

**2 Higgs pair production in the four bottom quarks final state**

**3**

**Ana Luísa Moreira de Carvalho**

**4**

Thesis to obtain the Master of Science Degree in

**5**

**Physics Engineering**

**6**

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

**7**

**Examination Committee**

**8**

Chairperson: Prof. Full Name

Supervisor: Prof. Full Name 1 (or 2)

Member of the Committee: Prof. Full Name 3

**9**

**Month Year**



Dedicated to someone special...



## **11 Acknowledgments**

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



<sup>14</sup> **Resumo**

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

- <sup>17</sup> **Palavras-chave:** palavra-chave1, palavra-chave2,...



<sup>18</sup> **Abstract**

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

<sup>20</sup> **Keywords:** keyword1, keyword2,...



# Contents

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

49	<b>5 Sample generation and analysis tools</b>	39
50	5.1 Fast simulation workflow . . . . .	39
51	5.1.1 Particle flow and calorimeter reconstruction in Delphes . . . . .	40
52	5.1.2 FCC-hh software . . . . .	41
53	5.2 Signal and background samples . . . . .	42
54	5.2.1 MadGraph . . . . .	42
55	5.2.2 Pythia . . . . .	43
56	5.2.3 Delphes . . . . .	45
57	<b>6 Analysis</b>	47
58	6.1 Overview of the $hh \rightarrow b\bar{b}b\bar{b}$ channel . . . . .	47
59	6.1.1 Event pre-selection . . . . .	48
60	6.1.2 Signal characterization . . . . .	49
61	6.1.3 Backgrounds . . . . .	50
62	6.2 Analysis strategy . . . . .	51
63	6.2.1 Implementation of b-tagging . . . . .	54
64	6.2.2 Baseline analysis . . . . .	55
65	6.2.3 Optimization . . . . .	55
66	<b>7 Results</b>	59
67	7.1 Di-Higgs discovery potential at the FCC-hh . . . . .	59
68	7.1.1 Statistical analysis . . . . .	59
69	7.2 Hadronic calorimeter granularity studies for future colliders . . . . .	59
70	7.3 Benchmarking against ATLAS analysis at $\sqrt{s} = 13$ TeV . . . . .	59
71	<b>8 Conclusions</b>	63
72	<b>Bibliography</b>	65
73	<b>A Jet radius discussion</b>	73
74	<b>B Additional background processes</b>	74
75	<b>C Softdrop mass</b>	75
76	<b>D Samples generation: parameters</b>	76
77	D.1 MadGraph . . . . .	76
78	D.2 Pythia . . . . .	76

# List of Tables

80	2.1 Summary of the particle content of the SM. . . . .	5
81	3.1 ATLAS tile and liquid argon hadronic calorimeters: summary of the pseudorapidity cover-	
82	ages and transversal segmentation (granularity). . . . .	21
83	3.2 Comparison between the working parameters of the LHC and of the FCC-hh. The val-	
84	ues of the number bunches and of the number of events per bunch crossing are given	
85	assuming a bunch spacing of 25 ns. . . . .	29
86	3.3 Hadronic calorimeter layout, granularity and energy resolution. . . . .	30
87	5.1 Summary of the cross sections, k factors and total number of events of the samples used	
88	in the analysis. . . . .	44
89	5.2 Summary of the benchmark granularity configurations of the HCAL. . . . .	45
90	6.1 b-Tagging (black), c (blue) and light (red) mistag probabilities as a function of $\eta$ and $p_T$ of	
91	the (sub)jet. The momentum dependent factor, $(1 - p_T/15000)$ , is common to the three	
92	probabilities. . . . .	54
93	7.1 Cumulative efficiency, in percentage, of each event selection criterion for the signal sam-	
94	ples (SM and 2HDM) and for the background sample ( $4b+j$ , $jj+0/1/2\ j$ and $t\bar{t}+0/1/2\ j$ ). The	
95	absolute value of expected events after some key selection cuts is shown in curved brack-	
96	ets. The number of expected events is normalized to $\mathcal{L} = 30\ ab^{-1}$ . The double horizontal	
97	line marks the pre-selection cuts. These results were obtained using the FCC-hh baseline	
98	detector design, as implemented in Delphes by the FCC-hh study group. . . . .	60
99	D.1 Generator (MadGraph5) level cuts for the signal and background samples. . . . .	77
100	D.2 Pythia settings for the signal and background samples. . . . .	77



# List of Figures

102	2.1	oi . . . . .	4
103	2.2	Postulated shape of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). . . . .	7
104	2.3	Feynman diagrams of Higgs pair production from gluon fusion. Triple vertex diagram (a) on the left and box diagram (b) on the right. The minus sign between the diagrams indicates that they interfere destructively. . . . .	10
105			
106	2.4	Total cross sections at the NLO in QCD for the six largest HH production channels at pp colliders. The thickness of the lines corresponds to the scale and PDF uncertainties added linearly. . . . .	12
107			
108	2.5	Here, $\alpha_1$ , $\alpha_2$ and $\alpha_3$ represent the coupling constants of $U(1)$ , $SU(2)$ and $SU(3)$ , respec- tively. What is shown is the variation of the inverse of the coupling with the energy: on the left for the SM and on the right for the MSSM.[REF] . . . . .	14
109			
110	2.6	BSM contributions to Higgs pair production. On the left (figure from [22]) $\eta$ represent the spin-2 Kaluza-Klein graviton that couples to a pair of SM Higgs bosons, $H$ , at tree level. On the right (figure from [23]) $\tilde{t}$ is the supersymmetric partner of the top quark that gives a contribution to the quantum loop in the Higgs pair production s-channel diagram. . . . .	16
111			
112	3.1	ATLAS detector. . . . .	22
113	3.2	FCC detector baseline concept. . . . .	30
114			
115	4.1	Exclusion limits on the mass of a spin 0 resonance from searches for Higgs pair production with the CMS detector. . . . .	35
116			
117	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 [67] (based on [66]). . . . .	37
118			
119	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) $0.1(0.025)\eta \times$ $5.6(1.4)\deg\phi$ , HCAL(ECAL) $0.05(0.012)\eta \times 2.8(0.7)\deg\phi$ and HCAL(ECAL) $0.025(0.006)\eta \times$ $1.4(0.35)\deg\phi$ . On the right, the root mean square (RMS) are also shown. Figures from [67]. . . . .	38
120			
121	4.4	$\tau_{32}$ resolution plots for QCD dijet events reconstructed using eflow jets (left) and tower jets (right). Figures from [68]. . . . .	38
122			
123			
124			
125			
126			
127			
128			
129			
130			

131	6.1 Higgs pairs branching ratios. . . . .	48
132	6.2 Event topology targeted by the boosted analysis region. The blob represents the in-	
133	teraction between the gluons and the Higg bosons that is represented by the Feynman	
134	diagrams shown in figure 2.3. . . . .	48
135	6.3 $\Delta R$ between the b quarks coming from the decay of the leading Higgs candidate, $h_1$ . . . . .	50
136	6.4 $\Delta\phi$ between the Higgs candidates. . . . .	50
137	6.5 Example of diagrams that contribute to the QCD multijet background: five final state jets,	
138	four of which are b-jets (left), three final state jets (middle) and four final state jets (right).	
139	Here, $q$ stands for a light quark/jet. . . . .	52
140	6.6 Dominant diagrams of $pp \rightarrow t\bar{t}$ at LO. . . . .	52
141	6.7 $p_T$ distributions for the leading (left) and sub leading (right) Higgs candidates. The signal	
142	is the SM $hh \rightarrow b\bar{b}$ process. The histograms are normalized to unit area. . . . .	52
143	6.8 $\eta$ distribution for the leading Higgs candidate (left) and softdrop mass distribution for the	
144	leading Higgs candidate (right). The signal is the SM $hh \rightarrow b\bar{b}$ process. The histograms	
145	are normalized to unit area. . . . .	53
146	6.9 oi . . . . .	53
147	6.10 Minimun $\Delta R$ between b quarks and subjets of the $R = 0.8$ jets. . . . .	54
148	6.11 $p_T$ distributions for the sub leading Higgs candidate (left) and for the Higgs pair (right).	
149	The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$ . . . . .	56
150	6.12 Distributions of the $\tau_{21}$ variable for the leading (left) and sub leading (right) Higgs candi-	
151	dates. The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$ . . . . .	56
152	6.13 $H_2$ variable for the leading Higgs candidate (left) and softdrop mass distribution for the	
153	leading Higgs candidate. The histograms are normalized to $\mathcal{L} = 30 \text{ ab}^{-1}$ . . . . .	56
154	7.1 Leading Higgs candidate softdrop mass (left) and $\tau_{21}$ (right) plots for the different detec-	
155	tor configurations for the SM signal sample. The plots contain all the signal events that	
156	passed all the cuts of the baseline analysis. For the plot on the left note that the x axis	
157	range is from 80 GeV to 160 GeV in order to make the differences between the histograms	
158	more clear. . . . .	60
159	7.2 Significance ( $S/\sqrt{B}$ ) as a function of the detector configuration for $\mathcal{L} = 30 \text{ ab}^{-1}$ (black)	
160	and $\mathcal{L} = 3 \text{ ab}^{-1}$ (blue) and for particle flow jets (squares) and pure calorimeter (HCAL)	
161	jets (triangles) for the SM signal sample. . . . .	61
162	7.3 Signal efficiency as a function of the detector configuration for particle flow jets (squares)	
163	and pure calorimeter (HCAL) jets (triangles). . . . .	62
164	A.1 $\Delta R(b, \bar{b})$ distributions for $p_T(h_1) > 200/300 \text{ GeV}$ (solid blue/black) and for $p_T(h_1, h_2) >$	
165	$200/300 \text{ GeV}$ (dashed blue/black). . . . .	73
166	C.1 Correlation between the maximum $\Delta R$ between the Higgs candidate jet and one of the b	
167	quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to $\Delta R = 0.8$ . . . . .	75

<sup>168</sup> **Chapter 1**

<sup>169</sup> **Introduction**

<sup>170</sup> NUMBERS? (luminosity, cross sections, BR,...)

---

<sup>172</sup>

<sup>173</sup> THESIS GOAL

<sup>174</sup> It is the ultimate goal of particle physics to discover and study all of Nature's fundamental particles  
<sup>175</sup> and to understand their interactions. Through a joint endeavor of theorists and experimentalists, models  
<sup>176</sup> that describe particle's dynamics and properties can be precisely probed at collider experiments such  
<sup>177</sup> as the Large Hadron Collider (LHC).

