

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

10 Acknowledgments

11 I would like to express my sincere thanks to my supervisor Ricardo Gonçalo for his support, invaluable
12 scientific advise and inspirational enthusiasm, without which this work would not have been concluded.

13 To my co-supervisor, Pedro Abreu, I thank for ...

14 I acknowledge the funding of this research by Fundação para a Ciência e Tecnologia - FCT, through

15 ...

16 I am very thankful to Patricia Conde and Rute Pedro for fruitful discussions and very helpful advise.

17 To my master colleagues in the LIP ATLAS group, Aidan Kelly, António Costa and Ricardo Barrué, I
18 thank for the companionship and motivation throughout this journey.

19 I am thankful to Duarte Azevedo, Pedro Ferreira and Rui Santos for being our on-call theoreticians,
20 always available to help.

21 I would like to thank all my colleagues in the LIP ATLAS group for making me feel very welcome.

22 I extend my gratitude to everybody at LIP for making it a fantastic place to work. I would also like to
23 acknowledge the LIP IT and secretariat for their patience and technical support.

24 I would like to thank my boyfriend, João Augusto, for his never ending patience and unconditional
25 support. To my friends of Chafarrica I express my most sincere thanks for always being a great support,
26 source of joy and laughter.

27 To my sister I thank for always welcoming me home with a smile, for the numerous time she offered
28 to help and for all the memes that lighten up some of my days.

29 To my parents, I will never be able to thank enough. I thank them for their love, unconditional support,
30 inspiration, help in times of need

Resumo

- ³¹ Inserir o resumo em Português aqui com o mximo de 250 palavras e acompanhado de 4 a 6 palavras-chave...

- ³⁴ **Palavras-chave:** palavra-chave1, palavra-chave2,...

³⁵ **Abstract**

³⁶ Insert your abstract here with a maximum of 250 words, followed by 4 to 6 keywords...

³⁷ **Keywords:** keyword1, keyword2,...

Contents

38	Acknowledgments	iii
39	Resumo	iv
40	Abstract	v
41	List of Tables	ix
42	List of Figures	xii
43	Nomenclature	1
44	Glossary	1
46	1 Introduction	1
47	2 The standard model and beyond	4
48	2.1 The Standard Model of Particle Physics	4
49	2.1.1 Higgs pair production	10
50	2.2 Going beyond	13
51	3 Collider experiments	17
52	3.1 Experimental aspects	18
53	3.2 The Large Hadron Collider	18
54	3.2.1 The ATLAS detector	20
55	3.3 Jet reconstruction	23
56	3.3.1 Jet properties and substructure observables	25
57	3.3.2 Jet grooming algorithms	26
58	3.4 Future Colliders	27
59	3.4.1 The hadronic Future Circular Collider	27
60	3.4.2 FCC-hh baseline detector	28
61	3.4.3 FCC-hh physics program	29
62	4 State of the art	31
63	4.1 Searches for Higgs pair production at the LHC	31
64	4.2 Feasibility studies for high-luminosity and future colliders	33
65	4.3 Hadronic calorimeter granularity studies	34

66	5 Sample generation and analysis tools	38
67	5.1 Fast simulation workflow	38
68	5.1.1 Particle flow and calorimeter reconstruction in Delphes	39
69	5.1.2 FCC-hh software	40
70	5.2 Signal and background samples	41
71	5.2.1 MadGraph	41
72	5.2.2 Pythia	43
73	5.2.3 Delphes	44
74	6 Analysis	47
75	6.1 Overview of the $hh \rightarrow b\bar{b}b\bar{b}$ channel	47
76	6.1.1 Event pre-selection	48
77	6.1.2 Signal characterization	48
78	6.1.3 Backgrounds	49
79	6.2 Analysis strategy	53
80	6.2.1 Implementation of b-tagging	53
81	6.2.2 Baseline analysis	54
82	6.2.3 Optimization	55
83	6.2.4 Handling of uncertainties	57
84	7 Results	60
85	7.1 Di-Higgs discovery potential at the FCC-hh	60
86	7.1.1 Statistical analysis	61
87	7.1.2 Comparing with the ATLAS detector	61
88	7.2 Hadronic calorimeter granularity studies for future colliders	64
89	7.2.1 Resolution of jet mass and N-subjetiness	64
90	7.2.2 Signal efficiency and significance	64
91	8 Conclusions	68
92	Bibliography	69
93	A Jet radius discussion	77
94	B Additional background processes	78
95	C Softdrop mass	79
96	D Samples generation: parameters	80
97	D.1 MadGraph	80
98	D.2 Pythia	80

List of Tables

99	3.1 ATLAS tile and liquid argon hadronic calorimeters: summary of the pseudorapidity coverages and transversal segmentation (granularity)	21
100	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.	28
101	3.3 Hadronic calorimeter layout, granularity and energy resolution.	29
102	5.1 Summary of the cross sections, k factors and total number of events of the samples used in the analysis.	44
103	5.2 Summary of the benchmark granularity configurations of the HCAL.	45
104	5.3 Summary of the benchmark granularity configurations of the HCAL.	46
105	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.	54
106	7.1 Cumulative efficiency, in percentage, of each event selection criterion for the signal samples (SM and 2HDM). 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.	61
107	7.2 Cumulative efficiency, in percentage, of each event selection criterion for the background samples ($4b + j$, $jj + 0/1/2j$ and $t\bar{t}+0/1/2j$). 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.	62

127	7.3 Cumulative efficiency, in percentage, of each event selection criterion for the signal sam-	
128	ples (SM and 2HDM). The absolute value of expected events after some key selection	
129	cuts is shown in curved brackets. The number of expected events is normalized to	
130	$\mathcal{L} = 30 \text{ ab}^{-1}$. The double horizontal line marks the pre-selection cuts. These results	
131	were obtained using the ATLAS detector design, as implemented in Delphes.	
132	62
133	7.4 Cumulative efficiency, in percentage, of each event selection criterion for the background	
134	samples ($4b + j$, $jj + 0/1/2j$ and $t\bar{t}+0/1/2 j$). The absolute value of expected events	
135	after some key selection cuts is shown in curved brackets. The number of expected	
136	events is normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$. The double horizontal line marks the pre-selection	
137	cuts. These results were obtained using the ATLAS detector design, as implemented in	
138	Delphes.	
139	63
140	7.5 Significances achieved with the ATLAS and FCC-hh detector configurations for $\mathcal{L} = 30 \text{ ab}^{-1}$	
141	for the three benchmark signal models: SM, 1 TeV DM mediator and 2HDM with $m_H = 900$	
142	GeV.	
143	63
144	D.1 Generator (MadGraph5) level cuts for the signal and background samples.	81
145	D.2 Pythia settings for the signal and background samples.	81

List of Figures

146	2.1	Schematic representation of the Standard Model of elementary particles. Particle are represented inside squares. The electric and color charge are shown in the right upper corner of each square. The spin is shown in the right lower corner. The mass of the particles is also shown, in electron Volt.	5
147	2.2	Postulated shape of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).	8
148	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.	11
149	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.	13
150	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. Plots from [23].	14
151	2.6	BSM contributions to Higgs pair production. On the left (figure from [26]) 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 [27]) \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
152	3.1	ATLAS detector.	22
153	3.2	FCC detector baseline concept.	30
154	4.1	Exclusion limits (at 95% CL) on the cross section of resonant Higgs pair production from the decay of heavy scalar particles. These results were obtained using proton-proton collision data collected with the CMS (left) and ATLAS (right) detector at a CM energy of $\sqrt{s} = 8$ TeV and result from the combination of different channels. Plots from [69] (left) and [73] (right).	33
155	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 [78] (based on [77]).	35

175	4.3 Jet mass (left) and jet mass resolution (right) plots in $t\bar{t}$ events for eflow jets with $p_T > 3$	36
176	TeV for three different HCAL and ECAL configurations: HCAL(ECAL) 0.1(0.025) $\eta \times$	
177	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) $\eta \times$	
178	1.4(0.35) deg ϕ . On the right, the root mean square (RMS) are also shown. Plots from [78].	
179	4.4 τ_{32} resolution plots for QCD dijet events reconstructed using eflow jets (left) and tower jets	37
180	(right). Plots from [79].	
181	4.5 τ_{21} variable for 20 TeV W (black) and QCD jets (red) for three HCAL granularities: $\Delta\eta \times$	
182	$\Delta\phi = 0.1 \times 0.1$ (left), $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$ (middle) and $\Delta\eta \times \Delta\phi = 0.005 \times 0.005$	
183	(right). Plots from [78].	37
184	6.1 Higgs pairs branching ratios (left) and event topology targeted by this analysis (right).	47
185	6.2 $\Delta\phi$ between the Higgs candidates, $\Delta\phi(hh)$, (left) and ΔR between the b quarks coming	
186	from the decay of the leading Higgs candidate, h_1 (right). To obtain this plot the cut	
187	$ M - 125 < 40$ GeV was applied in addition to the pre-selection cuts..	50
188	6.3 Example of diagrams that contribute to the QCD multijet background: five final state jets,	
189	four of which are b-jets (left), three final state jets (middle) and four final state jets (right).	
190	Here, q stands for a light quark/jet.	51
191	6.4 Dominant diagrams of $pp \rightarrow t\bar{t}$ at LO.	51
192	6.5 p_T distributions for the leading (left) and sub leading (right) Higgs candidates. The signal	
193	is the SM $hh \rightarrow b\bar{b}$ process. The histograms are normalized to unit area.	52
194	6.6 η distribution for the leading Higgs candidate (left) and softdrop mass distribution for the	
195	leading Higgs candidate (right). The signal is the SM $hh \rightarrow b\bar{b}$ process. The histograms	
196	are normalized to unit area.	52
197	6.7 Distributions of the $\Delta\eta$ between the Higgs candidates (left) and of the τ_{21} variable for the	
198	leading Higgs candidate (right). The histograms are normalized to unit area.	52
199	6.8 Minimun ΔR between b quarks and subjets of the $R = 0.8$ jets.	53
200	6.9 p_T distributions for the sub leading Higgs candidate (left) and for the Higgs pair (right).	
201	The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$.	55
202	6.10 Distributions of the τ_{21} variable for the leading (left) and sub leading (right) Higgs candi-	
203	dates. The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$.	56
204	6.11 H_2 variable for the leading Higgs candidate (left) and softdrop mass distribution for the	
205	leading Higgs candidate. The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$.	56
206	6.12 S/\sqrt{B} as a function of the cut on the p_T of the leading Higgs candidate, $p_T(h_1)$ (left) after	
207	the pre-selection cuts and of the sub leading Higgs candidate, $p_T(h_2)$ (right) after the cut	
208	on $p_T(h_1)$.	58
209	6.13 S/\sqrt{B} as a function of the cut on the p_T of the Higgs pair, $p_T(hh)$ (left) after the cut on	
210	$p_T(h_1)$ and as a function of the τ_{21} variable for leading Higgs candidate, $\tau_{21}(h_1)$ (right)	
211	after the cut on $p_T(hh)$.	58

212	6.14 S/\sqrt{B} as a function of the cut on the $\Delta\eta$ between the Higgs candidates, $\Delta\eta(hh)$ (left)	58
213	after the cuts on the transverse momentums and on the τ_{21} variables and as a function of	
214	the second Fox-Wolfram momentum for the leading Higgs candidate, $H_l(h_1)$ (right) after	
215	the cuts on the transverse momentums, τ_{21} and $\Delta\eta$	58
216	7.1 (a) Leading Higgs candidate softdrop mass after the pre-selection cuts, (b) The colors	
217	indicate the different detector configurations. The x axis range is from 80 GeV to 160 GeV	
218	in order to make the differences between the histograms more clear.	65
219	7.2	66
220	7.3 (a) Invariant mass of the Higgs pair for the 1 TeV DM mediator, (b) Invariant mass of the	
221	Higgs pair for the 2HDM with $m_H = 900$ GeV. The plots show include the events that pass	
222	all the analysis cuts. The colors indicate the different detector configurations.	66
223	7.4 (a) Leading Higgs candidate τ_{21} distribution for eflow jets, (b) Leading Higgs candidate τ_{21}	
224	distribution for HCAL jets. The colors indicate the different detector configurations. The	
225	distributions are shown for the signal (filled lines) and for the $4b + j$ background (dashed	
226	lines).	66
227	7.5 (a) Signal efficiency as a function of the detector configuration for particle flow jets, (b)	
228	Signal efficiency as a function of the detector configuration for calorimeter jets. Three	
229	signal models are shown: SM (filled squares), 1 TeV DM mediator (empty squares) and	
230	type II 2HDM with $m_H = 900$ GeV (crosses). The error bars are drawn but are smaller	
231	than the markers.	67
232	7.6 (a) Significance as a function of the detector configuration for the SM signal, (b) Sig-	
233	nificance as a function of the detector configuration for the DM mediator model. The	
234	significances are shown for $\mathcal{L} = 30 \text{ ab}^{-1}$ (filled) and $\mathcal{L} = 3 \text{ ab}^{-1}$ (empty) and for particle	
235	flow jets (squares) and pure calorimeter (HCAL) jets (triangles).	67
236	7.7 Significance as a function of the detector configuration for the 2HDM with $m_H = 900$ GeV.	
237	The significances are shown for $\mathcal{L} = 30 \text{ ab}^{-1}$ (filled) and $\mathcal{L} = 3 \text{ ab}^{-1}$ (empty) and for	
238	particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles).	67
239	A.1 $\Delta R(b, \bar{b})$ distributions for $p_T(h_1) > 200/300$ GeV (solid blue/black) and for $p_T(h_1, h_2) >$	
240	200/300 GeV (dashed blue/black).	77
241	C.1 Correlation between the maximum ΔR between the Higgs candidate jet and one of the b	
242	quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$	79

²⁴³ **Chapter 1**

²⁴⁴ **Introduction**

²⁴⁵ It is the ultimate goal of particle physics to discover and study all of Nature's fundamental particles and
²⁴⁶ to understand their interactions. Through a joint endeavor of theorists and experimentalists, models that
²⁴⁷ describe particle's dynamics and properties can be precisely probed at collider experiments such as the
²⁴⁸ Large Hadron Collider (LHC).

²⁴⁹ We know today that matter particles interact by means of four fundamental forces: electromagnetic,
²⁵⁰ weak, strong and gravitational, each associated with a mediator particle. We even know that a very
²⁵¹ special particle, the Higgs boson, is responsible for generating the mass of these particles through
²⁵² a mechanism called Electroweak Symmetry Breaking (EWSB). Through Yukawa couplings, the Higgs
²⁵³ boson is also responsible for the masses of matter particles. All of this knowledge is beautifully summa-
²⁵⁴ rized in the Standard Model of Particle Physics (SM) that was developed in the 1960's, long before many
²⁵⁵ of the particles it predicts were discovered. The extraordinary precision of the predictions it delivers
²⁵⁶ make it a very successful model. Its most recent prediction, the Higgs boson, was discovered in 2012
²⁵⁷ at the LHC which marks an important point in the history of particle physics: we have now found all the
²⁵⁸ particles predicted by the SM and yet we know that it cannot be the whole story. Mainly because there
²⁵⁹ are experimental evidences of physics that it cannot explain.

²⁶⁰ From the theoretical point of view, this is enough motivation to construct models that extend the SM
²⁶¹ but that can still deliver predictions that are compatible with experimental data. From the experimental
²⁶² standpoint, this is an indication that we need to keep increasing the precision of our measurements and
²⁶³ probing new kinematic regimes in the hope of finding some discrepancy with the SM or some hint that
²⁶⁴ some new phenomenon might be taking place.

²⁶⁵ A higher precision requires a larger integrated luminosity and the exploration of new kinematic
²⁶⁶ regimes asks for a larger center of mass energy. Very recently, work towards the upgrade of the LHC to
²⁶⁷ its High-Luminosity (HL) version has began. It is expected to work for a period of ten years between 2026
²⁶⁸ and 2036 and it will extend the experimental reach of the LHC. In order to keep extending the physics
²⁶⁹ reach of the LHC and HL-LHC, new colliders with unprecedentedly high center of mass (CM) energies
²⁷⁰ are currently being designed in the hope that they begin to deliver data shortly after the HL-LHC has
²⁷¹ reached its full discovery potential. One of these projects is the hadronic Future Circular Collider (FCC-

272 hh) that consists of a 100 km ring located in the Genve areas and it is expected to work at a CM energy
273 of 100 TeV. The FCC-hh will deliver a peak luminosity of $\mathcal{L} = 30 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ in its ultimate phase.
274 This will result in $O(30) \text{ ab}^{-1}$ per experiment which corresponds to ten times the expected integrated
275 luminosity by the end of the HL-LHC operation.

276 The next milestone for the FCC-hh project is the submission of a Conceptual Design Report by the
277 end of 2018. This document will be used as input for the next meeting of the European Strategy for
278 Particle Physics that will take place in the beginning of 2019. It should present a baseline design for the
279 detector, a first cost estimate and preliminary analysis for physics benchmark processes demonstrating
280 the physics reach of such a machine.

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

286 In this work we focus on the final state with four b quarks. We choose this final state because it
287 benefits from the large branching fraction of the Higgs boson to a pair of b quarks. However, in this
288 channel, the multijet production through Quantum Chromodynamics (QCD) interactions is extremely
289 overwhelming. Nonetheless, it is a well known feature of QCD interactions that the partons (and jets)
290 produced tend to have a low transverse momentum (p_T). This indicates that exploring a high p_T region
291 of the phase space could be the key to suppress this background. In this kinematic regime, traditional jet
292 reconstruction techniques, that establish a one-to-one correspondence between partons and jets, begin
293 to fail because the decay products of heavy, highly boosted particle become more collimated as the p_T
294 of the mother particle increases. State of the art jet reconstruction techniques make use of jets with a
295 larger radius parameter to reconstruct both decay products as a single jet that can then be used as a
296 proxy for the original particle. In order to recover as much information as possible from these jets, it is
297 important to analyze their intrinsic structure, referred to as substructure.

298 Jet substructure techniques explore the existence of subjets inside the large- R jets. For example, a
299 jet containing the two b quarks from the decay of a Higgs boson is expected to be more consistent with
300 having two subjets than a jet produced by a QCD process. From the standpoint of detector design for
301 the FCC-hh, as well as future upgrades of existent detectors, the ability to resolve the substructure of
302 large- R jets is a key requirement that is greatly influenced by the granularity of the hadronic calorimeter
303 (HCAL).

304 The main goal of this thesis is to use boosted di-Higgs production in the four b quarks final state to
305 study the influence of the granularity of the HCAL in the significance that can be achieved.

306 Chapter 2 presents an overview of the SM. It summarizes its particle content and interactions and
307 introduces the mathematical formulation of the EWSB breaking mechanism. The successes and short-
308 comings of the SM are presented and several BSM models are introduced and their motivations dis-
309 cussed. Finally, a theoretical description of the production of Higgs pairs is provided.

310 The FCC-hh baseline accelerator and detector were highly based on LHC and its current experi-

311 ments, namely ATLAS and CMS. In chapter 3, after a brief discussion of the general features of particle
312 accelerators, we introduce the LHC and the ATLAS experiment. A discussion of jet reconstruction is
313 included. We then introduce the FCC-hh accelerator and its current baseline detector design. Some key
314 detector parameters are compared to ATLAS.

315 In chapter 4 we present the state of the art of searches for Higgs pair production at the LHC and
316 of feasibility studies targeting searches for this process at future colliders (High Luminosity LHC and
317 FCC-hh). We also review previous studies that focus on the impact of the HCAL granularity on the jet
318 mass and jet substructure observables resolution.

319 In chapter 5 we describe in detail the Monte Carlo samples that were used in this work. In addition,
320 the different detector configurations that were tested are described. Chapter 6 entails the description
321 of the analysis strategy and its optimization. A comprehensive characterization of the signal and back-
322 grounds processes is provided.

323 In chapter 7 we present and discuss the results.

³²⁴ Chapter 2

³²⁵ The standard model and beyond

³²⁶ The Standard Model (SM) is the theoretical framework that summarizes our present knowledge of particle physics. In section 2.1, we provide an overview of this model, focusing on the Higgs mechanism. In ³²⁷ section 2.2, we motivate the need to explore models beyond the SM (BSM) and introduce some of the ³²⁸ most well known BSM models. In section 2.1.1, we provide a theoretical description of the production of ³²⁹ Higgs boson pairs which is the physical process that is under study throughout this work. ³³⁰

³³¹ 2.1 The Standard Model of Particle Physics

³³² The Standard Model of particle physics summarizes our present knowledge of fundamental particles ³³³ and their interactions. It is formulated in the framework of Quantum Field Theory (QFT) and describes ³³⁴ the subatomic world in terms of fields whose excitations are the particles we can detect. The particle ³³⁵ content of the SM is summarized in figure 2.1. Each particle is represented inside a square. The electric ³³⁶ and color charge are shown in the right upper corner and the spin in the right lower corner. The mass ³³⁷ of the particles are given in electron Volt on top of each square. There are two types of fundamental ³³⁸ particles: matter particles and force carriers.

³³⁹ Matter particles are the building blocks of all the matter in our world. They come in two groups, ³⁴⁰ leptons and quarks. Quarks make up atomic nuclei and leptons, namely electrons and muons, can ³⁴¹ orbit atomic nuclei forming atoms. Quarks and leptons are fermions which means they have half-integer ³⁴² spin. There are six quarks: three of the 'up type' (up, charm and top represented by u , d and t) with ³⁴³ electric charge of $+2/3$ and three of 'down type' (down, strange and bottom represented by d , s and b) ³⁴⁴ with electric charge of $-1/3$. Similarly, we have three leptons with charge -1 (electron, muon and tau ³⁴⁵ represented by e , μ and τ) and three neutral leptons (electron, muon and tau neutrinos represented by ³⁴⁶ ν with the symbol of the corresponding charged lepton as subscript) that are, within the SM, massless. ³⁴⁷ We can classify quarks and leptons in three generations, each composed of an up type and down type ³⁴⁸ quark or of a charged lepton and the corresponding neutrino.

³⁴⁹ The force carriers, technically called gauge bosons, are particles associated with the fundamental

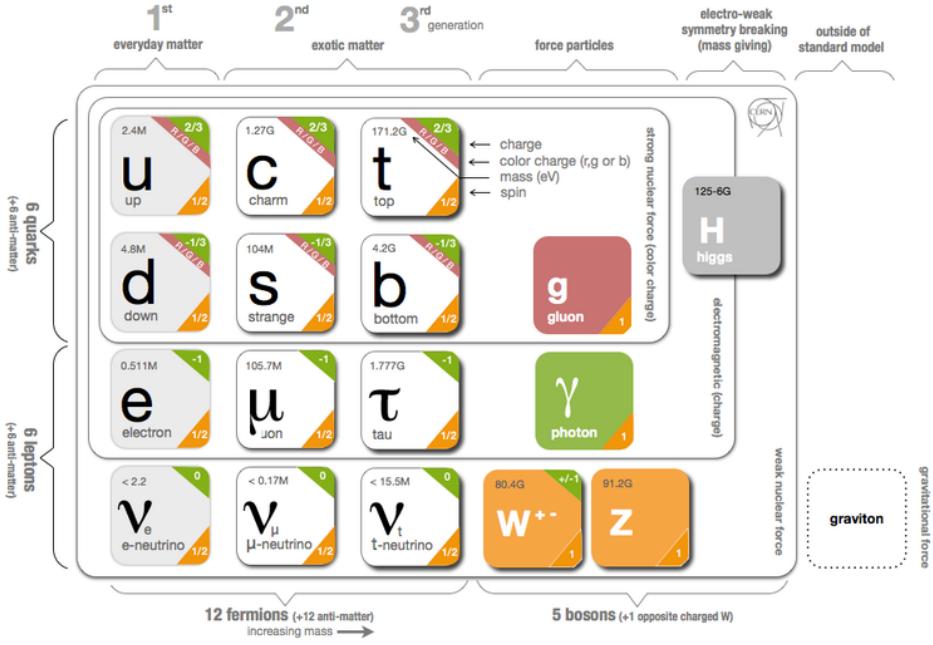


Figure 2.1: Schematic representation of the Standard Model of elementary particles. Particle are represented inside squares. The electric and color charge are shown in the right upper corner of each square. The spin is shown in the right lower corner. The mass of the particles is also shown, in electron Volt.

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.

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 theories could also be formulated for the weak and strong interactions. This is the context in which Yang-Mills theories emerged. They extend the concept of gauge theory from abelian groups, that lead to the development of QED, to non-abelian gauge groups. However, the quanta of the fields predicted by these theories must be massless in order to maintain gauge invariance. Therefore, they were set aside until the 1960's when the idea of particles acquiring mass through symmetry breaking in massless theories was put forward by Goldstone [1], Nambu and Jona-Lasinio [2]. In the following paragraphs we discuss in more detail the caveats of Yang-Mills theories and the phenomenon of Spontaneous Symmetry

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.

369 try Breaking (SSB) as the basis of the modern Higgs mechanism. We then describe this mechanism in
 370 the framework of the SM.

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

378 On the other hand, spontaneous symmetry breaking (SSB) is a phenomenon through which the
 379 invariance of a system under a certain symmetry group is destroyed [5]. The system may then be in-
 380 variant under a subgroup of the initial symmetry but the invariance under the original symmetry group is
 381 no longer present. In particle physics, this happens because the vacuum of the system (lowest energy
 382 states) does not share the symmetry of the Lagrangian. The SSB mechanism predicts the existence
 383 of scalar massless particles, the Nambu-Goldstone bosons, as a consequence of the Goldstone theo-
 384 rem [1] (the number depends on the number of generators of the original and final symmetry groups).
 385 Though, when considering this mechanism we get once again massless particles which does not seem
 386 to be a step in the right direction if we wish to describe weak interactions.

387 However, the real breakthrough occurs when we combine a theory with local gauge invariance with
 388 the mechanism of SSB. In this case the Nambu-Goldstone bosons do not appear and it is possible to
 389 give mass to the gauge bosons. This is the Higgs mechanism, proposed independently by P.W. Higgs
 390 [6], F. Englert and R. Brout [7] and by G. Guralnik, C. R. Hagen and T. Kibble [8] in 1964.

