynRules [79] (which is a Mathematica package for the imple-
1181 mentation of new physics models) model database and can be straightforwardly implemented in Mad-
1182 graph5. The parameters of the models, namely the masses of the new particles, can be changed by

1183 the user. In the case of this work, we want new particles to have a large mass so that they produce SM
1184 Higgs pairs with a high Lorentz boost.

Starting from [80], that constraints the parameter space of the CP-conserving 2HDM using ATLAS and CMS data collected during LHC's 7 and 8 TeV runs, we choose values for the masses of the additional Higgs bosons (h_2 , h_3 , h^\pm) that are as high as possible but that are within the allowed (not excluded by experimental data or theoretical constraints) phase space of the model. We set:

$$m_{h_2} = 600 \text{ GeV}, \quad m_{h_3} = 900 \text{ GeV}, \quad m_{h^\pm} = 360. \text{ GeV} \quad (5.2)$$

1185 As a safety check, we test the model with these parameters at a CM energy of 13 TeV. We obtain a value
1186 for the cross section of approximately 0.1 pb which smaller than the current experimental limit on the
1187 cross section.

1188 In the case of the DM model, the spin 0 mediator's mass is set at 1 TeV. The cross section for the
1189 signal generated with this model is smaller than the cross section of the SM signal (approximately 0.2
1190 pb *versus* 0.7 pb) and therefore this model is not excluded by experimental data.

1191 5.2.2 Pythia

1192 For the signal sample we simply turn off all other decays except $h \rightarrow b\bar{b}$ therefore forcing the Higgs to
1193 decay to a pair of b quarks leading to the desired final state with four b quarks. All other settings are left
1194 are not altered with respect to their default configuration.

1195 For the multijet and $t\bar{t}$ samples we have to perform jet matching because we require additional jets
1196 at the level of the matrix element. In addition to the partons generated in MadGraph and that produce a
1197 jet, Pythia may introduce extra jets that are usually soft and collinear (with the particle from which they
1198 were radiated) and result from the showering process. This could lead to the same process (meaning
1199 processes with the same final states) being counted twice (double counting). Take, for example, the
1200 processes $j j$ and $j j j$ at MadGraph level. It can happen that Pythia generates an extra jet for the first
1201 process but not for the second, leading to both processes having the same final state (three jets). Each
1202 process would then give its independent contribution to the total number of events but because they
1203 simply represent two distinct ways of achieving the same final state they should only be counted once.
1204 The goal of jet matching procedures is to avoid this problem.

1205 The cross section times branching ratio (when applicable), the k-factors and the total number of
1206 events used in the analysis are summarized in table 5.1, for all samples. The k-factor multiplies the cross
1207 section times branching ratio in order to reproduce known higher order results. This factor corresponds
1208 only to the ratio between the total cross section and it does not correct for possible difference that might
1209 exist between the differential cross sections.

| Sample | $\sigma \times BR$ [pb] | k-factor | Number of events |
|---|-------------------------|----------|------------------|
| hh ($hh \rightarrow b\bar{b}b\bar{b}$) - SM | 0.76 | 1.52 | |
| hh ($hh \rightarrow b\bar{b}b\bar{b}$) - DM mediator | 0.218 | 1.0 | |
| hh ($hh \rightarrow b\bar{b}b\bar{b}$) - 2HDM | 1.85 | 1.0 | |
| 4b+j (QCD, $200 < p_T^j < 500$) | 756.4 | 2.0 | |
| 4b+j (QCD, $p_T^j > 500$) | 57.71 | 2.0 | |
| 4b+j (QCD+EWK) | 6.204 | 1.0 | |
| 4b+j (EWK) | 0.07206 | 1.0 | |
| jj+0/1/2 j ($0 < H_T < 500$) | 6.52×10^9 | 1.0 | |
| jj+0/1/2 j ($500 < H_T < 1000$) | 1.64×10^7 | 1.0 | |
| jj+0/1/2 j ($1000 < H_T < 2000$) | 1.67×10^6 | 1.0 | |
| jj+0/1/2 j ($2000 < H_T < 4000$) | 1.32×10^5 | 1.0 | |
| jj+0/1/2 j ($4000 < H_T < 7200$) | 7.32×10^3 | 1.0 | |
| jj+0/1/2 j ($7200 < H_T < 15000$) | 4.75×10^2 | 1.0 | |
| jj+0/1/2 j ($15000 < H_T < 25000$) | 7.35 | 1.0 | |
| jj+0/1/2 j ($25000 < H_T < 35000$) | 0.176 | 1.0 | |
| jj+0/1/2 j ($35000 < H_T < 100000$) | 0.00765 | 1.0 | |
| t <bar>t>+0/1/2 j</bar> | 4.31×10^4 | 1.74 | |

Table 5.1: Summary of the cross sections, k factors and total number of events of the samples used in the analysis.

| Configuration | $\Delta\eta \times \Delta\phi$ | η range |
|-------------------------------------|--------------------------------|----------------------|
| 1 (ATLAS HCAL) | 0.1×0.1 | $ \eta < 2.5$ |
| | 0.2×0.2 | $2.5 < \eta < 5.0$ |
| 2 (ATLAS HCAL $\eta \times 4$) | 0.025×0.1 | $ \eta < 1.7$ |
| | 0.1×0.1 | $1.7 < \eta < 2.5$ |
| 3 (FCC HCAL $\phi/2$) | 0.025×0.05 | $ \eta < 2.5$ |
| | 0.05×0.1 | $2.5 < \eta < 6.0$ |
| 4 (FCC HCAL) | 0.025×0.025 | $ \eta < 2.5$ |
| | 0.05×0.05 | $2.5 < \eta < 6.0$ |
| 5 (FCC HCAL $\eta, \phi \times 2$) | 0.0125×0.0125 | $ \eta < 2.5$ |
| | 0.025×0.025 | $2.5 < \eta < 6.0$ |

Table 5.2: Summary of the benchmark granularity configurations of the HCAL.

5.2.3 Delphes

It is one of the main goals of this work to evaluate how the analysis sensitivity is influenced by the granularity of the hadronic calorimeter. We start from the same MadGraph level samples and pass them through Pythia and Delphes changing the setting of the Delphes card that correspond to the HCAL. All other detector's parameters were kept unchanged with respect to the FCC default Delphes card. We tested five benchmark granularity configurations:

1. ATLAS HCAL granularity (as implemented in the standard ATLAS Delphes card);
2. Starting from the ATLAS HCAL configuration we increase the granularity in $|\eta|$ by four, in the pseudo rapidity range $|\eta| < 1.7$ which corresponds to the TileCal region;
3. Starting from the FCC HCAL configuration we decrease the granularity in ϕ by two, in the entire pseudo rapidity range covered by the HCAL.
4. FCC HCAL default granularity (as implemented in the standard FCC Delphes card);
5. Starting from the FCC HCAL configuration we increase the granularity in $|\eta|$ and in ϕ by two, in the entire pseudo rapidity range covered by the HCAL.

The granularities of these five configurations are summarized in table 5.2. The values that are shown, as well as the corresponding pseudo rapidity regions, are exactly what is implemented in the Delphes' cards. In addition, we also passed the same generator level samples through the default ATLAS detector simulation in Delphes. The HCAL granularity is the one that is indicated in table 5.2 but the other detector's parameters, such as the radius, magnetic field, tracking resolutions are the ones that are implemented in the default ATLAS Delphes card.

1230 Chapter 6

1231 Analysis

1232 6.1 Overview of the $hh \rightarrow b\bar{b}b\bar{b}$ channel

- 1233 - Signal cross section and final state signature
 - 1234 - Pros and challenges of using the 4b final state
 - 1235 - Main backgrounds (the ones we consider) and respective cross sections (leave for appendix discussion
 - 1236 on other backgrounds, namely Higgs processes)
-
- 1237

1238 The searches for Higgs pair production in the four b quarks final state benefit from the large BR of
1239 $h \rightarrow b\bar{b}$. For the SM Higgs, with a mass around 125 GeV, the branching ratio of the $hh \rightarrow b\bar{b}b\bar{b}$ decay is
1240 approximately 33.6%, making it the most probable decay for Higgs pairs, as it is illustrate in Figure 6.1.

1241 However, in this channel, the main background is QCD multijet production that has a cross section
1242 several orders of magnitude larger than di-Higgs production in the SM, as Table 5.1 shows. Nonetheless,
1243 it is a well known feature of QCD that the majority of jets produced are soft. This means that the p_T
1244 distributions of this background have a very large yield close to zero and then fall very steeply while the
1245 signal has a much larger tail to high values of p_T . This indicates that searches targeting the boosted
1246 signal has a much larger tail to high values of p_T . This indicates that searches targeting the boosted
1247 kinematic regime may be the key to measure $hh \rightarrow b\bar{b}b\bar{b}$ using inclusive production.