<sup>178</sup> We know today that matter particles interact by means of four fundamental forces: electromagnetic,  
<sup>179</sup> weak, strong and gravitational, each associated with a mediator particle. We even know that a very  
<sup>180</sup> special particle, the Higgs boson, is responsible for generating the mass of all of these particles through  
<sup>181</sup> a mechanism called Electroweak Symmetry Breaking (EWSB). All of this knowledge is beautifully sum-  
<sup>182</sup> marized in the Standard Model of Particle Physics (SM) that was developed in the 1960's, long before  
<sup>183</sup> many of the particles it predicts were discovered. The extraordinary precision of the predictions it de-  
<sup>184</sup> livers make it a very successful model. Its most recent prediction, the Higgs boson, was discovered in  
<sup>185</sup> 2012 at the LHC which marks an important point in the history of particle physics: we have now found  
<sup>186</sup> all the particles predicted by the SM and yet we know that it cannot be the whole story. Mainly because  
<sup>187</sup> there are experimental evidences it cannot explain.

<sup>188</sup> From the theoretical point of view, this is enough motivation to construct models that extend the SM  
<sup>189</sup> but that can still deliver predictions that are compatible with experimental data. From the experimental  
<sup>190</sup> standpoint, this is an indication that we need to keep increasing the precision of our measurements and  
<sup>191</sup> probing new kinematic regimes in the hope of finding some discrepancy with the SM or some hint that  
<sup>192</sup> some new phenomenon might be taking place.

<sup>193</sup> A higher precision requires a larger integrated luminosity and the exploration of new kinematic  
<sup>194</sup> regimes ask for a larger center of mass energy. Very recently, the upgrade of the LHC to its High-  
<sup>195</sup> Luminosity (HL) version has began. It is expected to work for a period of ten years between 2026 and  
<sup>196</sup> 2036 and it will extend the experimental reach of the LHC. In order to keep extending the physics reach

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

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

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

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

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

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

## <sup>229</sup> Chapter 2

# <sup>230</sup> The standard model and beyond

<sup>231</sup> The Standard Model is the theoretical framework that summarizes our present knowledge of particle  
<sup>232</sup> physics. In section 2.1, we provide an overview of this model, focusing on the Higgs mechanism. In  
<sup>233</sup> section 2.2, we motivate the need to explore models beyond the SM and introduce some of the most  
<sup>234</sup> well known BSM models. In section 2.1.1, we provide a theoretical description of the production of Higgs  
<sup>235</sup> boson pairs.

## <sup>236</sup> 2.1 The Standard Model of Particle Physics

### <sup>237</sup> IDEAS FOR SECTION

- <sup>238</sup> - Historic development + biggest breakthroughs
- <sup>239</sup> - Theoretical description of SM: Symmetry group, Lagrangian, ...
- <sup>240</sup> - Biggest successes + experimental confirmations

<sup>241</sup>

- <sup>242</sup> - Theoretical description of Higgs mechanism: symmetry breaking + mass generation for gauge
- <sup>243</sup> bosons
- <sup>244</sup> - Yukawa couplings as a source of mass for leptons

---

<sup>245</sup>

<sup>246</sup> The Standard Model (SM) of particle physics summarizes our present knowledge of fundamental  
<sup>247</sup> particles and their interactions. It is formulated in the framework of Quantum Field Theory (QFT) and  
<sup>248</sup> describes the subatomic world in terms of fields whose excitations are the particles we can detect. The  
<sup>249</sup> particle content of the SM is summarized in figure 2.1. There are two types of fundamental particles:  
<sup>250</sup> matter particles and force carriers.

<sup>251</sup> Matter particles are the building blocks of all the matter in our world. They come in two groups,  
<sup>252</sup> leptons and quarks. Quarks make up atomic nuclei and leptons, namely electrons and muons can orbit  
<sup>253</sup> atomic nuclei forming atoms. Both quarks and leptons are fermions which means they have half-integer  
<sup>254</sup> spin. There are six quarks: three of the 'up type' (up, charm and top represented by  $u$ ,  $d$  and  $t$ ) with  
<sup>255</sup> electric charge of  $+2/3$  and three of 'down type' (down, strange and bottom represented by  $d$ ,  $s$  and  $b$ )

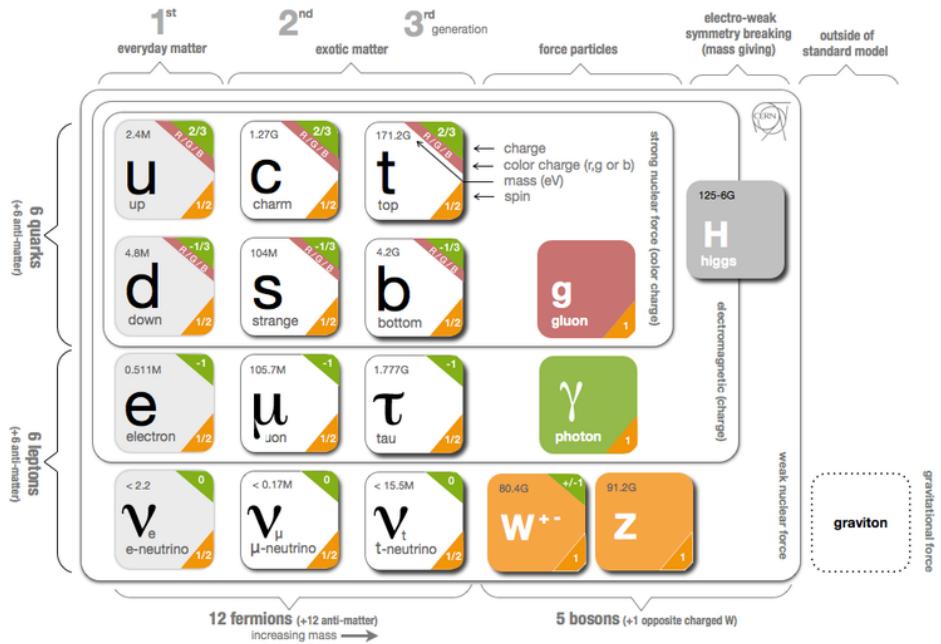


Figure 2.1: oi

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

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

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

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

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

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

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

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

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

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

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

301 However, the real breakthrough occurs when we combine a theory with local gauge invariance with  
 302 the mechanism of SSB. In this case the Nambu-Goldstone bosons do not appear and it is possible to  
 303 give mass to the gauge bosons. This is the Higgs mechanism, proposed independently by P.W. Higgs  
 304 [6], F. Englert and R. Brout [7] and by G. Guralnik, C. R. Hagen and T. Kibble [8] in 1964.

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

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

307 where the  $SU_{color}(3)$  group describes the strong interactions (QCD) and the  $SU_L(2) \times U_Y(1)$  group  
 308 describes the electroweak interactions. Here,  $L$  stands for left and  $Y$  stands for hypercharge. In the SM  
 309 the Higgs mechanism, which we now describe, is realized in the latter. The Lagrangian corresponding  
 310 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:

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

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

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

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

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

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

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

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

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

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

---

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

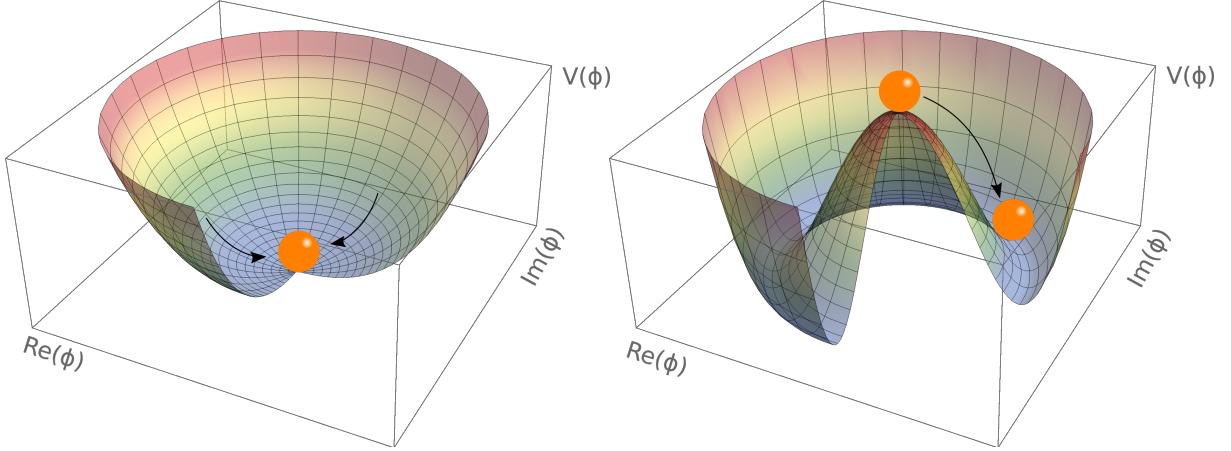


Figure 2.2: Postulated shape of the Higgs potential for  $\mu^2 > 0$  (left) and  $\mu^2 < 0$  (right).

321 be larger than zero to guarantee that the energy is bounded from below<sup>3</sup> because in Eq. 2.3  $\lambda$  is the  
 322 coefficient of the term with the highest power in  $\phi$  and therefore determines the concavity of the potential.

323 The shapes of the Higgs potential for  $\mu^2 > 0$  and  $\mu^2 < 0$  are shown in Figure 2.2 on the left and  
 324 right, respectively. For  $\mu^2 > 0$  (left) we have a single minimum located at  $\langle\phi\rangle = 0$ . For  $\mu^2 < 0$  (right)  
 325 the potential has the shape of a 'Mexican hat'. There is an infinite number of minima located in a  
 326 circumference centered at zero. In this case the minima occur for  $\langle\phi\rangle, \langle\phi^\dagger\rangle \neq 0$ . Therefore the fields  
 327 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):

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

328 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  
 329 consideration that  $W_\mu^a T^a$  represents a sum over all values of  $a$  we get

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

We see that for the  $W_\mu^1$  and  $W_\mu^2$  fields we have only terms that are quadratic in these fields. These correspond to mass terms. However, for the  $W_\mu^3$  and  $B_\mu$  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:

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

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

$A_\mu$  and  $Z_\mu$  are the physical fields that are related with  $W_\mu^3$  and  $B_\mu$  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  $\theta_W$  is the Weinberg angle. By inverting this relation we can write  $W_\mu^3$  and  $B_\mu$  as a function of  $A_\mu$  and  $Z_\mu$ . Replacing in 2.9 and imposing that the  $A_\mu$  field has zero mass we can determine  $\theta_W$ :  $\tan \theta_W = \frac{g'}{g}$ . The Lagrangian of 2.9 takes then the form

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

330 where we show only the mass terms for the gauge bosons. Note that, by construction, there is no mass  
 331 term for  $A_\mu$  which allows us to identify this field with the photon.  $W_\mu^1$  and  $W_\mu^2$  are related to the  $W^\pm$   
 332 boson and  $Z_\mu$  corresponds to the  $Z$  boson. We have shown that it is the fact that  $v \neq 0$  that allows for  
 333 the existence of non-zero mass terms for the  $W^\pm$  and  $Z$  bosons.