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

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

393 where the $SU_{color}(3)$ group describes the strong interactions (QCD) and the $SU_L(2) \times U_Y(1)$ group
 394 describes the electroweak interactions. Here, L stands for left and Y stands for hypercharge. In the SM
 395 the Higgs mechanism, which we now describe, is realized in the $SU_L(2) \times U_Y(1)$ group. The Lagrangian
 396 corresponding 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:

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

397 $W_{\mu\nu}^a$ and $B_{\mu\nu}$ are the field tensors, defined as a function of the gauge fields of $SU(2)$ and $U(1)$, respec-
 398 tively, W_μ^a ($a = 1, 2, 3$) and B_μ :

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

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

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

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

402 $T^a = \frac{\tau^a}{2}$ (where τ^a are the Pauli matrices) are the $SU(2)$ group generators in the fundamental representation
 403 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 zero². The values of v are determined by the minima of the potential:

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

404 For the equation on the right (for which we get $v \neq 0$) we only obtain a real value for v (which is a
 405 requirement for the VEV of a theory) if $\mu^2 < 0$. Therefore we conclude that the equation on the right
 406 corresponds to $\mu^2 \leq 0$ while the equation on the left corresponds to $\mu^2 \geq 0$. In both cases λ has to
 407 be larger than zero to guarantee that the energy is bounded from below³ because in Eq. 2.3 λ is the
 408 coefficient of the term with the highest power in ϕ and therefore determines the concavity of the potential.

409 The shapes of the Higgs potential for $\mu^2 > 0$ and $\mu^2 < 0$ are shown in Figure ?? on the left and
 410 right, respectively. For $\mu^2 > 0$ (left) we have a single minimum located at $\langle\phi\rangle = 0$. For $\mu^2 < 0$ (right)
 411 the potential has the shape of a 'Mexican hat'. There is an infinite number of minima located in a
 412 circumference centered at zero. In this case the minima occur for $\langle\phi\rangle, \langle\phi^\dagger\rangle \neq 0$. Therefore the fields
 413 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):

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

414 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
 415 consideration that $W_\mu^a T^a$ represents a sum over all values of a we get

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.

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

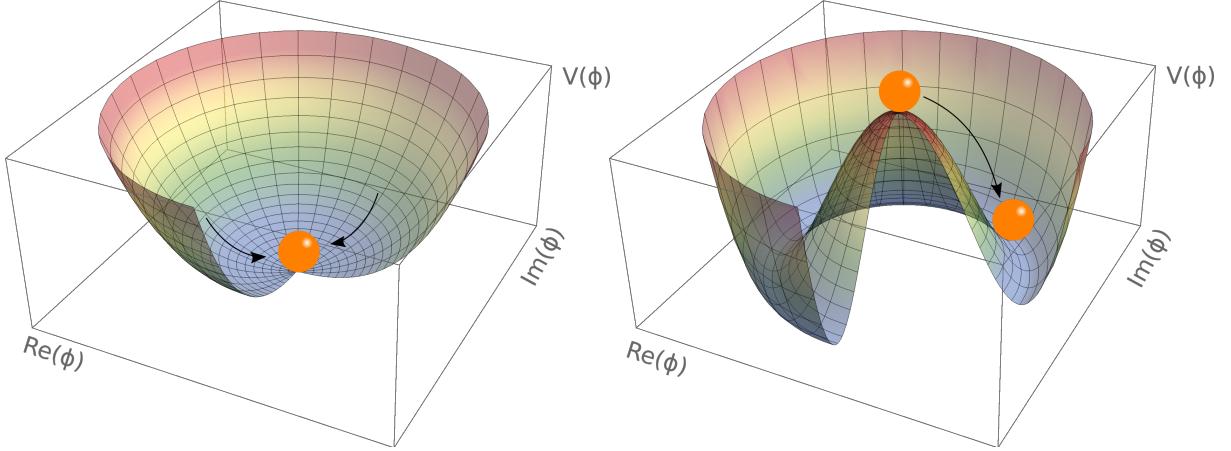


Figure 2.2: Postulated shape of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

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

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:

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

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 Eq. 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 Eq. 2.9 takes then the form

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

416 where we show only the mass terms for the gauge bosons. Note that, by construction, there is no mass
 417 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
 418 boson and Z_μ corresponds to the Z boson. We have shown that it is the fact that $v \neq 0$ that allows for
 419 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 (Eq. 2.3) using Eq. 2.8 we get, among other

terms,

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

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

In addition to being responsible for giving mass to the gauge bosons, the Higgs field is also responsible for the mass of the fermions. However, the mechanism through which this occurs is fundamentally different. In the case of the leptons the mass terms are placed explicitly in the Lagrangian:

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

where L denotes a left-handed fermion doublet and R denotes a right-handed fermion singlet. Here, left and right refer to helicity states. G_1 and G_2 are arbitrary coupling constants that can be written in terms of the fermion's mass and the VEV. ϕ is given by Eq. 2.8 and ϕ_c is given by (after the spontaneous symmetry 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, couple 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, R , are singlets, define as:

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

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

The first term in Eq. 2.14 gives mass to down type fermions (electron, muon, tau, down, strange and bottom quarks) and the second to up type fermions (up, charm and top quarks). In addition, these

436 terms give rise to the interaction terms between the Higgs field and the fermions. Take, as an example,
437 $\bar{L} = (\bar{t}, \bar{b})_L$ and $R = b_R$. For the first term of Eq. 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:

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

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

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

445 The SM has delivered extremely accurate predictions about the existence and properties of new
446 particles which make it a very successful theory. It predicted the existence of the W and Z bosons [12],
447 the gluon, the charm and top quarks and the Higgs boson [6–8]. In addition, the SM prediction for the
448 value of the anomalous magnetic dipole moment of the electron (calculated up to order α^5) agrees with
449 the measured value up to the 11th decimal place, making it the most precise measurement in science.

450 2.1.1 Higgs pair production

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

454 At the Large Hadron Collider (LHC), the main production process of Higgs pairs is gluon-gluon fusion
455 (ggF). Higgs pairs can also be produced through vector boson (V) fusion (VBF), in association with a
456 pair of top quarks ($t\bar{t}h$) or through Higgs strahlung (Vh) ⁴. At a center of mass (CM) energy of $\sqrt{s} = 13$
457 TeV, the ggF production process has a cross section approximately seventeen times larger than the next
458 most common production process which is VBF (approximately 30 fb *versus* 1.6 fb [13]). Therefore it
459 is the dominant contribution when we study inclusive production. For this reason we focus the following
460 discussion on this production mode. The leading order Feynman diagrams for Higgs pair production via
461 ggF are shown in figure 2.3.

462 The diagram on the left has an off-shell (virtual) Higgs boson, h^* , that couples to gluons by the usual
463 heavy quark triangle (same mechanism as in single Higgs production). h^* then decays to two on-shell
464 Higgs bosons. This diagram contains the three point interaction between Higgs bosons and therefore it
465 is the one that allows us to probe this coupling. In the diagram on the right, the two Higgs bosons couple

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

466 to the gluons by a box of heavy quarks and are directly radiated from a quark. The largest contributions
467 for these quantum loops come from heavy quarks, such as the top and bottom, because the coupling
468 constant of the Higgs boson to fermions is directly proportional to the fermions mass (see section 2.1).

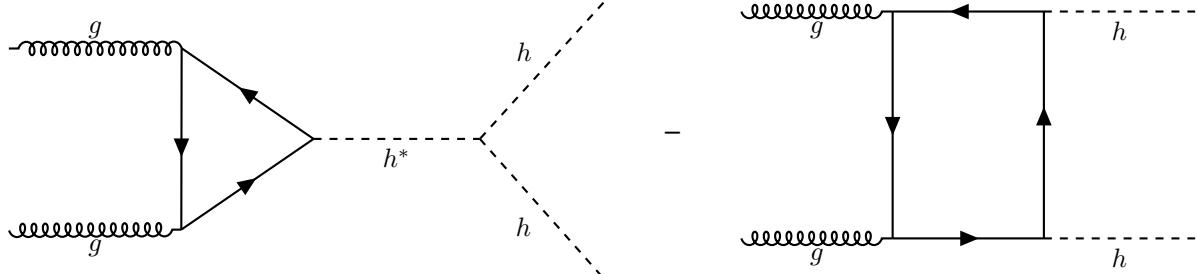


Figure 2.3: Feynman diagrams of Higgs pair production from gluon fusion. Triple vertex diagram (a) on the left and box diagram (b) on the right. The minus sign between the diagrams indicates that they interfere destructively.

The amplitudes for the box, \mathcal{M}_\square , and triangle, \mathcal{M}_Δ , diagrams scale as [14]:

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

469 where \hat{s} is the CM energy, y_t is the Yukawa coupling of the top quark, α_s is the electroweak coupling
470 constant, λ_{hhh} is the coupling constant of the Higgs boson three point self-interaction and m_h, m_t are
471 the masses of the Higgs boson and top quark.

472 At a CM energy of $\sqrt{s} = 13$ TeV the cross section for Higgs pair production, as predicted by NLO
473 calculations, is very small, approximately 30 fb [13]. It is suppressed due to the destructive interference
474 between the LO diagrams that leads to a $\sim 50\%$ suppression of the total cross section [14]. Furthermore,
475 the cross section of the triangle diagram is smaller than the one of the box diagram, approximately 4 fb
476 compared to 30 fb ⁵, and it is strongly suppressed for larger values of the CM energy which can be seen
477 directly from the expression of the amplitude in Eq. 2.20. This means that the Higgs trilinear coupling
478 mostly affects the Higgs pair production at threshold, in particular, the m_{hh} distribution. The tail of this
479 distribution (high invariant mass of the Higgs pair), however, is mostly determined by the box diagram
480 contribution [14].

The LO calculation for the cross section of Higgs pair production has been performed, for example, in [15]. 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 [16]:

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

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.

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

484 This result shows that the introduction of NLO contributions produces a significantly different result.
 485 Therefore the inclusion of such effects is necessary if we wish to obtain an accurate result that can be
 486 compared to experimental values. The analytical expressions for the NLO cross section are long and
 487 complex so we abstain from reproducing them here. Nonetheless, the LO cross section can be written in
 488 a compact form and it allows us to discuss some key features of the process. Therefore, we will present
 489 it here, based on [15] and [17].

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

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

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 $\beta_h^2 = 1 - 4\frac{m_h^2}{\hat{s}}$ and m_h is the mass of the Higgs boson [17]. F_Δ , F_\square and G_\square are form factors whose full expressions can be found, for example, in [15]. C_Δ and C_\square can 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 m_Z is the mass of the Z boson. 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

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

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

496 Secondly, although this is not evident from Eq. 2.25, the cross section for di-Higgs production in-
 497 creases with \hat{s} . This can be seen in figure 2.4 that shows the variation of the total (integrated) NLO cross
 498 section with the CM energy for the six largest production channels. Note that increasing the CM energy
 499 from 13 to 100 TeV increases the inclusive cross section by approximately two orders of magnitude which

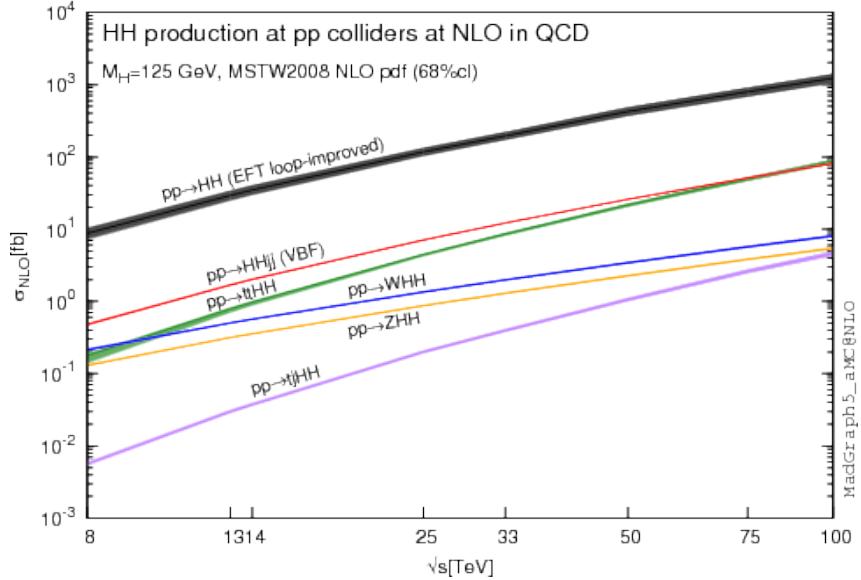


Figure 2.4: Total cross sections at the NLO in QCD for the six largest HH production channels at pp colliders. The thickness of the lines corresponds to the scale and PDF uncertainties added linearly.

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

2.2 Going beyond

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

On the one hand, there are several pieces of experimental evidence that the SM cannot explain. These include the nature of dark matter, postulated to explain the experimental observations of the velocity of far away galaxies [18], the asymmetry between matter and anti-matter in the present Universe ⁶ and the fact the neutrinos oscillate between flavors which implies that they have a non-zero mass. This phenomenon was measured independently by two collaborations, the Sudbury Neutrino Observatory (SNO) and the Super-Kamiokande, in 1998 and 2001-2002, respectively, [19–21].

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

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

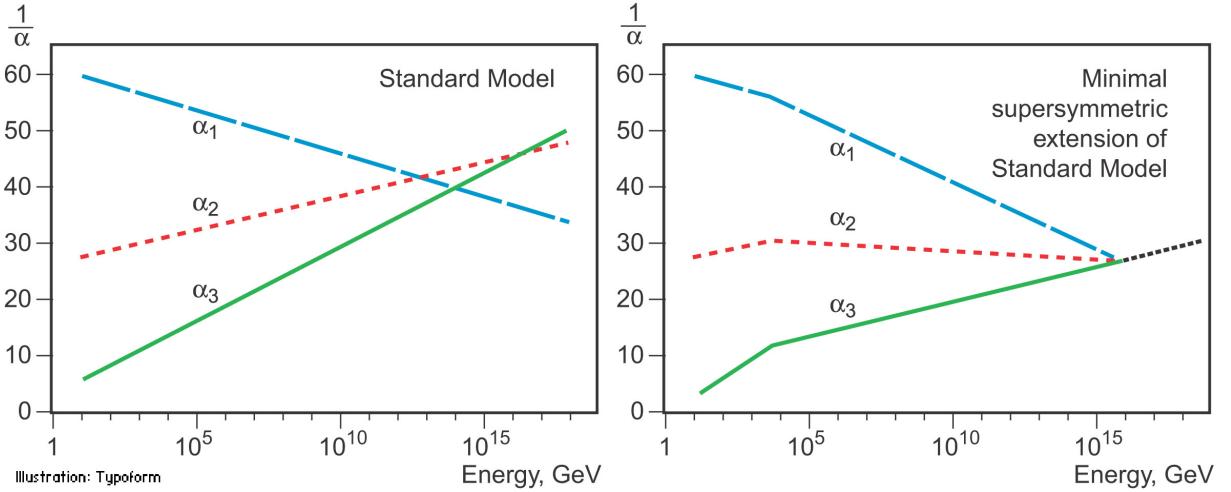


Figure 2.5: Here, α_1 , α_2 and α_3 represent the coupling constants of $U(1)$, $SU(2)$ and $SU(3)$, respectively. What is shown is the variation of the inverse of the coupling with the energy: on the left for the SM and on the right for the MSSM. Plots from [23].

When faced with these weaknesses, or rather hints of incompleteness, the theoretical community put a great effort into the development of models that add new ingredients to the SM but that simultaneous lead to predictions that are compatible with existent experimental measurements. In the following paragraphs we will introduce and briefly describe some of the most well studied (both theoretically and experimentally) BSM models. We follow the discussion presented in [22] as a starting point.

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

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

An early proposal of a theory that could unify gravity with electromagnetism was given by Theodore Kaluza in 1921. In particular, he showed that these two forces could stem from a single tensor with the introduction of an extra space dimension. In 1926, Oscar Klein offered an explanation for this extra dimension; he proposed that it had a circular topology such that at each point of the four dimensional

space-time we would have a circle with a small radius. This theory has more degrees of freedom (because it is formulated in a higher dimensional space-time) and therefore it predicts new particles that are usually known as Kaluza-Klein gravitons (and their excited states). Nonetheless, the Kaluza-Klein does not provide a satisfactory explanation for the hierarchy problem. Therefore, in 1999, Lisa Randall and Raman Sundrum introduced a new model that does. This model introduces only two new particles: a spin 2 graviton (and its Kaluza-Klein excitations) and a radion, that is a spin 0 neutral particle.

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

The most general renormalizable 2HDM scalar potential is written as [24]

$$\begin{aligned} V = & m_{11}^2 |\Phi_1|^2 + m_{22}^2 |\Phi_2|^2 - (m_{12}^2 \Phi_1^\dagger \Phi_2 + h.c.) \\ & + \frac{1}{2} \lambda_1 |\Phi_1|^4 + \frac{1}{2} \lambda_2 |\Phi_2|^4 + \lambda_3 |\Phi_1|^2 |\Phi_2|^2 + \lambda_4 |\Phi_1^\dagger \Phi_2|^2 \\ & + \left[\frac{1}{2} \lambda_5 (\Phi_1^\dagger \Phi_2)^2 + \lambda_6 |\Phi_1|^2 (\Phi_1^\dagger \Phi_2)^2 + \lambda_7 |\Phi_2|^2 (\Phi_1^\dagger \Phi_2)^2 + h.c. \right], \end{aligned} \quad (2.26)$$

where Φ_1 and Φ_2 are hypercharge doublets and the coefficients m_{12}^2 and $\lambda_{5,6,7}$ can be complex. However, when including the Yukawa interactions, the most general lagrangian leads to tree-level flavor changing neutral currents (FCNC) in the Yukawa sector. These FCNC are very tightly constrained by experimental data and should be avoided. They can be eliminated by imposing a \mathbb{Z}_2 symmetry. Although, usually, this symmetry is allowed to be softly broken in order to allow the theory to have a decoupling limit [25] where the mass of all the scalars other than the SM-like one can be made very large. In addition to the softly-broken \mathbb{Z}_2 symmetry, which leads to $\lambda_{6,7} = 0$, we also impose CP conservation which makes all possible complex phases vanish. In this case, we obtain five Higgs bosons that are CP eigenstates. Three of them are neutral, h , H and A , and the other two are charged, H^\pm . h and H are CP-even states while A is CP-odd. h is usually taken to be the SM Higgs boson and its mass is set to 125 GeV.

There are several types of 2HDM classified according to their fermion-scalar interactions. We highlight the type II (the one used in this work), where all right-handed up-type quarks couple to Φ_2 and right-handed down-type quarks and charged leptons couple to Φ_1 . This type of couplings is analogous to what happens in SUSY models.

Instead of the parameters in Eq. 2.26, we can describe the model in terms of the four physical masses, m_h , m_H , m_A and m_{H^\pm} , the angles α and β , the VEV $v = 246$ GeV and a further parameter, chosen to be m_{12}^2 [24]. The quartic couplings of the potential can then be written in terms of these parameters, as can be found in Eq. 11 of [24].

Simplified dark matter (DM) models are based on the exchange of a single particle between DM and SM particles and try to explain the nature of DM and how it interacts with the SM. The particle exchanged is called a mediator and, depending on the specific model, it can be neutral or electrically charged and

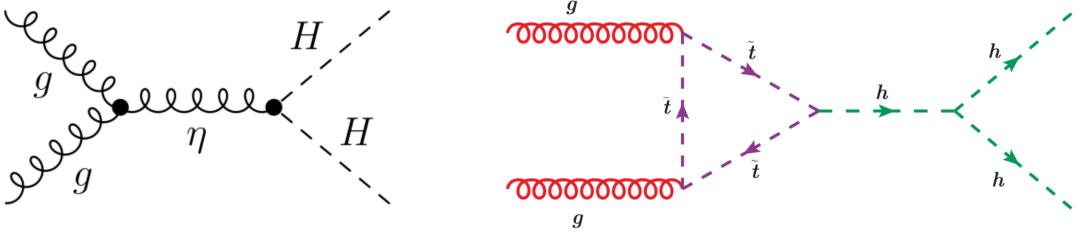


Figure 2.6: BSM contributions to Higgs pair production. On the left (figure from [26]) η represent the spin-2 Kaluza-Klein graviton that couples to a pair of SM Higgs bosons, H , at tree level. On the right (figure from [27]) \tilde{t} is the supersymmetric partner of the top quark that gives a contribution to the quantum loop in the Higgs pair production s-channel diagram.

579 have spin 0, 1 or 2. The simplest possible scenario is a neutral scalar mediator.

580 The CP-conserving 2HDM and dark matter model with a spin 0 mediator are explored in this work
 581 as sources of alternative di-Higgs production processes. For these models, the main modification with
 582 respect to the SM occurs because new heavy particles, namely, H and the DM mediator can couple
 583 to the Higgs bosons through the s-channel diagram. This corresponds to replacing the off-shell Higgs
 584 boson by one of these particles in the Feynman diagram on the left in figure 2.3.

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

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

601 Chapter 3

602 Collider experiments

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

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

611 In section 3.4 we shift the focus to future collider experiments and accelerators and motivate their
612 need. In sections 3.4.1 and 3.4.2 we introduce the hadronic Future Circular Collider (FCC-hh) and
613 describe the baseline detector design, respectively.

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

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

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

633 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:

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

634 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
635 characterize the transverse beam size in the horizontal and vertical directions.

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

639 While the cross sections of most physics processes increase when the CM energy goes from 13 to
640 100 TeV, many BSM models predict new processes, or new contributions to existing processes, whose
641 cross sections increase more rapidly than the SM backgrounds. In addition, by conservation of energy,
642 a larger CM energy implies that particles with larger mass can be created. Based on Eq. 3.2, we can
643 tune its parameters to obtain the highest possible luminosity. Nonetheless, because we are dealing with
644 charged particles, there is a limit on how close the bunches can be and on how many protons we can
645 pack in a bunch. Moreover, the beam's transverse dimensions cannot be infinitely reduced. A smarter
646 way to increase the number of collisions that an accelerator can produce is to run for a longer time,
647 therefore increasing the integration time in Eq. 3.1.

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

650 3.2 The Large Hadron Collider

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

655 science.

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

662 The acceleration of charged particles at the LHC is based on radio frequency (RF) cavities. These
663 cavities are shaped to sustain a resonant electromagnetic field that oscillates at a frequency of 400 MHz.
664 During the acceleration stage, charged particles passing through the cavities feel an overall force that
665 propels them forward. When the LHC is running at full energy, a perfectly timed proton with exactly the
666 right energy feels a zero net force when passing the cavities. Protons with a slightly different energies
667 arriving slightly earlier or later are decelerated or accelerated in order to keep the beam sorted in discrete
668 packages with the same energy. These are called bunches. There are 2808 bunches circulating at the
669 same time, each containing approximately 10^{11} protons. The bunches are spaced by 25 ns. Furthermore,
670 the successful operation of the LHC also relies on superconducting magnets made of Niobium-Titanium
671 filaments chilled to -271.3° C and on an ultra high vacuum (of the order of $10^{-10} - 10^{-11}$ mbar) inside
672 the beam pipes. The magnets are placed along the LHC ring and produce dipole and quadrupole
673 electromagnetic fields. The dipole magnets create a nominal field of 8.3 T and bend the beam along the
674 tunnel. The quadrupole magnets focus the beam at the interaction points. The ultra high vacuum greatly
675 reduces the probability that the beam interacts with any particle. It is crucial to keep a stable beam to
676 continuously maintain collisions during long runs.

677 One of the main research goals of the LHC was to discover the Higgs boson. This was achieved
678 in 2012 when ATLAS and CMS reported the discovery of a particle consistent with the boson predicted
679 by the Higgs mechanism, with a mass of 125 GeV [28],[29]. Ever since, efforts have been directed
680 to measuring its mass, couplings, spin-parity properties with increasing precision using different decay
681 channels and production modes.

682 ATLAS and CMS reported the observation (measurement with a significance greater than five sigma)
683 of the Higgs decaying to $b\bar{b}$ [30, 31]. The searches targeted the VH production mode. It offers the best
684 sensitivity to the $hb\bar{b}$ Yukawa coupling because requiring a vector boson helps reduce the SM back-
685 grounds, namely the ones from QCD interactions. The observation of $h \rightarrow b\bar{b}$ means we have now
686 observed all of the Higgs boson's decay modes predicted by the SM. CMS reported the first observation
687 of the Higgs boson decaying to a pair of tau leptons [32]. In addition, the observation of Higgs boson
688 production in association with a $t\bar{t}$ pair was very recently reported by both collaborations [33, 34]. More-
689 over, precision measurements of the masses of the Higgs [35, 36] and W [37] bosons and of the top
690 quark [38, 39] were also performed. So far, no conclusive signs of new physics were seen at the LHC.

691 Future prospects for the LHC include its upgrade to the High Luminosity-LHC (HL-LHC) after the
692 scheduled long shutdown of 2024-2026. This upgrade will increase the size of the dataset to 3000 fb^{-1}
693 over the course of ten years [40]. During the shutdown, the ATLAS detector will be upgraded.

694 In the Higgs sector, the high value of the integrated luminosity will improve the statistical precision of
695 already measured channels and the discovery potential of rare processes [41].

696 **3.2.1 The ATLAS detector**

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

701 A combination of cartesian and cylindrical coordinates is used to describe the detector. In both
702 cases, the origin is defined to coincide with the interaction point. The Cartesian system is right-handed
703 and the z axis is defined to be the direction of the beam. The x-axis points from the interaction point
704 to the center of the LHC ring and the y-axis points upwards. The azimuthal angle, ϕ , is measured
705 around the beam axis and the polar angle, θ , from the beam line. The pseudorapidity is defined as
706 $\eta = -\ln \tan(\theta/2)$. Another commonly used quantity is the rapidity, y , defined as a function of a particle's
707 energy, E , and longitudinal 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
708 negligible with respect to its momentum the pseudorapidity converges to the definition of rapidity. In
709 addition, the angular distance between two points, ΔR , is defined as $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where
710 rapidity can also be used instead of the pseudorapidity.

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

715 The ID covers the pseudorapidity range $|\eta| < 2.5$ and it makes up the innermost layer of the detector.
716 It consists of silicon pixel, silicon micro-strip, and straw tube transition radiation tracking detectors. It is
717 contained in a solenoid magnet with a central field of 2 T. The tracker provides precision measurements
718 of the positions and momenta of charged particles. As a charged particle transverses the several layers
719 of the ID it ionizes the medium creating electrical signals that can be read out. These individual electrical
720 signals are then combined to reconstruct the trajectory of the particle.

721 Lead/Liquid-Argon (LAr) sampling electromagnetic (EM) calorimeters cover the pseudorapidity range
722 $|\eta| < 3.2$. The EM calorimeter has an accordion like structure with layers of showering material (lead)
723 interleaved with layers of active material (liquid argon). These calorimeters provide measurements of
724 the energy of electrons and photons. The interaction of these particles with the lead layers induces the
725 production of an EM shower whose energy is measured in the liquid argon layers. The granularity of the
726 EM calorimeter strongly depends on the longitudinal layer and on the pseudorapidity region.

727 The hadronic calorimetry in the pseudorapidity range $|\eta| < 1.7$ is provided by a scintillator-tile
728 calorimeter (TileCal) which is divided in a central barrel and two smaller end-cap barrels, one on each
729 side of the central barrel. The active components are scintillator tiles made of polystyrene that are in-
730 terleaved with steel plates as the passive material. The scintillation light emitted by the tiles when an

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).