1248 In this work we target the boosted regime. In this kinematic region, the final state of the signal is
1249 characterized by two jets with a large R parameter. Each jet is expected to contain the two b quarks
1250 originated from the decay of a Higgs boson, as Figure 6.4 illustrates. Extra jets are also expected to be
1251 reconstructed as a consequence of QCD activity.

1252 The analysis presented here is performed using the Monte-Carlo samples described in chapter 5. All
1253 samples assume $m_h = 125$ GeV. The event selection criteria are designed to optimize the significance,
1254 given by S/\sqrt{B} , where S and B represent the number of signal and background events, respectively.

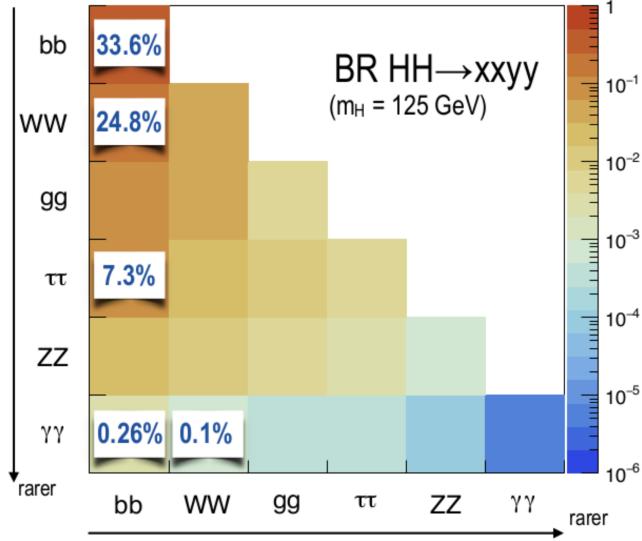


Figure 6.1: Higgs pair branching ratios.

6.1.1 Signal event characterization

The cross section for di-Higgs inclusive production from pp collisions at 100 TeV is 1.2 pb [51]. This value is computed at NLO accuracy. The dominant process is gluon-gluon fusion that has been extensively discussed in section 2.1.1. The signal process comprehends the di-Higgs inclusive production and the $hh \rightarrow b\bar{b}b\bar{b}$ subsequent decays. The total process cross section is then given by $\sigma(pp \rightarrow hh) \times BR(hh \rightarrow b\bar{b}b\bar{b})$ and it amounts to approximately 0.4 pb [NOT VERY COHERENT WITH MG VALUES - CHECK]. For an integrated luminosity of 3000fb^{-1} of proton collisions at $\sqrt{s} = 100$ TeV we expect X events.

The Higgs candidates are reconstructed using large R jets that can be directly measured using information from the calorimeters and tracking systems. EVENT TOPOLOGY: TWO JETS BACK TO BACK, EXTRA JET???. Each wide jet is expected to be consistent with having two subjets associated with the two b quarks from the $h \rightarrow b\bar{b}$ decay. In the Higgs rest frame, the two subjets are produced back-to-back conserving momentum. In the laboratory frame, the ΔR between the subjets depends on the momentum of the Higgs boson, with larger ΔR corresponding to Higgs bosons with a lower momentum.

6.1.2 Backgrounds

The relevant backgrounds for this analysis are QCD (light?) multijet production, $t\bar{t}$ and $b\bar{b}b\bar{b}$. Although the $b\bar{b}b\bar{b}$ background is a particular case of a QCD multijet production process we consider it separately because it constitutes the irreducible background and therefore it will have a higher efficiency in the analysis. The cross sections for these processes are several orders of magnitude larger than the cross section for the signal, as Table 5.1 shows. In addition, in the case of the $t\bar{t}$ background, the event topology is expected to be similar to the signal because it also consists in the production of a pair of particles with the same mass.

The assumption that QCD multijet production and $t\bar{t}$ two main backgrounds is corroborated by the

1277 ATLAS di-Higgs search performed in the same channel where these backgrounds are found to be the
1278 dominant ones and that all other sources of backgrounds, including processes involving Higgs bosons,
1279 are found to be negligible [53]. In appendix B we discuss and evaluate the importance of some back-
1280 grounds that include Higgs bosons to our analysis.

1281 Figure 6.2 shows examples of LO Feynman diagrams that contribute to $b\bar{b}b\bar{b}+j$ (left), three (middle)
1282 and four (right) light jets production. The $b\bar{b}b\bar{b}$ background is generated with an extra light jet at generator
1283 level. This jet boosts the four b quarks and has a minimum p_T of 200 GeV. This increases the probability
1284 of two b quarks being reconstructed as single large R jet therefore emulating the signal's final state
1285 signature. The QCD multijet background consists of two, three and four jet events (represented as
1286 $jj+0/1/2 j$), at generator level. The jets can originated from light and b quarks and from gluons. A jet
1287 matching procedure is implemented in Pythia in order to avoid double counting.

1288 HOW MULTIJETS PENETRATE ANALYSIS (MISS TAGGING). CROSS SECTION FOR THESE
1289 PROCESSES.

1290 Figure 6.3 shows examples of LO Feynman diagrams that contribute to $t\bar{t}$ production through $q\bar{q}$ (left)
1291 and gg (middle and right) fusion. The $t\bar{t}$ background is simulated with additional zero, one or two jets
1292 (represented as $t\bar{t}+0/1/2 j$), at generator level. The extra jets can originated from light and b quarks
1293 and from gluons. A jet matching procedure is implemented in Pythia in order to avoid double counting.
1294 The top quark has a very short life time, predicted to be 5×10^{-25} s, such that it decays before it can
1295 hadronize. This sample is inclusive in the top quark decay modes. However, the most favoured decay of
1296 the top quark is $t \rightarrow Wb$ with a branching ratio close to 96% [81]. Therefore, $t\bar{t}$ events will, most of the
1297 times, result in the $WWbb$ final state. We do not specify any decay mode for the W such that the sample
1298 is also inclusive in the W decay modes. The W decays to hadrons ($W^+ \rightarrow q\bar{q}$) and leptons ($W^+ \rightarrow l^+\nu$)
1299 with $\text{BR}(W^+ \rightarrow q\bar{q}) \sim 68\%$ and $\text{BR}(W^+ \rightarrow l^+\nu) \sim 10\%$. If one (or both) W bosons decays to hadrons
1300 then there will be additional jets in the final state. These can be b jets or can be misidentified as such. If
1301 both W bosons decay to leptons there will still be at least two b jets in the final state, coming from the $t\bar{t}$
1302 decay.

1303 TOP DECAYS, RESULTING FINAL STATES AND HOW THEY CAN PENETRATE THE ANALYSIS.
1304 CROSS SECTION.

1305 6.2 Analysis strategy

-
- 1306
- 1307
- 1308 - Event topology: two boosted jets each corresponding to a Higgs
 - 1309 - Physics objects: particle flow anti-kT R=0.8 jets (discussion about jet radius in appendix)
 - 1310 - Selection criteria
 - 1311 - Substructure variables
 - 1312 - Optimization (efficiency plots, correlations, MVA...)
-

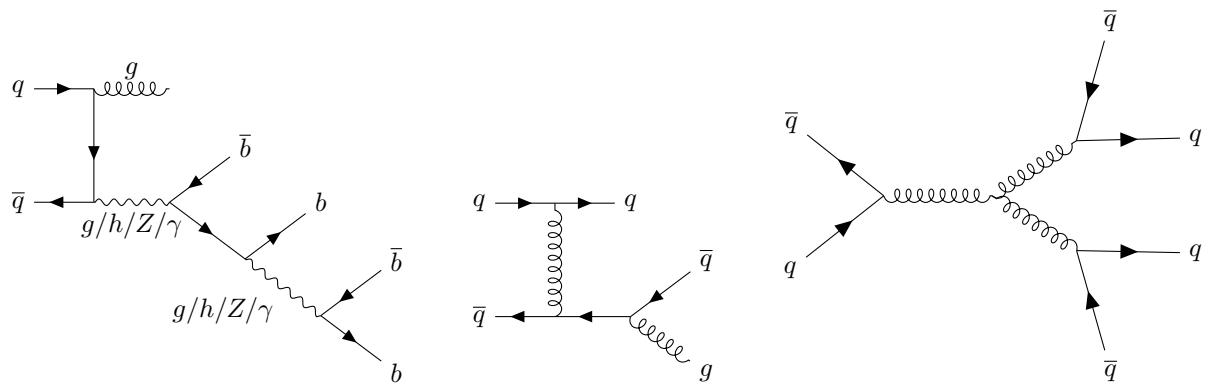


Figure 6.2: Example of diagrams that contribute to the QCD multijet background: five final state jets, four of which are b-jets (left), three final state jets (middle) and four final state jets (right). Here, q stands for a light quark/jet.