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

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

334 These terms encode the Higgs self interactions and represent, respectively, the three and four point  
 335 interactions. We see that the coupling constants of these interactions depend on the parameters of the  
 336 Higgs potential,  $\mu^2$  and  $\lambda$ .

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

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

341 where  $L$  denotes a left-handed fermion doublet and  $R$  denotes a right-handed fermion singlet. Here, left  
 342 and right refer to helicity states.  $G_1$  and  $G_2$  are arbitrary coupling constants that can be written in terms  
 343 of the fermion's mass and the VEV.  $\phi$  is given by 2.8 and  $\phi_c$  is given by (after the spontaneous symmetry  
 344 breaking and in the unitary gauge):

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

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

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

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

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

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

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

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

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

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

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

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

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

360 The SM has delivered extremely accurate predictions about the existence and properties of new  
361 particles which makes it a very successful theory. It predicted the existence of the W and Z bosons [12],  
362 the gluon, the charm and top quarks and the Higgs boson. [REFS] In addition, the SM prediction for the  
363 value of the anomalous magnetic dipole moment of the electron (calculated up to order  $\alpha^5$ ) agrees with  
364 the measured value up to the 11<sup>th</sup> decimal place, making it the most precise measurement in science.

## 365 2.1.1 Higgs pair production

### 366 IDEAS FOR SECTION

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

371  
 372 Within the SM there are still some processes that have not been measured. One of these is the  
 373 production of pairs of Higgs bosons. The experimental challenges and efforts related to this process are  
 374 discussed in section 4.1. Here we provide a theoretical description of the process.

375 At the LHC, the main production process of Higgs pairs is gluon-gluon fusion (ggF). Higgs pairs can  
 376 also be produced through vector boson ( $V$ ) fusion (VBF), namely of  $W$  and  $Z$  bosons, in association with  
 377 a pair of top quarks ( $t\bar{t}h$ ) or through Higgs strahlung ( $Vh$ )<sup>4</sup>. The ggF production process has a cross  
 378 section ten times larger than the next most common production process which is VBF (approximately  
 379 50 pb *versus* 4 pb) and therefore it is the dominant contribution when we study inclusive production.  
 380 For this reason we focus the following discussion on this production mode. The leading order Feynman  
 381 diagrams for Higgs pair production via ggF are shown in figure 2.3.

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

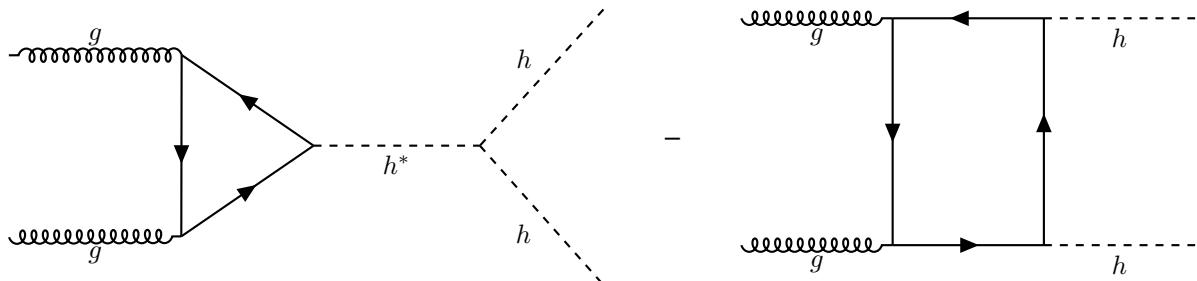


Figure 2.3: Feynman diagrams of Higgs pair production from gluon fusion. Triple vertex diagram (a) on the left and box diagram (b) on the right. The minus sign between the diagrams indicates that they interfere destructively.

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

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

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

<sup>4</sup>In this process, at LO, a Higgs boson is radiated from a vector boson

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

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

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

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

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

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

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

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

---

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

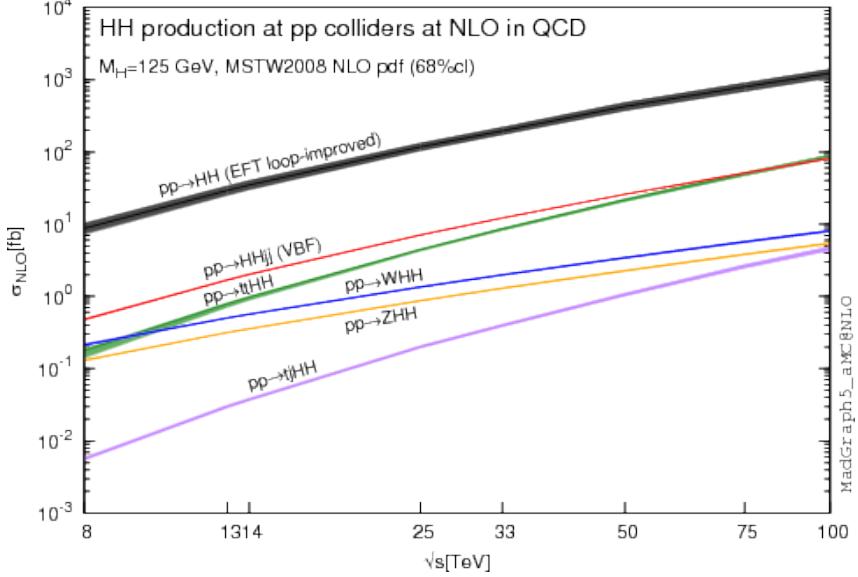


Figure 2.4: Total cross sections at the NLO in QCD for the six largest HH production channels at pp colliders. The thickness of the lines corresponds to the scale and PDF uncertainties added linearly.

be interpreted as generalized couplings and are given by

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

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

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

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

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

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

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

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

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

## 422 2.2 Going beyond

### 423 IDEAS FOR SECTION

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

428

---

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

433 On the one hand, there are several pieces of experimental evidence that the SM cannot explain.  
434 These include the nature of dark matter, postulated to explain the experimental observations of the  
435 velocity of far away galaxies [17], the asymmetry between matter and anti-matter in the present Universe  
436<sup>6</sup> and the fact the neutrinos oscillate between flavors which implies that they have a non-zero mass. This  
437 phenomenon was measured independently by two collaborations, the Sudbury Neutrino Observatory  
438 (SNO) and the Super-Kamiokande, in 1998 and 2001-2002, respectively, [18–20].

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

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

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

---

<sup>6</sup>Or why do we live in an Universe made entirely out of matter?

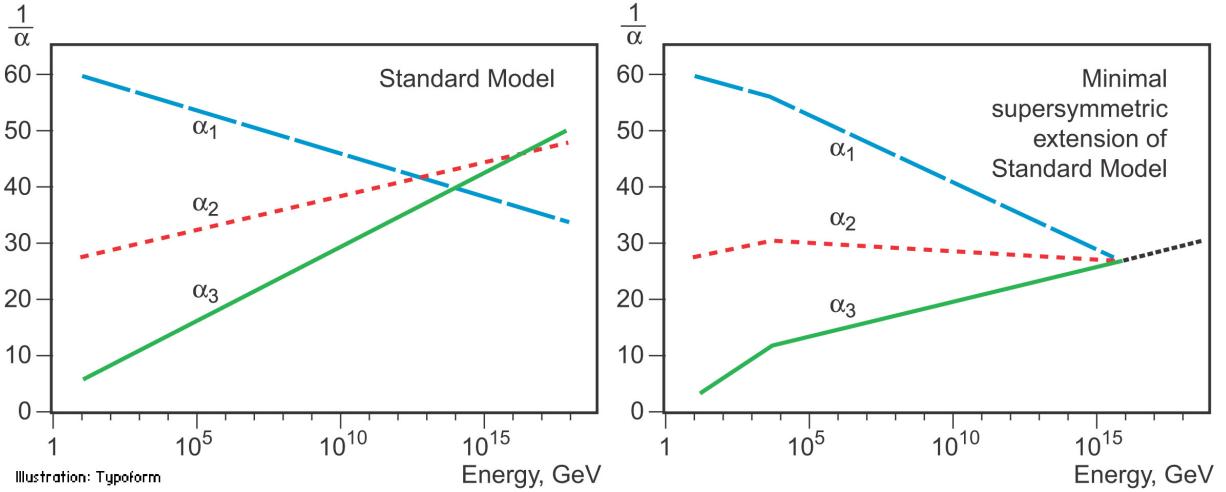


Figure 2.5: Here,  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  represent the coupling constants of  $U(1)$ ,  $SU(2)$  and  $SU(3)$ , respectively. What is shown is the variation of the inverse of the coupling with the energy: on the left for the SM and on the right for the MSSM.[REF]

456 symmetry group.

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

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

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

481 Universe.

482 When studying the 2HDM there are several restrictions that can be placed on the parameters of  
483 the model so as to reduce its complexity. A generally well accepted assumption is that there is no CP  
484 violation coming, at tree-level, from the Higgs sector. This is the so-called CP-conserving 2HDM. In this  
485 case, we obtain five Higgs bosons that are CP eigenstates. Three of them are neutral,  $h_1$ ,  $h_2$  and  $h_3$ ,  
486 and the other two are charged,  $h^\pm$ .  $h_1$  is usually taken to be the SM Higgs boson and its mass is set to  
487 125 GeV.