731 ionising particle crosses the calorimeter is collected on both ends of the tiles by wavelength-shifting
 732 optical fibers. The light signal emitted is proportional to the particle's energy. The TileCal is composed of
 733 several cells with transverse segmentation $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ¹ [42].

734 For $|\eta| > 1.5$ LAr calorimeters extend the pseudorapidity range to $|\eta| = 4.9$. The LAr calorimeter is
 735 divided in end-cap and forward. These cover the pseudorapidity ranges $1.5 < |\eta| < 3.2$ and $3.2 < |\eta| <$
 736 4.9, respectively. The active material is liquid-argon and the absorbers are copper and tungsten for the
 737 end-cap and forward calorimeters, respectively. In the end-cap LAr calorimeters the segmentation is
 738 $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $1.5 < |\eta| < 2.5$ and 0.2×0.2 for $2.5 < |\eta| < 3.2$. In the forward LAr calorimeter
 739 the segmentation is $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$.

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

741 The hadronic calorimeters provide measurements of the energy of hadrons, jets, τ leptons and miss-
 742 ing transverse energy (E_T^{miss}). Approximately one third of the energy of jets is deposited in this layer.
 743 In the TileCal, the jet energy resolution is given by $\sigma/E \sim 50\%/\sqrt{E} + 3\%$ [42], where the first term is the
 744 stochastic term that derives from sampling fluctuations and follows a Poisson distribution and the second
 745 term is a constant that depends on the characteristics of the calorimeter. For the LAr calorimeter, the jet
 746 energy resolution is given by $\sigma/E \sim 60\%/\sqrt{E} + 2\%$ [43].

747 The MS is the outermost layer of the detector and it is dedicated to detecting muons that travel
 748 through the previous layers almost without interacting. This layer provides measurements of the muons
 749 transverse momenta. It is composed of Monitored Drift Tubes (MDT) and Cathode Strip Chambers
 750 (CSC) that provide high precision measurements of the muons' momentum in the pseudorapidity range
 751 $|\eta| < 2.7$ and of Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) dedicated to triggering
 752 purposes for $|\eta| < 2.4$.

753 **b-Tagging**

754 Each collision produces a large number of hadronic jets (we refer to section 3.3 for a detailed description
 755 of jets and how they are reconstructed). For this work, jets initiated by a b quark (b jets) are particularly
 756 important: we are looking for a Higgs pair decaying to four b quarks which leads to an experimental

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.

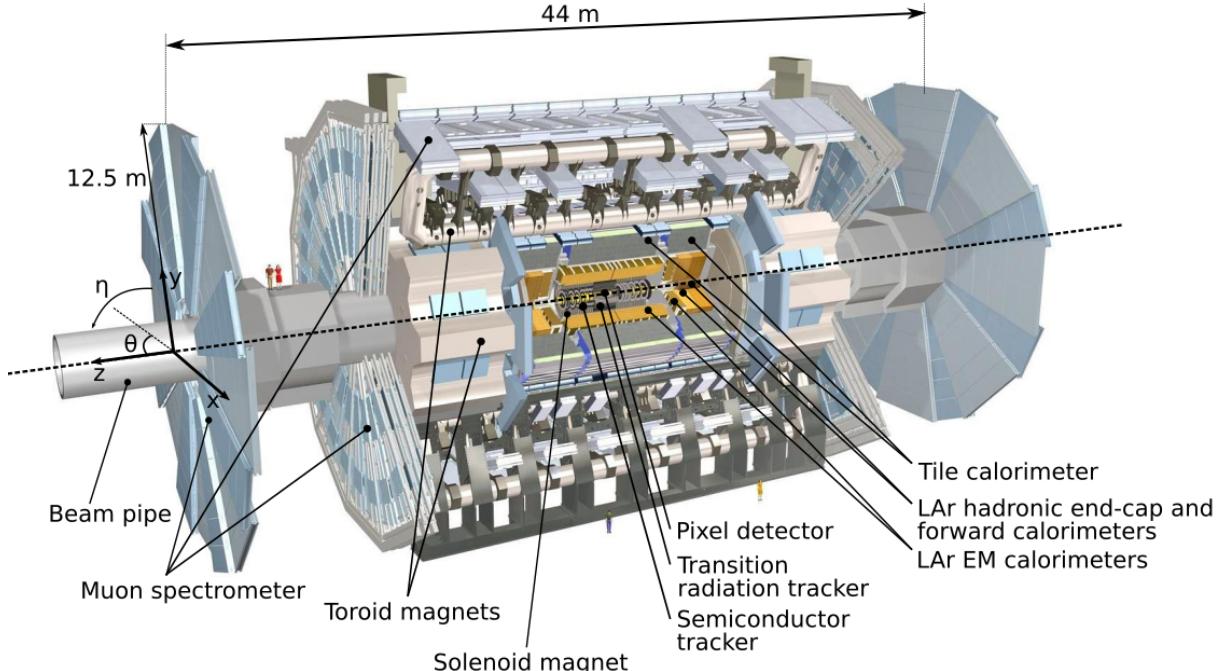


Figure 3.1: ATLAS detector.

757 signature that consists of four b jets. b-Tagging algorithms determine, with a given probability, if a jet
 758 was originated by a b quark.

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

767 In ATLAS, b-tagging algorithms are applied to the sub-set of tracks that are associated with a given
 768 jet. The matching between tracks and calorimeter-based jets is performed using the ghost association
 769 technique [44]². The identification of b-jets in ATLAS is based on distinct strategies encoded on three
 770 b-tagging algorithms: impact parameter-based algorithms, an inclusive secondary vertex reconstruction
 771 algorithm and a decay chain multi-vertex reconstruction algorithm. The output of these algorithms are
 772 combined in a multivariate discriminant based on a Boosted Decision Tree (BDT) which provides the
 773 best discrimination between the different jet flavors [45].

774 The impact parameter-based algorithms [46], IP2D and IP3D, use as discriminant variables the trans-
 775 verse impact parameter significance and the transverse and longitudinal impact parameter significance,
 776 respectively. The secondary vertex finding algorithm [46], SV, explicitly reconstructs a displaced sec-

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.

777 secondary vertex inside the jet by trying to find pairs of tracks with a common origin. The decay chain
778 multi-vertex reconstruction algorithm [47], JetFitter, tries to reconstruct the full b-hadron decay chain.
779 This approach allows to resolve b- and c-hadrons vertices even if there are not two tracks associated
780 with them.

781 **Trigger and data acquisition**

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

787 The trigger is responsible for selecting events with interesting experimental signatures. The trigger
788 system in Run-2 consists of a hardware-based first level trigger (Level-1) and a software-based high
789 level trigger (HLT). The Level-1 trigger takes as input coarse granularity calorimeter and muon detector
790 information and reduces the event rate to 100 kHz. The HLT uses full granularity detector information
791 and reduces the rate to approximately 1 kHz [48].

792 **3.3 Jet reconstruction**

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

799 At a particle detector we are interested in reconstructing jets. These are the objects that are used in
800 the physics analysis. Working with jets instead of hadrons is an advantage because it greatly reduces the
801 number of objects we need to analyze per event. In addition, they work as a *proxy* for the fundamental
802 partons produced in the event.

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

808 The most widely used jet finding algorithms are the k_T [49], the Cambridge/Aachen (C/A) [50] and
809 the anti- k_T [51]. They are based on distance measurements:

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

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

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

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

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

830 Boosted kinematic regime

831 With the increase of the CM energy of particle colliders the production of particles with a transverse
 832 momentum much larger than their mass became a reality. In this kinematic regime (referred to as boosted
 833 due to the high Lorentz boost of the particles) traditional jet reconstruction algorithms, that rely on a one-
 834 to-one correspondence between jets and partons, begin to fail [52].

Due to the high Lorentz boosts, decay products of heavy resonances get more collimated. The angular separation of the decay products is approximately [53]:

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

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

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

843 Higgs boson is already similar to the default jet diameter, which can jeopardize the ability to resolve the
844 individual decay products.

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

851 **3.3.1 Jet properties and substructure observables**

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

858 Heavy resonances decaying to two(three) particles will produce large- R jets that are consistent with
859 the existence of two(three) energy clusters. Examples of such topologies are the Higgs boson and top
860 quark decays: the Higgs boson always decays to pairs of particles (leptons, quarks or bosons) and the
861 top quark decays, with a probability close to 100%, to a b quark and a W boson that then decays to a
862 pair of leptons or quarks. These topologies are usually referred to as two or three prong. In contrast, jets
863 initiated by a gluon or quark splitting are not expected to have a meaningful substructure. The energy
864 is expected to be concentrated around the jet axis following an isotropic distribution and to become less
865 dense as we approach the jet's border. This is the signature of a one-prong topology. The previous
866 discussion is valid at LO and captures the generic features of jets that are targeted by jet substructure
867 techniques. Nonetheless, other effects may come into play. Take, for example, a highly virtual gluon with
868 a high p_T . If it splits into two quarks the resulting jet may have two subjets and thus mimic the topology
869 of a heavy resonance decay.

The N-subjetiness variable [54], τ_N , may be used to identify jets compatible with N subjets. It is given
by

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

870 where the index k runs over all particles in a jet, the indexes 1 to N identify the number of axis inside the
871 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
872 (the ones we are most interested in) are clustered around the axis (because ΔR will be small) and will
873 have a larger value otherwise. A jet with small τ_N is considered to be consistent with having N or fewer
874 subjets because all its constituents are aligned with the axis.

875 We usually use ratios of τ_N variables ($\tau_{MN} = \tau_M / \tau_N$). Of particular interest for this work are the τ_{21} ,
876 τ_{31} and τ_{32} ratios. These observables can take values between zero and one. For τ_{21} , for example, a

877 small value (close to zero) indicates that the jet is more compatible with two subjets than with one and
 878 therefore it can help discriminate between two-prong and one-prong jets. The same applies to every
 879 other ratio.

The Fox-Wolfram moments, H_l , can be used to identify jets that have a structure of two back-to-back subjets in their rest frames [55]. They are given by:

$$H_l = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{E^2} P_l(\cos(\theta_{i,j})) \quad (3.6)$$

880 where θ_{ij} is the angle between energy clusters i and j , E is the total energy of the clusters in the jet rest
 881 frame and $P_l(x)$ are the Legendre polynomials. For a jet that has a structure of two back-to-back subjets
 882 in its rest frame, $H_1 = 0$, $H_l \sim 1$ for even l , and $H_l \sim 0$ for odd l . This is what we expect for jets coming
 883 from the decay of a boosted resonance. In this work we make use of H_2 .

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:

$$M = \left(\sum_i E_i \right)^2 - \left(\sum_i \vec{p}_i \right)^2 \quad (3.7)$$

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

888 3.3.2 Jet grooming algorithms

889 The main goal of jet grooming algorithms is to remove contamination of softer jet constituents from pileup
 890 or underlying event and to leave behind the hard substructure. The main advantage of such algorithms is
 891 that they improve the mass resolution of jets. These features are of particular interest in high luminosity
 892 environments such as the HL-LHC and future high energy colliders.

893 There are three main jet grooming algorithms: trimming, pruning and mass drop filtering. In this
 894 work we do not use pruning or trimming, although these techniques might be worth exploring in future,
 895 more comprehensive, studies. In particular, they can be useful to help reject pileup contributions. The
 896 mass drop filtering procedure isolates relatively symmetric subjets within a jet, each with a significantly
 897 smaller mass than the original jet [53]. This technique was developed and optimized using C/A jets for
 898 the search of Higgs decaying to $b\bar{b}$ pairs [56]. It works as follows:

- The last step of the C/A clustering is undone such that the jet is split in to 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} : $m_{j_1}/m_{jet} < \mu_{frac}$, where μ_{frac} is a parameter of the algorithm. In addition the splitting

is required to be relatively symmetric:

$$\frac{\min[(p_T^{j_1})^2, (p_T^{j_2})^2]}{(m_{\text{jet}})^2} \times \Delta R_{j_1, j_2}^2 > y_{\text{cut}} \quad (3.8)$$

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

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

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

3.4 Future Colliders

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

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

3.4.1 The hadronic Future Circular Collider

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

The FCC-hh baseline design consists of a proton-proton circular collider with a maximum CM energy of 100 TeV housed by a 100 km tunnel in the area of Geneva. This machine will extend the research program of the LHC (and of the HL-LHC) after these have reached their full discovery potential, by around 2040. In addition, it will allow for the exploration of an entirely new kinematic regime, probing energy scales where new physics may come into play. A possible way of defining the target luminosity of

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.

930 this machine is to require that within the first year of operation it surpasses the exploration potential of the
 931 LHC [57]. Comprehensive studies [57, 58] indicate that this can be achieved with an integrated luminosity
 932 of the order of 10 ab^{-1} per experiment. Considering a reasonable operation period of 10 years this leads
 933 to integrated luminosity per experiment of the order of 1 ab^{-1} per year. [LUMINOSITY TARGET?]

934 The FCC-hh is expected to work with a dipole field of 16 T and to provide a peak instantaneous lumi-
 935 nosity thirty times larger than the LHC. The number of bunches is expected to be almost a factor of four
 936 larger than for the LHC and the number of events per bunch crossing is expected to be approximately
 937 1000. The latter brings a lot of technical challenges because it means that the mean pileup expected
 938 for the FCC-hh is almost 40 times larger than for the LHC. This requires the development (or improve-
 939 ment of already existing techniques) of techniques and algorithms that allow us to further reject pileup
 940 contributions.

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

942 3.4.2 FCC-hh baseline detector

943 The design of the FCC-hh baseline detector, which we describe in detail in this section, has been greatly
 944 based on that of the ATLAS and CMS experiments, in particular the central barrel. The layers and sub
 945 detectors are arranged in the same order and perform very similar roles. The geometry is cylindrical
 946 and therefore we can use exactly the same coordinate system that was introduced in section 3.2.1. The
 947 dimensions are very close to the ones of ATLAS: 25 meters in height (diameter) and 48 m in length [59].
 948 A schematic representation of the FCC detector is shown in figure 3.2.

949 The detector consists of trackers, EM and hadronic calorimeters and MS. The magnet configuration
 950 consists of three solenoid magnets (one central-barrel and two forward) that surround the central bar-
 951 rel calorimeters and the forward trackers. The center solenoid delivers a magnetic field of 4 T at the
 952 interaction point [60].

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

957 The EM calorimeter [61, 62] covers the pseudorapidity range $|\eta| < 6$. It is divided in barrel, end-cap

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.

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 steal 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 [63] 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.

3.4.3 FCC-hh physics program

The FCC-hh physics program is vast and diverse. Within the SM, it includes the study of gauge bosons pair production and heavy flavor production, the measurement of the top quark's properties and the study of the EWSB mechanism via multi-Higgs production. In addition, heavy ions collisions will allow for a deeper understanding of the Quark-Gluon Plasma. From a BSM standpoint, searches for supersymmetric and dark matter particles in new kinematic regimes can be pursued. A comprehensive review of the physics potential of the FCC-hh can be found here [14]. This document collects the results of the many studies that have been carried out since the beginning of the FCC initiative, in 2014.

Regarding Higgs pair production, a maximum precision on the SM cross section of 3% is expected to be achieved using the $b\bar{b}\gamma\gamma$ final state. This would allow to constraint the Higgs triple coupling to be $\lambda_{hhh} \in [0.97, 1.03]$. The $hh \rightarrow b\bar{b}b\bar{b}$ would allow for a 5% precision on the SM cross section and to constraint the triple coupling to be $\lambda_{hhh} \in [0.9, 1.5]$. In spite of the larger background yield, the $b\bar{b}b\bar{b}$ channel provides a reasonable number of events in the tail of the m_{hh} distribution. As discussed in

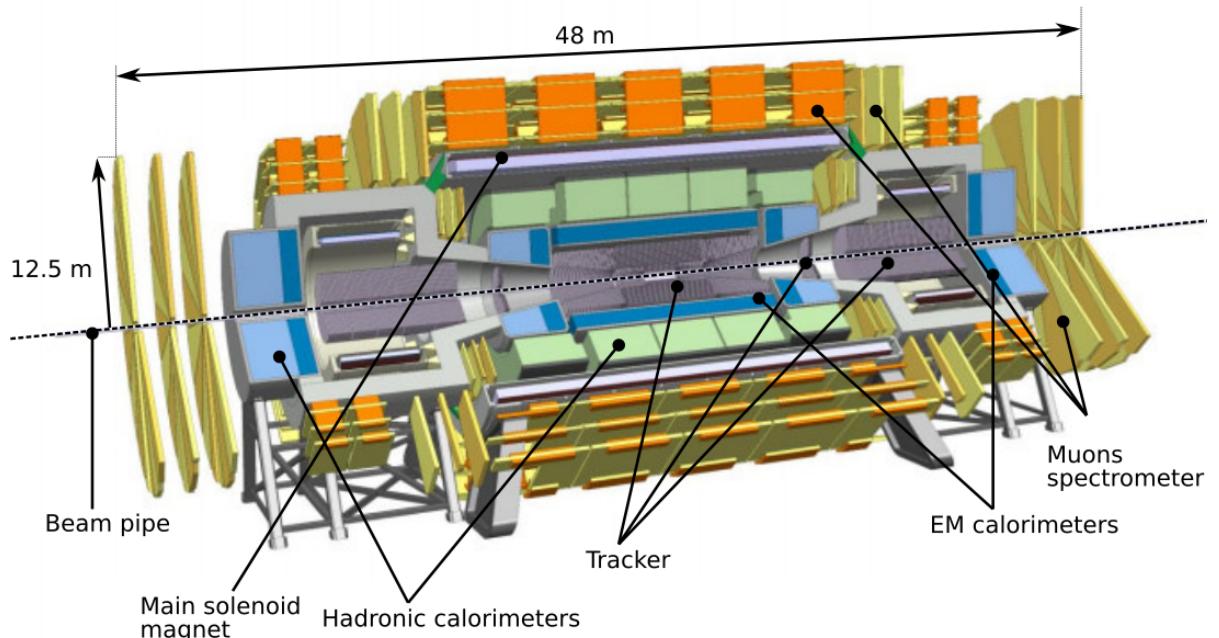


Figure 3.2: FCC detector baseline concept.

988 section 2.1.1, the tail of this distribution does not have a large contribution from the triangle diagram.
 989 Therefore, the sensitivity to the Higgs triple coupling is expected to be smaller. However, the high
 990 energy regime can be more sensitive to new physics contributions which makes the $b\bar{b}b\bar{b}$ channel a very
 991 interesting one.

992 The Higgs quartic coupling could be probed through triple Higgs production. In this case the most
 993 promising final state seems to be $b\bar{b}b\bar{b}\gamma\gamma$. This channel could constrain the Higgs quartic coupling to be
 994 $\lambda_{hhh} \in [-4, 16]$.

⁹⁹⁵ Chapter 4

⁹⁹⁶ State of the art

⁹⁹⁷ In this chapter we present the state of the art of searches for Higgs pairs production at the LHC and
⁹⁹⁸ feasibility studies targeting searches for this process at future colliders. In addition, previous studies
⁹⁹⁹ focusing on the granularity of the hadronic calorimeter as a key detector parameter are also reviewed.

¹⁰⁰⁰ In section 4.1, we review the searches that have been conducted at the LHC by the ATLAS and CMS
¹⁰⁰¹ experiments. We include discussions on the different final states that were targeted and report on the
¹⁰⁰² constraints that were derived for the di-Higgs cross section and trilinear coupling. A brief overview of
¹⁰⁰³ the current constraints on some BSM models is also presented. In section 4.2, we present the results
¹⁰⁰⁴ obtained from feasibility studies that access the discovery potential for this process at the HL-LHC and
¹⁰⁰⁵ at the FCC-hh. In section 4.3, we present previous studies on the impact of the granularity of the
¹⁰⁰⁶ hadronic calorimeter on the spatial resolution of hadrons and on the resolution of the jet mass and of jet
¹⁰⁰⁷ substructure observables.

¹⁰⁰⁸ 4.1 Searches for Higgs pair production at the LHC

¹⁰⁰⁹ The discovery of the Higgs boson is a strong evidence that the Higgs mechanism operates in Nature.
¹⁰¹⁰ However, by itself, it does not guarantee that the shape of the Higgs potential is the one depicted in
¹⁰¹¹ figure ??, on the right. In order to reconstruct the Higgs potential and gain a deeper understanding of
¹⁰¹² the mechanism that leads to the breaking of the electroweak symmetry one must measure the Higgs
¹⁰¹³ boson self-couplings, namely its three and four point interactions, whose strengths depend on the values
¹⁰¹⁴ of the parameters of the Higgs potential, as shown in section 2.13.

¹⁰¹⁵ However, in the SM, the cross section for the production of Higgs pairs through ggF is extremely
¹⁰¹⁶ small: ~ 30 fb at the current CM energy achieved at the LHC (value computed at Next to Leading
¹⁰¹⁷ Order (NLO) accuracy) [64]. This value is approximately three orders of magnitude smaller than the
¹⁰¹⁸ production cross section of a single Higgs boson. In addition, this value has to be multiplied by the
¹⁰¹⁹ branching fraction of the chosen decay channel which further reduces the effective cross section of the
¹⁰²⁰ full process. Nonetheless, ATLAS and CMS have conducted searches for this process whose results
¹⁰²¹ are summarized in this section.

1022 The searches performed so far covered different decay channels and targeted not only the SM pro-
 1023 cess but also some BSM scenarios where di-Higgs production is enhanced. Neither could achieve
 1024 enough statistical significance to declare the measurement of this process in the SM nor have found any
 1025 significant deviation from the expected values. These searches resulted in upper limits for the cross sec-
 1026 tion of di-higgs production in the SM and for the values of the parameters of BSM benchmark theories.
 1027 From the limits on the cross section it is also possible to constraint the values of the Higgs self coupling,
 1028 $k_\lambda = \lambda_{hhh}/\lambda_{hhh}^{SM}$.

1029 The $hh \rightarrow b\bar{b}b\bar{b}$ channel [65, 66] benefits from the large branching fraction of $h \rightarrow b\bar{b}$ ($\sim 58\%$). In
 1030 addition, ATLAS showed that this is the most sensitive channel to resonance masses over 500 GeV [67].
 1031 However, this channel suffers with an overwhelming multijet background which drives the need for very
 1032 tight trigger level cuts in order to bring the event rates down to manageable values.

1033 The $hh \rightarrow b\bar{b}\tau\tau$ analysis [67, 68] benefits from a sizable branching fraction of $h \rightarrow \tau^+\tau^-$ ($\sim 7.3\%$)
 1034 and from a relatively small background contribution from other SM processes. These searches target
 1035 the semi-leptonic decay of the $\tau\tau$ pair to reduce contamination from QCD processes.

1036 The $hh \rightarrow b\bar{b}\gamma\gamma$ [69, 70], $WW^*\gamma\gamma$ [67, 71], $ZZ\gamma\gamma$ [71] analysis can make use of very efficient diphoton
 1037 triggers and isolation criteria that greatly reduce multijet background. In addition the excellent mass
 1038 resolution of $h \rightarrow \gamma\gamma$ can be exploited. The $hh \rightarrow b\bar{b}WW^*$ channel also benefits from the large branching
 1039 fraction of $h \rightarrow W^+W^-$ ($\sim 21\%$).

1040 The most stringent upper limit on the cross section of Higgs pair production comes from a combina-
 1041 tion of three searches using up to 36.1 fb^{-1} of proton-proton collision data at a CM energy of $\sqrt{s} = 13$
 1042 TeV collected with the ATLAS detector [72]. The combination is based on the three most sensitive decay
 1043 channels of Higgs boson pairs, i.e., $hh \rightarrow b\bar{b}\gamma\gamma$, $hh \rightarrow b\bar{b}\tau^+\tau^-$ and $hh \rightarrow b\bar{b}b\bar{b}$. The combined observed
 1044 limit on the non-resonant Higgs boson pair cross-section is 0.223 pb at 95% CL, which corresponds to
 1045 6.7 times the cross-section predicted by the SM. The ratio of the Higgs boson self-coupling to its SM
 1046 expectation, k_λ , is constrained, at 95% CL, to be $-5.0 < k_\lambda < 12.1$.

1047 The upper limits on the cross section (at 95% CL) of Higgs pair production at 8 TeV as a function of the
 1048 mass of a spin 0 resonance are summarized in figure 4.1 on the left [69]. The limits were obtained using
 1049 data collected with the CMS detector and come from searches using different final states, namely, $b\bar{b}\gamma\gamma$ (blue),
 1050 $b\bar{b}b\bar{b}$ (red and pink) and $b\bar{b}\tau\tau$ (green). Depending on the analysis, the corresponding integrated
 1051 luminosity varies between 17.9 fb^{-1} and 19.7 fb^{-1} . The results are usually interpreted in the framework
 1052 of the Randall-Sundrum models such that the spin 0 resonance corresponds to the radion that decays
 1053 to a pair of Higgs bosons. For a mass of 1 TeV the upper limit on the cross section is approximately 12
 1054 pb.

1055 The upper limits on the cross section (at 95% CL) of Higgs pair production as a function of the
 1056 mass of a heavy scalar Higgs, m_H , are summarized in figure 4.1 on the right [73]. The limits were
 1057 obtained through a combination of searches performed in the $hh \rightarrow b\bar{b}\gamma\gamma$, $hh \rightarrow b\bar{b}b\bar{b}$, $hh \rightarrow b\bar{b}\tau^+\tau^-$ and
 1058 $hh \rightarrow \gamma\gamma W^+W^-$ channels using proton-proton collision data corresponding to an integrated luminosity
 1059 of 20.3 fb^{-1} collected with the ATLAS detector at a CM energy of $\sqrt{s} = 8 \text{ TeV}$. The improvement above
 1060 $m_H = 500 \text{ GeV}$ is due to the sensitivity of the $hh \rightarrow b\bar{b}b\bar{b}$ analysis. For $m_H = 900 \text{ GeV}$ the observed limit

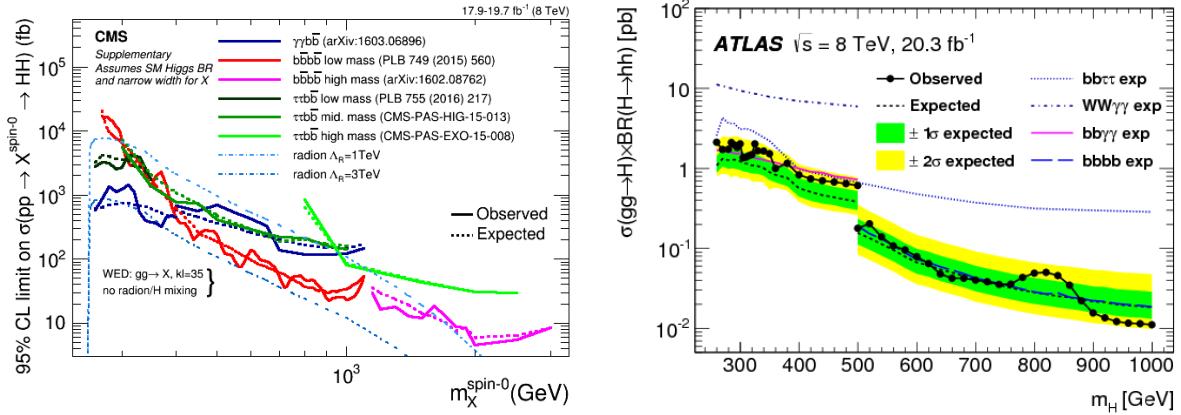


Figure 4.1: Exclusion limits (at 95% CL) on the cross section of resonant Higgs pair production from the decay of heavy scalar particles. These results were obtained using proton-proton collision data collected with the CMS (left) and ATLAS (right) detector at a CM energy of $\sqrt{s} = 8$ TeV and result from the combination of different channels. Plots from [69] (left) and [73] (right).

is 0.015 pb.