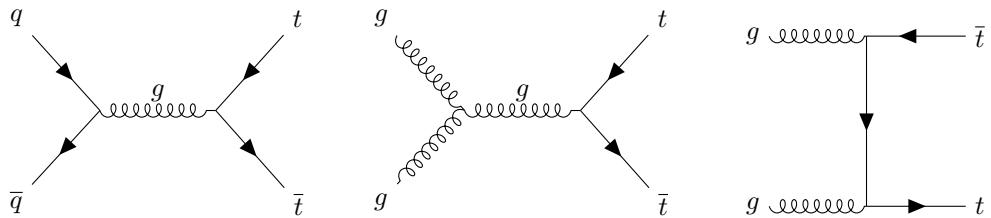


Figure 6.3: Dominant diagrams of $pp \rightarrow t\bar{t}$ at LO.

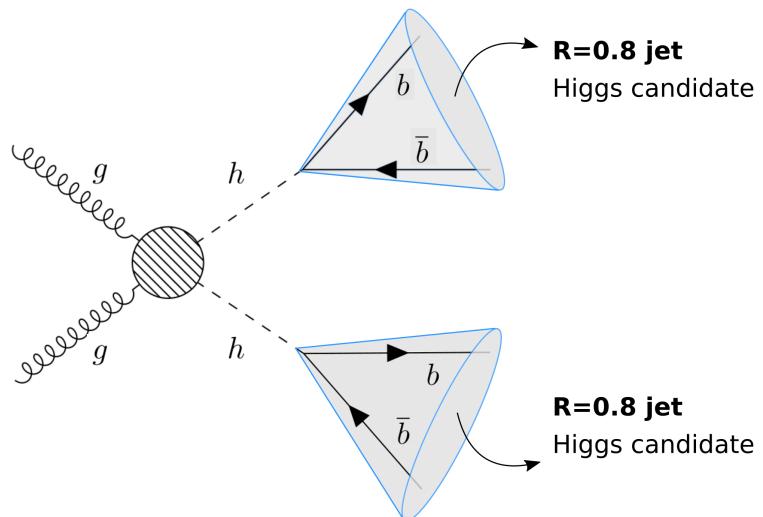


Figure 6.4: Event topology targeted by the boosted analysis region. The blob represents the interaction between the gluons and the Higg bosons that is represented by the Feynman diagrams shown in figure 2.3.

| p_T | $10 < p_T < 500$ | $500 < p_T < 15000$ |
|----------------------|---------------------------|---|
| η | | |
| $ \eta < 2.5$ | 0.85; 0.05; 0.01 | (0.85; 0.05; 0.01) $\times (1 - p_T/15000)$ |
| $2.5 < \eta < 4.0$ | 0.64; 0.03; 0.0075 | (0.64; 0.03; 0.0075) $\times (1 - p_T/15000)$ |

Table 6.1: b-Tagging (black), c (blue) and light (red) mistag probabilities as a function of η and p_T of the (sub)jet. The momentum dependent factor, $(1 - p_T/15000)$, is common to the three probabilities.

This analysis targets events in which at least two large R jets are reconstructed. The jet with the highest momentum is assumed to correspond to the leading Higgs candidate and the jet with the second highest momentum to the sub leading one. Both the leading and sub leading jets must be b-tagged in order for the event to be accepted.

The events are reconstructed using particle flow (or eflow) or pure calorimeter jets (we explored both approaches) with $R = 0.8^1$, clustered with the anti- k_T algorithm. We perform the b-tagging of jets using truth level information as it is described in the following section. Jets with a large R parameter cannot be b-tagged using Delphes default algorithm because the tagging of large R jets is an ambiguous task that can be performed in several different ways. Therefore, we need to implement our own b-tagging algorithm/strategy.

6.2.1 Implementation of b-tagging

For each jet, the two hardest subjets are found using the mass drop procedure. It might happen that there are not two subjets because the algorithm's criteria are not met. In that case, the jet is rejected. We compute the ΔR distance between all b and c quarks in the event with Pythia 8 status equal to 23 and with $p_T > 10$ GeV and each subjet, $\Delta R(\text{subjet,parton})$. According to the Pythia manual, particles with status 23 result directly from the hardest subprocess. We consider that a subjet is matched to a given quark if $\Delta R(\text{subjet,parton}) < 0.3$, as indicated by the plot in Figure 6.5. In this plot there is a peak for $\Delta R < 0.3$ that corresponds to the If the subjet is matched to at least a b quark, we b-tag the subjet with a given probability. If the subjet is not matched to any b quark but it is matched to at least one c quark we apply a c mistag rate. If the subjet is not matched to any b or c quark we apply a light mistag rate. The b-tag probability and mistag rates were obtained from the Delphes FCC-hh card. They depend on the momentum of the jet and on its η coordinate. They are summarized in Table 6.1. Note that a jet cannot be b-tagged if $|\eta| > 4$ or if its momentum is smaller than 10 GeV or larger than 15000 GeV.

6.2.2 Pre-selection

TO INCLUDE:

- discussion on trigger requirements: important in real analysis
- show and discuss some distributions: invariant mass (h, hh), momentum, tau21, delta eta (hh), more

¹Find discussion about jet radius in the appendix.

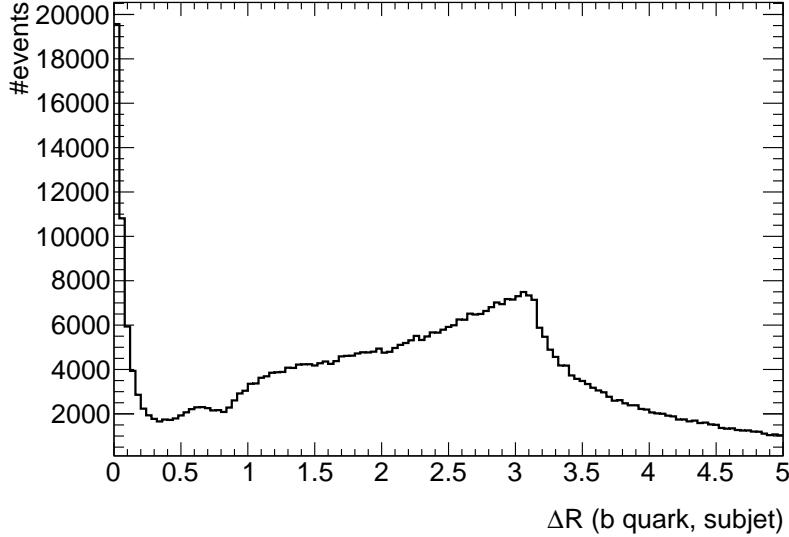


Figure 6.5: Minimum ΔR between b quarks and subjets of the $R = 0.8$ jets. There is a minimum around $\Delta R = 0.3$. It corresponds to the maximum ΔR for which the b quark and the subjet are matched.

1341 substructure variables (may be all that we considered relevant)

1342 - compare with other articles/analysis ?

1343

1344

1345 We require at least two b-tagged jets. In addition, the leading and sub leading jets must have $p_T \geq$
 1346 200 GeV in order for the event to be accepted. Due to the b-tagging efficiency formulas, there is a natural
 1347 cutoff at $|\eta| = 4$ so we do not place any additional cut in η . From now on, these cuts are referred to as
 1348 pre-selection cuts.

1349 6.2.3 Baseline analysis

1350 As a first step, we implemented a baseline analysis based on the

1351 6.2.4 Optimization

1352 We explored different methods to optimize the baseline analysis in order to try to increase the achieved
 1353 significance.

1354 TO INCLUDE:

1355 - Mention study on correlations ?

1356 - Optimization of cut-based analysis based on S/\sqrt{B} plots

1357 - MVA analysis ?

1358 **Chapter 7**

1359 **Results**

1360 In this chapter we describe the main results of the search for $hh \rightarrow b\bar{b}b\bar{b}$ at the FCC-hh using two
1361 benchmark luminosities, 3 ab^{-1} and 30 ab^{-1} (section 7.1). The statistical analysis used to extract the
1362 signal strength and to set limits on the Higgs boson triple coupling is also discussed. In section 7.2 we
1363 show how the significance of the analysis varies as a function of the granularity of the HCAL and/or the
1364 detector configuration. We also compare the results obtained using particle flow and pure calorimeter
1365 jets.