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

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

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

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

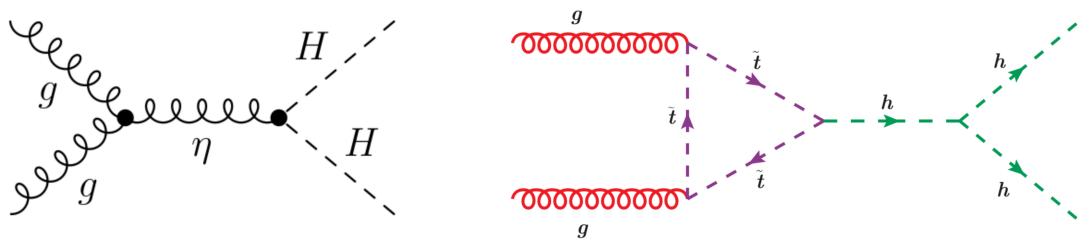


Figure 2.6: BSM contributions to Higgs pair production. On the left (figure from [22])  $\eta$  represent the spin-2 Kaluza-Klein graviton that couples to a pair of SM Higgs bosons,  $H$ , at tree level. On the right (figure from [23])  $\tilde{t}$  is the supersymmetric partner of the top quark that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

## 512 Chapter 3

# 513 Collider experiments

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

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

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

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

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

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

### 544 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  $\sigma$ , 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:

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

545 where  $f_{coll}$  is the collision frequency,  $n_1$  and  $n_2$  are the number of protons in each bunch and  $\sigma_x$  and  $\sigma_y$   
546 characterize the transverse beam size in the horizontal and vertical directions.

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

550 While the cross sections of most physics processes increase when the CM energy goes from 13 to  
551 100 TeV, many BSM models predict new processes, or new contributions to existing processes, whose  
552 cross sections increase more rapidly than the SM backgrounds. In addition, by conservation of energy,  
553 a larger CM energy implies that particles with larger mass can be created. Based on Eq. 3.2, we can  
554 tune its parameters to obtain the highest possible luminosity. Nonetheless, because we are dealing with  
555 charged particles, there is a limit on how close the bunches can be and on how many protons we can  
556 pack in a bunch. Moreover, the beam's transverse dimensions cannot be infinitely reduced. A smarter  
557 way to increase the number of collisions that an accelerator can produce is to run for a longer time,  
558 therefore increasing the integration time in Eq. 3.1.

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

### 561 3.2 The Large Hadron Collider

#### 562 IDEAS FOR SECTION

- 563 - The LHC: goals and physics programs
- 564 - Successes, discoveries
- 565 - Upgrades and outlook

566

---

567 The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. It  
568 is housed by the European Organization for Nuclear Research (CERN) which focuses on fundamental  
569 particle physics with the goal of probing matter's most elementary structure. Ever since its creation, in  
570 1954, CERN has housed many accelerators and experiments and played a key role in the development  
571 of fundamental and applied science.

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

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

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

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

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

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

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

### 612 3.2.1 The ATLAS detector

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

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

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

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

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

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

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

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

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

650 For  $|\eta| > 1.5$  LAr calorimeters extend the pseudorapidity range to  $|\eta| = 4.9$ . The LAr calorimeter is  
651 divided in end-cap and forward. These cover the pseudorapidity ranges  $1.5 < |\eta| < 3.2$  and  $3.2 < |\eta| <$   
652 4.9, respectively. The active material is liquid-argon and the absorbers are copper and tungsten for the  
653 end-cap and forward calorimeters, respectively. In the end-cap LAr calorimeters the segmentation is  
654  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  for  $1.5 < |\eta| < 2.5$  and  $0.2 \times 0.2$  for  $2.5 < |\eta| < 3.2$ . In the forward LAr calorimeter  
655 the segmentation is  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ .

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

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

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

<sup>1</sup>The TileCal is composed of three longitudinal layers. Only the first two have a segmentation equal to  $0.1 \times 0.1$ . In the third layer the segmentation is  $0.2 \times 0.1$ . 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.

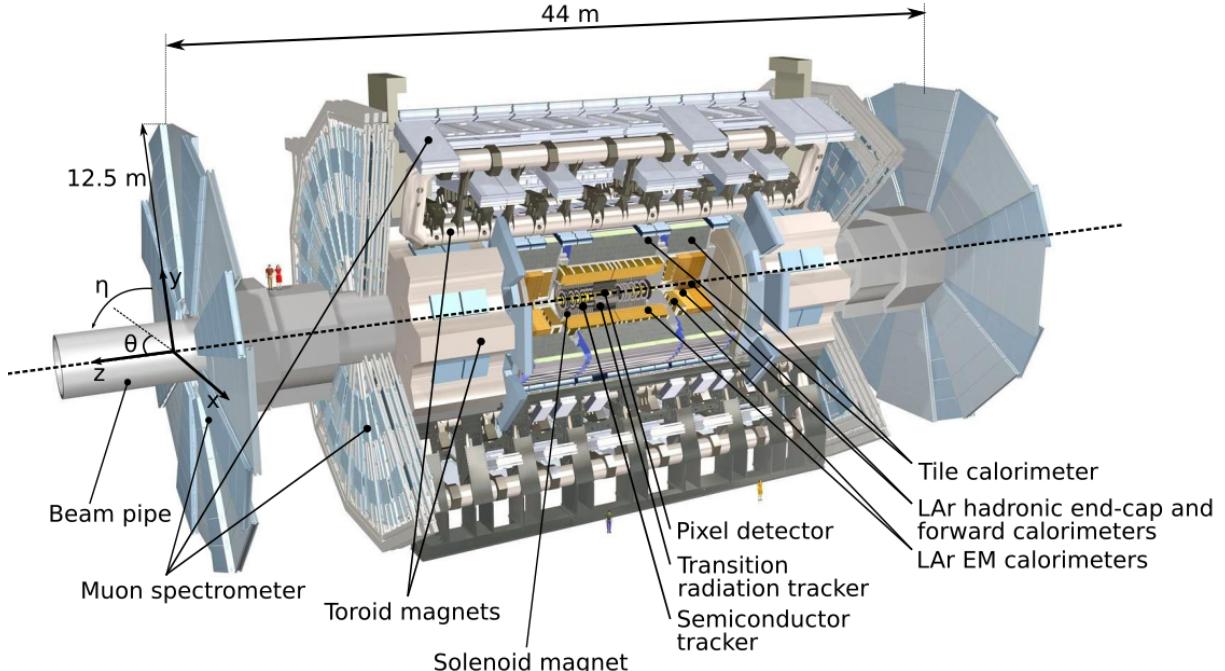


Figure 3.1: ATLAS detector.

## 669 b-Tagging

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

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

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

<sup>2</sup>This 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.

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

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

## 696 **Trigger and data acquisition**

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

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

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

## 714 **3.3 Jet reconstruction**

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

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

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

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

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

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

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

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

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

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

## 752 Boosted kinematic regime

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

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

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

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

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

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

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

### 3.3.1 Jet properties and substructure observables

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

- tau\_N
- energy correlation functions and ratios
- Fox Wolfram
- ...
- Jet mass definition!

---

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

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

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

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

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

800 where the index  $k$  runs over all particles in a jet, the indexes 1 to  $N$  identify the number of axis inside the  
 801 jet and  $R_0$  is the radius of the jet. This variable will have a small value if the particles with the highest  $p_T$   
 802 (the ones we are most interested in) are clustered around the axis (because  $\Delta R$  will be small) and will  
 803 have a larger value otherwise. A jet with small  $\tau_N$  is considered to be consistent with having  $N$  or fewer  
 804 subjets because all its constituents are aligned with the axis.

805 We usually use ratios of  $\tau_N$  variables ( $\tau_{MN} = \tau_M / \tau_N$ ). Of particular interest for this work are the  $\tau_{21}$ ,  
 806  $\tau_{31}$  and  $\tau_{32}$  ratios. These observables can take values between zero and one. For  $\tau_{21}$ , for example, a  
 807 small value (close to zero) indicates that the jet is more compatible with two subjets than with one and  
 808 therefore it can help discriminate between two-prong and one-prong jets. The same applies to every  
 809 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 [49]. They are given by:

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

810 where  $\theta_{ij}$  is the angle between energy clusters i and j,  $E$  is the total energy of the clusters in the jet rest  
 811 frame, the  $P_l(x)$  are the Legendre polynomials. For a jet that has a structure of two back-to-back subjets  
 812 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  
 813 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:

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

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

### 818 **3.3.2 Jet grooming algorithms**

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

823 There are three main jet grooming algorithms: trimming, pruning and mass drop filtering. In this  
 824 work we do not use pruning or trimming, although these techniques might be worth exploring in future,  
 825 more comprehensive, studies. In particular, they can be useful to help reject pileup contributions. The  
 826 mass drop filtering procedure isolates relatively symmetric subjets within a jet, each with a significantly  
 827 smaller mass than the original jet [47]. This technique was developed and optimized using C/A jets for  
 828 the search of Higgs decaying to  $b\bar{b}$  pairs [50]. 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 between the mass of the original jet,  $m_{jet}$ , and  $m_{j_1}$ :  $m_{j_1}/m_{jet} < \mu_{frac}$ , where  $\mu_{frac}$  is a parameter of the algorithm. In addition the splitting is required to be relatively symmetric:

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

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

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

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

## 837 **3.4 Future Colliders**

838 As we already argued in the beginning of this chapter, a larger CM energy is one of the factors driving  
 839 the discovery potential of an accelerator. As far as we know today, proton-proton colliders are the main,  
 840 and possibly only, man-made experimental tool available to explore particle physics in the energy range  
 841 on tens of TeV. With this in mind, new hadronic colliders with CM energies of the order of tens of TeV

842 have been proposed. The main projects are the hadronic Future Circular Collider (FCC-hh) lead by  
843 CERN and the Super Proton-Proton Collider (SPPC) proposed by China.

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

### 846 3.4.1 The hadronic Future Circular Collider

#### 847 IDEAS FOR SECTION

- 848 - Center of mass energy ✓
- 849 - Basic design: circular, circumference, location, ... ✓
- 850 - CERN FCC study group: creation, mission, main tasks ✓
- 851 - The role of the hadronic calorimeter
- 852 - Basic structure: tracker, calorimeters, muon chambers ✓
- 853 - Technical specifications: nb. bunches, magnets, ... ✓
- 854 - Technical dimensions and materials (predictes) of each layer ✓
- 855 - Focus on hadronic calorimeter: material being studied, resolution, granularity ✓
- 856 - What processes we want to study at the FCC
- 857 - Focus on Higgs pair production

858

859

---

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

868 The FCC-hh baseline design consists of a proton-proton circular collider with a maximum CM energy  
869 of 100 TeV housed by a 100 km tunnel in the area of Geneva. This machine will extend the research  
870 program of the LHC (and of the HL-LHC) after these have reached their full discovery potential, by  
871 around 2040. In addition, it will allow for the exploration of an entirely new kinematic regime, probing  
872 energy scales where new physics may come into play. A possible way of defining the target luminosity of  
873 this machine is to require that within the first year of operation it surpasses the exploration potential of the  
874 LHC [51]. Comprehensive studies [51, 52] indicate that this can be achieved with an integrated luminosity  
875 of the order of  $10 \text{ ab}^{-1}$  per experiment. Considering a reasonable operation period of 10 years this leads  
876 to integrated luminosity per experiment of the order of  $1 \text{ ab}^{-1}$  per year.