4.2 Feasibility studies for high-luminosity and future colliders

Without any BSM contribution, the discovery of Higgs pairs production in the four b quarks final state at the LHC is highly unlikely even considering the total expected integrated luminosity of 300 fb^{-1} . There might be evidence for this process but retrieving useful information about the value of the Higgs trilinear coupling will remain out of reach. Nonetheless, the HL-LHC as well as future colliders pose a good opportunity for the discovery and precision studies of this process. Therefore, Monte Carlo studies assessing the feasibility of searches for $hh \rightarrow b\bar{b}b\bar{b}$ at the HL-LHC and at the FCC-hh have been performed.

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 $pp \rightarrow t\bar{t} \rightarrow b\bar{b}jjjj$ backgrounds reports a significance (S/\sqrt{B}) of 4 (1.3) for an integrated luminosity of 3000 (300) fb^{-1} [74]. The analysis is performed in three orthogonal regions (boosted, intermediate and resolved) and the reported significance is obtained from the combination of these three regions. The highest significance ($S/\sqrt{B} = 2.9$) is achieved in the boosted category. In addition, the impact of pile up is evaluated. Considering a mean pile up of 80 and making use of jet grooming techniques, a significance of 3.1 (1.0) for an integrated luminosity of 3000 (300) fb^{-1} is reported (also considering the three analysis regions). This work makes use of artificial neural networks (ANN's) as well as of jet substructure observables to further increase the signal-background separation. Furthermore, single Higgs backgrounds, namely $Z(\rightarrow b\bar{b})h(\rightarrow b\bar{b})$, $t\bar{t}h(\rightarrow b\bar{b})$ and $b\bar{b}h(\rightarrow b\bar{b})$, are shown to be negligible for the analysis, when compared to the dominant QCD multijet background.

ATLAS and CMS also carried out preliminary studies for the sensitivity on the trilinear coupling at the HL-LHC. Several channels have been investigated: $hh \rightarrow b\bar{b}\gamma\gamma$, $hh \rightarrow b\bar{b}\tau^+\tau^-$, $hh \rightarrow b\bar{b}W^+W^-$ 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

1084 reported significances of 1.05σ and 1.6σ , respectively, for an integrated luminosity of 3000 fb^{-1} . Taking,
1085 as an example, the ATLAS result, the achieved significance translates to an upper limit on the total
1086 di-Higgs cross section of approximately twice the SM value. This corresponds to an exclusion limit of
1087 $-0.8 < k_\lambda < 7.7$. The analysis conducted by ATLAS in this channel is done at generator level. The
1088 energy and momenta of the particles are smeared to simulate the detector's response. A mean pile up
1089 of 200 is considered. No MVA techniques are employed in the analysis.

1090 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 can
1091 be excluded at 95% CL, with systematic uncertainties being taken into consideration. The Higgs trilinear
1092 coupling is expected to be constrained to $-3.5 < k_\lambda < 11$. These results are based on the extrapolation
1093 of the current results obtained with the 2016 dataset, comprising an integrated luminosity of 10.1 fb^{-1} .
1094 This study is based only on the resolved analysis documented in [65].

1095 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
1096 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} ,
1097 considering an analysis that targets the boosted region [75]. The extra jet in the signal sample has
1098 $p_T > 200 \text{ GeV}$ which provides the Higgs pair with a large Lorentz boost. This enhances the sensitivity
1099 to the Higgs boson self coupling because it favors highly boosted virtual Higgs bosons decaying to a
1100 pair of Higgs bosons. The analysis relies on the jet substructure observable τ_{21} and on a tight mass cut
1101 around the Higgs mass. No multivariate techniques are employed.

1102 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 at
1103 the FCC-hh was presented in [76]. Only the irreducible background is considered. For a boosted region,
1104 cut based analysis, a significance of 1.1 is reported for an integrated luminosity of 3 ab^{-1} at $\sqrt{s} = 14 \text{ TeV}$
1105 (HL-LHC). For an integrated luminosity of 10 ab^{-1} at $\sqrt{s} = 100 \text{ TeV}$, which corresponds to the FCC-hh,
1106 this number is 5.7. While the significance is large for the FCC-hh, the reported signal to background ratio
1107 (S/B) is approximately one order of magnitude smaller, which means the measurement might be more
1108 sensitive to systematic uncertainties on the backgrounds.

1109 4.3 Hadronic calorimeter granularity studies

1110 Even before the baseline design for the FCC-hh was established, there were studies regarding the
1111 impact of the granularity of the calorimeters in the spatial resolving power of hadronic showers and on
1112 the resolution of jet mass and substructure variables. These studies targeted the development of future
1113 colliders and greatly influenced the baseline design of the FCC-hh.

1114 The granularity of hadronic calorimeters is a key parameter for future collider detectors because it
1115 determines how well we can resolve energy deposits from pileup vertices and highly-boosted jet topolo-
1116 gies. In [77], the use of smaller calorimeter cells (smaller than the ones that are used in currently
1117 operating hadronic calorimeters) to resolve individual hadrons is investigated. For two Kaons (K_L^0) with
1118 an energy of 100 GeV each and with a truth level separation (ΔR) equal to 0.035 the energy deposited in
1119 the ECAL (blue) and HCAL (red) is shown as a function of ΔR for HCAL cells with sizes $20 \text{ cm} \times 20 \text{ cm}$
1120 ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$) 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 left and

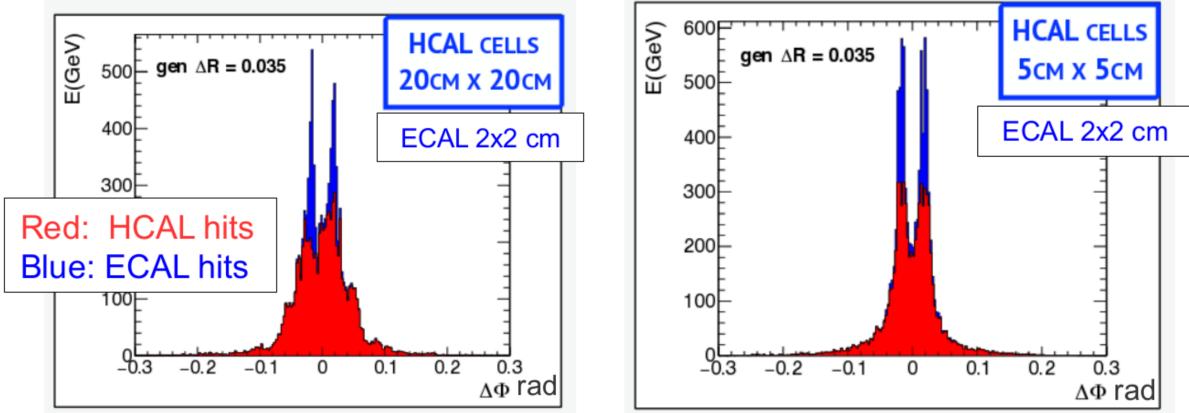


Figure 4.2: Azimuthal distribution of the energy deposits in the ECAL and HCAL for a pair of K_L^0 with $E = 100$ GeV for an hadronic calorimeter with $20\text{ cm} \times 20\text{ cm}$ (left) and $5\text{ cm} \times 5\text{ cm}$ (right). Figures from [78] (based on [77]).

right, respectively. The ECAL segmentation is equal to $2\text{ cm} \times 2\text{ cm}$. It is shown that for a granularity of $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$ both hadrons can be resolved in the HCAL.

Additional studies focusing on the jet mass resolution and on the resolution of jet substructure observables were performed for three calorimeter configurations (HCAL and ECAL). 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 [78–80]. The calorimeter configurations tested are: HCAL(ECAL) $0.1(0.025) \eta \times 5.6(1.4) \deg \phi$, HCAL(ECAL) $0.05(0.012) \eta \times 2.8(0.7) \deg \phi$ and HCAL(ECAL) $0.025(0.006) \eta \times 1.4(0.35) \deg \phi$.

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 resolution¹ is shown in the same figure, on the right. 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.

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. Regardless of the type of jets used (eflow or tower jets), the τ_{32} resolution increases as the HCAL resolution increases. In addition, it is important to note that the resolution is a lot worse when using tower jets which is an indication that exploiting tracking information is vital to achieve a good resolution in substructure variables.

Another interesting result, presented in [78], has to do with the overlap between the τ_{21} distribution for QCD and W jets. Jets with of 2.5 TeV show a reduction in overlap ($80\% \rightarrow 60\%$) going from $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ to $\Delta\eta \times \Delta\phi = 0.005 \times 0.005$ HCAL cells. For 5(10) TeV jets the overlap goes from 88(91)% to 78(85)%. However, for 20 TeV jets, the change in the HCAL granularity does not modify the overlap, as is shown in figure 4.5.

In summary, previous studies focus on the impact of granularity in the resolution of jet mass and jet substructure observables. We did not find any studies that explored the change in the significance of a

¹The resolution of a given observable, X , is given by $(X^{\text{reco}} - X^{\text{true}})/X^{\text{true}}$.

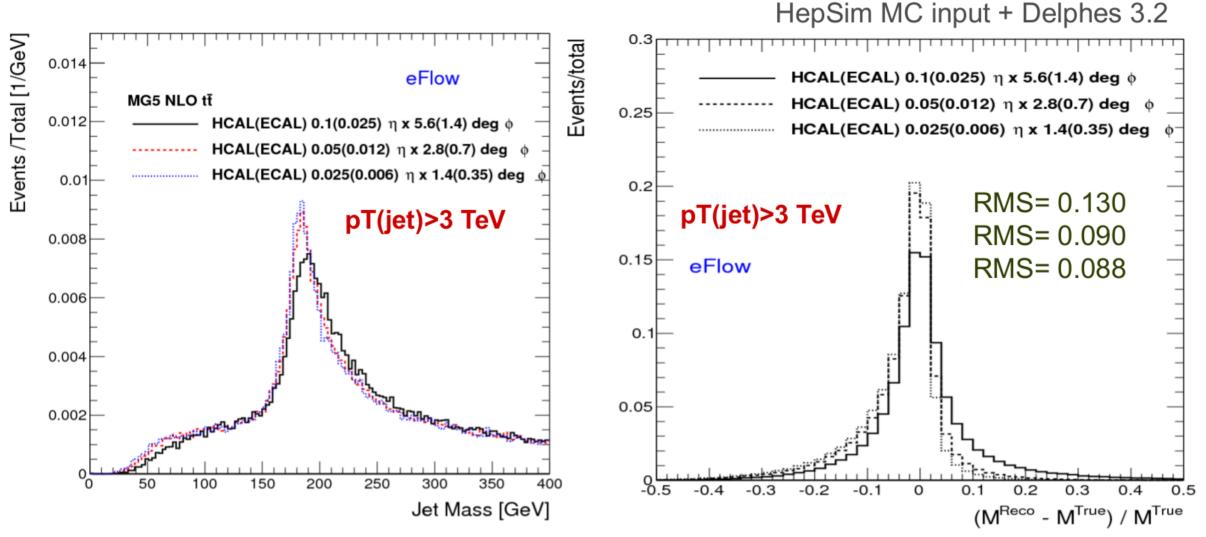


Figure 4.3: Jet mass (left) and jet mass resolution (right) plots in $t\bar{t}$ events for eflow jets with $p_T > 3$ TeV for three different HCAL and ECAL configurations: HCAL(ECAL) 0.1(0.025) $\eta \times 5.6(1.4)$ deg ϕ , HCAL(ECAL) 0.05(0.012) $\eta \times 2.8(0.7)$ deg ϕ and HCAL(ECAL) 0.025(0.006) $\eta \times 1.4(0.35)$ deg ϕ . On the right, the root mean square (RMS) are also shown. Plots from [78].

¹¹⁴⁷ given analysis as a function of the granularity or of the detector configuration.

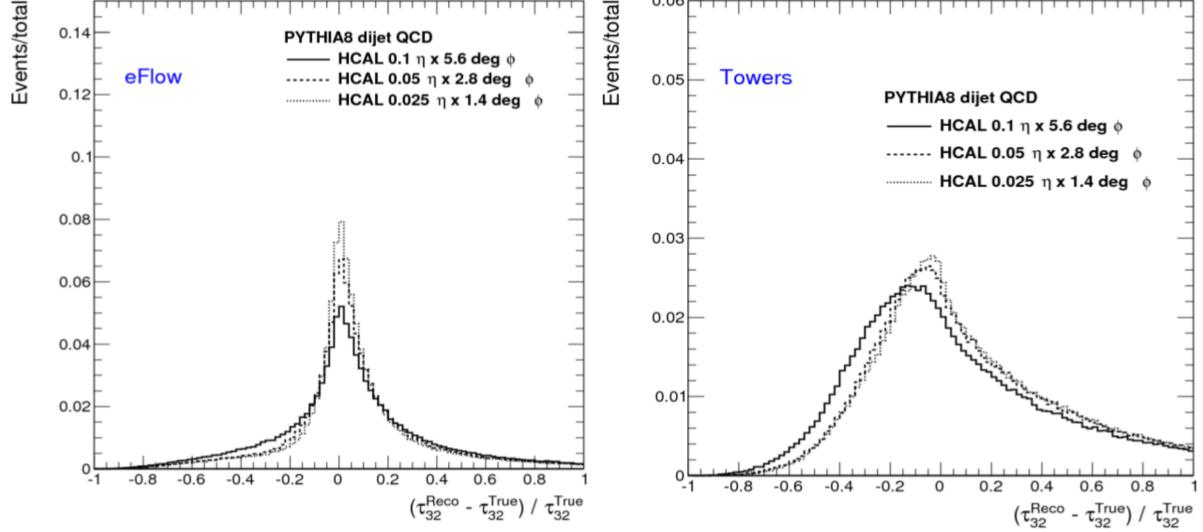


Figure 4.4: τ_{32} resolution plots for QCD dijet events reconstructed using eflow jets (left) and tower jets (right). Plots from [79].

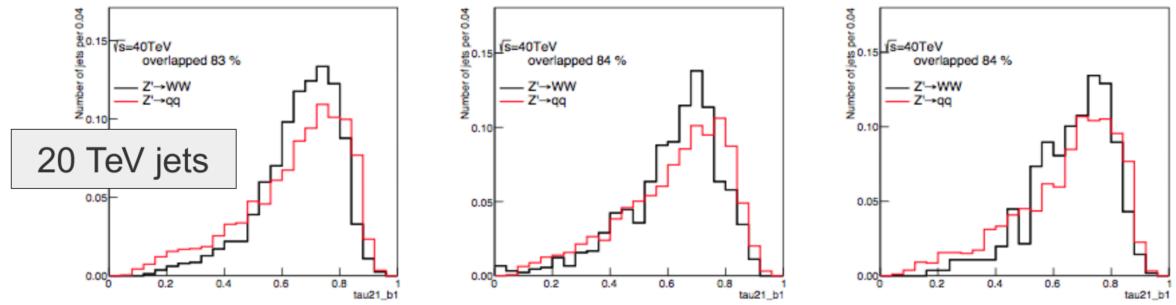


Figure 4.5: τ_{21} variable for 20 TeV W (black) and QCD jets (red) for three HCAL granularities: $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ (left), $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$ (middle) and $\Delta\eta \times \Delta\phi = 0.005 \times 0.005$ (right). Plots from [78].

1148 **Chapter 5**

1149 **Sample generation and analysis tools**

1150 A crucial component of this work is the generation of the Monte Carlo samples that are used in the anal-
1151 ysis. They are produced using fast simulation. We use the software and machinery that was developed
1152 by the FCC study group at CERN. The simulation work flow and Monte Carlo samples are described in
1153 this chapter.

1154 In section 5.1 we introduce the generators used to produce the Monte Carlo samples and briefly
1155 describe their working principles and functionalities. We then focus on the FCC-hh software, in subsec-
1156 tion 5.1.2, and explain how the previously described simulation work flow is implemented. A detailed
1157 description of the technical settings used to produce the samples is provided in sections 5.2 and 5.2.3.

1158 **5.1 Fast simulation workflow**

1159 The Monte Carlo samples used in this study were generated using MadGraph5_aMC@NLO [81]. The
1160 showering and hadronization are simulated using Pythia 8 [82] and the detector response is parametrized
1161 using Delphes 3 [83]. These samples are available at the CERN EOS storage.

1162 MadGraph5 is a matrix element generator for high energy physics processes, such as decays and
1163 scatterings. The user specifies the desired process in terms of the initial and final states and can impose
1164 additional constraints such as allowing for a number of refined criteria, including forced or forbidden s-
1165 channel resonances, excluded internal particles, and forced decay chains of final state particles [81].
1166 As a result, MadGraph automatically generates the corresponding Feynman diagrams and creates the
1167 necessary code to compute the matrix element at a given point of the phase space. The output file is
1168 written in the Les Houches Accord [84].

1169 Pythia8 is frequently used for event generation in high energy physics. In this work, however, we use
1170 it only to simulate the showering and hadronization process and not the parton level hard process which
1171 is simulated in MadGraph5. The Les Houches file that is produced by MadGraph5 is used as input to a
1172 Pythia8 program that can decay unstable particles, simulate initial and final state showers as well as the
1173 hadronization of coloured particles, such as quarks and gluons. The desired settings can be specified
1174 in an additional file, a card, that is used as input for a Pythia8 run.

1175 Delphes3 allows for a quick and simple simulation of the detector's response [83]. Its goal is to allow
1176 the fast simulation of a multipurpose detector for phenomenological studies. The simulation includes a
1177 track propagation system, electromagnetic and hadronic calorimeters and a muon identification system.
1178 Low level physics objects, such as tracks and energy deposits, and high level physics objects, such as
1179 leptons and jets, are reconstructed from the detector's response and can be used to perform physics
1180 analysis. In the following paragraphs we briefly describe how the detector response is simulated and
1181 how jet reconstruction is performed.

1182 The magnetic field is uniform, axial and parallel to the beam direction. Charged particles follow a
1183 helicoidal trajectory from the interaction point to the calorimeter while neutral particles have a straight
1184 line trajectory. The probability of a charged particle being reconstructed as a track is defined by the
1185 user. Only a smearing in the modulus of the transverse momentum is applied (not to the direction).
1186 The tracking efficiency as well as energy and momentum resolutions are specified by the user and may
1187 include a dependency on the particle type, momentum and pseudorapidity. The calorimeters have a
1188 finite segmentation in the (η, ϕ) plane and the cell size can be defined in the configuration file. The
1189 amount of energy deposited in the calorimeters by each particle type can be defined by the user. By
1190 default, stable hadrons deposit all their energy in the HCAL although in a real detector a significant
1191 fraction of their energy is deposited in the ECAL. The energy resolution of the ECAL and HCAL are
1192 parameterized as a function of η and include stochastic, noise and constant terms [FORMULA?]. The
1193 electromagnetic and hadronic energy deposits are independently smeared by log-normal distributions.

1194 Jets can be produced using generator level long-lived particles after showering and hadronization,
1195 tracks, calorimeter towers or particle-flow tracks and towers. These are referred to as generator, track,
1196 calorimeter or particle-flow jets, respectively. For generator level jets no detector simulation nor recon-
1197 struction are taken into account. In spite of the type of jet the user can choose the jet clustering algorithm
1198 and the values of its parameters as well as the minimum transverse momentum of the jets that are stored
1199 in the final collection. Delphes integrates the FastJet package [85] and therefore allows jet reconstruc-
1200 tion with the most popular jet reconstruction algorithms, namely, anti- k_T , k_T and Cambridge-Aachen
1201 (C/A). Jets resulting from the hadronization of a b quark (known as b jets), are identified if a b quark is
1202 found within a ΔR distance from the jet's axis. The tagging efficiency and mis-tagging probabilities can
1203 be defined by the user.

1204 5.1.1 Particle flow and calorimeter reconstruction in Delphes

1205 In Delphes, hadronic jets can be reconstructed using only the information from the HCAL towers or using
1206 a particle flow algorithm that combines information from the tracking system and from the HCAL towers.
1207 These two approaches create jets that are referred to as calorimeter and particle flow jets, respectively.
1208 The latter can also be referred to as energy flow jets (eflow jets in short). In this work we performed the
1209 analysis using both sets of jets and compare the results. Therefore we briefly describe them here.

1210 Calorimeter jets are very simple. They are reconstructed using as input for the jet clustering algorithm
1211 the 4-vectors associated with the calorimeter towers, after a cell energy smearing has been applied.

1212 Therefore the energy resolution is limited by the transverse segmentation of the calorimeters.

1213 The goal of the particle flow approach is to make use of all the available information provided by the
1214 various sub-detectors for reconstructing an event [83]. This approach is used by some experimental
1215 collaborations [86, 87] but the exact implementation depends on the specificities of the experiment. If
1216 the momentum resolution of the tracking system is better than the energy resolution of the calorimeters
1217 it might be convenient to use the tracking information to estimate the momentum of charged particles.
1218 In real experiments, the tracking resolution is only better than the calorimeter's energy resolution up to
1219 some energy threshold. However, in Delphes, it is assumed that it is always convenient to estimate the
1220 momentum of charged particles via the tracker.

1221 The particle flow algorithm works as follows [83]. For each calorimeter tower it counts:

- 1222
- 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
1223 has been reconstructed, $E_{\text{ECAL,trk}}$ and $E_{\text{HCAL,trk}}$, respectively.
1224

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

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

1225 All reconstructed tracks result in a particle flow track. If $E_{\text{Tower}}^{\text{eflow}} > 0$ a particle flow tower is created with
1226 energy $E_{\text{Tower}}^{\text{eflow}}$. The particle flow tracks and jets are then used as input for the jet clustering algorithms.

1227 5.1.2 FCC-hh software

1228 FCC software (FCCSW) [88], common to all FCC studies (electron-electron, electron-hadron and hadron-
1229 hadron) has been developed and is maintained by the FCC study group. The software is based on Gaudi
1230 [89]. An FCC Event Data Model based on Podio [90] was also developed. It consists in specific classes
1231 that encode the information about the events.

1232 The FCC-hh study group is responsible for the generation of Monte Carlo quick simulation samples
1233 for the main benchmark processes for the FCC-hh. The samples are generated using the workflow
1234 described in the previous section. CERN users can request rights to run the EventProducer package
1235 [91] and produce samples for any desired process using the machinery that is already implemented. In
1236 addition, the FCC Event Data Model classes are directly accessible and can be used to read the ROOT
1237 files that are produced after the events are passed through Delphes.

1238 The machinery to submit jobs to CERN's batch system is also implemented for both the generation
1239 (MadGraph5) and reconstruction (Pythia8 plus Delphes3) levels.

1240 In this work we make use of this software in order to produce the necessary samples.

1241 **5.2 Signal and background samples**

1242 In this work we focus on the $hh \rightarrow b\bar{b}b\bar{b}$ channel and perform an analysis targeting the boosted kinematic
1243 region. The main backgrounds are multijet and $t\bar{t}$ production and the irreducible $pp \rightarrow b\bar{b}b\bar{b}$ process. This
1244 was introduced and motivated in chapter 1 and it is discussed in detail in chapter 6. Here we provide a
1245 technical description of how the signal and background samples were generated.

1246 **5.2.1 MadGraph**

1247 The irreducible background is generated with an extra jet with a high p_T at generator level, $4b+j$, where j
1248 stands for a light jet (initiated by a gluon or u, d, c, s quarks). This is referred to as the four flavor scheme.
1249 A high- p_T extra jet forces the four b quarks to have a high Lorentz boost and therefore increases the
1250 probability of the events being reconstructed with two large- R jets with a two-prong substructure. For this
1251 process, QCD and electroweak contributions are considered separately, i.e, we include three different
1252 types of samples: one in which only QCD processes are considered, one in which only electroweak
1253 processes are considered and one in which both process are considered simultaneously. The samples
1254 do not overlap. The $4b+j$ QCD sample is constituted by two independent samples that have a different
1255 generator level cut in the minimum p_T of the light jets, namely, $200 < p_{T,j}^{\min} < 500$ GeV and $p_{T,j}^{\min} > 500$
1256 GeV. This allows for a more efficient generation.

1257 The multijet background is simulated through $jj+0/1/2 j$ where j stands for a light or b jet (five flavor
1258 scheme). This background is divided into several individual samples that are produced in different H_T
1259 regions, where H_T is the scalar sum of the p_T of all partons at generator level. The minimum(maximum)
1260 allowed H_T is 500(100,000) GeV. Since these backgrounds are QCD processes, the p_T distribution of
1261 the final state jets falls very steeply as the p_T increases. Therefore, if one were to generate events for
1262 these processes without restricting the phase space, most events would consist of jets with a very low
1263 p_T which are exactly the type of events that are rejected the most by a boosted analysis (see chapter
1264 6 for more details on the event topology that is targeted and on the analysis strategy). As we move to
1265 regions of the phase with a higher H_T the cross section decreases meaning we need fewer MC events
1266 to properly simulate the background in that region.

1267 Note that for the $4b+j$ (QCD) and $jj+0/1/2 j$ samples we do not take into account the regions of the
1268 phase space with $p_{T,j} < 200$ GeV and with $0 < H_T < 500$ GeV. We assume that we can reject most of
1269 the events (if not all) with these kinematic characteristics by going to a sufficiently boosted region of the
1270 phase space. In addition, note that after the showering procedure there could be some overlap between
1271 the $4b+j$ and $jj+0/1/2 j$. This is taken care of in our analysis code: if an event from the $jj+0/1/2 j$
1272 has four b quarks at truth level [GIVE PERCENTAGE] then we do not consider it because it will certainly
1273 overlap with an event from the $4b+j$ sample.

1274 The $t\bar{t}$ background sample is generated with extra jets at generator level, $t\bar{t}+0/1/2 j$, using the five
1275 flavor scheme. We consider an inclusive sample for this background, meaning that we do not force any
1276 particular decay of the top quark or of the subsequent particles.