1366 **7.1 Di-Higgs discovery potential at the FCC-hh**

1367 **7.1.1 Statistical analysis**

1368 **7.2 Hadronic calorimeter granularity studies for future colliders**

1369 **7.3 Benchmarking against ATLAS analysis at $\sqrt{s} = 13 \text{ TeV}$**

| Selection | SM | 2HDM | 4b+j | jj+0/1/2 j | t̄t+0/1/2 j |
|--|-------------------|------------------|--------------------|----------------------|---------------------|
| Gen level | 100 (3.46e7) | 100 (5.55e7) | 100 (4.90e10) | 100 (5.47e14) | 100 (2.25e12) |
| N(b-tags) ≥ 4 | 92.5 | 91.8 | 75.8 | 3.963 | 53.5 |
| $p_T(j_1, j_2) \geq 200 \text{ GeV}$ | 16.6 (5.75e6) | 30.6 (1.70e7) | 17.8 (8.73e9) | 0.74 (4.06e12) | 1.06 (2.38e10) |
| $p_T(j_1) \geq 400 \text{ GeV}$ | 8.623 | 18.2 | 7.0 | 0.18 | 0.446 |
| $p_T(j_2) \geq 350 \text{ GeV}$ | 5.7 | 11.2 | 3.9 | 0.121 | 0.263 |
| $p_T(j_1 + j_2) \geq 100 \text{ GeV}$ | 4.648 | 8.16 | 3.3 | 0.07 | 0.223 |
| $\tau_{21}(j_1, j_2) < 0.55$ | 1.7 | 3.3 | 0.54 | 0.005 | 0.069 |
| $FW2(j_1) > 0.2$ | 1.28 | 2.7 | 0.32 | 0.003 | 0.020 |
| $(100 < M_{SD}(j_1, j_2) < 135) \text{ GeV}$ | 0.422 (1.46e5) | 1.09 (6.07e5) | 0.0136 (6.66e6) | 0.000008 (4.38e8) | 0.00078 (1.75e7) |

Table 7.1: Cumulative efficiency, in percentage, of each event selection criterion for the signal samples (SM and 2HDM) and for the background sample (4b+j, jj+0/1/2 j and t̄t+0/1/2 j). The absolute value of expected events after some key selection cuts is shown in curved brackets. The number of expected events is normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$. The double horizontal line marks the pre-selection cuts. These results were obtained using the FCC-hh baseline detector design, as implemented in Delphes by the FCC-hh study group.