877 The FCC-hh is expected to work with a dipole field of 16 T and to provide a peak instantaneous lumi-  
878 nosity thirty times larger than the LHC. The number of bunches is expected to be almost a factor of four

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.

879 larger than for the LHC and the number of events per bunch crossing is expected to be approximately  
 880 1000. The latter brings a lot of technical challenges because it means that the mean pileup expected  
 881 for the FCC-hh is almost 40 times larger than for the LHC. This requires the development (or improve-  
 882 ment of already existing techniques) of techniques and algorithms that allow us to further reject pileup  
 883 contributions.

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

### 885 3.4.2 FCC-hh baseline detector

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

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

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

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

904 The hadronic calorimeter covers the pseudorapidity range  $|\eta| < 6$ . It is also divided in barrel, end-  
 905 cap and forward that cover the pseudorapidity ranges  $|\eta| < 1.3$ ,  $1.0 < |\eta| < 1.8$  and  $2.3 < |\eta| < 6.0$ ,  
 906 respectively. The proposed layout for the barrel calorimeter consists of scintillator tiles interleaved with

Parameter	Barrel	End-cap	Forward
$\eta$ 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 ( $\Delta\eta \times \Delta\phi$ )	$0.025 \times 0.025$	$0.025 \times 0.025$	$0.05 \times 0.05$
Energy resolution ( $\sigma_E/E$ )	$40\%/\sqrt{E} \oplus 2.5\%$	$50\%/\sqrt{E} \oplus 3\%$	$100\%/\sqrt{E} \oplus 5\%$

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

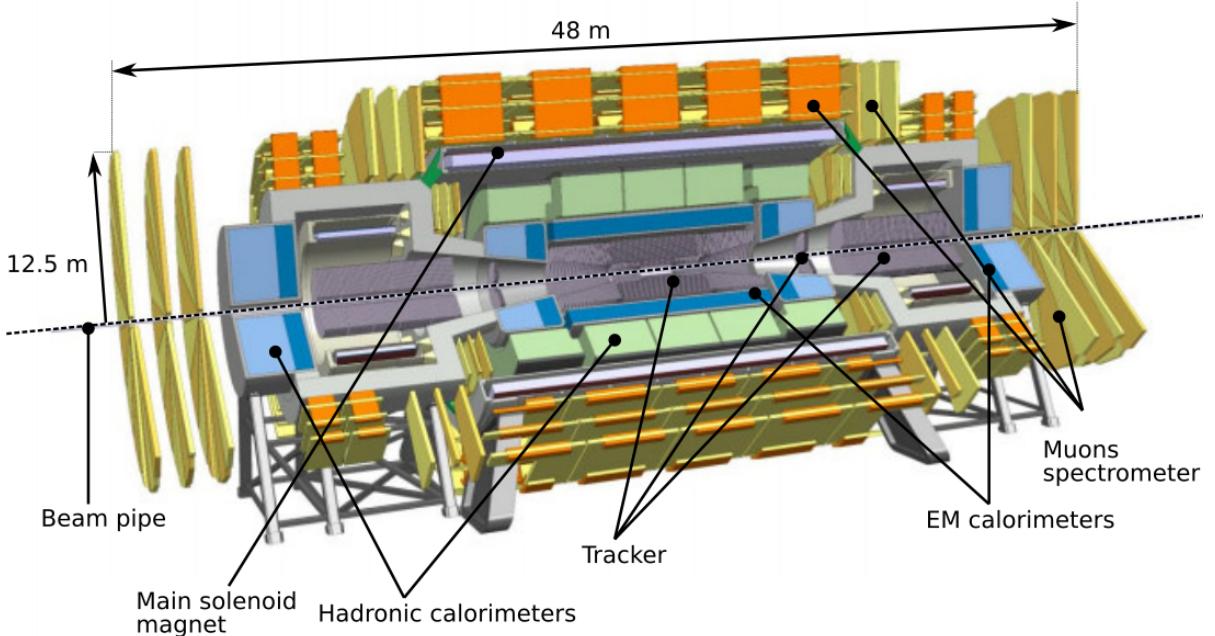


Figure 3.2: FCC detector baseline concept.

907 lead and stainless steel plates as absorbers. The end-cap and forward calorimeters are expected to be  
 908 based on liquid argon with copper plates as absorbers. For the barrel and end-cap calorimeters, the  
 909 expected segmentation  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  while for the forward calorimeter it is  $\Delta\eta \times \Delta\phi = 0.05 \times$   
 910 0.05. Overall, this corresponds to approximately four times the ATLAS HCAL granularity. The energy  
 911 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-  
 912 cap and forward calorimeters, respectively. The pseudorapidity coverage, layout, granularity and energy  
 913 resolution of the hadronic calorimeters are summarized in table 3.3.

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

### 917 3.4.3 FCC-hh physics program

918 The FCC-hh physics program is vast and diverse. Within the SM, it includes the study of gauge bosons  
 919 pair production and heavy flavor production, the measurement of the top quark's properties and the

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

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



# <sup>937</sup> Chapter 4

## <sup>938</sup> State of the art

<sup>939</sup> In this chapter we present the state of the art of the searches for Higgs pairs production both at the  
<sup>940</sup> LHC and in future colliders. In section 4.1, we review the searches that have been conducted at the  
<sup>941</sup> LHC by the ATLAS and CMS experiments. We include discussions on the different final states that were  
<sup>942</sup> target and report on the constraints that were derived for the cross section and trilinear coupling. A brief  
<sup>943</sup> overview of the current constraints on some BSM models is also presented. In section 4.2, we present  
<sup>944</sup> the results obtained from feasibility studies that access the discovery potential for this process at the  
<sup>945</sup> HL-LHC and at the FCC-hh.

### <sup>946</sup> 4.1 Searches for Higgs pair production at the LHC

#### <sup>947</sup> IDEAS FOR SECTION

- <sup>948</sup> - Previous searches by ATLAS and CMS at the LHC ✓
  - <sup>949</sup> - Different decay channels used: pros and cons + compare achieved sensitivities ✓
  - <sup>950</sup> Channels: bbbb, bb tau tau, bb gamma gamma, gamma gamma WW, bb ZZ ✓
  - <sup>951</sup> - Models investigated and motivation
  - <sup>952</sup> - Exclusion limits
- 

<sup>953</sup>  
<sup>954</sup> The discovery of the Higgs boson is a strong evidence that the Higgs mechanism operates as predicted.  
<sup>955</sup> However, by itself, it does not guarantee that the shape of the Higgs potential is the one depicted in  
<sup>956</sup> figure 2.2, on the right. In order to reconstruct the Higgs potential and gain a deeper understanding of  
<sup>957</sup> the mechanism that leads to the breaking of the electroweak symmetry one must measure the Higgs  
<sup>958</sup> boson self-couplings, namely its three and four point interactions, whose strengths depend on the values  
<sup>959</sup> of the parameters of the Higgs potential, as shown in 2.13.

<sup>960</sup> However, in the SM, the cross section for the production of Higgs pairs through ggF is extremely  
<sup>961</sup> small:  $\sim 30$  fb at the current CM energy achieved at the LHC (value computed at Next to Leading  
<sup>962</sup> Order (NLO) accuracy) [53]. This value is approximately three orders of magnitude smaller than the  
<sup>963</sup> production cross section of a single Higgs boson. In addition, this value has to be multiplied by the

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

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

The  $h \rightarrow bbbb$  channel [54, 55] benefits from the large branching fraction of  $h \rightarrow bb$  ( $\sim 58\%$ ). In addition, ATLAS showed that this is the most sensitive channel to resonance masses over 500 GeV [56]. However, this channel suffers with an overwhelming multijet background which drives the need for very tight trigger level cuts in order to bring the event rates down to manageable values.

The  $h \rightarrow b\bar{b}\tau\tau$  analysis [56, 57] benefits from a sizable branching fraction ( $\sim 7.3\%$ ) and from a relatively small background contribution from other SM processes. This searches target the semi-leptonic decay of the  $\tau\tau$  pair.

The  $hh \rightarrow bb\gamma\gamma$  [58, 59],  $WW^*\gamma\gamma$  [56, 60],  $ZZ\gamma\gamma$  [60] analysis can make use of very efficient diphoton triggers and isolation criteria that greatly reduce multijet background. In addition the excellent mass resolution of  $h \rightarrow \gamma\gamma$  can be exploited. The  $hh \rightarrow bbWW^*$  channel also benefits from the large branching fraction of  $h \rightarrow WW$  ( $\sim 21\%$ ).

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

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

## 4.2 Feasibility studies for high-luminosity and future colliders

IDEAS FOR SECTION

- HL-LHC feasibility study [61],[62],[63]
- FCC hh to 4b (Selvaggi et al.) [64], [65]

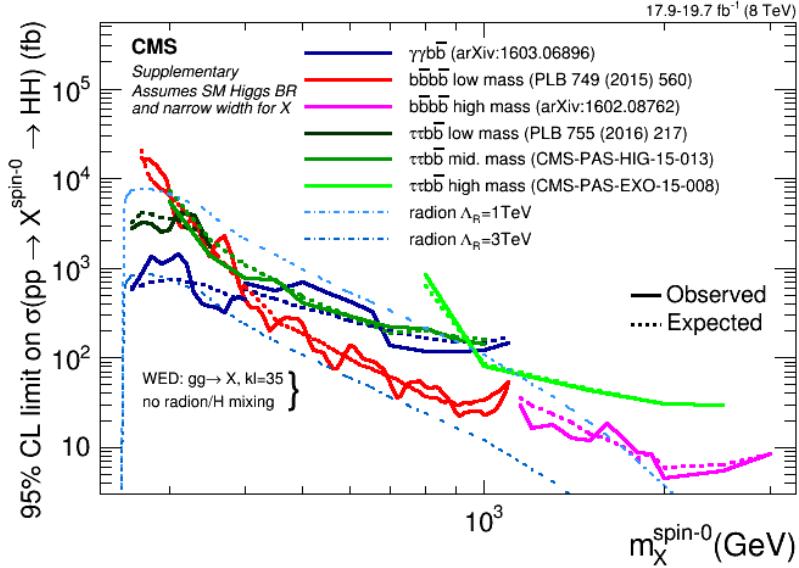


Figure 4.1: Exclusion limits on the mass of a spin 0 resonance from searches for Higgs pair production with the CMS detector.