1277 For the hh SM signal sample, one of the Higgs is decayed to $b\bar{b}$ in MadGraph. The reasoning behind

1278 this choice is that most searches for Higgs pair production make use of a final state that includes at
 1279 least two b quarks in order to keep the cross section times BR of the process large enough. In addition,
 1280 this method allows the same generator level samples to be used to perform different analysis, simply by
 1281 choosing the decay channel of the remaining Higgs boson. The decay of the remaining Higgs boson
 1282 can be implemented in Pythia, in the case of this work, to $b\bar{b}$.

1283 The generator level cuts for the signal and background samples are summarized in table ?? in
 1284 appendix D.

1285 Signal samples - BSM

1286 In addition to the SM di-Higgs signal, we also explore the signature and analysis sensitivity for di-Higgs
 1287 signals produced by two benchmark BSM models: CP-conserving type II 2HDM and a simplified dark
 1288 matter model (DM) with a spin 0 mediator. These models were described in section 2.2.

1289 Both models are readily available in FeynRules [92] (which is a Mathematica package for the imple-
 1290 mentation of new physics models) model database and can be straightforwardly implemented in Mad-
 1291 graph5. The parameters of the models, namely the masses of the new particles, can be changed by
 1292 the user. In the case of this work, we want new particles to have a large mass so that they produce SM
 1293 Higgs pairs with a high Lorentz boost.

1294 For the DM model [18], the spin 0 mediator's mass is set at 1 TeV. The cross section for the signal
 1295 generated with this model is smaller than the cross section of the SM signal (approximately 0.2 pb *versus*
 1296 0.7 pb) and therefore this model is not excluded by experimental data.

1297 For the 2HDM [93, 94], a finer tuning of the parameters is required because the model is very general
 1298 and has a lot external free parameters. In addition, it is written in the Higgs basis such that we need to
 1299 convert the parameters from the physical basis (introduced in 2.2) to this basis. The free parameters
 1300 of the model that we are interested in are the masses of the neutral, $mh1, mh2, mh3$, and charged
 1301 scalars, mhc , the neutral scalars mixing angles, $mixh, mixh2, mixh3$, the real and imaginary parts of
 1302 the up, down and charged lepton 3×3 Yukawa matrices, $GUR, GUI, GDR, GDI, GLR, GLI$ and the
 1303 values (real and imaginary parts) of the second, third and seventh quartic couplings, $l_{2,3,7},.$. In the CP-
 1304 conserving model $h_{1,2,3}$ can be identified with the physical states h, H and A , respectively.

We start from the following parameters (in the physical basis):

$$\begin{aligned}
 m_h &= 125 \text{ GeV}, & m_H &= 900 \text{ GeV}, & m_A &= 850 \text{ GeV}, & m_{H^\pm} &= 800 \text{ GeV} \\
 \beta &= \frac{\pi}{4}, & \alpha &= -0.75 \\
 m_{12}^2 &= [(m_H^2 + m_A^2 + m_{H^\pm}^2)/3] \cos(\beta) \sin(\beta) \simeq 181041.6667.
 \end{aligned} \tag{5.2}$$

In the Higgs basis, the input parameters for the model are

$$\begin{aligned}
m_{h_1} &= 125 \text{ GeV}, & m_{h_2} &= 900 \text{ GeV}, & m_{h_3} &= 850 \text{ GeV}, & m_{hc} &= 800 \text{ GeV} \\
mix_{h_2} &= mix_{h_3} = 0 \\
mix_{h_1} &= \frac{\pi}{2} - (\beta - \alpha) \simeq 0.035 \\
l_2 &\simeq 0.27, & l_3 &\simeq 9.46, & l_7 &\simeq 0.46.
\end{aligned} \tag{5.3}$$

The masses of the scalars cannot change when going from one basis to the other. Therefore we set them to the exact same values. $mix_{h_2}, mix_{h_3} = 0$ because we are considering the CP-conserving model. To obtain the values of $l_{2,3,7}$ we use Eq. 11 from [24] we write the λ parameters of the scalar potential in terms of the parameters defined in Eq. 5.2. Then we use Eq. 47,48 and 50 from [95] to obtain the values of the quartic coupling in Higgs basis.

Regarding the Yukawa interactions, we work in the type II model. Following the type II restriction file found in [93] we set

$$\begin{aligned}
GLR &= GLI = GDI = GUI = 0 \\
GDR &= \text{diag} \left(0, 0, \frac{m_b \sqrt{2} \tan(\beta)}{v} \right) \\
GDR &= \text{diag} \left(0, 0, -\frac{m_t \sqrt{2}}{v \tan(\beta)} \right)
\end{aligned} \tag{5.4}$$

where $m_b = 4.7$ GeV and $m_t = 172$ GeV are the masses of the bottom and top quarks.

5.2.2 Pythia

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

For the $jj + 0/1/2$ and $t\bar{t} + 0/1/2 j$ samples we have to perform jet matching because we require additional jets at the level of the matrix element. In addition to the partons generated in MadGraph and that can produce a jet, Pythia may introduce extra jets that are usually soft and collinear (with the particle from which they were radiated) and result from the showering process. This could lead to the same process (meaning processes with the same final states) being counted twice (double counting). Take, for example, the processes jj and $j jj$ at MadGraph level. It can happen that Pythia generates an extra jet for the first process but not for the second, leading to both processes having the same final state (three jets). Each process would then give its independent contribution to the total number of events but because they simply represent two distinct ways of achieving the same final state they should only be counted once. The goal of jet matching procedures is to avoid this problem.

The cross section times branching ratio (when applicable), the k-factors and the total number of events used in the analysis are summarized in table 5.1, for all samples. The k-factor multiplies the

Sample	$\sigma \times BR$ [pb]	k-factor	Number of events
$hh \rightarrow b\bar{b}b\bar{b}$ - SM	0.76	1.52	
$hh \rightarrow b\bar{b}b\bar{b}$ - DM mediator	0.218	1.0	
$hh \rightarrow b\bar{b}b\bar{b}$ - 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	
$jj+0/1/2 j$ ($500 < H_T < 1000$)	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	
$t\bar{t}+0/1/2 j$	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.

1327 cross section times branching ratio in order to reproduce known higher order results. It corresponds
 1328 only to the ratio between the total cross sections and it does not correct for possible difference that
 1329 might exist between the differential cross sections.

1330 The parameters used in Pythia 8 can be found in table ?? in appendix D.

1331 5.2.3 Delphes

1332 It is one of the main goals of this work to evaluate how the analysis sensitivity is influenced by the
 1333 granularity of the hadronic calorimeter. We start from the same MadGraph level samples and pass them
 1334 through Pythia and Delphes changing the settings of the Delphes card that correspond to the HCAL. All
 1335 other detector's parameters were kept unchanged with respect to the FCC default Delphes card. We
 1336 tested five benchmark granularity configurations:

- 1337 1. ATLAS HCAL granularity (as implemented in the standard ATLAS Delphes card);
- 1338 2. Starting from the ATLAS HCAL configuration we increase the granularity in $|\eta|$ by a factor of four,
 1339 in the pseudo rapidity range $|\eta| < 1.7$ which corresponds to the TileCal region;
- 1340 3. Starting from the FCC HCAL configuration we decrease the granularity in ϕ by a factor of two, in
 1341 the entire pseudo rapidity range covered by the HCAL.

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.

- 1342 4. FCC HCAL default granularity (as implemented in the standard FCC Delphes card);
 1343 5. Starting from the FCC HCAL configuration we increase the granularity in $|\eta|$ and in ϕ by a factor of
 1344 two, in the entire pseudo rapidity range covered by the HCAL.

1345 The granularities of these five configurations are summarized in table 5.2. The values that are shown,
 1346 as well as the corresponding pseudorapidity regions, are exactly what is implemented in Delphes.

1347 In addition, we also passed the same generator level samples through the default ATLAS detector
 1348 simulation in Delphes. The HCAL granularity is the one that is indicated in table 5.2 but the other detector
 1349 parameters, such as the radius, magnetic field, tracking resolutions are the ones that are implemented
 1350 in the default ATLAS Delphes card. This additional detector configuration gives us an extra point in the
 1351 space of parameters that we are trying to explore. Furthermore, in a first, very crude, approximation,
 1352 it allow us to compare the results obtained at $\sqrt{s} = 100$ TeV to the ones obtained at $\sqrt{s} = 13$ TeV
 1353 by ATLAS. For completion we summarize, in table 5.3, the values of some key detector parameters as
 1354 implemented in Delphes for the ATLAS and FCC-hh detectors. The magnetic field is twice as strong for
 1355 the FCC-hh. The charged hadrons tracking efficiency and the HCAL resolution are fairly similar between
 1356 the two detector configurations.

Parameter	FCC-hh	ATLAS
Radius of magnetic field coverage [m]	1.5	1.15
Half length [m]	5	3.51
Magnetic field [T]	4	2
Charged hadrons tracking efficiency	90% for $2.5 < \eta < 4.0$, $p_T > 1.0$	85% for $1.5 < \eta < 2.5$, $p_T > 1.0$
HCAL resolution	$\sqrt{E^2(0.03)^2 + E(0.60)^2}$ for $1.7 < \eta < 4.0$	$\sqrt{E^2(0.05)^2 + E(0.706)^2}$ for $1.7 < \eta < 3.2$

Table 5.3: Summary of the benchmark granularity configurations of the HCAL.

1357

Chapter 6

1358

Analysis

1359

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

1360 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 approx-
1361 imately 33.6%, making it the most probable decay for Higgs pairs, as is illustrated in figure ?? on the
1362 left. However, in this channel, the main background is QCD multijet production that has a cross section
1363 several orders of magnitude larger than di-Higgs production in the SM, as table 5.1 shows. Nonetheless,
1364 the jet p_T distributions in this background have a very large yield close to zero and then fall very steeply
1365 while the signal has a much larger tail to high values of p_T . This indicates that searches targeting the
1366 boosted kinematic regime may be the key to measure $hh \rightarrow b\bar{b}b\bar{b}$ using inclusive production.

1367 In this work we target the boosted regime. In this kinematic region, the final state of the signal is
1368 characterized by two jets with a large radius parameter. Each jet is expected to contain the two b quarks
1369 originated from the decay of a Higgs boson, as the scheme on the right of figure ?? illustrates. Extra
1370 jets are also expected to be reconstructed as a consequence of QCD radiation.

1371 The analysis presented here is performed using the Monte-Carlo samples described in chapter 5. All

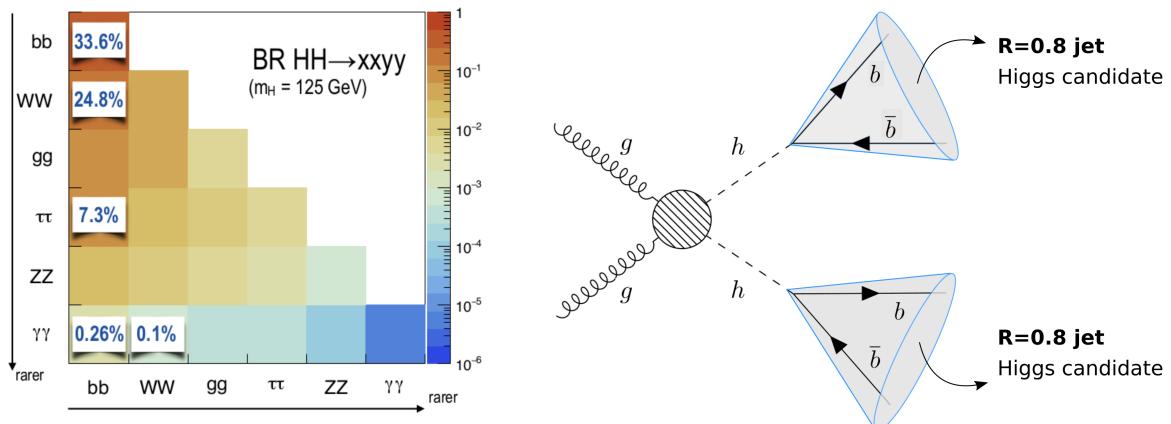


Figure 6.1: Higgs pairs branching ratios (left) and event topology targeted by this analysis (right).

1372 samples assume $m_h = 125$ GeV. The event selection criteria are designed to optimize the significance,
1373 given by S/\sqrt{B} , where S and B represent the number of signal and background events, respectively.

1374 6.1.1 Event pre-selection

1375 We start by applying very simple and loose cuts that target events consistent with the boosted topol-
1376 ogy. Jets are reconstructed with the anti- k_T algorithm with an radius parameter $R = 0.8$. They are
1377 reconstructed from particle flow objects or calorimeter towers. Both approaches were explored in this
1378 analysis.

1379 We require at least two b-tagged $R = 0.8$ jets which corresponds to at least four b-tagged subjets,
1380 at least two in each $R = 0.8$ jet. The subjets are found through the soft drop mass technique that was
1381 described in section 3.3.2. The leading and sub leading jets must have $p_T \geq 200$ GeV in order for the
1382 event to be accepted. From Eq. 3.4, a Higgs boson with, e.g., $p_T = 200$ GeV leads to a pair of b quarks
1383 with $\Delta R \sim \frac{2m}{p_T} = 1.25$. Therefore, considering jet with $R = 0.8$ (diameter equal to 1.6) and $p_T \geq 200$ GeV
1384 is a reasonable choice. For a more complete description about the jet radius see appendix A. Due to the
1385 b-tagging parameterization (described in detail in section 6.2.1), there is a natural cutoff at $|\eta| = 4$ so we
1386 do not place any additional cut in η . From now on, these cuts are referred to as pre-selection cuts.

1387 After the pre-selection cuts we can plot all the relevant distributions that allow us to characterize the
1388 signal and background events. These are shown in the next two sections. In addition, these distributions
1389 provide a first insight into the discriminating power of the kinematic and substructure variables that are
1390 explored.

1391 6.1.2 Signal characterization

1392 The cross section for di-Higgs inclusive production from pp collisions at 100 TeV is 1.2 pb [58]. This value
1393 is computed at NLO accuracy. The dominant process is gluon-gluon fusion that has been extensively
1394 discussed in section 2.1.1. The signal process comprehends the di-Higgs inclusive production and the
1395 $hh \rightarrow b\bar{b}b\bar{b}$ subsequent decays. The total process cross section is then given by $\sigma(pp \rightarrow hh) \times BR(hh \rightarrow$
1396 $b\bar{b}b\bar{b})$ and it amounts to approximately 0.4 pb [NOT VERY COHERENT WITH MG VALUES - CHECK].
1397 For an integrated luminosity of 3000 fb^{-1} of proton collisions at $\sqrt{s} = 100$ TeV we expect X events.

1398 The Higgs candidates are reconstructed using large R jets that can be directly measured using
1399 information from the calorimeters and tracking systems. The plot on the left of figure ?? shows the $\Delta\phi$
1400 between the two leading jets (Higgs candidates), $\Delta\phi(hh)$. For a large fraction of the signal events (red
1401 curve) $\Delta\phi(hh) \sim \pi$ which indicates that the jets are produced back-to-back in the transverse plane.
1402 There is also a small fraction of events for which the jets have $\Delta\phi(hh) \sim 1$. For these events there must
1403 be at least a third jet that balances the momentum of the jets pair.

1404 Each wide jet is expected to be consistent with having two subjets associated with the two b quarks
1405 from the $h \rightarrow b\bar{b}$ decay. In the Higgs rest frame, the two subjets are produced back-to-back conserving
1406 momentum. In the laboratory frame, the ΔR between the subjets depends on the momentum of the
1407 Higgs boson, with larger ΔR corresponding to Higgs bosons with a lower momentum. This correla-

1408 tion between the ΔR between the subjets of the leading Higgs candidate ($\Delta R(b\bar{b})$) and its momentum
1409 ($p_T(h_1)$) is shown in figure ??, on the right.

1410 The p_T distributions of the leading and sub leading jets (Higgs candidates) are shown in figure ??
1411 on the left and right, respectively. The histograms are normalized to unit area in order to allow for shape
1412 comparison between them. For the leading jet, the signal p_T spectrum peaks at approximately 350 GeV
1413 and has a long tail for large values of p_T , in particular, longer than any of the backgrounds, at least up to
1414 $p_T > 1000$ GeV. For the sub leading jet the spectrum is softer, as expected, but the tail of the distribution
1415 is still longer than for any of the backgrounds. These distributions show that the signal process produces
1416 jets with larger transverse momenta than the jets produce by the background processes which indicates
1417 that a boosted analysis might perform well.

1418 The η distribution of the leading jet is shown in figure ?? on the left. This distribution is limited to
1419 $|\eta| < 4$ because the b-tagging efficiency goes to zero for $|\eta| > 4$ (see section 6.2.1 for more details). The
1420 number of events decreases as we go to larger values of η which is due to the detector's acceptance.
1421 The η distribution of the sub leading jet is very similar to this one so we abstain from showing it here.

1422 The softdrop mass distribution of the leading jet is shown in figure ?? on the right. For the signal (red
1423 curve) there is a clear peak at approximately 120 GeV which corresponds to the SM Higgs boson mass
1424 peak. The peak is quite broad which means the mass resolution in this channel is not very good. On the
1425 one hand, this is because we are using large- R jets to reconstruct the Higgs candidates. These objects
1426 have a worse mass resolution than other cleaner objects such as photons or electrons. On the other
1427 hand, the cuts applied before plotting these distributions are very loose. As we place additional cuts the
1428 mass peak can be made slightly narrower. Another interesting feature of the signal mass spectrum is
1429 the existence of a peak close to zero. The explanation is the following: some jets do not contain both b
1430 quarks from the Higgs decay such that when applying the mass drop procedure only one b quark with a
1431 mass of approximately 5 GeV remains inside the jet, creating the peak at lower masses. A more detailed
1432 discussion, as well as a plot that supports this interpretation, can be found in appendix C.

1433 In addition to the basic kinematic distributions that we have just described, there are a multitude of
1434 other variables one can look at. Here we show the $\Delta\eta$ between the Higgs pairs, figure ?? on the left,
1435 and the τ_{21} variable for the leading Higgs candidate, figure ?? on the right. The first distribution shows
1436 that the Higgs pair tends to have a $\Delta\eta$ close to zero, which indicates that the pair is highly boosted in the
1437 longitudinal plane. The τ_{21} distribution shows that for the signal this variable takes values close to zero
1438 which means the jet is consistent with having two subjets.

1439 6.1.3 Backgrounds

1440 The relevant backgrounds for this analysis are QCD multijet production, $t\bar{t}$ and $b\bar{b}b\bar{b}$. Although the $b\bar{b}b\bar{b}$
1441 background is a particular case of a QCD multijet production process we consider it separately because
1442 it constitutes the irreducible background and therefore will have a higher efficiency in the analysis. The
1443 cross sections for these processes are several orders of magnitude larger than the cross section for the
1444 signal, as table 5.1 shows. In addition, in the case of the $t\bar{t}$ background, the event topology is expected

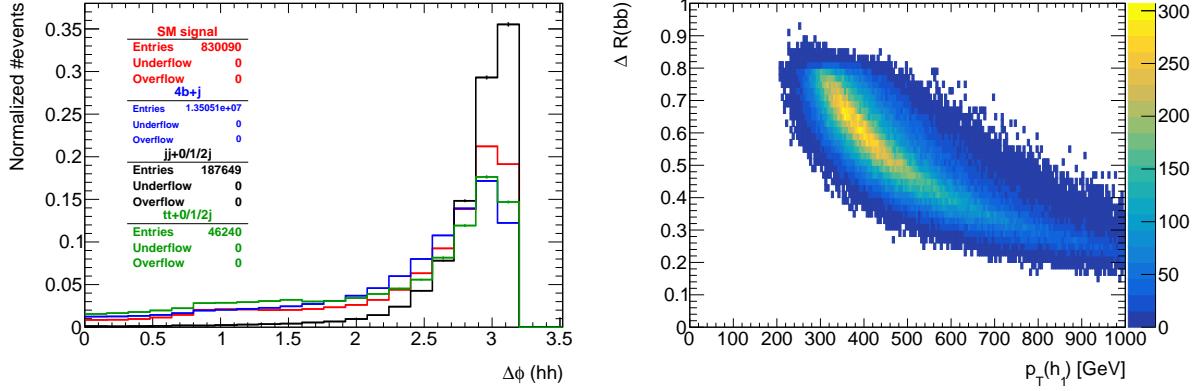


Figure 6.2: $\Delta\phi$ between the Higgs candidates, $\Delta\phi(hh)$, (left) and ΔR between the b quarks coming from the decay of the leading Higgs candidate, h_1 (right). To obtain this plot the cut $|M - 125| < 40$ GeV was applied in addition to the pre-selection cuts..

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 $t\bar{t}$ are the two main backgrounds is corroborated by the ATLAS di-Higgs search performed in the same channel, where these backgrounds are found to be the dominant ones. All other sources of backgrounds, including processes involving Higgs bosons, are found to be negligible [65]. In appendix B we discuss and evaluate the importance of some backgrounds that include Higgs bosons in our analysis.

Figure ?? shows examples of LO Feynman diagrams that contribute to $4b + j$ (left), three (middle) and four (right) light jets production. The $b\bar{b}b\bar{b}$ background is generated with an extra light jet at generator level, as discussed in chapter 5. This jet boosts the four b quarks and has a minimum p_T of 200 GeV. This increases the probability of two b quarks being reconstructed as single large- R jet therefore emulating the signal's final state signature, i.e, high p_T jets compatible with a two-prong substructure. The QCD $4b + j$ sample with $(200 < p_T < 500)$ GeV has a cross section of approximately 756 pb. In addition, as indicated in table 5.1, we multiply this cross section by a k-factor of 2 in order to account for unknown high order contributions.

The QCD multijet background consists of two, three and four jet events (represented as $jj + 0/1/2 j$), at generator level. The jets can originate from light and b quarks and from gluons. A jet matching procedure is implemented in Pythia in order to avoid double counting. Due to mis-tagging probabilities, light and c jets can be identified as b jets. Although these probabilities are relatively small when compared to the b-tagging efficiency (see section 6.2.1) the cross section of multijet processes is very large such that this background becomes dominant. The $jj + 0/1/2 j$ sample with $(500 < H_T < 1000)$ GeV has a cross section of the order of 10^7 GeV. As anticipated, the p_T spectrum of the multijet background is a lot softer than for the signal and remaining backgrounds, as it is shown on both plots of figure ???. In addition, for this background the τ_{21} variable takes values close to one, as it is shown in figure ?? on the right. This indicates that the jets are not consistent with two subjects.

Figure ?? shows examples of LO Feynman diagrams that contribute to $t\bar{t}$ production through $q\bar{q}$ (left) and gg (middle and right) fusion. The $t\bar{t}$ background is simulated with additional zero, one or two jets

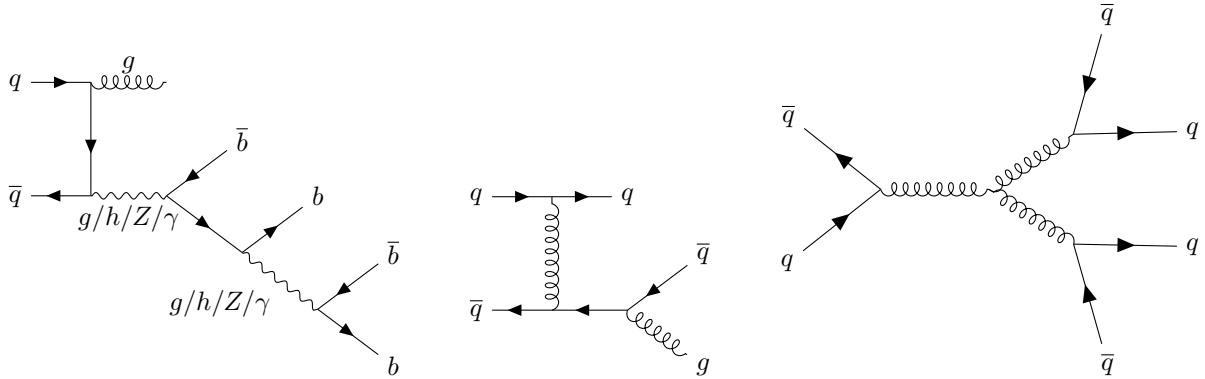


Figure 6.3: Example of diagrams that contribute to the QCD multijet background: five final state jets, four of which are b-jets (left), three final state jets (middle) and four final state jets (right). Here, q stands for a light quark/jet.

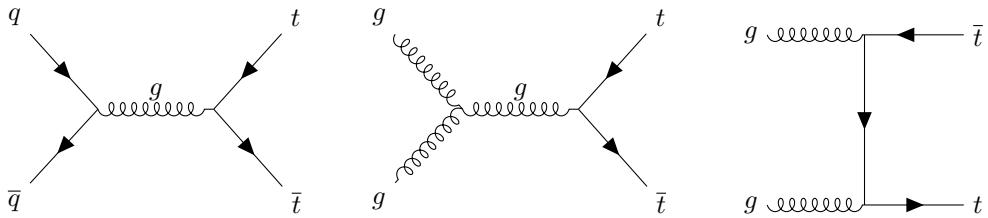


Figure 6.4: Dominant diagrams of $pp \rightarrow t\bar{t}$ at LO.

1472 (represented as $t\bar{t}+0/1/2$ j), at generator level. The extra jets can originate from light and b quarks
 1473 and from gluons. A jet matching procedure is implemented in Pythia in order to avoid double counting.
 1474 The top quark has a very short life time, predicted to be 5×10^{-25} s, such that it decays before it can
 1475 hadronize. This sample is inclusive in the top quark decay modes. However, the most favoured decay
 1476 of the top quark is $t \rightarrow Wb$ with a branching ratio close to 96% [96]. Therefore, $t\bar{t}$ events will, most of
 1477 the times, result in the $W^+W^-b\bar{b}$ final state. We do not specify any decay mode for the W such that
 1478 the sample is also inclusive in the W decay modes. The W decays to hadrons ($W^+ \rightarrow q\bar{q}$) and leptons
 1479 ($W^+ \rightarrow l^+\nu$) 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
 1480 to hadrons then there will be additional jets in the final state. These can be b jets or can be misidentified
 1481 as such. If both W bosons decay to leptons there will still be at least two b jets in the final state, coming
 1482 from the $t\bar{t}$ decay. The cross section of the $t\bar{t} + 0/1/2$ j is approximately 4.3×10^4 pb. A k-factor of 1.74
 1483 is applied to account for higher order effects.

1484 It is interesting to note that in the softdrop mass plot for the leading Higgs candidate (figure ?? on the
 1485 right) the $t\bar{t}$ background (green line) shows a small peak around 170 GeV. In this region, all the decay
 1486 products of the top quark are contained inside the $R = 0.8$ jet such that the jet mass corresponds to the
 1487 mass of the original top quark.