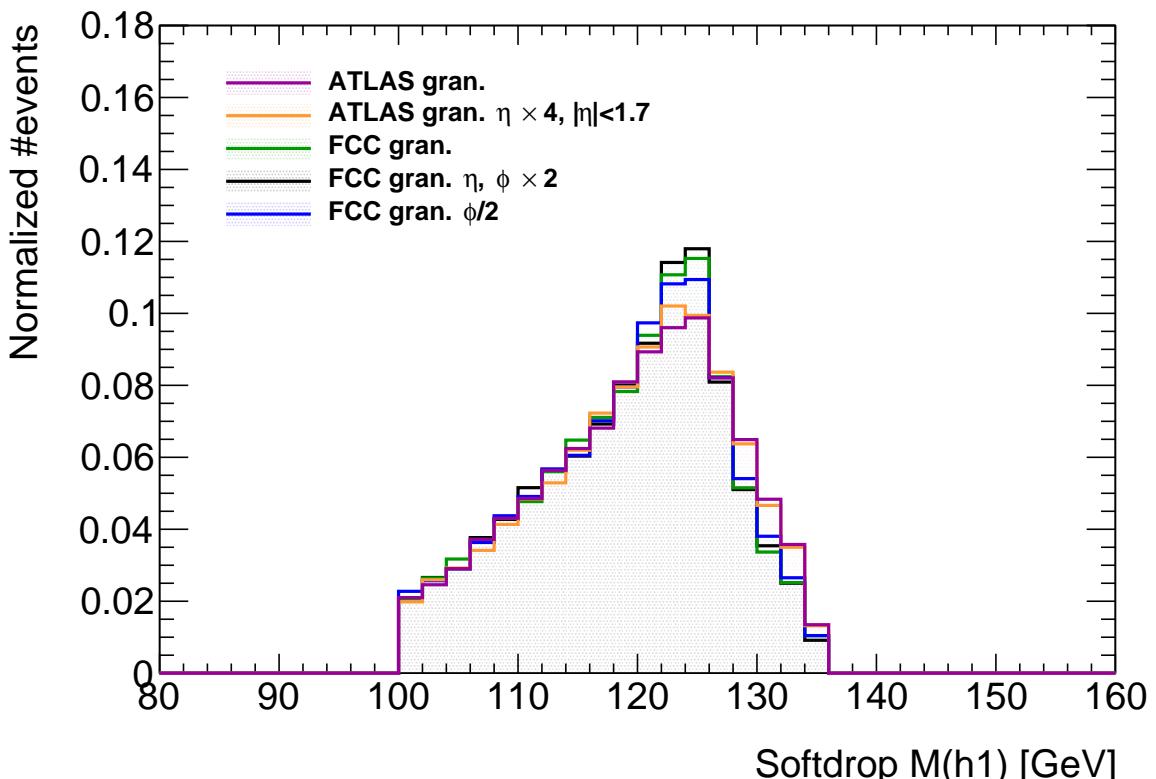


Figure 7.1: oi

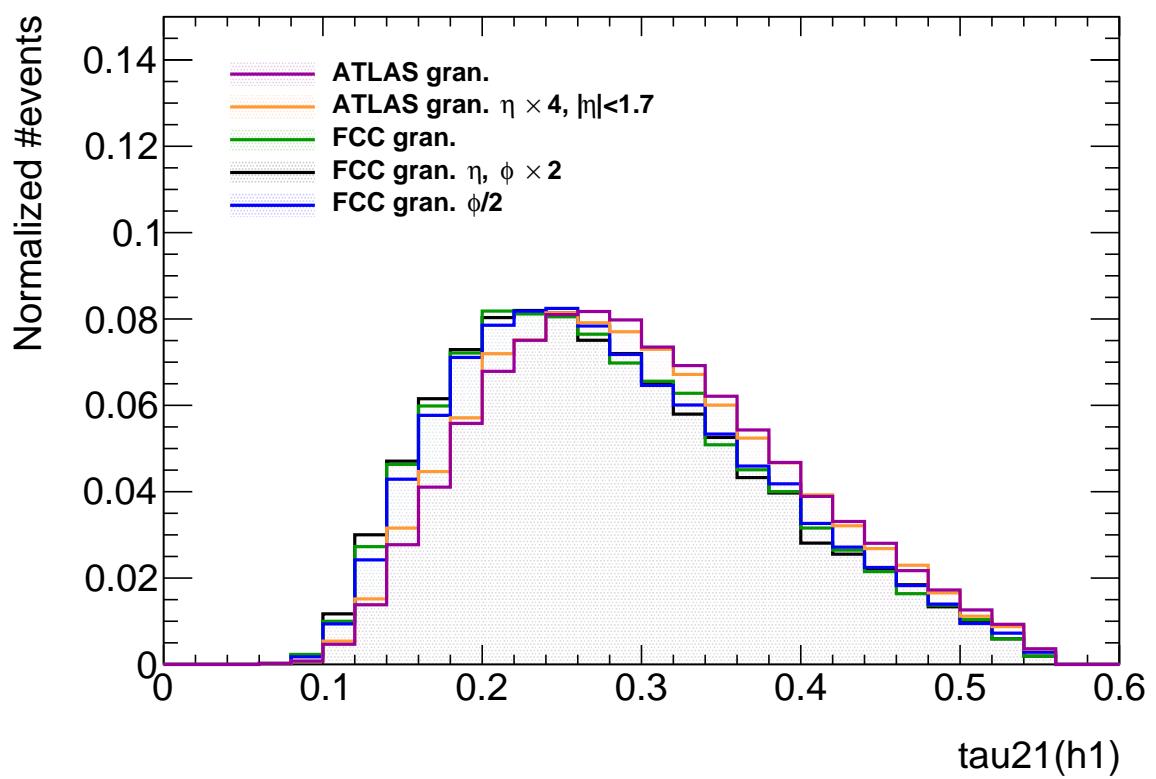


Figure 7.2: SUPERIMPOSE ONE BKG? MAYBE NOT

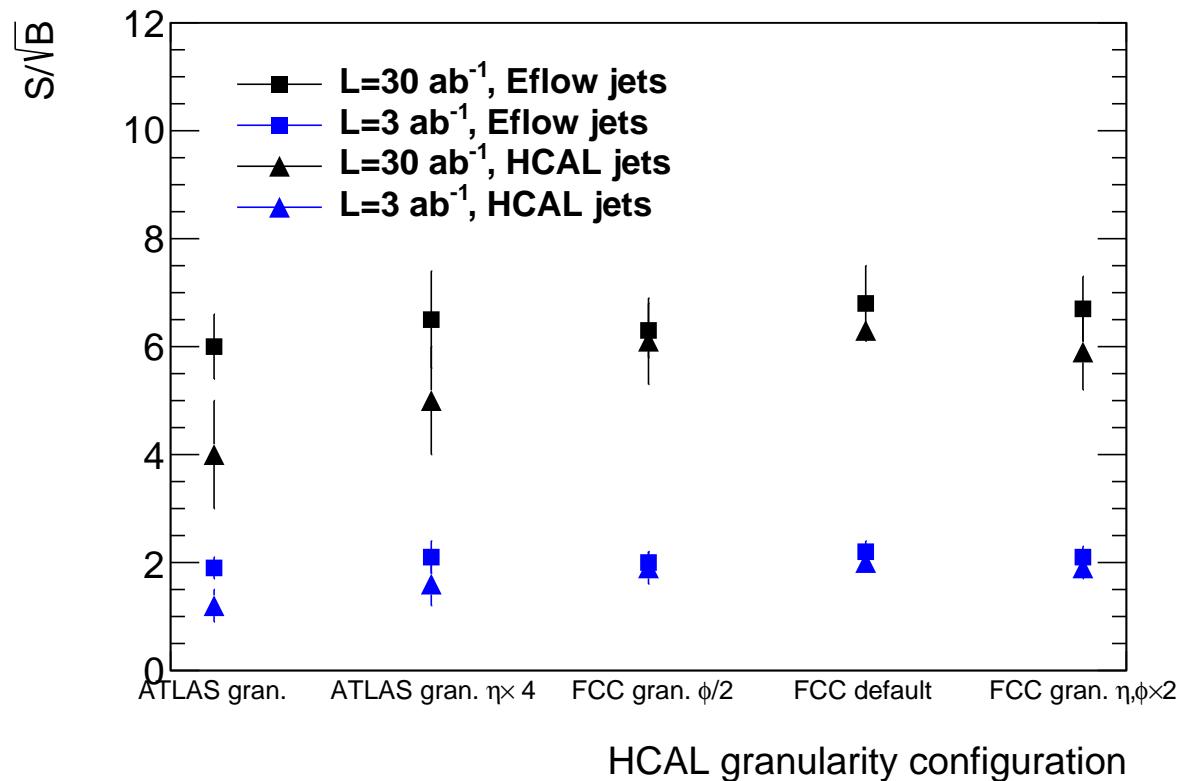


Figure 7.3: S/\sqrt{B} as a function of the detector configuration for $\mathcal{L} = 30 \text{ ab}^{-1}$ (black) and $\mathcal{L} = 3 \text{ ab}^{-1}$ (blue) and for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles).

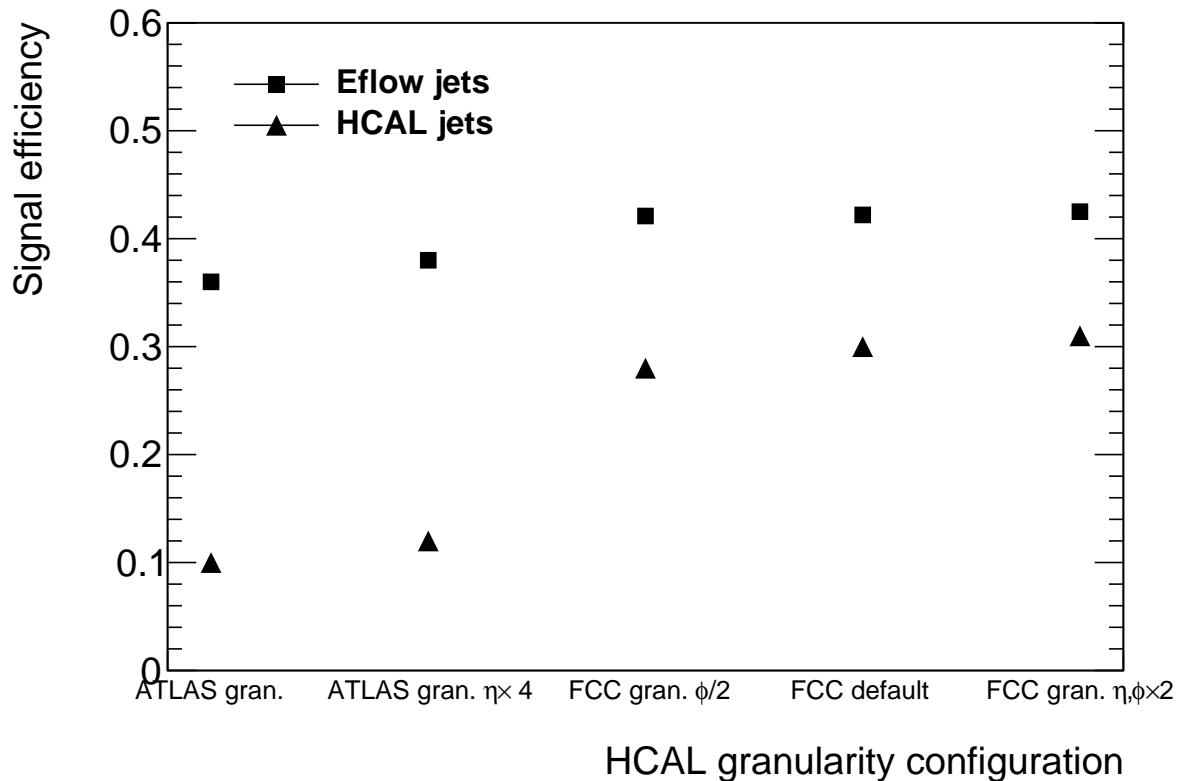


Figure 7.4: Signal efficiency as a function of the detector configuration for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles).

1370 **Chapter 8**

1371 **Conclusions**

1372 **BOOSTED H TO BB OBSERVATION AT THE FCC-HH**

- 1373 - Using a simple cut based analysis we achieved a S/\sqrt{B} of approximately 6 for $\mathcal{L} = 30 \text{ ab}^{-1}$. This is
1374 above the observation threshold (5σ). This result indicates that using the full dataset collected during
1375 the operation of the FCC-hh could lead to the observation of di-Higgs production, using the final state
1376 with four b quarks.
- 1377 - The sensitivity to the Higgs triple coupling is ...
- 1378 - For $\mathcal{L} = 3 \text{ ab}^{-1}$ the achieved significance is approximately 2 which is still below the evidence threshold
1379 (3σ).

1380

1381 **GRANULARITY STUDIES - SM**

- 1382 - Based only on the plots of the Higgs invariant mass and of some substructure variables for the sig-
1383 nal sample we can see a clear difference between the different granularity configurations. In particular,
1384 the mass resolution increases slightly and for the tau_21 we see a shift to the left. These plots work as
1385 good safety checks and highlight the impact of the different granularity configurations. However they are
1386 based only on the signal sample and therefore do not provide any information regarding the change in
1387 significance.
- 1388 - Using Eflow (particle flow) jets there is not a big change in S/\sqrt{B} as we change the detector configu-
1389 rations. The largest difference is of X %.
- 1390 - Based on this observation we redid the analysis using pure calorimeter jets (reconstructed using only
1391 information from the HCAL). The difference in S/\sqrt{B} over the detector configurations tested increases.
1392 The largest difference is of Y %.
- 1393 - Eflow jets use information from both the tracking and the hadronic calorimeter (in the case of hadronic
1394 jets). Based on the fact that S/\sqrt{B} varies very little when using eflow jets and that its variation increases
1395 if we use calorimeter jets we conclude that the resolution of the tracking system is so high that it domi-
1396 nates the jet reconstruction.
- 1397 - Nonetheless the signal efficiency (computed as the ratio between the number of signal events after all
1398 the analysis cuts and the total number of events) increases as the granularity of the detector increases

1399 when using eflow and calorimeter jets. This increment is more accentuated in the case of calorimeter
1400 jets where the maximum difference is $Z\%$. For Eflow jets the maximum variation is $W\%$.

1401

1402

1403 GRANULARITY STUDIES - BSM

1404 - In addition to the SM production of Higgs pairs we also analyzed two benchmark models in which new
1405 heavy particles couple to the SM Higgs boson through the s-channel diagram.

¹⁴⁰⁶ Bibliography

- ¹⁴⁰⁷ [1] J. Goldstone. Field Theories with Superconductor Solutions. *Nuovo Cim.*, 19:154–164, 1961. doi:
¹⁴⁰⁸ 10.1007/BF02812722.
- ¹⁴⁰⁹ [2] Y. Nambu and G. Jona-Lasinio. Dynamical Model of Elementary Particles Based on an Analogy with
¹⁴¹⁰ Superconductivity. 1. *Phys. Rev.*, 122:345–358, 1961. doi: 10.1103/PhysRev.122.345. [,127(1961)].
- ¹⁴¹¹ [3] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2
¹⁴¹² at the CERN SPS Collider. *Phys. Lett.*, B126:398–410, 1983. doi: 10.1016/0370-2693(83)90188-0.
¹⁴¹³ [,7.55(1983)].
- ¹⁴¹⁴ [4] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with
¹⁴¹⁵ Associated Missing Energy at $s^{**}(1/2) = 540\text{-GeV}$. *Phys. Lett.*, B122:103–116, 1983. doi: 10.1016/
¹⁴¹⁶ 0370-2693(83)91177-2. [,611(1983)].
- ¹⁴¹⁷ [5] F. Wilczek. Spontaneous symmetry breaking: General. <http://web.mit.edu/8.701/www/LectureNotes/8.701spontaneousSymmetryBreaking-GeneralFA13.pdf>, 2013. Accessed 2018-
¹⁴¹⁸ 01-11.
- ¹⁴²⁰ [6] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509,
¹⁴²¹ 1964. doi: 10.1103/PhysRevLett.13.508. [,160(1964)].
- ¹⁴²² [7] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev.*
¹⁴²³ *Lett.*, 13:321–323, 1964. doi: 10.1103/PhysRevLett.13.321. [,157(1964)].
- ¹⁴²⁴ [8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles.
¹⁴²⁵ *Phys. Rev. Lett.*, 13:585–587, 1964. doi: 10.1103/PhysRevLett.13.585. [,162(1964)].
- ¹⁴²⁶ [9] R. Costa Batalha Pedro, J. Maneira, and P. Conde Muio. Search for the Higgs boson at ATLAS/LHC
¹⁴²⁷ in WH associated production and decay to b quark pairs, Apr 2017. URL <https://cds.cern.ch/record/2304102>. Presented 09 Nov 2017.
- ¹⁴²⁹ [10] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity
¹⁴³⁰ conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. doi: 10.1103/PhysRev.105.
¹⁴³¹ 1413. URL <https://link.aps.org/doi/10.1103/PhysRev.105.1413>.