Without any BSM contribution the discovery (or even evidence) 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  $300 \text{ fb}^{-1}$ . 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)  $\text{fb}^{-1}$  [61]. The analysis is performed in three orthogonal regions (boosted, intermediate and resolved)<sup>1</sup> 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)  $\text{fb}^{-1}$  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 increases 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 reported significances of  $1.05\sigma$  and  $1.6\sigma$ , respectively, for an integrated luminosity of 3000  $\text{fb}^{-1}$ . Taking, as an example, the ATLAS result, the achieved significance translates to an upper limit on the total

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

1024 di-Higgs cross section of approximately twice the SM value. This corresponds to an exclusion limit of  
1025  $-0.8 < \lambda_{hhh}/\lambda_{SM} < 7.7$ . The analysis conducted by ATLAS in this channel is done at generator level.  
1026 The energy and momenta of the particles are smeared to simulate the detector's response. A mean pile  
1027 up of 200 is considered. No MVA techniques are employed in the analysis.

1028 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  
1029 can be excluded at 95% CL, with systematics uncertainties being taken into consideration. The Higgs  
1030 trilinear coupling is expected to be constrained to  $-3.5 < \lambda_{hhh}/\lambda_{SM} < 11$ . These results are based  
1031 on the extrapolation of the current results obtained with the 2016 dataset, comprising an integrated  
1032 luminosity of  $10.1 \text{ fb}^{-1}$ . This study is based only on the resolved analysis documented in [54].

1033 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  
1034 only the irreducible  $pp \rightarrow b\bar{b}b\bar{b}j$  reports a significance of 6.61 for an integrated luminosity of  $30 \text{ ab}^{-1}$ ,  
1035 considering an analysis that targets the boosted region [64]. The extra jet in the signal sample has  $p_T >$   
1036 200 GeV which provides the Higgs pair with a large Lorentz boost. EXPLAIN WHY THIS INCREASES  
1037 SENSITIVITY TO HIGGS TRIPLE COUPLING. The analysis relies on the jet substructure observable  
1038  $\tau_{21}$  and on a tight mass cut around the Higgs mass. No multivariate techniques are employed.

1039 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  
1040 at the FCC-hh was presented in [65]. Only the irreducible background is considered. For a boosted cut  
1041 based analysis, a significance of 1.1 is reported for an integrated luminosity of  $3 \text{ ab}^{-1}$  at  $\sqrt{s} = 14 \text{ TeV}$   
1042 (HL-LHC). For an integrated luminosity of  $10 \text{ ab}^{-1}$  at  $\sqrt{s} = 100 \text{ TeV}$ , which corresponds to the FCC-hh,  
1043 this number is 5.7. While the significance is large for the FCC-hh, the reported signal to background ratio  
1044 ( $S/B$ ) is approximately one order of magnitude smaller which means the measurement might be more  
1045 sensitive to systematic uncertainties on the backgrounds.

### 1046 4.3 Hadronic calorimeter granularity studies

1047 Even before the baseline design for the FCC-hh was established, there were studies regarding the  
1048 impact of the granularity of the calorimeters in the spatial resolving power of hadronic showers and on  
1049 the resolution of jet mass and substructure variables. These studies targeted the development of future  
1050 colliders and greatly influenced the baseline design of the FCC-hh.

1051 The granularity of hadronic calorimeters is a key parameter for future collider detectors because it  
1052 determines how well we can resolve energy deposits from pileup vertices and highly-boosted jet topolo-  
1053 gies. In [66], the use of smaller calorimeter cells (smaller than the ones that are used in currently  
1054 operating hadronic calorimeters) to resolve individual hadrons is investigated. For two Kaons ( $K_L^0$ ) with  
1055 an energy of 100 GeV each and with a truth level separation ( $\Delta R$ ) equal to 0.035 the energy deposited in  
1056 the ECAL (blue) and HCAL (red) is shown as a function of  $\Delta R$  for HCAL cells with sizes  $20 \text{ cm} \times 20 \text{ cm}$   
1057 ( $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ) 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  
1058 right, respectively. The ECAL segmentation is equal to  $2 \text{ cm} \times 2 \text{ cm}$ . It is shown that for a granularity of  
1059  $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$  both hadrons can be resolved in the HCAL.

1060 Additional studies focusing on the jet mass resolution and jet substructure observables resolution

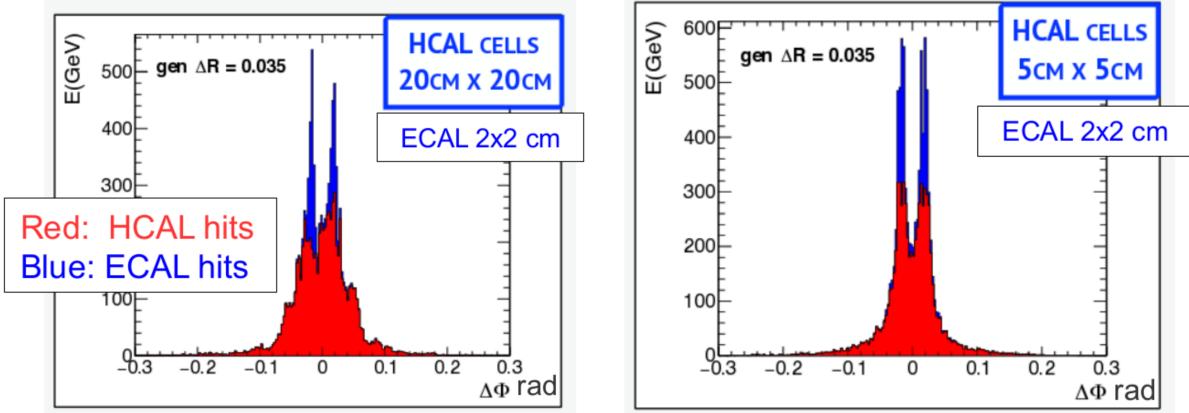


Figure 4.2: Azimuthal distribution of the energy deposits in the ECAL and HCAL for a pair of  $K_L^0$  with  $E = 100$  GeV for an hadronic calorimeter with  $20\text{ cm} \times 20\text{ cm}$  (left) and  $5\text{ cm} \times 5\text{ cm}$  (right). Figures from [67] (based on [66]).

were performed for three calorimeter (HCAL and ECAL) configurations. Some of these studies were presented in major conferences focused on future colliders, namely FCC week 2015 and 2016 and BOOST 2017. Here, we show results presented in [67–69]. 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 resolution<sup>2</sup> 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  $\tau_{32}$  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  $\tau_{32}$  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.

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

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

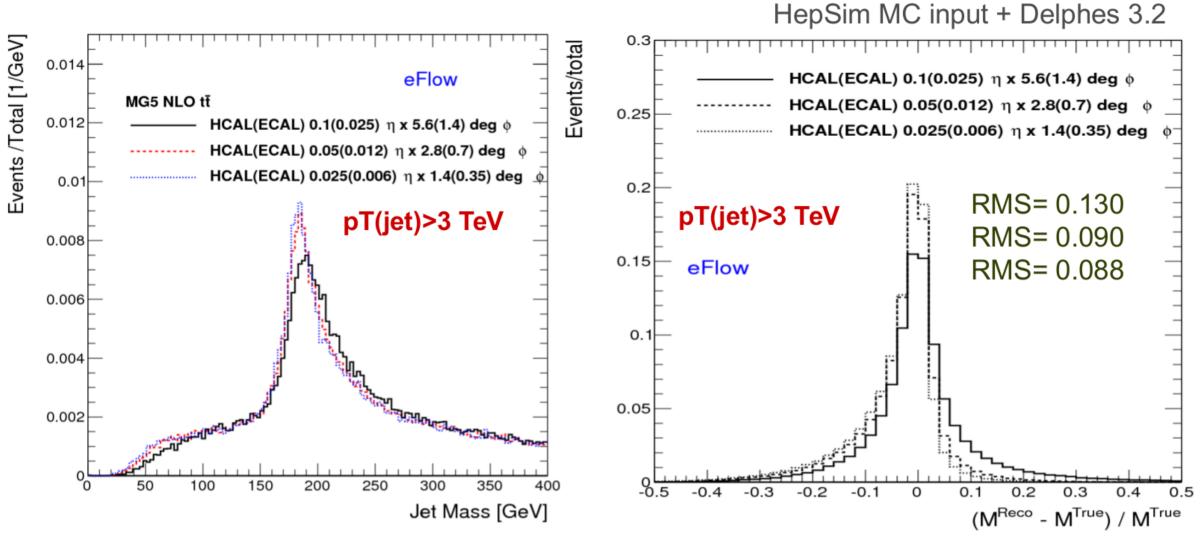


Figure 4.3: Jet mass (left) and jet mass resolution (right) plots in  $t\bar{t}$  events for eflow jets with  $p_T > 3$  TeV for three different HCAL and ECAL configurations: HCAL(ECAL) 0.1(0.025)  $\eta \times 5.6(1.4)$  deg  $\phi$ , HCAL(ECAL) 0.05(0.012)  $\eta \times 2.8(0.7)$  deg  $\phi$  and HCAL(ECAL) 0.025(0.006)  $\eta \times 1.4(0.35)$  deg  $\phi$ . On the right, the root mean square (RMS) are also shown. Figures from [67].