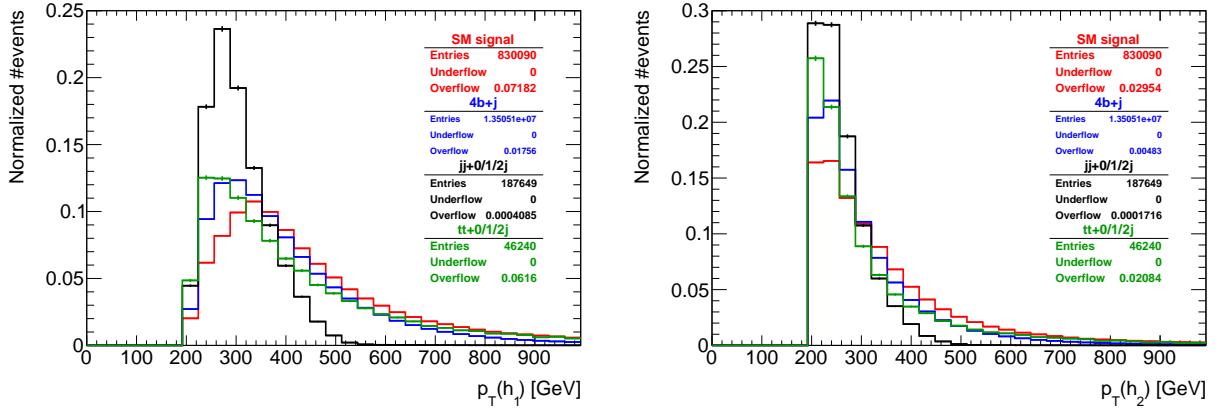


Figure 6.5: p_T distributions for the leading (left) and sub-leading (right) Higgs candidates. The signal is the SM $hh \rightarrow b\bar{b}$ process. The histograms are normalized to unit area.

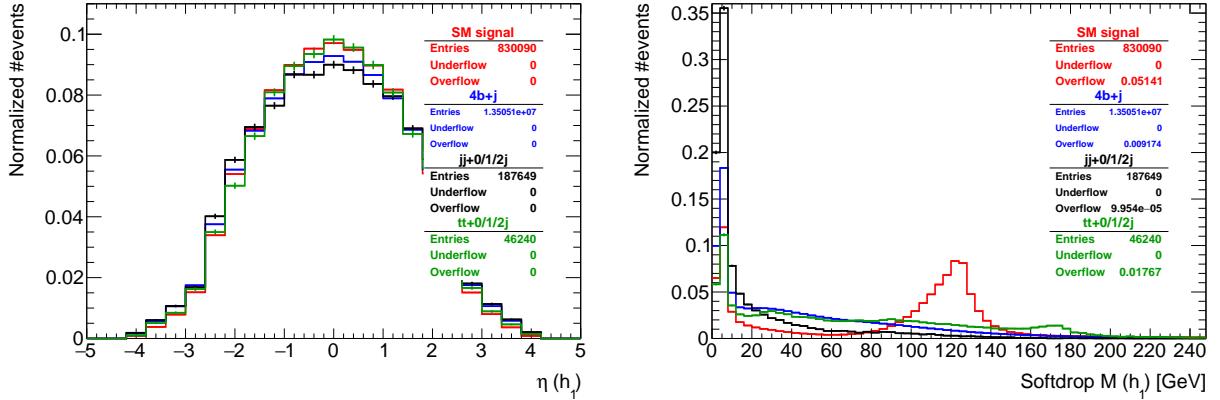


Figure 6.6: η distribution for the leading Higgs candidate (left) and softdrop mass distribution for the leading Higgs candidate (right). The signal is the SM $hh \rightarrow b\bar{b}$ process. The histograms are normalized to unit area.

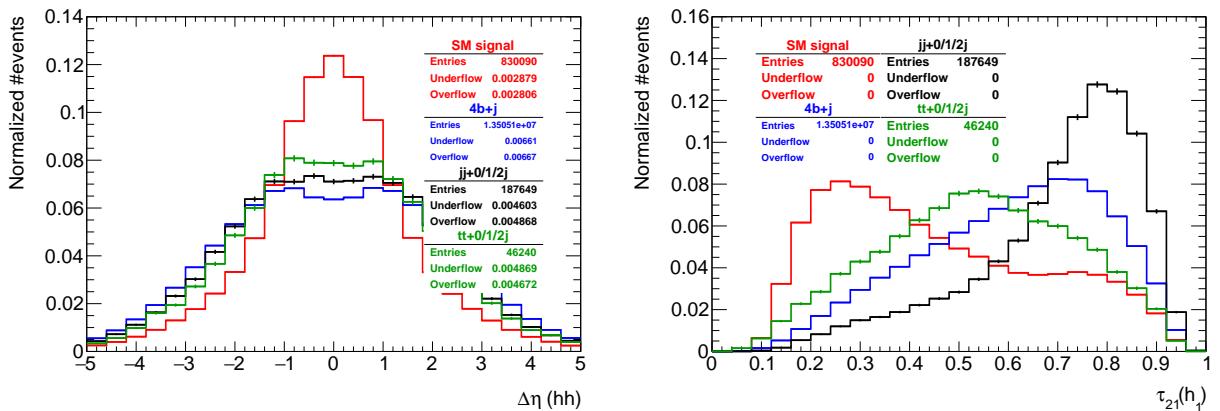


Figure 6.7: Distributions of the $\Delta\eta$ between the Higgs candidates (left) and of the τ_{21} variable for the leading Higgs candidate (right). The histograms are normalized to unit area.

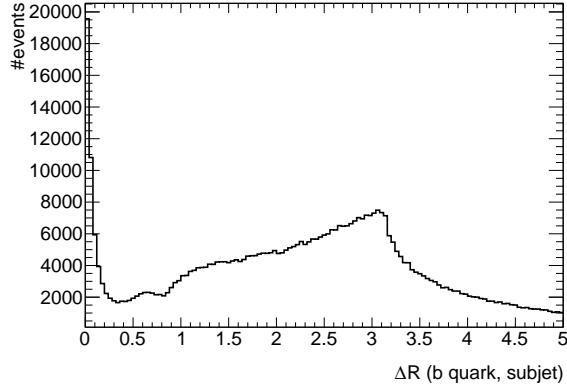


Figure 6.8: Minimum ΔR between b quarks and subjets of the $R = 0.8$ jets.

6.2 Analysis strategy

As already discussed, 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 $R = 0.8$, 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 implemented our own b-tagging algorithm that is described in section 6.2.1.

In this section, we present the baseline analysis based on cuts on the kinematic and substructure variables. An optimized version of the analysis is also presented and compared to the baseline. The optimization is based on plots of the significance (S/\sqrt{B}) as a function of the cut on a given variable. We also briefly describe how statistical uncertainties are taken into account.

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 6.8. 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 mis tag 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

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.

1514 summarized in table 6.1. The b-tagging probabilities are given in black and the c and light mis tagging
 1515 probabilities are given in blue and red, respectively. Note that a jet cannot be b-tagged if $|\eta| > 4$ or if its
 1516 momentum is smaller than 10 GeV or larger than 15000 GeV.

1517 In terms of the technical implementation, the b-tagging algorithm works as follows: for each subjet we
 1518 look for a truth-level b quark within $\Delta R = 0.3$ of the subjet and calculate the b-tagging and mis-tagging
 1519 efficiencies using the expressions in table 6.1, where p_T and η refer to the subjet. If a b quark is found we
 1520 generate a random number between 0 and 1. If it smaller than the b-tagging efficiency we consider the
 1521 subjet to be b-tagged. If the subjet is not b-tagged we look for truth-level c quarks within $\Delta R = 0.3$ of the
 1522 subjet. If one is found we generate a new random number between 0 and 1 and if the number is smaller
 1523 than the c mis-tag probability we consider the subjet to be b-tagged. If the subjet is not b-tagged we
 1524 generate a random number between 0 and 1 and consider the subjet b-tagged if the number is smaller
 1525 than the light mis-tag probability.

1526 6.2.2 Baseline analysis

1527 The analysis described in this section and later was developed using the sample simulated with the
 1528 default FCC-hh detector implementation. The same analysis selection was then applied to the samples
 1529 generated using the different detector configurations. This allows for a straightforward comparison of the
 1530 results in terms of the significances.

As a first step, we implemented a baseline analysis based on rectangular cuts on kinematic and substructure variables. Firstly, we apply cuts on the transverse momenta of the leading and sub leading Higgs candidates, $p_T(h_1)$ and $p_T(h_2)$, and of the Higgs pair, $p_T(hh)$:

$$1531 \quad p_T(h_1) > 400 \text{ GeV}, \quad p_T(h_2) > 350 \text{ GeV}, \quad p_T(hh) > 100 \text{ GeV}. \quad (6.1)$$

1532 These cuts follow from both distributions shown in figure ?? and guarantee that we are choosing events
 1533 for which the Higgs candidates are sufficiently boosted. In addition, a high threshold for the p_T of the
 1534 jets suppresses $t\bar{t}$ events because the decay products of the top quark are reconstructed in a single jet.
 In a more realistic analysis, these cuts would also be dictated by the trigger threshold.

From figure ??, on the left, we then apply a cut on the τ_{21} of the leading Higgs candidate to select

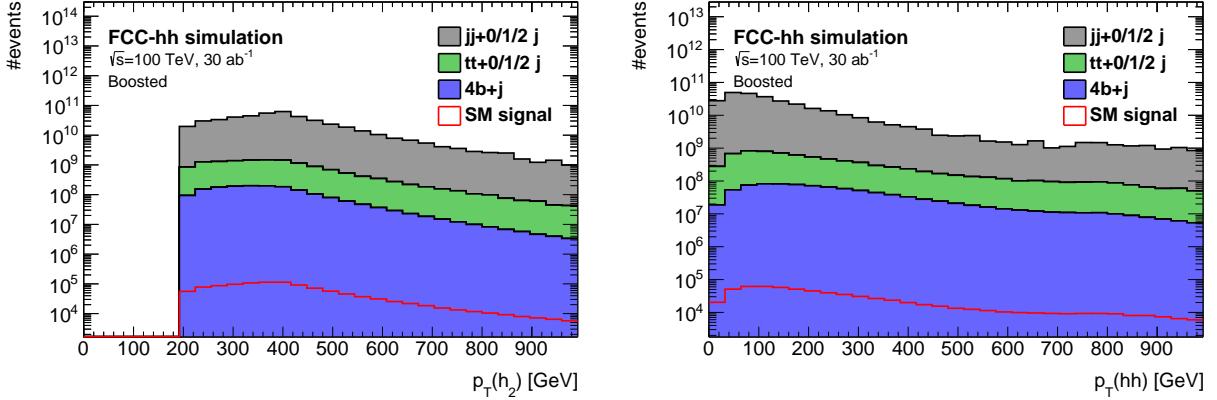


Figure 6.9: p_T distributions for the sub leading Higgs candidate (left) and for the Higgs pair (right). The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$.

jets that are more compatible with a two-prong structure:

$$\tau_{21}(h_1) < 0.55. \quad (6.2)$$

The cut on this variable works as a Higgs tagging method. Therefore we apply the same cut on the τ_{21} of the sub leading Higgs candidate although that does not necessarily follow from the distribution on the right of figure ???. Then, from the distribution on the right of figure ??, we place a cut on the second Fox-Wolfram momentum (defined in section 3.3.1) of the leading Higgs candidate, $H_2(h_1)$:

$$H_2(h_1) > 0.2 \quad (6.3)$$

1535 This substructure variable is particularly interesting because it helps suppress the $t\bar{t}$ background.

Finally, based on the distribution shown in figure ?? on the right, we apply a cut on the softdrop mass of both Higgs candidates, $M_{SD}(h_1, h_2)$. This cut is placed in a window around the nominal SM Higgs mass:

$$(100 \leq M_{SD}(h_1, h_2) \leq 135) \text{ GeV}. \quad (6.4)$$

1536 Using this analysis, we achieve a significance, S/\sqrt{B} , of $6.8 \pm 0.7(2.15 \pm 0.22)$ for an integrated luminosity of $30(3) \text{ ab}^{-1}$.
1537

1538 6.2.3 Optimization

The first approach to the optimization consists in placing successive cuts in the most relevant kinematic variables. The value of each cut is chosen in order to optimize the significance after that cut. As for the baseline analysis we start by looking at the transverse momenta of the leading and sub leading Higgs candidates. After the pre-selection cuts are applied, we scan the histograms of the p_T of the leading Higgs candidate for signal and backgrounds by placing a lower cut on this variable. For each value of the cut we integrate upwards in order to obtain the expected number of signal and background events after the cut. Using these numbers we calculate the significance, S/\sqrt{B} . The significance as a function

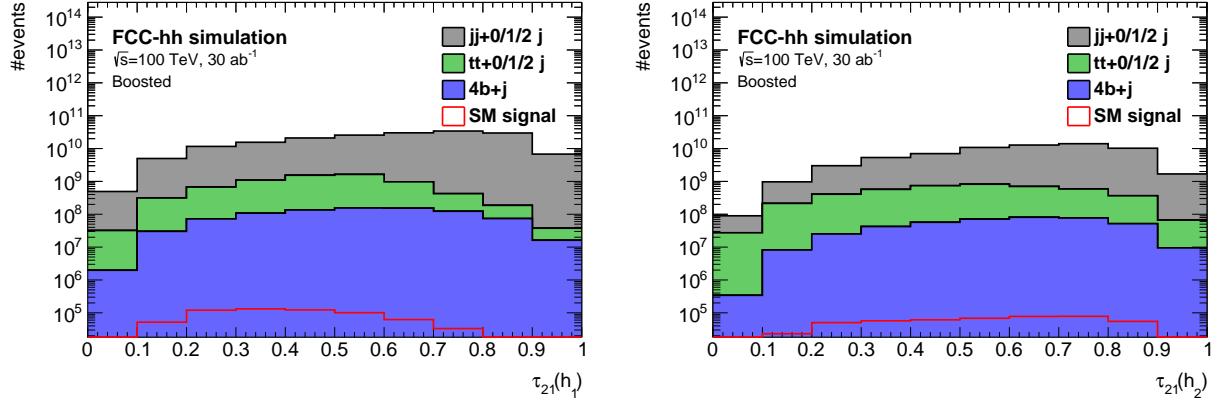


Figure 6.10: Distributions of the τ_{21} variable for the leading (left) and sub leading (right) Higgs candidates. The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$.

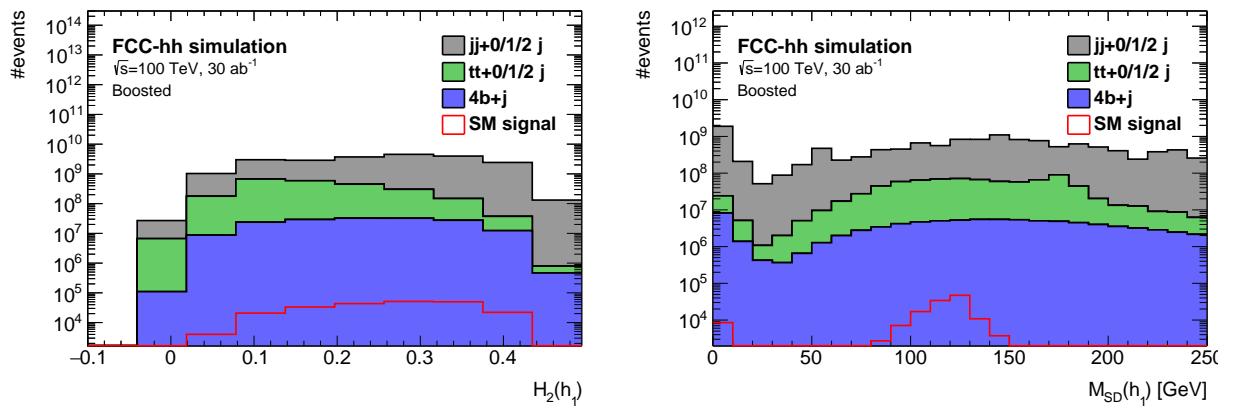


Figure 6.11: H_2 variable for the leading Higgs candidate (left) and softdrop mass distribution for the leading Higgs candidate. The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$.

of the lower cut on $p_T(h_1)$ is shown in figure ?? on the left. Based on this plot we choose the cut:

$$p_T(h_1) > 300 \text{ GeV}. \quad (6.5)$$

After placing the cut on $p_T(h_1)$ we do the same plot for the p_T of the sub leading Higgs candidate and of the Higgs pair. These are shown in figures ?? on the right and ?? on the left, respectively. From figure ?? on the right we see that a cut on $p_T(h_2)$ above 200 GeV is not favorable. Therefore, we do not apply any other cut on this variable. Based on the plot shown on the left of figure ?? we choose the cut:

$$p_T(hh) > 100 \text{ GeV}. \quad (6.6)$$

Following the cuts on the momentums we place a cut on the τ_{21} variable for both Higgs candidates. From all the variables considered during the optimization process this was the one that lead to the highest increase in the significance. From the definition of the τ_{21} variable we expect the signal to take lower values than the background. Therefore we optimize the cut on this variable by placing an upper cut and integrating the distribution backwards. The plot is shown in figure ?? on the right. From this plot we place the following cut on the τ_{21} of the leading Higgs candidate:

$$\tau_{21}(h_1) < 0.4. \quad (6.7)$$

1539 We apply exactly the same cut on the τ_{21} of the sub leading Higgs candidate. [JUSTIFY? HIGGS
1540 TAGGING].

Next we apply cuts on the $\Delta\eta$ between the Higgs candidates and on the second Fox-Wolfram momentum of the leading Higgs candidates, in this order:

$$|\Delta\eta(hh)| < 1.5 \quad H_2(h_1) > 0.2. \quad (6.8)$$

1541 These cut follow from the plots on figure ?? . The optimization plot for $\Delta\eta(hh)$ is obtained placing a cut
1542 in a window around zero which follows directly from the shape of the distributions in figure ?? on the left.

1543 For the $H_2(h_1)$ variable the cut is placed on the lower value.

1544 Finally we apply the mass cuts on the leading and sub leading Higgs candidates. These are the
1545 same that were applied in the baseline analysis. [OPTIMIZE THEM?].

1546 Using the optimized analysis we obtain a significance, S/\sqrt{B} , of $10 \pm 2(3.2 \pm 0.6)$ for an integrated
1547 luminosity of $30(3) \text{ ab}^{-1}$. This corresponds to an improvement of approximately 47% with respect to the
1548 baseline analysis.

6.2.4 Handling of uncertainties

1549 In this work, only statistical uncertainties are considered. This is due to the exploratory nature of this
1550 analysis but also because systematic uncertainties are usually related to the detector and object recon-
1551 struction specificities making them difficult to list and quantify for a detector that is not yet built.

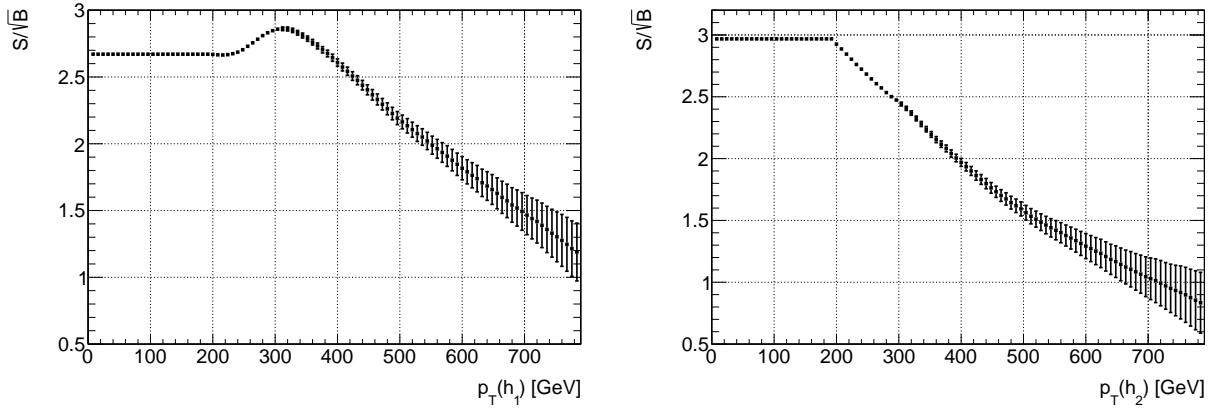


Figure 6.12: S/\sqrt{B} as a function of the cut on the p_T of the leading Higgs candidate, $p_T(h_1)$ (left) after the pre-selection cuts and of the sub leading Higgs candidate, $p_T(h_2)$ (right) after the cut on $p_T(h_1)$.

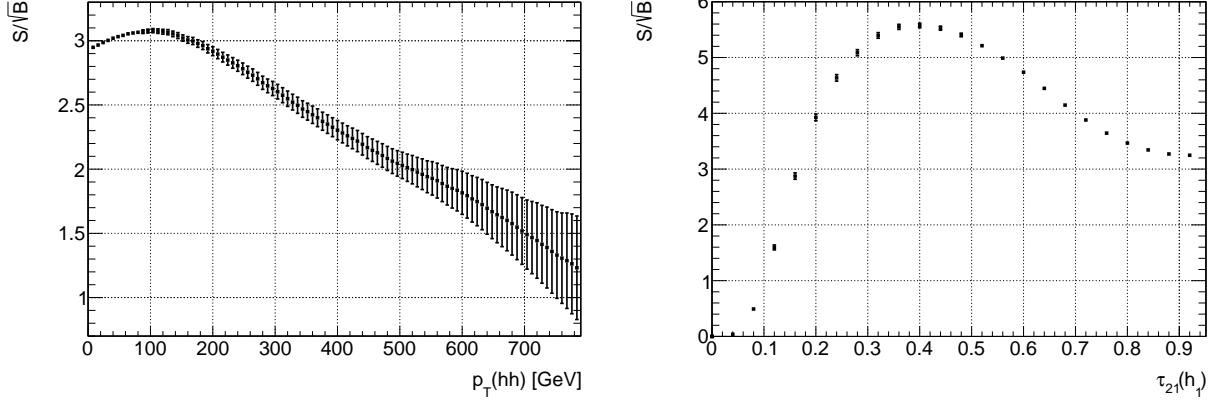


Figure 6.13: S/\sqrt{B} as a function of the cut on the p_T of the Higgs pair, $p_T(hh)$ (left) after the cut on $p_T(h_1)$ and as a function of the τ_{21} variable for leading Higgs candidate, $\tau_{21}(h_1)$ (right) after the cut on $p_T(hh)$.

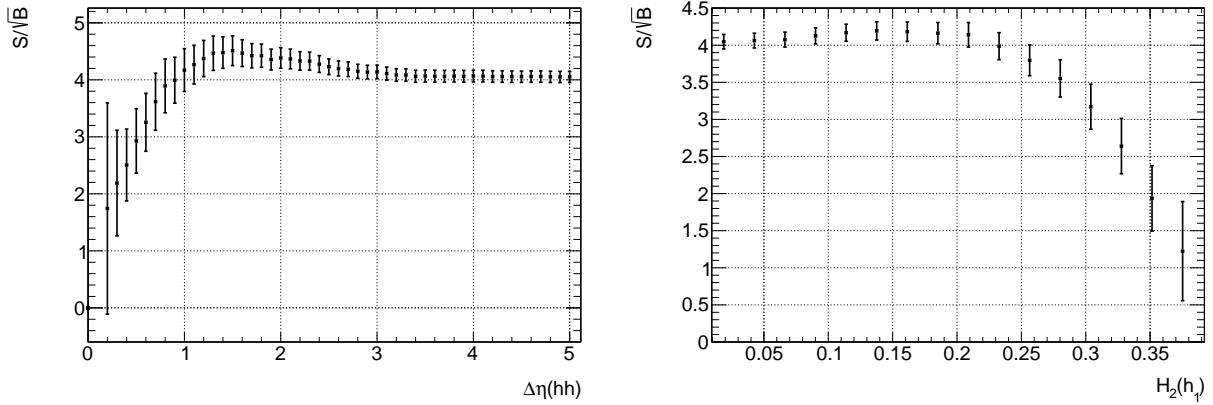


Figure 6.14: S/\sqrt{B} as a function of the cut on the $\Delta\eta$ between the Higgs candidates, $\Delta\eta(hh)$ (left) after the cuts on the transverse momentums and on the τ_{21} variables and as a function of the second Fox-Wolfram momentum for the leading Higgs candidate, $H_2(h_1)$ (right) after the cuts on the transverse momentums, τ_{21} and $\Delta\eta$.

For challenging analyses such as this one, the backgrounds usually have a much larger cross section than the signal meaning that we need a lot of Monte Carlo events to properly simulate them. Most of the times it is not feasible to generate as much events as we would need for a given luminosity. This is the case with this work, namely because we are targeting very high luminosities (of the order of tens of ab^{-1}). Take for example the multijet sample with $500 < H_T < 1000$. The cross section is approximately 10^7 pb which means that we expect a total of 3×10^{14} events for an integrated luminosity of 30 ab^{-1} . It is not feasible, within the time frame of this work and with the available computational resources, to generate this number of events. Therefore, we apply a weight, w , to each event given by:

$$w = \frac{\mathcal{L} \times \sigma}{N} \quad (6.9)$$

1553 where \mathcal{L} is the target integrated luminosity, σ is the cross section of the sample and N is the number of
1554 MC events generated.

If we assume that the number of events follows a Poisson distribution then the standard deviation (or uncertainty) is given by \sqrt{N} where N is the mean number of events. In order to normalize the number of events to a given luminosity we need to multiply the mean value and the uncertainty by the respective weight such that the number of normalized events, including the uncertainty, $N_{\text{norm}} \pm \Delta N_{\text{norm}}$, is given by:

$$N_{\text{norm}} \pm \Delta N_{\text{norm}} = w \times (N \pm \sqrt{N}) . \quad (6.10)$$

ΔN_{norm} can then be used in standard error propagation to compute the statistical uncertainty associated with any expression, in particular, with S/\sqrt{B} . The error associated with S/\sqrt{B} is given by:

$$\Delta \left(\frac{S}{\sqrt{B}} \right) = \left| \frac{1}{\sqrt{B}} \right|^2 (\Delta S)^2 + \left| -\frac{S}{B^{3/2}} \right|^2 (\Delta B)^2 \quad (6.11)$$

where ΔS and ΔB are the uncertainties associated with the number of signal and background events. ΔB is given by:

$$(\Delta B)^2 = \sum_i (\Delta B_i)^2 \quad (6.12)$$

1555 where the index i runs over all independent backgrounds.