- 1432 [11] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Phys. Rev.*, 109:1015–1017,
 1433 Feb 1958. doi: 10.1103/PhysRev.109.1015. URL <https://link.aps.org/doi/10.1103/PhysRev.109.1015>.
- 1434
- 1435 [12] S. M. Bilenky and J. Hosek. GLASHOW-WEINBERG-SALAM THEORY OF ELECTROWEAK IN-
 1436 TERATIONS AND THE NEUTRAL CURRENTS. *Phys. Rept.*, 90:73–157, 1982. doi: 10.1016/0370-1573(82)90016-3.
- 1437
- 1438 [13] M. Mangano. *Physics at the FCC-hh, a 100 TeV pp collider*. CERN Yellow Reports: Monographs.
 1439 CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2270978>.
- 1440 [14] T. Plehn, M. Spira, and P. M. Zerwas. Pair production of neutral Higgs particles in gluon-
 1441 gluon collisions. *Nucl. Phys.*, B479:46–64, 1996. doi: 10.1016/0550-3213(96)00418-X,10.1016/0550-3213(98)00406-4. [Erratum: Nucl. Phys.B531,655(1998)].
- 1442
- 1443 [15] S. Borowka, N. Greiner, G. Heinrich, S. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke. Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence. *Phys. Rev. Lett.*, 117(1):012001, 2016. doi: 10.1103/PhysRevLett.117.079901,10.1103/PhysRevLett.117.012001. [Erratum: Phys. Rev. Lett.117,no.7,079901(2016)].
- 1444
- 1445
- 1446
- 1447 [16] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke. Full top quark mass dependence in Higgs boson pair production at NLO. *JHEP*, 10:107, 2016. doi: 10.1007/JHEP10(2016)107.
- 1448
- 1449
- 1450 [17] H. C. van de Hulst, E. Raimond, and H. van Woerden. Rotation and density distribution of the Andromeda nebula derived from observations of the 21-cm line. 14:1, Nov. 1957.
- 1451
- 1452 [18] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998. doi: 10.1103/PhysRevLett.81.1562.
- 1453
- 1454 [19] Q. R. Ahmad et al. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, 2001. doi: 10.1103/PhysRevLett.87.071301.
- 1455
- 1456
- 1457 [20] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002. doi: 10.1103/PhysRevLett.89.011301.
- 1458
- 1459
- 1460 [21] B. Gripaios. Lectures on Physics Beyond the Standard Model. 2015.
- 1461
- 1462 [22] H. Sun, Y.-J. Zhou, and H. Chen. Constraints on large-extra-dimensions model through 125-GeV Higgs pair production at the LHC. *Eur. Phys. J.*, C72:2011, 2012. doi: 10.1140/epjc/s10052-012-2011-4.
- 1463
- 1464 [23] P. Huang, A. Joglekar, M. Li, and C. E. M. Wagner. Corrections to di-Higgs boson production with light stops and modified Higgs couplings. *Phys. Rev.*, D97(7):075001, 2018. doi: 10.1103/PhysRevD.97.075001.
- 1465
- 1466

- 1467 [24] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with
 1468 the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- 1469 [25] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment
 1470 at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
- 1471 [26] M. Aaboud et al. Evidence for the $H \rightarrow b\bar{b}$ decay with the ATLAS detector. 2017.
- 1472 [27] Evidence for the higgs boson decay to a bottom quarkantiquark pair. *Physics Letters B*, 780:501
 1473 – 532, 2018. ISSN 0370-2693. doi: <https://doi.org/10.1016/j.physletb.2018.02.050>. URL <http://www.sciencedirect.com/science/article/pii/S0370269318301618>.
- 1475 [28] A. M. Sirunyan et al. Observation of the Higgs boson decay to a pair of τ leptons with the CMS
 1476 detector. *Phys. Lett.*, B779:283–316, 2018. doi: 10.1016/j.physletb.2018.02.004.
- 1477 [29] A. M. Sirunyan et al. Observation of $t\bar{t}H$ production. *Phys. Rev. Lett.*, 120:231801, 2018. doi:
 1478 [10.1103/PhysRevLett.120.231801](https://doi.org/10.1103/PhysRevLett.120.231801).
- 1479 [30] M. Aaboud et al. Observation of Higgs boson production in association with a top quark pair at the
 1480 LHC with the ATLAS detector. 2018.
- 1481 [31] Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels with
 1482 $\sqrt{s}=13\text{TeV}$ pp collisions using the ATLAS detector. Technical Report ATLAS-CONF-2017-046,
 1483 CERN, Geneva, Jul 2017. URL <http://cds.cern.ch/record/2273853>.
- 1484 [32] G. Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and
 1485 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015. doi: 10.1103/
 1486 [PhysRevLett.114.191803](https://doi.org/10.1103/PhysRevLett.114.191803).
- 1487 [33] M. Aaboud et al. Measurement of the W -boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS
 1488 detector. *Eur. Phys. J.*, C78(2):110, 2018. doi: 10.1140/epjc/s10052-017-5475-4.
- 1489 [34] R. Nisius. Measurements of the top quark mass with the ATLAS detector. *PoS*, EPS-HEP2017:
 1490 453, 2017. doi: 10.22323/1.314.0453.
- 1491 [35] Measurement of the top quark mass with muon+jets final states in pp collisions at $\sqrt{s} = 13$ TeV.
 1492 Technical Report CMS-PAS-TOP-16-022, CERN, Geneva, 2017. URL [http://cds.cern.ch/](http://cds.cern.ch/record/2255834)
 1493 record/2255834.
- 1494 [36] CERN. The high-luminosity lhc. <http://cds.cern.ch/record/2114693>, 2015. Accessed
 1495 2018/01/11.
- 1496 [37] T. Koffas. ATLAS Higgs physics prospects at the high luminosity LHC. *PoS*, ICHEP2016:426, 2017.
- 1497 [38] A. M. Henriques Correia. The ATLAS Tile Calorimeter. Technical Report ATL-TILECAL-PROC-
 1498 2015-002, CERN, Geneva, Mar 2015. URL <https://cds.cern.ch/record/2004868>.
- 1499 [39] ATLAS liquid argon calorimeter: Technical design report. 1996.