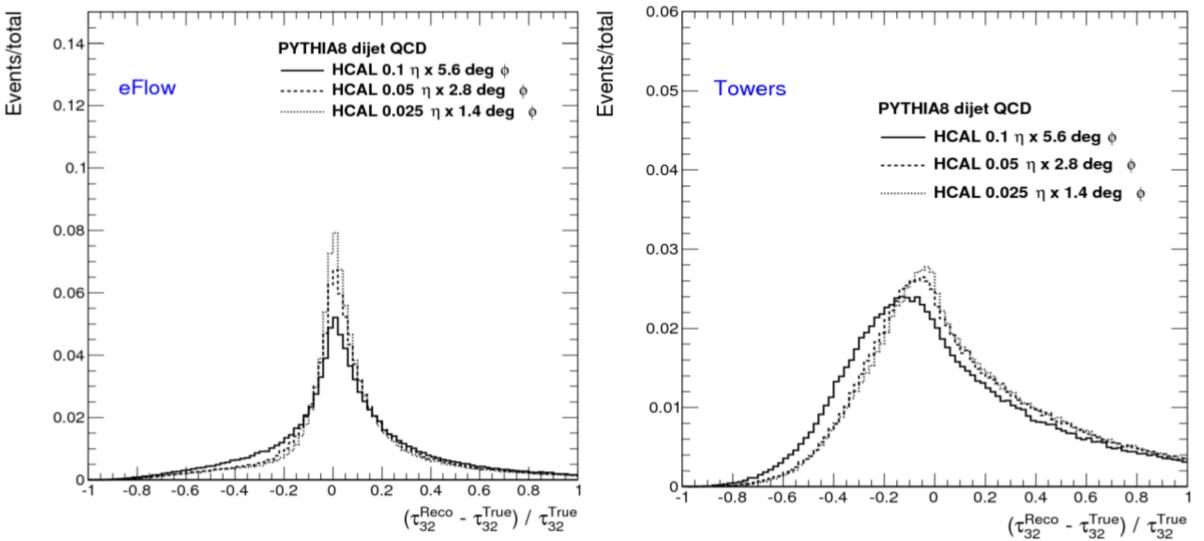


Figure 4.4:  $\tau_{32}$  resolution plots for QCD dijet events reconstructed using eflow jets (left) and tower jets (right). Figures from [68].

1080 **Chapter 5**

1081 **Sample generation and analysis tools**

1082 A crucial component of this work is the generation of the Monte Carlo samples that are used in the anal-  
1083 ysis. They are produced using fast simulation. We use the software and machinery that was developed  
1084 by the FCC study group at CERN. The simulation work flow and Monte Carlo samples are described in  
1085 this chapter.

1086 In section 5.1 we introduce the generators used to produce the Monte Carlo samples and briefly  
1087 describe their working principles and functionalities. We then focus on the FCC-hh software, in sub-  
1088 section 5.1.2, and explain how the implementation of the previously described simulation work flow is  
1089 implemented. A detailed description of the technical settings used to produce the samples is provided  
1090 in sections 5.2 and 5.2.3.

1091 **5.1 Fast simulation workflow**

- 1092 - Samples used for signal and backgrounds  
1093 - Number of events, cross section, how they were generated (gen level cuts)
- 

1094  
1095 The Monte Carlo samples used in this study were generated using MadGraph5\_aMC@NLO [70]. The  
1096 showering and hadronization are simulated using Pythia 8 [71] and the detector response is parametrized  
1097 using Delphes 3 [72] with the FCC-hh card. These samples are available at the CERN EOS storage.

1098 MadGraph5 is a matrix element generator for high energy physics processes, such as decays and  
1099 scatterings. The user specifies the desired process in terms of the initial and final states and can impose  
1100 additional constraints such as allowing for a number of refined criteria, including forced or forbidden s-  
1101 channel resonances, excluded internal particles, and forced decay chains of final state particles [70].  
1102 As a result, MadGraph automatically generates the corresponding Feynman diagrams and creates the  
1103 necessary code to compute the matrix element at a given point of the phase space.

1104 The Pythia8 program is frequently used for event generation in high energy physics. In this work,  
1105 however, we use it only to simulate the showering and hadronization process and not the parton level  
1106 hard process which is simulated in MadGraph5. The Les Houches file that is produced by MadGraph5

1107 is used as input to a Pythia8 program that can decay unstable particles, simulate initial and final state  
1108 showers as well as the hadronization of coloured particles, such as quarks and gluons. The desired  
1109 settings can be specified in an additional file, a card, that is used as input for a Pythia8 run.

1110 Delphes3 allows for a quick and simple simulation of the detector's response [72]. Its goal is to allow  
1111 the fast simulation of a multipurpose detector for phenomenological studies. The simulation includes a  
1112 track propagation system, electromagnetic and hadronic calorimeters and a muon identification system.  
1113 Low level physics objects, such as tracks and energy deposits, and high level physics objects, such as  
1114 leptons and jets, are reconstructed from the detector's response and can be used to perform physics  
1115 analysis. In the following paragraphs we briefly describe how the detector response is simulated and  
1116 how jet reconstruction is performed.

1117 The magnetic field is uniform, axial and parallel to the beam direction. Charged particles follow a  
1118 helicoidal trajectory from the interaction point to the calorimeter while neutral particles have a straight  
1119 line trajectory. The probability of a charged particle being reconstructed as a track is defined by the  
1120 user. Only a smear in the modulus of the transverse momentum is applied (not to the direction). The  
1121 tracking efficiency as well as energy and momentum resolutions are specified by the user and may  
1122 include a dependency on the particle type, momentum and pseudorapidity. The calorimeters have a  
1123 finite segmentation in the  $(\eta, \phi)$  plane and the cell size can be defined in the configuration file. For  
1124 simplicity, the segmentation is the same for the ECAL and in the HCAL and they are perfectly overlaid.  
1125 The amount of energy deposited in the calorimeters by each particle type can be defined by the user.  
1126 By default, stable hadrons deposit all their energy in the HCAL although in a real detector a significant  
1127 fraction of their energy is deposited in the ECAL. The energy resolution of the ECAL and HCAL are  
1128 parameterized as a function of  $\eta$  and include stochastic, noise and constant terms. The electromagnetic  
1129 and hadronic energy deposits are independently smeared by a log-normal distribution.

1130 Jets can be produced using generator level long-lived particles after showering and hadronization,  
1131 tracks, calorimeter towers or particle-flow tracks and towers. These are referred to as generator, track,  
1132 calorimeter or particle-flow jets, respectively. For generator jets no detector simulation nor reconstruction  
1133 are taken into account. In spite of the type of jet the user can choose the jet clustering algorithm and  
1134 the values of its parameters as well as the minimum transverse momentum of the jets that are stored in  
1135 the final collection. Delphes integrates the FastJet package [73] and therefore allows jet reconstruction  
1136 with the most popular jet reconstruction algorithms, namely, anti- $k_T$ ,  $k_T$  and C/A. Jets resulting from the  
1137 hadronization of a b quark, b jets, are identified if a b quark is found within a  $\Delta R$  distance from the jet's  
1138 axis. The tagging efficiency and mis-tagging probabilities can be defined by the user.

### 1139 **5.1.1 Particle flow and calorimeter reconstruction in Delphes**

1140 In Delphes, hadronic jets can be reconstructed using only the information from the HCAL towers or using  
1141 a particle flow algorithm that combines information from the tracking system and from the HCAL towers.  
1142 These two approaches create jets that are referred to as calorimeter and particle flow jets, respectively.  
1143 The latter can also be referred to as energy flow jets (eflow jets in short). In this work we performed the

1144 analysis using both sets of jets and compare the results. Therefore we briefly describe them here.

1145 Calorimeter jets are very simple. They are reconstructed using as input for the jet clustering algo-  
1146 rithm the 4-vectors associated with the calorimeter towers, after an energy smearing has been applied.  
1147 Therefore the energy resolution is limited by the transverse segmentation of the calorimeters.

1148 The goal of the particle flow approach is to make use of all the available information provided by the  
1149 various sub-detectors for reconstructing an event [72]. This approach is used by some experimental  
1150 collaborations [74, 75] but the exact implementation depends on the specificities of the experiment. If  
1151 the momentum resolution of the tracking system is better than the energy resolution of the calorimeters  
1152 it might be convenient to use the tracking information to estimate the momentum of charged particles.  
1153 In real experiments, the tracking resolution is only better than the calorimeter's energy resolution up to  
1154 some energy threshold. However, in Delphes, it is assumed that it is always convenient to estimate the  
1155 momentum of charged particles via the tracker.

1156 The particle flow algorithm works as follows [72]. For each calorimeter tower it counts:

- 1157
- the total energy deposited in ECAL and HCAL,  $E_{\text{ECAL}}$  and  $E_{\text{HCAL}}$ , respectively;
  - the total energy deposited in ECAL and HCAL originating from charged particles for which a track  
1159 has been reconstructed,  $E_{\text{ECAL,trk}}$  and  $E_{\text{HCAL,trk}}$ , respectively.

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

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

1160 All reconstructed tracks result in a particle flow track. If  $E_{\text{Tower}}^{\text{eflow}} > 0$  a particle flow tower is created with  
1161 energy  $E_{\text{Tower}}^{\text{eflow}}$ . The particle flow tracks and jets are then used as input for the jet clustering algorithms.

## 1162 5.1.2 FCC-hh software

1163 FCC software (FCCSW) [76], common to all FCC experiments (electron-electron, electron-hadron and  
1164 hadron-hadron) has been developed and is maintained by the FCC study group. The software is based  
1165 on Gaudi [77]. An FCC Event Data Model based on podio [78] was also developed. It consists in specific  
1166 classes that encode the information about the events.

1167 The FCC-hh study group is responsible for the generation of Monte Carlo quick simulation samples  
1168 for the main benchmark processes for the FCC-hh. The samples are generated using the workflow  
1169 described in the previous section. CERN users can request rights to run the EventProducer package  
1170 [79] and produce samples for any desired process using the machinery that is already implemented. In  
1171 addition, the FCC Event Data Model classes are directly accessible and can be used to read the ROOT  
1172 files that are produced after the events are pass through Delphes.

1173 The machinery to submit jobs to CERN's batch system is also implemented for both the generation  
1174 (MadGraph5) and reconstruction (Pythia8 plus Delphes3) levels.