1556

Chapter 7

1557

Results

1558 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
1559 benchmark luminosities, 30 ab^{-1} and 3 ab^{-1} (section 7.1). The statistical analysis used to extract the
1560 signal strength and to set limits on the Higgs boson triple coupling is also discussed. In section 7.2 we
1561 show how the significance of the analysis varies as a function of the granularity of the HCAL and/or the
1562 detector configuration. We also compare the results obtained using particle flow and pure calorimeter
1563 jets.

1564

7.1 Di-Higgs discovery potential at the FCC-hh

1565 The event selection of the baseline analysis for the search for $hh \rightarrow b\bar{b}b\bar{b}$ at the FCC with the baseline
1566 detector design is summarized in tables 7.1 and 7.2 for the signal samples (SM, DM mediator and type
1567 II 2HDM) and for the background samples ($4b + j$, $jj + 0/1/2j$ and $t\bar{t}$), respectively.

1568 From table 7.1, we see that for the BSM models the signal efficiency is higher than for the SM. It
1569 varies from 0.422 in the SM, to 0.487 in the DM mediator model and to 1.342 in the type II 2HDM.

Considering the SM production of Higgs pairs, the achieved significance is

$$S/\sqrt{B} = 6.8 \pm 0.7 \quad (2.15 \pm 0.27) \quad (7.1)$$

1570 for an integrated luminosity of $30(3) \text{ ab}^{-1}$. For $\mathcal{L} = 30 \text{ ab}^{-1}$, the significance is above the 5σ threshold
1571 while for $\mathcal{L} = 3 \text{ ab}^{-1}$ it is above the 3σ threshold. These results indicated that with the entire dataset that
1572 is expected to be accumulated by the FCC-hh detector it should be possible to observe the production
1573 of Higgs pairs.

1574 For a signal model that includes a 1 TeV dark matter mediator that can decay to pairs of SM Higgs
1575 bosons the achieve significance is $1.48 \pm 0.15(0.47 \pm 0.05)$ for an integrated luminosity of $30(3) \text{ ab}^{-1}$.
1576 The significance is well bellow the 3σ threshold for both luminosities. Therefore, we do not expect to
1577 be able to detect or exclude this signal at the FCC-hh. In this model, the coupling of the DM mediator
1578 to the Higgs pairs is small which means that the contribution from the box diagram dominates over the

Table 7.1: Cumulative efficiency, in percentage, of each event selection criterion for the signal samples (SM and 2HDM). 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.

Selection [FCC-hh]	SM	DM mediator	2HDM type II
Gen level	100 $(34638 \pm 16) \times 10^3$	100 $(65400 \pm 29) \times 10^2$	100 $(13977 \pm 7) \times 10^3$
$N(\text{b-tags}) \geq 4$	92.488	92.593	93.430
$p_T(j_1, j_2) \geq 200 \text{ GeV}$	16.6602 $(5751 \pm 6) \times 10^3$	17.033 $(11140 \pm 12) \times 10^2$	58.860 $8227 \pm 5 \times 10^3$
<hr/>			
$p_T(j_1) \geq 400 \text{ GeV}$	8.623	9.156	21.041
$p_T(j_2) \geq 350 \text{ GeV}$	5.709	6.161	13.202
$p_T(j_1 + j_2) \geq 100 \text{ GeV}$	4.648	4.968	9.624
$\tau_{21}(j_1, j_2) < 0.55$	1.705	1.878	4.057
$FW2(j_1) > 0.2$	1.281	1.421	3.267
$(100 < M_{SD}(j_1, j_2) < 135) \text{ GeV}$	0.422 $(1463 \pm 10) \times 10^2$	0.487 $(3188 \pm 20) \times 10$	1.342 $(1876 \pm 8) \times 10^2$

1579 resonant production (s-channel diagram), just like in the SM.

For the type II 2HDM the achieved significance is

$$S/\sqrt{B} = 8.7 \pm 0.9(2.76 \pm 0.28) \quad (7.2)$$

1580 for an integrated luminosity of $30(3) \text{ ab}^{-1}$. These results indicate that it can be observed (or excluded) at
1581 the FCC-hh. The high efficiency of this signal sample through the cuts, reflected in the high significance
1582 that is achieved, make it a very exciting and achievable benchmark.

1583 7.1.1 Statistical analysis

1584 7.1.2 Comparing with the ATLAS detector

1585 The event selection of the search for $hh \rightarrow b\bar{b}b\bar{b}$ with the ATLAS detector at a CM energy of 100 TeV is
1586 summarized in tables 7.3 and 7.4 for the signal and background samples, respectively.

1587 It is interesting to compare the results obtained with the FCC-hh default detector simulation with the
1588 ones obtained using the simulation of the ATLAS detector. These are summarized in table 7.5 in terms
1589 of the achieved significance for an integrated luminosity of 30 ab^{-1} .

1590 For all the signal models, the significance increases approximately 20% going from the ATLAS de-
1591 tector to the FCC-hh. For the SM signal, Nonetheless, using the ATLAS default detector configuration
1592 the achieved significance is already above 5σ : $S/\sqrt{B} = 5.6 \pm 0.6$.

Table 7.2: Cumulative efficiency, in percentage, of each event selection criterion for the background samples ($4b + j$, $jj + 0/1/2j$ and $t\bar{t} + 0/1/2j$). 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.

Selection [FCC-hh]	$4b + j$	$jj + 0/1/2j$	$t\bar{t}$
Gen level	100 $(49035 \pm 12) \times 10^6$	100 $(54698 \pm 17) \times 10^{10}$	100 $(22503 \pm 11) \times 10^8$
$N(\text{b-tags}) \geq 4$	75.819	3.963	53.495
$p_T(j_1, j_2) \geq 200 \text{ GeV}$	17.811 $(8734 \pm 5) \times 10^6$	0.742 $(4058 \pm 14) \times 10^9$	1.056 $(2377 \pm 11) \times 10^7$
<hr/>			
$p_T(j_1) \geq 400 \text{ GeV}$	7.008	0.183	0.446
$p_T(j_2) \geq 350 \text{ GeV}$	3.928	0.121	0.263
$p_T(j_1 + j_2) \geq 100 \text{ GeV}$	3.311	0.070	0.223
$\tau_{21}(j_1, j_2) < 0.55$	0.540	0.005	0.069
$FW2(j_1) > 0.2$	0.320	0.003	0.020
$(100 < M_{SD}(j_1, j_2) < 135) \text{ GeV}$	0.014 $(666 \pm 13) \times 10^4$	0.00008 $(4 \pm 4) \times 10^8$	0.0008 $(175 \pm 30) \times 10^5$

Table 7.3: Cumulative efficiency, in percentage, of each event selection criterion for the signal samples (SM and 2HDM). 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 ATLAS detector design, as implemented in Delphes.

Selection [ATLAS]	SM	DM mediator	2HDM type II
Gen level	100 $(34638 \pm 16) \times 10^3$	100 $(65400 \pm 29) \times 10^2$	100
$N(\text{b-tags}) \geq 4$	88.691	88.787	
$p_T(j_1, j_2) \geq 200 \text{ GeV}$	15.534 $(5381 \pm 6) \times 10^3$	15.94 $(10426 \pm 12) \times 10^2$	
<hr/>			
$p_T(j_1) \geq 400 \text{ GeV}$	7.997	8.499	
$p_T(j_2) \geq 350 \text{ GeV}$	5.283	5.704	
$p_T(j_1 + j_2) \geq 100 \text{ GeV}$	4.305	4.604	
$\tau_{21}(j_1, j_2) < 0.55$	1.648	1.809	
$FW2(j_1) > 0.2$	1.125	1.243	
$(100 < M_{SD}(j_1, j_2) < 135) \text{ GeV}$	0.323 $(1119 \pm 9) \times 10^2$	0.374 $(2446 \pm 18) \times 10$	

Table 7.4: Cumulative efficiency, in percentage, of each event selection criterion for the background samples ($4b + j$, $jj + 0/1/2j$ and $t\bar{t}+0/1/2 j$). 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 ATLAS detector design, as implemented in Delphes.

Selection [ATLAS]	$4b + j$	$jj + 0/1/2j$	$t\bar{t}$
Gen level	100 $(49035 \pm 15) \times 10^6$	100 $(54698 \pm 15) \times 10^{10}$	100 $(22503 \pm 9) \times 10^8$
$N(\text{b-tags}) \geq 4$	71.617	3.747	51.782
$p_T(j_1, j_2) \geq 200 \text{ GeV}$	16.301 $(7993 \pm 6) \times 10^6$	0.767 $(4193 \pm 12) \times 10^9$	0.984 $(2215 \pm 9) \times 10^7$
<hr/>			
$p_T(j_1) \geq 400 \text{ GeV}$	6.378	0.170	0.416
$p_T(j_2) \geq 350 \text{ GeV}$	3.560	0.112	0.245
$p_T(j_1 + j_2) \geq 100 \text{ GeV}$	3.000	0.064	0.206
$\tau_{21}(j_1, j_2) < 0.55$	0.545	0.008	0.064
$FW2(j_1) > 0.2$	0.272	0.003	0.016
$(100 < M_{SD}(j1, j2) < 135) \text{ GeV}$	0.010 $(496 \pm 14) \times 10^4$	0.00007 $(37 \pm 30) \times 10^7$	0.0007 $(165 \pm 25) \times 10^5$

Table 7.5: Significances achieved with the ATLAS and FCC-hh detector configurations for $\mathcal{L} = 30 \text{ ab}^{-1}$ for the three benchmark signal models: SM, 1 TeV DM mediator and 2HDM with $m_H = 900 \text{ GeV}$.

Signal sample	ATLAS	FCC-hh
SM	5.6 ± 0.6	6.8 ± 0.7
1 TeV DM mediator	1.23 ± 0.12	1.48 ± 0.15
2HDM type II	7.2 ± 0.7	8.7 ± 0.9

1593 **7.2 Hadronic calorimeter granularity studies for future colliders**

1594 **7.2.1 Resolution of jet mass and N-subjetiness**

1595 In this section we present the results that allow us to compare the different detector configurations.

1596 The softdrop mass of the leading Higgs candidate for the SM signal sample is shown in figure 7.1(a)
1597 for the different detector configurations. The same plot is shown in figure 7.1(b) for calorimeter jets. We
1598 see that the mass resolution increases as we increase the granularity. This effect is more pronounced
1599 when using pure calorimeter jets. [QUANTIFY]

1600 The invariant mass of the Higgs pair is shown in figure 7.2 for the SM signal. The same distribution
1601 is shown for the DM mediator and 2HDM signals in figures 7.3(a) and 7.3(b), respectively. For the SM,
1602 the mass spectrum of the Higgs pair does not show any resonance. On the one hand, the production
1603 is dominated by the box diagram. On the other hand, even if we only had the contribution from the
1604 triangle diagram, the Higgs that decays to a pair of Higgs bosons has to be off shell, therefore we
1605 expect a broad spectrum. For the DM mediator model, there is an enhancement with respect to the
1606 SM for $M(hh) \sim 1$ TeV. This is due to the contribution of the 1 TeV DM mediator decaying to pairs
1607 of SM Higgs bosons. Nonetheless, the spectrum is still quite broad because the box diagram stills
1608 gives a very large contribution. For the 2HDM, there is a clear peak at $M(hh) \sim 900$ GeV because
1609 resonant production dominates. The mass resolution of the resonances does not seem to increase as
1610 the granularity increases [?]. [CHECK WITH HCAL JETS]

1611 The τ_{21} variable for the leading Higgs candidate for the SM signal (filled lines) and for the $4b + j$
1612 background (dashed lines) is shown in figure 7.4(a) for the different detector configurations for particle
1613 flow jets. The same plot is show in figure 7.4(b) for calorimeter jets. The separation between signal
1614 and background increases. This was expected because an increase in the granularity of the hadronic
1615 calorimeter should help resolve better the substructure of boosted jets. This effect is more evident when
1616 using pure calorimenter jets. For particle flow jets, the distance between the maximum points of the
1617 two distributions varies from 0.4 for the ATLAS detector, ATLAS HCAL and ATLAS HCAL with $\eta \times 4$ to
1618 0.44 for the FCC HCAL with $\phi/2$ and to 0.46 for the remaining configurations. For calorimeter jets the
1619 distance between the maximum is always smaller than when using particle flow jets, considering the
1620 same detector configuration. Therefore, the separation is always bettter when using particle flow jets.
1621 The separation is the largest between the signal and the $jj + 0/1/2 j$ background, as expected. For this
1622 background, the distance between the maximums of the distributions varies from 0.48(24) to 0.56(0.48),
1623 for particle flow (calorimeter) jets.

1624 **7.2.2 Signal efficiency and significance**

1625 Figure 7.5(a) shows the signal efficiency for three signal models: SM (filled squares), 1 TeV DM mediator
1626 (empty squares) and type II 2HDM with $m_H = 900$ GeV, for eflow jets. The same plot is shown in figure
1627 7.5(b) for calorimeter jets. For eflow jets the efficiency increases as we increase the granularity, for all
1628 signal models. For HCAL jets there is a more complex dependence [WHY?]. The efficiency is higher

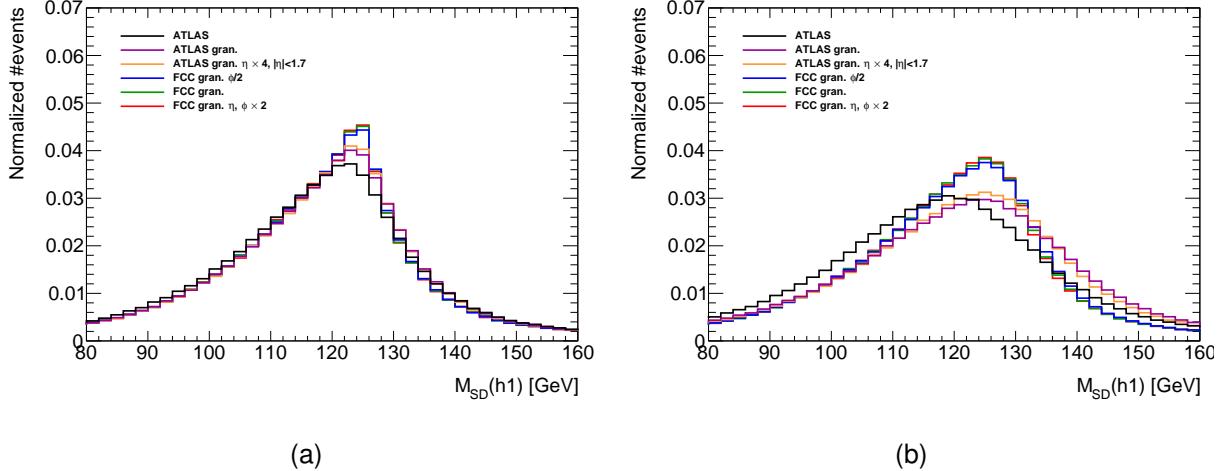


Figure 7.1: (a) Leading Higgs candidate softdrop mass after the pre-selection cuts, (b) The colors indicate the different detector configurations. The x axis range is from 80 GeV to 160 GeV in order to make the differences between the histograms more clear.

for both BSM models than for the SM. This is because these models were chosen to have very heavy particles decaying to a pair of highly boosted Higgs pairs.

The significances achieved with the optimized analysis as a function of the detector configuration for the SM signal, the DM mediator model and the type II 2HDM are shown in figures 7.6(a), 7.6(b) and 7.7, respectively. The uncertainty associated with each value of the significance is computed using standard error propagation. Only the statistical error is taken into account. Motivated by the small change in significance over the range of configurations that were tested we implemented exactly the same analysis but using HCAL jets instead of eflow jets. The results are shown in the same plots using triangular markers. On the one hand, the achieved significance is always smaller when using HCAL jets because we are not making use of the tracking information. On the other hand, when using HCAL jets, the significance changes a lot more over the configuration range. In particular, it increases as the HCAL granularity is increased. [QUANTIFY]

The small change in the significance when using eflow jets and the fact the change increases when using HCAL jets indicate that in the FCC-hh baseline detector design the resolution of the tracking system is so good, in particular, so much better than the HCAL resolution [NUMBERS] that it is the limiting factor.

[COMPARE THE DIFFERENT SIGNAL SAMPLES IN TERMS OF THE CHANGE IN SIGNIFICANCE]

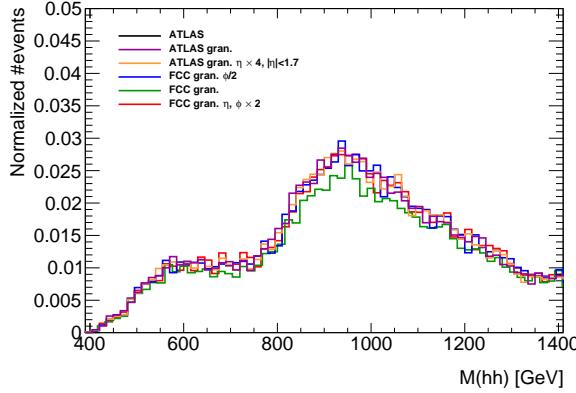


Figure 7.2

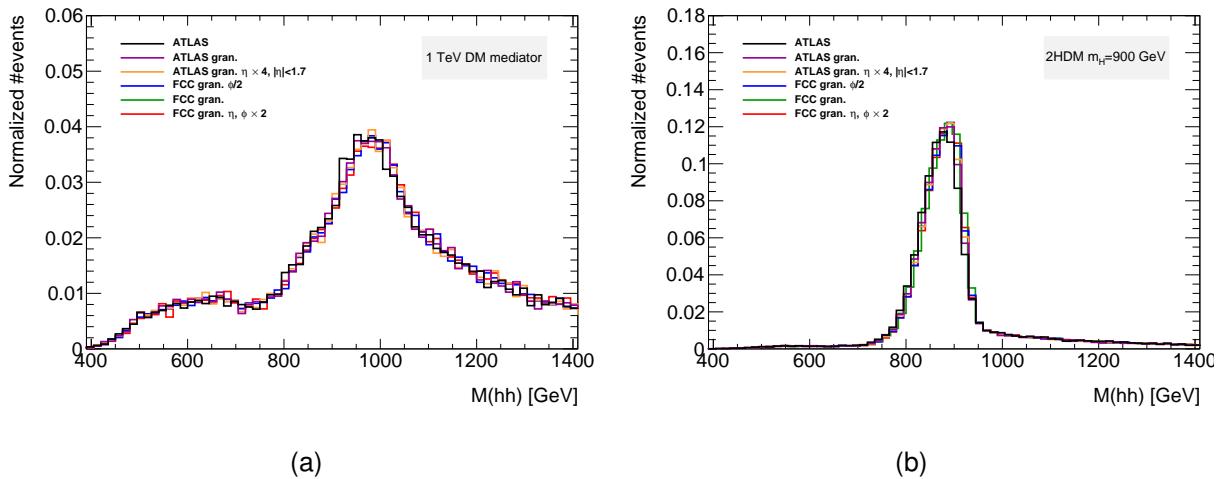


Figure 7.3: (a) Invariant mass of the Higgs pair for the 1 TeV DM mediator, (b) Invariant mass of the Higgs pair for the 2HDM with $m_H = 900$ GeV. The plots show include the events that pass all the analysis cuts. The colors indicate the different detector configurations.

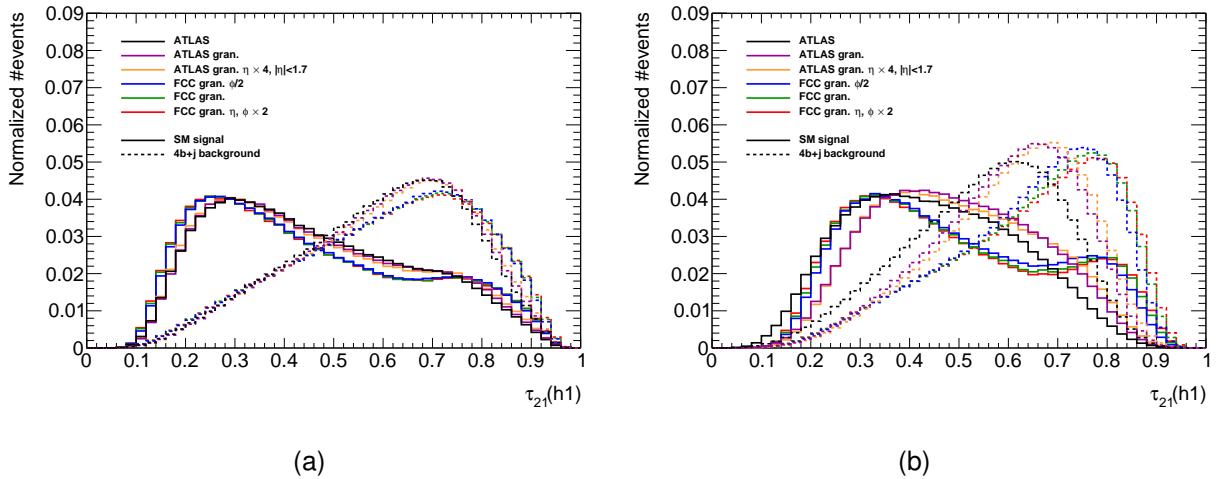


Figure 7.4: (a) Leading Higgs candidate τ_{21} distribution for eflow jets, (b) Leading Higgs candidate τ_{21} distribution for HCAL jets. The colors indicate the different detector configurations. The distributions are shown for the signal (filled lines) and for the $4b + j$ background (dashed lines).

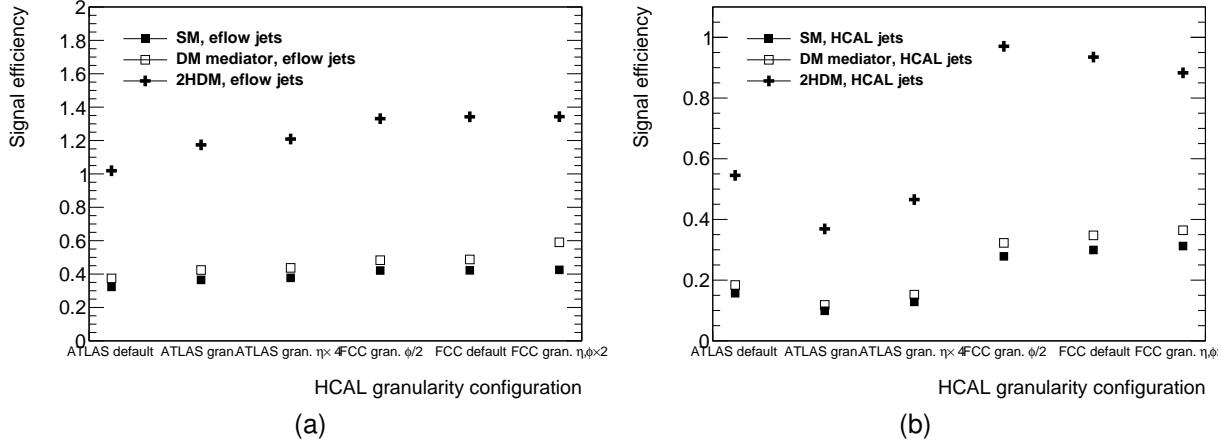


Figure 7.5: (a) Signal efficiency as a function of the detector configuration for particle flow jets, (b) Signal efficiency as a function of the detector configuration for calorimeter jets. Three signal models are shown: SM (filled squares), 1 TeV DM mediator (empty squares) and type II 2HDM with $m_H = 900$ GeV (crosses). The error bars are drawn but are smaller than the markers.

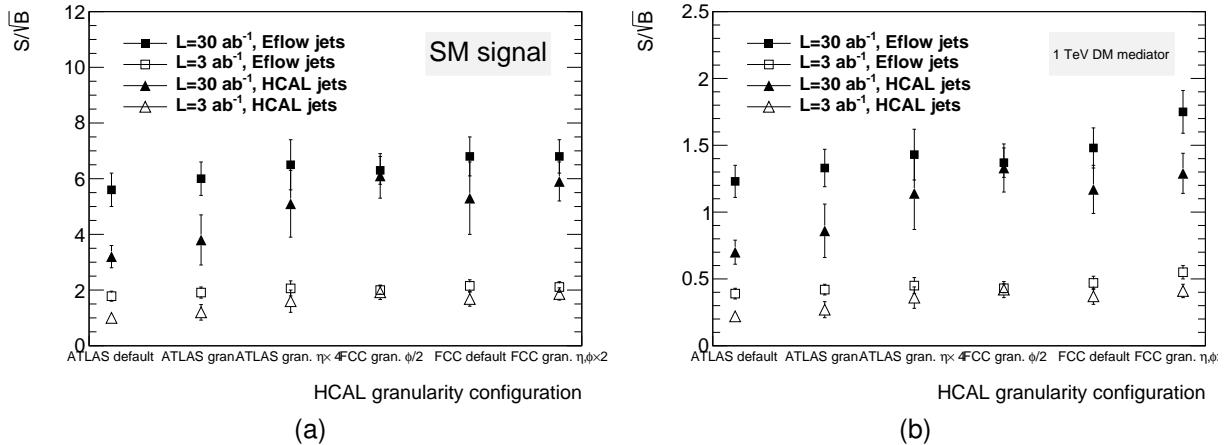


Figure 7.6: (a) Significance as a function of the detector configuration for the SM signal, (b) Significance as a function of the detector configuration for the DM mediator model. The significances are shown for $\mathcal{L} = 30 \text{ ab}^{-1}$ (filled) and $\mathcal{L} = 3 \text{ ab}^{-1}$ (empty) and for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles).

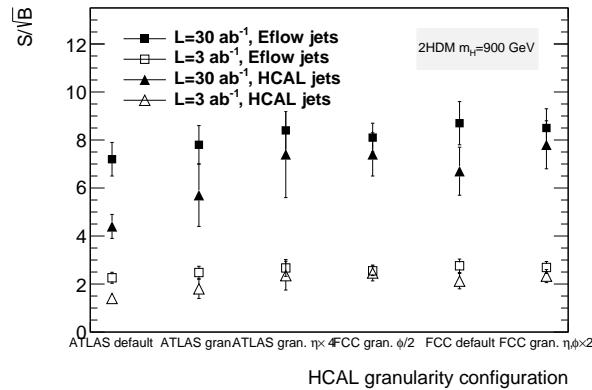


Figure 7.7: Significance as a function of the detector configuration for the 2HDM with $m_H = 900$ GeV. The significances are shown for $\mathcal{L} = 30 \text{ ab}^{-1}$ (filled) and $\mathcal{L} = 3 \text{ ab}^{-1}$ (empty) and for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles).

1646 **Chapter 8**

1647 **Conclusions**

1648 This thesis presented a feasibility study targeting the search for Higgs pair production in the $pp \rightarrow hh \rightarrow$
1649 $b\bar{b}b\bar{b}$ channel at a center of mass energy of $\sqrt{s} = 100$ TeV. The analysis targeted a boosted kinematic
1650 regime. The impact of the granularity of the hadronic calorimeter in the achieved significance was
1651 analyzed.