- 1500 [40] M. Cacciari, G. P. Salam, and G. Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008. doi:
1501 10.1088/1126-6708/2008/04/005.
- 1502 [41] Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run. Technical Report ATL-
1503 PHYS-PUB-2016-012, CERN, Geneva, Jun 2016. URL <https://cds.cern.ch/record/2160731>.
- 1504 [42] A. Ruiz-Martinez and A. Collaboration. The Run-2 ATLAS Trigger System. Technical Report ATL-
1505 DAQ-PROC-2016-003, CERN, Geneva, Feb 2016. URL <https://cds.cern.ch/record/2133909>.
- 1506 [43] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*,
1507 D48:3160–3166, 1993. doi: 10.1103/PhysRevD.48.3160.
- 1508 [44] M. Wobisch and T. Wengler. Hadronization corrections to jet cross-sections in deep inelastic scat-
1509 tering. In *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany,*
1510 1998-1999, pages 270–279, 1998. URL https://inspirehep.net/record/484872/files/arXiv:hep-ph_9907280.pdf.
- 1512 [45] M. Cacciari, G. P. Salam, and G. Soyez. The Anti- $k(t)$ jet clustering algorithm. *JHEP*, 04:063, 2008.
1513 doi: 10.1088/1126-6708/2008/04/063.
- 1514 [46] A. Altheimer et al. Boosted objects and jet substructure at the LHC. Report of BOOST2012, held
1515 at IFIC Valencia, 23rd-27th of July 2012. *Eur. Phys. J.*, C74(3):2792, 2014. doi: 10.1140/epjc/
1516 s10052-014-2792-8.
- 1517 [47] G. A. et al. Performance of jet substructure techniques for large- r jets in proton-proton collisions
1518 at $\sqrt{s} = 7$ tev using the atlas detector. *Journal of High Energy Physics*, 2013(9), Sep 2013. doi:
1519 10.1007/JHEP09(2013)076. URL [https://doi.org/10.1007/JHEP09\(2013\)076](https://doi.org/10.1007/JHEP09(2013)076).
- 1520 [48] J. Thaler and K. Van Tilburg. Identifying Boosted Objects with N-subjettiness. *JHEP*, 03:015, 2011.
1521 doi: 10.1007/JHEP03(2011)015.
- 1522 [49] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam. Jet substructure as a new Higgs
1523 search channel at the LHC. *Phys. Rev. Lett.*, 100:242001, 2008. doi: 10.1103/PhysRevLett.100.
1524 242001.
- 1525 [50] I. Hinchliffe, A. Kotwal, M. L. Mangano, C. Quigg, and L.-T. Wang. Luminosity goals for a 100-TeV
1526 pp collider. *Int. J. Mod. Phys.*, A30(23):1544002, 2015. doi: 10.1142/S0217751X15440029.
- 1527 [51] N. Arkani-Hamed, T. Han, M. Mangano, and L.-T. Wang. Physics opportunities of a 100 TeV pro-
1528 tonproton collider. *Phys. Rept.*, 652:1–49, 2016. doi: 10.1016/j.physrep.2016.07.004.
- 1529 [52] J. Baglio, A. Djouadi, R. Grber, M. M. Mhlleitner, J. Quevillon, and M. Spira. The measurement
1530 of the Higgs self-coupling at the LHC: theoretical status. *JHEP*, 04:151, 2013. doi: 10.1007/
1531 JHEP04(2013)151.

- 1532 [53] M. Aaboud et al. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–
 1533 proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev.*, D94(5):052002, 2016. doi:
 1534 10.1103/PhysRevD.94.052002.
- 1535 [54] A. M. Sirunyan et al. Search for a massive resonance decaying to a pair of Higgs bosons in the
 1536 four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- 1537 [55] G. Aad et al. Searches for Higgs boson pair production in the $hh \rightarrow b\bar{b}\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$
 1538 channels with the ATLAS detector. *Phys. Rev.*, D92:092004, 2015. doi: 10.1103/PhysRevD.92.
 1539 092004.
- 1540 [56] A. M. Sirunyan et al. Search for Higgs boson pair production in events with two bottom quarks and
 1541 two tau leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- 1542 [57] V. Khachatryan et al. Search for two Higgs bosons in final states containing two photons and
 1543 two bottom quarks in proton-proton collisions at 8 TeV. *Phys. Rev.*, D94(5):052012, 2016. doi:
 1544 10.1103/PhysRevD.94.052012.
- 1545 [58] G. Aad et al. Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision
 1546 Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector. *Phys. Rev. Lett.*, 114(8):081802, 2015. doi: 10.1103/
 1547 PhysRevLett.114.081802.
- 1548 [59] A. M. Sirunyan et al. Search for resonant and nonresonant Higgs boson pair production in the
 1549 $b\bar{b}lnlnu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- 1550 [60] J. K. Behr, D. Bortoletto, J. A. Frost, N. P. Hartland, C. Issever, and J. Rojo. Boosting Higgs pair
 1551 production in the $b\bar{b}b\bar{b}$ final state with multivariate techniques. *Eur. Phys. J.*, C76(7):386, 2016. doi:
 1552 10.1140/epjc/s10052-016-4215-5.
- 1553 [61] D. Wardrope, E. Jansen, N. Konstantinidis, B. Cooper, R. Falla, and N. Norjoharuddeen. Non-
 1554 resonant Higgs-pair production in the $b\bar{b} b\bar{b}$ final state at the LHC. *Eur. Phys. J.*, C75(5):219, 2015.
 1555 doi: 10.1140/epjc/s10052-015-3439-0.
- 1556 [62] D. E. Ferreira de Lima, A. Papaefstathiou, and M. Spannowsky. Standard model Higgs boson pair
 1557 production in the $(b\bar{b})(b\bar{b})$ final state. *JHEP*, 08:030, 2014. doi: 10.1007/JHEP08(2014)030.
- 1558 [63] S. Banerjee, C. Englert, M. L. Mangano, M. Selvaggi, and M. Spannowsky. $hh +$ jet production at
 1559 100 TeV. *Eur. Phys. J.*, C78(4):322, 2018. doi: 10.1140/epjc/s10052-018-5788-y.
- 1560 [64] N. P. Hartland. Double Higgs production from $HH \rightarrow (b\bar{b})(b\bar{b})$ at a 100 TeV hadron collider. *PoS*,
 1561 DIS2016:115, 2016.
- 1562 [65] S. V. Chekanov, M. Beydler, A. V. Kotwal, L. Gray, S. Sen, N. V. Tran, S. S. Yu, and J. Zuzelski. Initial
 1563 performance studies of a general-purpose detector for multi-TeV physics at a 100 TeV pp collider.
 1564 *JINST*, 12(06):P06009, 2017. doi: 10.1088/1748-0221/12/06/P06009.

- 1565 [66] S. Chekanov. Boosted jets in high-granularity calorimeter at a 100 TeV pp collider. BOOST 2017,
1566 2017.
- 1567 [67] S. Chekanov. Physics requirements for a Hadron Calorimeter for a 100 TeV proton-proton collider.
1568 FCC week 2015, 2015.
- 1569 [68] S. Chekanov. Simulation of a high-granular hadronic calorimeter for multi-TeV physics. FCC week
1570 2016, 2016.
- 1571 [69] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer.
- 1572 [70] T. Sjostrand, S. Mrenna, and P. Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys.*
1573 *Commun.*, 178:852–867, 2008. doi: 10.1016/j.cpc.2008.01.036.
- 1574 [71] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lematre, A. Mertens, and M. Selvaggi.
1575 DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:
1576 057, 2014. doi: 10.1007/JHEP02(2014)057.
- 1577 [72] M. Cacciari, G. P. Salam, and G. Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012. doi:
1578 10.1140/epjc/s10052-012-1896-2.
- 1579 [73] D. Buskulic et al. Performance of the ALEPH detector at LEP. *Nucl. Instrum. Meth.*, A360:481–506,
1580 1995. doi: 10.1016/0168-9002(95)00138-7.
- 1581 [74] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical
1582 Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009. URL <http://cds.cern.ch/record/1194487>.
- 1583 [75] FCC study group. Fccsw. <https://github.com/HEP-FCC/FCCSW>, 2018.
- 1584 [76] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing
1585 applications. *Comput. Phys. Commun.*, 140:45–55, 2001. doi: 10.1016/S0010-4655(01)00254-5.
- 1586 [77] F. Gaede, B. Hegner, and P. Mato. Podio: An event-data-model toolkit for high energy physics
1587 experiments. *Journal of Physics: Conference Series*, 898(7):072039, 2017. URL <http://stacks.iop.org/1742-6596/898/i=7/a=072039>.
- 1588 [78] FCC-hh study group. Eventproducer. <https://github.com/FCC-hh-framework/EventProducer>,
1589 2018.
- 1590 [79] N. D. Christensen and C. Duhr. FeynRules - Feynman rules made easy. *Comput. Phys. Commun.*,
1591 180:1614–1641, 2009. doi: 10.1016/j.cpc.2009.02.018.
- 1592 [80] P. M. Ferreira, R. Santos, M. Sher, and J. P. Silva. 2HDM confronting LHC data. In *Proceedings,*
1593 *48th Rencontres de Moriond on QCD and High Energy Interactions: La Thuile, Italy, March 9-16, 2013*, pages
1594 67–70, 2013. URL <https://inspirehep.net/record/1234236/files/arXiv:1305.4587.pdf>.

₁₅₉₈ [81] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016. doi: 10.1088/
₁₅₉₉ 1674-1137/40/10/100001.

1600 **Appendix A**

1601 **Jet radius discussion**

1602 It is important to discuss if the radius of the jet that is used to reconstruct the boosted Higgs boson
1603 candidates ($R = 0.8$) is the appropriate one. The question arises because in ATLAS boosted objects
1604 with $p_T \sim 200$ GeV are reconstructed using $R = 1.0$ jets. In this work, we work with a CM energy of 100
1605 TeV and require that the two leading jets have $p_T > 200$ GeV and use jets with $R = 0.8$. It is necessary
1606 to understand if these jets are large enough to fully reconstruct the Higgs candidates.

1607 As a first approximation we compute the angle between the b quarks produced by the decay of
1608 a Higgs boson. We assume that the b quarks are massless and that the Higgs moves only in the
1609 transverse plane (perpendicular to the beam pipe) such that it has no longitudinal momentum and the
1610 angle between the b quarks is given by $\Delta\phi$. For $p_T(\text{Higgs}) = 200$ GeV, we get $\Delta\phi(b, \bar{b}) = 1.1$ which is
1611 smaller than the jet's diameter (1.6) and therefore the two b quarks can both be contained inside the jet
1612 and the Higgs boson fully reconstructed.