1652 The analysis consisted of a cut-based event selection designed to efficiently select signal events
1653 consistent with a boosted topology. In addition to standard kinematic variables, jet substructure variables
1654 are used to further suppress the background. The analysis required the existence of at least two b-
1655 tagged large- R jets consistent with having two subjets. ... Multijet production through QCD interactions
1656 and $t\bar{t}$ production constitute the main backgrounds.

1657 In addition to SM Higgs pair production, two BSM benchmark processes were considered: a 1
1658 TeV dark matter mediator and a heavy Higgs boson with $m_H = 900$ GeV in the framework of the
1659 CP-conserving type II 2HDM, both decaying to pairs of SM Higgs bosons. The existence of heavy
1660 resonances is expected to enhance Higgs pair production with respect to the SM.

1661 For the hadronic Future Circular Collider (FCC-hh), the significance obtained was $S/\sqrt{B} = \dots(\dots)$
1662 for an integrated luminosity of $\mathcal{L} = 30$ (3) ab^{-1} . For $\mathcal{L} = 30 \text{ ab}^{-1}$, the value is above the observation
1663 threshold (5σ) which indicates that the full dataset that is expected to be collected during the operation of
1664 the FCC-hh should be enough to claim the observation of Higgs pair production, assuming SM produc-
1665 tion. With the ATLAS detector simulation, the achieved significance is $S/\sqrt{B} = \dots(\dots)$ for an integrated
1666 luminosity of $\mathcal{L} = 30$ (3) ab^{-1} .

Bibliography

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

- 1693 [11] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Phys. Rev.*, 109:1015–1017,
 1694 Feb 1958. doi: 10.1103/PhysRev.109.1015. URL <https://link.aps.org/doi/10.1103/PhysRev.109.1015>.
- 1695
- 1696 [12] S. M. Bilenky and J. Hosek. GLASHOW-WEINBERG-SALAM THEORY OF ELECTROWEAK IN-
 1697 TERATIONS AND THE NEUTRAL CURRENTS. *Phys. Rept.*, 90:73–157, 1982. doi: 10.1016/0370-1573(82)90016-3.
- 1698
- 1699 [13] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, P. Torrielli, E. Vryonidou, and M. Zaro.
 1700 Higgs pair production at the LHC with NLO and parton-shower effects. *Phys. Lett.*, B732:142–149,
 1701 2014. doi: 10.1016/j.physletb.2014.03.026.
- 1702
- 1703 [14] M. Mangano. *Physics at the FCC-hh, a 100 TeV pp collider*. CERN Yellow Reports: Monographs.
 CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2270978>.
- 1704
- 1705 [15] T. Plehn, M. Spira, and P. M. Zerwas. Pair production of neutral Higgs particles in gluon-
 1706 gluon collisions. *Nucl. Phys.*, B479:46–64, 1996. doi: 10.1016/0550-3213(96)00418-X, 10.1016/S0550-3213(98)00406-4. [Erratum: Nucl. Phys.B531,655(1998)].
- 1707
- 1708 [16] S. Borowka, N. Greiner, G. Heinrich, S. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke.
 1709 Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass
 1710 Dependence. *Phys. Rev. Lett.*, 117(1):012001, 2016. doi: 10.1103/PhysRevLett.117.079901, 10.
 1711 1103/PhysRevLett.117.012001. [Erratum: Phys. Rev. Lett.117,no.7,079901(2016)].
- 1712
- 1713 [17] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke. Full top quark
 1714 mass dependence in Higgs boson pair production at NLO. *JHEP*, 10:107, 2016. doi: 10.1007/
 1715 JHEP10(2016)107.
- 1716
- 1717 [18] H. C. van de Hulst, E. Raimond, and H. van Woerden. Rotation and density distribution of the
 1718 Andromeda nebula derived from observations of the 21-cm line. 14:1, Nov. 1957.
- 1719
- 1720 [19] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567,
 1721 1998. doi: 10.1103/PhysRevLett.81.1562.
- 1722
- 1723 [20] 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.
- 1724
- 1725 [21] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current
 1726 interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002. doi:
 10.1103/PhysRevLett.89.011301.
- 1727
- 1728 [22] B. Gripaios. Lectures on Physics Beyond the Standard Model. 2015.
- 1729
- 1730 [23] N. M. A. 2018. Popular information, 2004. URL <https://www.nobelprize.org/prizes/physics/2004/popular-information/>.
- 1731

- 1727 [24] V. Branchina, F. Contino, and P. M. Ferreira. Electroweak vacuum lifetime in two Higgs doublet
 1728 models. 2018.
- 1729 [25] J. F. Gunion and H. E. Haber. The CP conserving two Higgs doublet model: The Approach to the
 1730 decoupling limit. *Phys. Rev.*, D67:075019, 2003. doi: 10.1103/PhysRevD.67.075019.
- 1731 [26] H. Sun, Y.-J. Zhou, and H. Chen. Constraints on large-extra-dimensions model through 125-
 1732 GeV Higgs pair production at the LHC. *Eur. Phys. J.*, C72:2011, 2012. doi: 10.1140/epjc/
 1733 s10052-012-2011-4.
- 1734 [27] P. Huang, A. Joglekar, M. Li, and C. E. M. Wagner. Corrections to di-Higgs boson production
 1735 with light stops and modified Higgs couplings. *Phys. Rev.*, D97(7):075001, 2018. doi: 10.1103/
 1736 PhysRevD.97.075001.
- 1737 [28] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with
 1738 the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- 1739 [29] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment
 1740 at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
- 1741 [30] Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector. Technical Report
 1742 ATLAS-CONF-2018-036, CERN, Geneva, Jul 2018. URL <http://cds.cern.ch/record/2630338>.
- 1743 [31] Observation of Higgs boson decay to bottom quarks. Technical Report arXiv:1808.08242. CMS-
 1744 HIG-18-016-003, CERN, Geneva, Aug 2018. URL <https://cds.cern.ch/record/2636067>. * Tem-
 1745 porary entry *.
- 1746 [32] A. M. Sirunyan et al. Observation of the Higgs boson decay to a pair of τ leptons with the CMS
 1747 detector. *Phys. Lett.*, B779:283–316, 2018. doi: 10.1016/j.physletb.2018.02.004.
- 1748 [33] A. M. Sirunyan et al. Observation of $t\bar{t}H$ production. *Phys. Rev. Lett.*, 120:231801, 2018. doi:
 1749 10.1103/PhysRevLett.120.231801.
- 1750 [34] M. Aaboud et al. Observation of Higgs boson production in association with a top quark pair at the
 1751 LHC with the ATLAS detector. 2018.
- 1752 [35] Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels with
 1753 $\sqrt{s}=13\text{TeV}$ pp collisions using the ATLAS detector. Technical Report ATLAS-CONF-2017-046,
 1754 CERN, Geneva, Jul 2017. URL <http://cds.cern.ch/record/2273853>.
- 1755 [36] G. Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and
 1756 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015. doi: 10.1103/
 1757 PhysRevLett.114.191803.
- 1758 [37] M. Aaboud et al. Measurement of the W -boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS
 1759 detector. *Eur. Phys. J.*, C78(2):110, 2018. doi: 10.1140/epjc/s10052-017-5475-4.

- 1760 [38] R. Nisius. Measurements of the top quark mass with the ATLAS detector. *PoS*, EPS-HEP2017:
1761 453, 2017. doi: 10.22323/1.314.0453.
- 1762 [39] Measurement of the top quark mass with muon+jets final states in pp collisions at $\sqrt{s} = 13$ TeV.
1763 Technical Report CMS-PAS-TOP-16-022, CERN, Geneva, 2017. URL [http://cds.cern.ch/](http://cds.cern.ch/record/2255834)
1764 record/2255834.
- 1765 [40] CERN. The high-luminosity lhc. <http://cds.cern.ch/record/2114693>, 2015. Accessed
1766 2018/01/11.
- 1767 [41] T. Koffas. ATLAS Higgs physics prospects at the high luminosity LHC. *PoS*, ICHEP2016:426, 2017.
- 1768 [42] A. M. Henriques Correia. The ATLAS Tile Calorimeter. Technical Report ATL-TILECAL-PROC-
1769 2015-002, CERN, Geneva, Mar 2015. URL <https://cds.cern.ch/record/2004868>.
- 1770 [43] ATLAS liquid argon calorimeter: Technical design report. 1996.
- 1771 [44] M. Cacciari, G. P. Salam, and G. Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008. doi:
1772 10.1088/1126-6708/2008/04/005.
- 1773 [45] Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run. Technical Report ATL-
1774 PHYS-PUB-2016-012, CERN, Geneva, Jun 2016. URL <https://cds.cern.ch/record/2160731>.
- 1775 [46] Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data.
1776 Technical Report ATLAS-CONF-2011-102, CERN, Geneva, Jul 2011. URL [https://cds.cern.](https://cds.cern.ch/record/1369219)
1777 ch/record/1369219.
- 1778 [47] G. Aad et al. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics.
1779 2009.
- 1780 [48] A. Ruiz-Martinez and A. Collaboration. The Run-2 ATLAS Trigger System. Technical Report ATL-
1781 DAQ-PROC-2016-003, CERN, Geneva, Feb 2016. URL <https://cds.cern.ch/record/2133909>.
- 1782 [49] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*,
1783 D48:3160–3166, 1993. doi: 10.1103/PhysRevD.48.3160.
- 1784 [50] M. Wobisch and T. Wengler. Hadronization corrections to jet cross-sections in deep inelastic scat-
1785 tering. In *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany,*
1786 1998-1999, pages 270–279, 1998. URL [https://inspirehep.net/record/484872/files/arXiv:](https://inspirehep.net/record/484872/files/arXiv-hep-ph_9907280.pdf)
1787 [hep-ph_9907280.pdf](https://inspirehep.net/record/484872/files/hep-ph_9907280.pdf).
- 1788 [51] M. Cacciari, G. P. Salam, and G. Soyez. The Anti- $k(t)$ jet clustering algorithm. *JHEP*, 04:063, 2008.
1789 doi: 10.1088/1126-6708/2008/04/063.
- 1790 [52] A. Altheimer et al. Boosted objects and jet substructure at the LHC. Report of BOOST2012, held
1791 at IFIC Valencia, 23rd-27th of July 2012. *Eur. Phys. J.*, C74(3):2792, 2014. doi: 10.1140/epjc/
1792 s10052-014-2792-8.

- 1793 [53] G. A. et al. Performance of jet substructure techniques for large-r jets in proton-proton collisions
 1794 at $\sqrt{s} = 7$ tev using the atlas detector. *Journal of High Energy Physics*, 2013(9), Sep 2013. doi:
 1795 10.1007/JHEP09(2013)076. URL [https://doi.org/10.1007/JHEP09\(2013\)076](https://doi.org/10.1007/JHEP09(2013)076).
- 1796 [54] J. Thaler and K. Van Tilburg. Identifying Boosted Objects with N-subjettiness. *JHEP*, 03:015, 2011.
 1797 doi: 10.1007/JHEP03(2011)015.
- 1798 [55] C. Chen. New approach to identifying boosted hadronically-decaying particle using jet substructure
 1799 in its center-of-mass frame. *Phys. Rev.*, D85:034007, 2012. doi: 10.1103/PhysRevD.85.034007.
- 1800 [56] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam. Jet substructure as a new Higgs
 1801 search channel at the LHC. *Phys. Rev. Lett.*, 100:242001, 2008. doi: 10.1103/PhysRevLett.100.
 1802 242001.
- 1803 [57] I. Hinchliffe, A. Kotwal, M. L. Mangano, C. Quigg, and L.-T. Wang. Luminosity goals for a 100-TeV
 1804 pp collider. *Int. J. Mod. Phys.*, A30(23):1544002, 2015. doi: 10.1142/S0217751X15440029.
- 1805 [58] N. Arkani-Hamed, T. Han, M. Mangano, and L.-T. Wang. Physics opportunities of a 100 TeV pro-
 1806 tonproton collider. *Phys. Rept.*, 652:1–49, 2016. doi: 10.1016/j.physrep.2016.07.004.
- 1807 [59] C. Helsens. Fcc/he-lhc physics/performance studies towards the cdr. URL: https://indico.cern.ch/event/675588/contributions/2765045/attachments/1572548/2484309/Clement_FCC.pdf,
 1808 2017.
- 1810 [60] T. an FCC-hh Detector Magnet System. E. bielert, 2018. URL <https://ep-news.web.cern.ch/content/towards-fcc-hh-detector-magnet-system>.
- 1811 [61] J. Faltova. Electromagnetic calorimeter using lar technology for fcc-hh. URL: https://indico.cern.ch/event/556692/contributions/2465166/attachments/1468150/2272205/ecal_fccBerlin.pdf, 2017.
- 1812 [62] A. Zaborowska. Electromagnetic calorimetry based on liquid argon for the fcc-hh experiments.
 1813 URL: https://indico.cern.ch/event/656491/contributions/2923535/attachments/1631506/2601159/FCChh_LArCalo_Zaborowska.pdf, 2018.
- 1814 [63] C. Neubser. Performance studies and requirements on the calorimeters for a fcc-hh
 1815 experiment. URL: <https://indico.ihep.ac.cn/event/6387/session/14/contribution/117/material/slides/0.pdf>, 2017.
- 1816 [64] J. Baglio, A. Djouadi, R. Grber, M. M. Mhlleitner, J. Quevillon, and M. Spira. The measurement
 1817 of the Higgs self-coupling at the LHC: theoretical status. *JHEP*, 04:151, 2013. doi: 10.1007/
 1818 JHEP04(2013)151.
- 1819 [65] M. Aaboud et al. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–
 1820 proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev.*, D94(5):052002, 2016. doi:
 1821 10.1103/PhysRevD.94.052002.

- 1827 [66] A. M. Sirunyan et al. Search for a massive resonance decaying to a pair of Higgs bosons in the
 1828 four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- 1829 [67] G. Aad et al. Searches for Higgs boson pair production in the $hh \rightarrow bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$
 1830 channels with the ATLAS detector. *Phys. Rev.*, D92:092004, 2015. doi: 10.1103/PhysRevD.92.
 1831 092004.
- 1832 [68] A. M. Sirunyan et al. Search for Higgs boson pair production in events with two bottom quarks and
 1833 two tau leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- 1834 [69] V. Khachatryan et al. Search for two Higgs bosons in final states containing two photons and
 1835 two bottom quarks in proton-proton collisions at 8 TeV. *Phys. Rev.*, D94(5):052012, 2016. doi:
 1836 10.1103/PhysRevD.94.052012.
- 1837 [70] G. Aad et al. Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision
 1838 Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector. *Phys. Rev. Lett.*, 114(8):081802, 2015. doi: 10.1103/
 1839 PhysRevLett.114.081802.
- 1840 [71] A. M. Sirunyan et al. Search for resonant and nonresonant Higgs boson pair production in the
 1841 $b\bar{b}lnlnu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2017.
- 1842 [72] A. Collaboration. Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the
 1843 ATLAS experiment. Technical Report ATLAS-COM-CONF-2018-048, CERN, Geneva, Aug 2018.
 1844 URL <https://cds.cern.ch/record/2634999>.
- 1845 [73] G. Aad and t. . others.
- 1846 [74] J. K. Behr, D. Bortoletto, J. A. Frost, N. P. Hartland, C. Issever, and J. Rojo. Boosting Higgs pair
 1847 production in the $b\bar{b}b\bar{b}$ final state with multivariate techniques. *Eur. Phys. J.*, C76(7):386, 2016. doi:
 1848 10.1140/epjc/s10052-016-4215-5.
- 1849 [75] S. Banerjee, C. Englert, M. L. Mangano, M. Selvaggi, and M. Spannowsky. $hh +$ jet production at
 1850 100 TeV. *Eur. Phys. J.*, C78(4):322, 2018. doi: 10.1140/epjc/s10052-018-5788-y.
- 1851 [76] N. P. Hartland. Double Higgs production from $HH \rightarrow (b\bar{b})(b\bar{b})$ at a 100 TeV hadron collider. *PoS*,
 1852 DIS2016:115, 2016.
- 1853 [77] S. V. Chekanov, M. Beydler, A. V. Kotwal, L. Gray, S. Sen, N. V. Tran, S. S. Yu, and J. Zuzelski. Initial
 1854 performance studies of a general-purpose detector for multi-TeV physics at a 100 TeV pp collider.
 1855 *JINST*, 12(06):P06009, 2017. doi: 10.1088/1748-0221/12/06/P06009.
- 1856 [78] S. Chekanov. Boosted jets in high-granularity calorimeter at a 100 TeV pp collider. BOOST 2017,
 1857 2017.
- 1858 [79] S. Chekanov. Physics requirements for a Hadron Calorimeter for a 100 TeV proton-proton collider.
 1859 FCC week 2015, 2015.

- 1860 [80] S. Chekanov. Simulation of a high-granular hadronic calorimeter for multi-TeV physics. FCC week
1861 2016, 2016.
- 1862 [81] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer.
- 1863 [82] T. Sjostrand, S. Mrenna, and P. Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys.*
1864 *Commun.*, 178:852–867, 2008. doi: 10.1016/j.cpc.2008.01.036.
- 1865 [83] J. de Favereau, C. Delaere, P. Demin, A. Giammancò, V. Lematre, A. Mertens, and M. Selvaggi.
1866 DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:
1867 057, 2014. doi: 10.1007/JHEP02(2014)057.
- 1868 [84] E. Boos et al. Generic user process interface for event generators. In *Physics at TeV colliders.*
1869 *Proceedings, Euro Summer School, Les Houches, France, May 21-June 1, 2001*, 2001. URL
1870 <http://lss.fnal.gov/archive/preprint/fermilab-conf-01-496-t.shtml>.
- 1871 [85] M. Cacciari, G. P. Salam, and G. Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012. doi:
1872 10.1140/epjc/s10052-012-1896-2.
- 1873 [86] D. Buskulic et al. Performance of the ALEPH detector at LEP. *Nucl. Instrum. Meth.*, A360:481–506,
1874 1995. doi: 10.1016/0168-9002(95)00138-7.
- 1875 [87] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical
1876 Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009. URL <http://cds.cern.ch/record/1194487>.
- 1878 [88] FCC study group. Fccsw. <https://github.com/HEP-FCC/FCCSW>, 2018.
- 1879 [89] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing
1880 applications. *Comput. Phys. Commun.*, 140:45–55, 2001. doi: 10.1016/S0010-4655(01)00254-5.
- 1881 [90] F. Gaede, B. Hegner, and P. Mato. Podio: An event-data-model toolkit for high energy physics
1882 experiments. *Journal of Physics: Conference Series*, 898(7):072039, 2017. URL <http://stacks.iop.org/1742-6596/898/i=7/a=072039>.
- 1884 [91] FCC-hh study group. Eventproducer. <https://github.com/FCC-hh-framework/EventProducer>,
1885 2018.
- 1886 [92] N. D. Christensen and C. Duhr. FeynRules - Feynman rules made easy. *Comput. Phys. Commun.*,
1887 180:1614–1641, 2009. doi: 10.1016/j.cpc.2009.02.018.
- 1888 [93] T. general Two-Higgs-Doublet Model. C. degrande, c. duhr, m. herquet, 2018. URL <http://feynrules.irmp.ucl.ac.be/wiki/2HDM>.
- 1890 [94] C. Degrande. Automatic evaluation of UV and R2 terms for beyond the Standard Model La-
1891 grangians: a proof-of-principle. *Comput. Phys. Commun.*, 197:239–262, 2015. doi: 10.1016/j.
1892 cpc.2015.08.015.

₁₈₉₃ [95] H. E. Haber and O. Stl. New LHC benchmarks for the \mathcal{CP} -conserving two-Higgs-doublet model. *Eur. Phys. J.*, C75(10):491, 2015. doi: 10.1140/epjc/s10052-015-3697-x, 10.1140/epjc/s10052-016-4151-4. [Erratum: *Eur. Phys. J.* C76,no.6,312(2016)].

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

1898 **Appendix A**

1899 **Jet radius discussion**

1900 It is important to discuss if the radius of the jet that is used to reconstruct the boosted Higgs boson
1901 candidates ($R = 0.8$) is the appropriate one. The question arises because in ATLAS boosted objects
1902 with $p_T \sim 200$ GeV are reconstructed using $R = 1.0$ jets. In this work, we work with a CM energy of 100
1903 TeV and require that the two leading jets have $p_T > 200$ GeV and use jets with $R = 0.8$. It is necessary
1904 to understand if these jets are large enough to fully reconstruct the Higgs candidates.

1905 As a first approximation we compute the angle between the b quarks produced by the decay of
1906 a Higgs boson. We assume that the b quarks are massless and that the Higgs moves only in the
1907 transverse plane (perpendicular to the beam pipe) such that it has no longitudinal momentum and the
1908 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
1909 smaller than the jet's diameter (1.6) and therefore the two b quarks can both be contained inside the jet
1910 and the Higgs boson fully reconstructed.

1911 Another test we can make is to compute the
1912 ΔR between the b quarks coming from the lead-
1913 ing Higgs candidate with p_T larger than a given
1914 value using truth level information. In figure A.1
1915 we show the distribution of $\Delta R(b, \bar{b})$ for p_T of
1916 the leading Higgs candidate larger than 200 GeV
1917 (solid blue) and 300 GeV (solid black). To obtain
1918 the dashed histograms we apply the p_T to both
1919 Higgs candidates. The integral of the histograms
1920 between 0 and 1.6 gives an estimate of the frac-
1921 tion of signal we keep if we apply these p_T cuts.
1922 For $p_T(h_1) > 200(300)$ GeV we get that 93(98) %
1923 of the signal has b quarks with $\Delta R < 1.6$ and
1924 therefore can be fully reconstructed using a jet with $R = 0.8$.

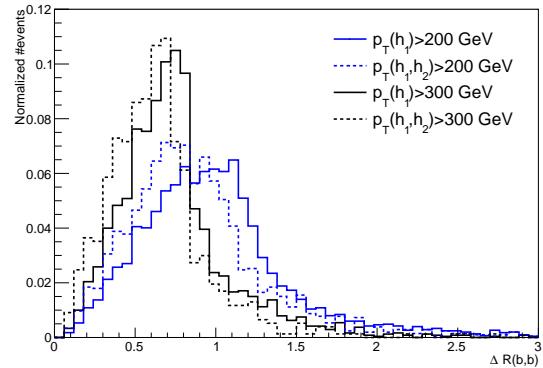


Figure A.1: $\Delta R(b, \bar{b})$ distributions for $p_T(h_1) > 200/300$ GeV (solid blue/black) and for $p_T(h_1, h_2) > 200/300$ GeV (dashed blue/black).

1925 **Appendix B**

1926 Additional background processes

1927 In this section we discuss the importance of additional background processes and estimate how they
1928 influence the analysis. In particular, we investigate backgrounds involving Higgs bosons.

1929 **Appendix C**

1930 **Softdrop mass**

1931 It is noticeable in the softdrop mass spectrum of both Higgs candidates (for signal and backgrounds) the
1932 existence an atypical peak close to zero. In this section we explain the origin of this feature.

1933 We believe that the peak close to zero corresponds to Higgs candidates ($R = 0.8$ jets) that do not
1934 contain both b quarks from the Higgs decay. The plot that support this conclusion is shown in figure
1935 C.1. It shows the correlation between the maximum ΔR between the Higgs candidate jet and one of
1936 the b quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$. For
1937 small masses (< 80 GeV) the maximum ΔR is usually larger than 0.8 which means that at least one
1938 of the b quarks is not contained in the jet's cone. When applying the soft drop procedure to these jets,
1939 soft radiation is removed and we are left with a single b quark. The mass of b quarks is ~ 5 GeV and
1940 therefore we get a peak at this mass.

1941 In practice, this does not affect our analysis because we place a mass window cut around the Higgs
1942 boson mass and therefore the low mass peak is removed. Nonetheless, the study presented here is
1943 extremely important because it helps rule out possible malfunctions of the soft drop algorithm and gives
1944 us confidence that we understand exactly what is happening.

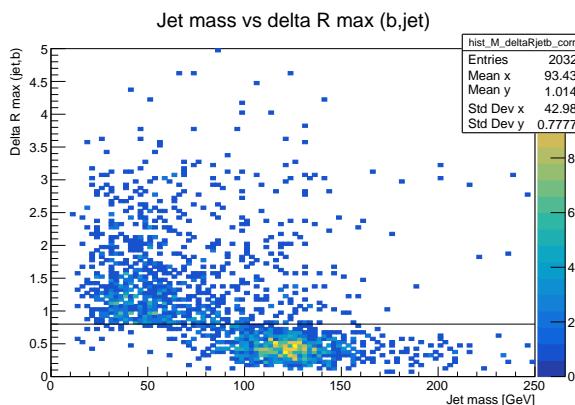


Figure C.1: Correlation between the maximum ΔR between the Higgs candidate jet and one of the b quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$.

1945 **Appendix D**

1946 **Samples generation: parameters**

1947 In this appendix we provide we list the values of parameters used in MadGraph (section D.1) and Pythia
1948 (section D.2) to generate the samples used in this work.

1949 **D.1 MadGraph**

1950 The MadGraph5 level cuts are summarized in table ?? . We show only the most relevant cuts for this
1951 analysis: the minimum p_T of light and b quarks, $p_{T,j}^{\min}$ and $p_{T,b}^{\min}$, the maximum pseudorapidity range for
1952 light and b quarks, η_j^{\max} and η_b^{\max} and the ΔR separation between two light quarks, $\Delta R(jj)$, two b
1953 quarks, $\Delta R(bb)$, and between a light and b quarks, $\Delta R(jb)$. The $xqcut$ parameter is a measure of the
1954 required parton separation at Madgraph level. Whenever MadGraph produces two partons, i and j ,
1955 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
1956 of this expression is smaller than the specified value of $xqcut$ then we do not generate the event. The
1957 $bwcutoff$ parameter defines what is considered to be on-shell s-channel resonances. The H_T variable
1958 is the scalar sum of the p_T of all truth level partons, including b quarks.

1959 **D.2 Pythia**

1960 The settings for jet matching can be found in table ?? under the corresponding samples' columns. We
1961 perform the jet matching procedure (merge=on) using the MLM matching scheme and the appropriate
1962 algorithm for a parton level process generated in MadGraph (scheme=1). We do not read the matching
1963 parameters from the MadGraph file (setMad=off) because this option is not available for these files. The
1964 size of the cone drawn around the jet's center, the maximum pseudorapidity and the maximum number
1965 of jets to be matched are given by coneRadius, etaJetMax and nJetMax, respectively. The cone radius
1966 is set to one. The maximum allowed pseudorapidity of jets is ten which is a much loser cut than the
1967 acceptance of any current detector. The maximum number of jets is set to four for the jj+0/1/2 j and to
1968 two for the t̄t+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.