1613 Another test we can make is to compute the
1614 ΔR between the b quarks coming from the lead-
1615 ing Higgs candidate with p_T larger than a given
1616 value using truth level information. In figure A.1
1617 we show the distribution of $\Delta R(b, \bar{b})$ for p_T of
1618 the leading Higgs candidate larger than 200 GeV
1619 (solid blue) and 300 GeV (solid black). To obtain
1620 the dashed histograms we apply the p_T to both
1621 Higgs candidates. The integral of the histograms
1622 between 0 and 1.6 gives an estimate of the frac-
1623 tion of signal we keep if we apply these p_T cuts.
1624 For $p_T(h_1) > 200(300)$ GeV we get that 93(98) %
1625 of the signal has b quarks with $\Delta R < 1.6$ and
1626 therefore can be fully reconstructed using a jet
1627 with $R = 0.8$.

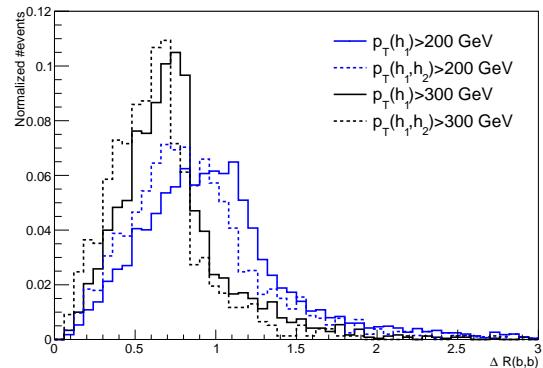


Figure A.1: $\Delta R(b, \bar{b})$ distributions for $p_T(h_1) > 200/300$ GeV (solid blue/black) and for $p_T(h_1, h_2) > 200/300$ GeV (dashed blue/black).

1628 **Appendix B**

1629 **Additional background processes**

1630 In this section we discuss the importance of additional background processes and estimate how they
1631 influence the analysis. In particular, we investigate backgrounds involving Higgs bosons.

₁₆₃₂ **Appendix C**

₁₆₃₃ **Softdrop mass**

₁₆₃₄ It is noticeable in the softdrop mass spectrum of both Higgs candidates (for signal and backgrounds) the
₁₆₃₅ existence an atypical peak close to zero. In this section we explain the origin of this feature.

₁₆₃₆ We believe that the peak close to zero corresponds to Higgs candidates ($R = 0.8$ jets) that do not
₁₆₃₇ contain both b quarks from the Higgs decay. The plot that support this conclusion is shown in figure
₁₆₃₈ C.1. It shows the correlation between the maximum ΔR between the Higgs candidate jet and one of
₁₆₃₉ the b quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$. For
₁₆₄₀ small masses (< 80 GeV) the maximum ΔR is usually larger than 0.8 which means that at least one
₁₆₄₁ of the b quarks is not contained in the jet's cone. When applying the soft drop procedure to these jets,
₁₆₄₂ soft radiation is removed and we are left with a single b quark. The mass of b quarks is ~ 5 GeV and
₁₆₄₃ therefore we get a peak at this mass.

₁₆₄₄ In practice, this does not affect our analysis because we place a mass window cut around the Higgs
₁₆₄₅ boson mass and therefore the low mass peak is removed. Nonetheless, the study presented here is
₁₆₄₆ extremely important because it helps rule out possible malfunctions of the soft drop algorithm and gives
₁₆₄₇ us confidence that we understand exactly what is happening.

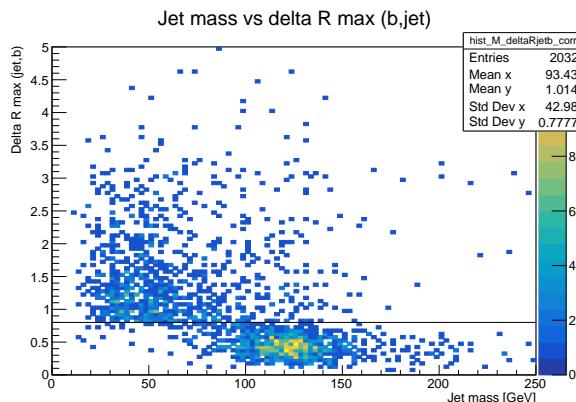


Figure C.1: Correlation between the maximum ΔR between the Higgs candidate jet and one of the b quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$.

1648 **Appendix D**

1649 **Samples generation: parameters**

1650 In this appendix we provide we list the values of parameters used in MadGraph (section D.1) and Pythia
1651 (section D.2) to generate the samples used in this work.

1652 **D.1 MadGraph**

1653 The MadGraph5 level cuts are summarized in table D.1. We show only the most relevant cuts for this
1654 analysis: the minimum p_T of light and b quarks, $p_{T,j}^{\min}$ and $p_{T,b}^{\min}$, the maximum pseudorapidity range for
1655 light and b quarks, η_j^{\max} and η_b^{\max} and the ΔR separation between two light quarks, $\Delta R(jj)$, two b
1656 quarks, $\Delta R(bb)$, and between a light and b quarks, $\Delta R(jb)$. The $xqcut$ parameter is a measure of the
1657 required parton separation at Madgraph level. Whenever MadGraph produces two partons, i and j ,
1658 we define the distance between them as $\sqrt{2 * \min(p_{T,i}, p_{T,j}) [\cosh(\eta_i - \eta_j) - \cos(\phi_i - \phi_j)]}$. If the value
1659 of this expression is smaller than the specified value of $xqcut$ then we do not generate the event. The
1660 $bwcutoff$ parameter defines what is considered to be on-shell s-channel resonances. The H_T variable
1661 is the scalar sum of the p_T of all truth level partons, including b quarks.

1662 **D.2 Pythia**

1663 The settings for jet matching can be found in table D.2 under the corresponding samples' columns. We
1664 perform the jet matching procedure (merge=on) using the MLM matching scheme and the appropriate
1665 algorithm for a parton level process generated in MadGraph (scheme=1). We do not read the matching
1666 parameters from the MadGraph file (setMad=off) because this option is not available for these files. The
1667 size of the cone drawn around the jet's center, the maximum pseudorapidity and the maximum number
1668 of jets to be matched are given by coneRadius, etaJetMax and nJetMax, respectively. The cone radius
1669 is set to one. The maximum allowed pseudorapidity of jets is ten which is a much loser cut than the
1670 acceptance of any current detector. The maximum number of jets is set to four for the jj+0/1/2 j and to
1671 two for the $t\bar{t}$ +0/1/2 j. The qCut parameter defines the k_T scale for merging shower products into jets.

| MadGraph5 | SM ($h \rightarrow b\bar{b}$) & BSM hh | hh & (QCD) | 4b+j (QCD+EWK) | 4b+j (EWK) | jj+0/1/2 j | tt+0/1/2 j |
|--|--|--------------------|-------------------|---------------|--|------------|
| $p_{T,j}^{\min}, p_{T,b}^{\min}$ [GeV] | 0 | 200; 30 500; 30 | 20; 15 | 20; 15 | 20; 5 | 5; 5 |
| $\eta_j^{\max}, \eta_b^{\max}$ | — | 5; 5 | 5; 3 | 5; 3 | 8; 8 | 8; 8 |
| $\Delta R(jj), \Delta R(bb), \Delta R(jb)$ | 0.001 | 0.4; 0.1; 0.3 | 0.4; 0.2; 0.4 | 0.4; 0.2; 0.4 | 0; 0.001; 0.001 | 0.001 |
| xqcut [GeV] | 0 | 0 | 0 | 0 | 20 | 60 |
| bwcutoff [GeV] | 30 | 15 | 15 | 15 | 30 | 30 |
| H_T | — | — | — | — | 0 – 500 500 – 1k 1k – 2k 2k – 4k 4k – 7.2k 7.2k – 15k 15k – 25k 25k – 35k 35k – 100k | — |

Table D.1: Generator (MadGraph5) level cuts for the signal and background samples.

| Pythia | hh ($h \rightarrow b\bar{b}$) | 4b+j | jj+0/1/2 j | tt+0/1/2 j |
|-------------------|---|------|--|--|
| Relevant settings | 25:onMode=off 25:onIfAny= 5 -5 | — | Jet matching: merge=on scheme=1 setMad=off coneRadius=1.0 etaJetMax=10 nJetMax=4 qCut=30 | Jet matching: merge=on scheme=1 setMad=off coneRadius=1.0 etaJetMax=10 nJetMax=2 qCut=60 |
| Description | Turn on the $h \rightarrow b\bar{b}$ decay for the undecayed Higgs, in the case of the SM sample, or for both Higgs in the case of the BSM samples. | — | Set the parameters for jet matching (a detailed description can be found in the text). | |

Table D.2: Pythia settings for the signal and background samples.