1175 In this work we make use of this software in order to produce the necessary samples.

1176 **5.2 Signal and background samples**

1177 **5.2.1 MadGraph**

1178 The 4b+j samples are generated with an extra jet with a high  $p_T$  at generator level in order to force the  
1179 four b quarks to have a high Lorentz boost and therefore increases the probability of the events being  
1180 reconstructed with two large-R jets with a two-prong substructure. The 4b+j QCD sample is constituted  
1181 by two independent samples that have a different generator level cut in the minimum  $p_T$  of the light jets,  
1182 namely,  $200 < p_{T,j}^{\min} < 500$  GeV and  $p_{T,j}^{\min} > 500$  GeV. The jj+0/1/2 j is divided into several individual  
1183 samples that are produced in different  $H_T$  regions, where  $H_T$  is the scalar sum of the  $p_T$  of all partons  
1184 at generator level. The minimum allowed  $H_T$  is 500 GeV. Since these backgrounds are QCD processes,  
1185 the  $p_T$  distribution of the final state jets falls very steeply as the  $p_T$  increases. Therefore, if one were  
1186 to generate events for these processes without restricting the phase space, most events would consist  
1187 of jets with a very low  $p_T$  which are exactly the type of events that are rejected the most by a boosted  
1188 analysis (see chapter 6 for more details on the event topology that is targeted and on the analysis  
1189 strategy). Also note that for the 4b+j (QCD) and jj+0/1/2 j samples we do not take into account the  
1190 regions of the phase space with  $p_{T,j} < 200$  GeV and with  $0 < H_T < 500$  GeV. We assume that we can  
1191 reject most of the events (if not all) events with these kinematic characteristics by going to a sufficiently  
1192 boosted region of the phase space.

1193 The 4b+j is generated using the four flavor scheme meaning that the extra parton can be a gluon  
1194 or a quark from the first he second geerations. For the jj+0/1/2 j and  $t\bar{t}$ +0/1/2 j samples the five flavor  
1195 scheme is used and therefore the extra partons can also be b quarks. Notice that after the showering  
1196 procedure there could be some overlapp between the 4b+j and jj+0/1/2 j. This is taken care of in our  
1197 analysis code: if an event from the jj+0/1/2 j has four b quarks at truth level [GIVE PERCENTAGE] then  
1198 we do not consider it because it will certainly overlapp with an event from the 4b+j sample.

1199 For the  $hh$  SM sample, one of the Higgs is decayed to  $b\bar{b}$  in MadGraph. The reasoning behind this  
1200 choice is that most searches for Higgs pair production make use of final channel that include at least two  
1201 b quarks in order to keep the cross section times BR of the process large enough an this way the same  
1202 generator level sample can be used to perform different analysis. The decay of the remaining Higgs  
1203 boson can then be implemented in Pythia, in the case of this work, to  $b\bar{b}$ . For the  $hh$  BSM samples both  
1204 Higgs are decayed only in Pythia.

1205 **Signal samples - BSM**

1206 In addtion to the SM di-Higgs signal, we also explore the signature and analysis sensitivity for di-Higgs  
1207 signals produced by two benchmark BSM models: CP-conserving 2HDM and a simplified dark matter  
1208 model (DM) with a spin 0 mediator. These models were described in section 2.2.

1209 Both models are readily available in FeynRules [80] (which is a Mathematica package for the imple-  
1210 mentation of new physics models) model database and can be straightforwardly implemented in Mad-  
1211 graph5. The parameters of the models, namely the masses of the new particles, can be changed by

1212 the user. In the case of this work, we want new particles to have a large mass so that they produce SM  
1213 Higgs pairs with a high Lorentz boost.

Starting from [81], that constraints the parameter space of the CP-conserving 2HDM using ATLAS and CMS data collected during LHC's 7 and 8 TeV runs, we choose values for the masses of the additional Higgs bosons ( $h_2$ ,  $h_3$ ,  $h^\pm$ ) that are as high as possible but that are within the allowed (not excluded by experimental data or theoretical constraints) phase space of the model. We set:

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

1214 As a safety check, we test the model with these parameters at a CM energy of 13 TeV. We obtain a value  
1215 for the cross section of approximately 0.1 pb which smaller than the current experimental limit on the  
1216 cross section.

1217 In the case of the DM model, the spin 0 mediator's mass is set at 1 TeV. The cross section for the  
1218 signal generated with this model is smaller than the cross section of the SM signal (approximately 0.2  
1219 pb *versus* 0.7 pb) and therefore this model is not excluded by experimental data.

## 1220 5.2.2 Pythia

1221 For the signal sample we simply turn off all other decays except  $h \rightarrow b\bar{b}$  therefore forcing the Higgs to  
1222 decay to a pair of b quarks leading to the desired final state with four b quarks. All other settings are left  
1223 are not altered with respect to their default configuration.

1224 For the multijet and  $t\bar{t}$  samples we have to perform jet matching because we require additional jets  
1225 at the level of the matrix element. In addition to the partons generated in MadGraph and that produce a  
1226 jet, Pythia may introduce extra jets that are usually soft and collinear (with the particle from which they  
1227 were radiated) and result from the showering process. This could lead to the same process (meaning  
1228 processes with the same final states) being counted twice (double counting). Take, for example, the  
1229 processes  $j j$  and  $j j j$  at MadGraph level. It can happen that Pythia generates an extra jet for the first  
1230 process but not for the second, leading to both processes having the same final state (three jets). Each  
1231 process would then give its independent contribution to the total number of events but because they  
1232 simply represent two distinct ways of achieving the same final state they should only be counted once.  
1233 The goal of jet matching procedures is to avoid this problem.

1234 The cross section times branching ratio (when applicable), the k-factors and the total number of  
1235 events used in the analysis are summarized in table 5.1, for all samples. The k-factor multiplies the cross  
1236 section times branching ratio in order to reproduce known higher order results. This factor corresponds  
1237 only to the ratio between the total cross section and it does not correct for possible difference that might  
1238 exist between the differential cross sections.

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

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

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

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

### 5.2.3 Delphes

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

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

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



1259 

# Chapter 6

1260 

# Analysis

1261 

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

- 1262 - Signal cross section and final state signature  
1263 - Pros and challenges of using the 4b final state  
1264 - Main backgrounds (the ones we consider) and respective cross sections (leave for appendix discussion  
1265 on other backgrounds, namely Higgs processes)
- 

1266  
1267  
1268 The searches for Higgs pair production in the four b quarks final state benefit from the large BR of  
1269  $h \rightarrow b\bar{b}$ . For the SM Higgs, with a mass around 125 GeV, the branching ratio of the  $hh \rightarrow b\bar{b}b\bar{b}$  decay is  
1270 approximately 33.6%, making it the most probable decay for Higgs pairs, as it is illustrate in Figure ??.

1271 However, in this channel, the main background is QCD multijet production that has a cross section  
1272 several orders of magnitude larger than di-Higgs production in the SM, as Table 5.1 shows. Nonetheless,  
1273 it is a well known feature of QCD that the majority of jets produced are soft. This means that the  $p_T$   
1274 distributions of this background have a very large yield close to zero and then fall very steeply while the  
1275 signal has a much larger tail to high values of  $p_T$ . This indicates that searches targeting the boosted  
1276 kinematic regime may be the key to measure  $hh \rightarrow b\bar{b}b\bar{b}$  using inclusive production.

1277 In this work we target the boosted regime. In this kinematic region, the final state of the signal is  
1278 characterized by two jets with a large R parameter. Each jet is expected to contain the two b quarks  
1279 originated from the decay of a Higgs boson, as Figure 6.2 illustrates. Extra jets are also expected to be  
1280 reconstructed as a consequence of QCD activity.

1281 The analysis presented here is performed using the Monte-Carlo samples described in chapter 5. All  
1282 samples assume  $m_h = 125$  GeV. The event selection criteria are designed to optimize the significance,  
1283 given by  $S/\sqrt{B}$ , where  $S$  and  $B$  represent the number of signal and background events, respectively.

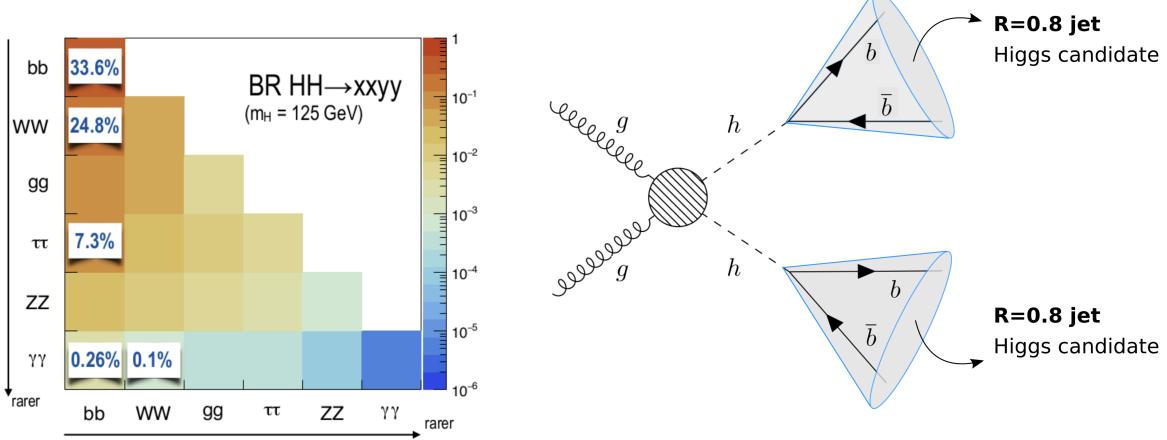


Figure 6.1: Higgs pairs branching ratios.

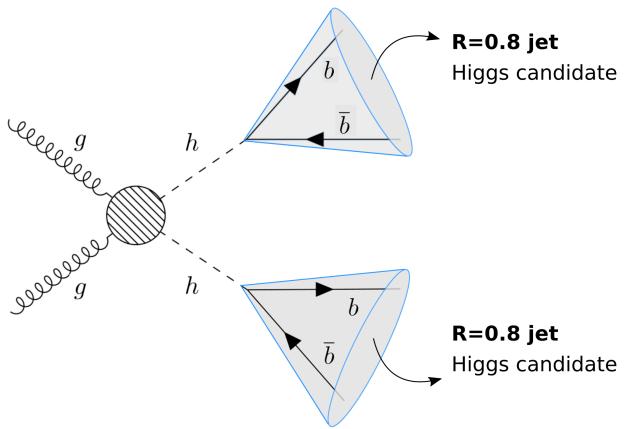


Figure 6.2: Event topology targeted by the boosted analysis region. The blob represents the interaction between the gluons and the Higg bosons that is represented by the Feynman diagrams shown in figure 2.3.

### 1284 6.1.1 Event pre-selection

1285 TO INCLUDE:

- 1286 - discussion on trigger requirements: important in real analysis  
 1287 - show and discuss some distributions: invariant mass (h, hh), momentum, tau21, delta eta (hh), more  
 1288 substructure variables (may be all that we considered relevant)  
 1289 - compare with other articles/analysis ?

1290

1291

1292 We start by applying very simple and loose cuts that target events consistent with the boosted topology.

1293 We require at least two b-tagged  $R = 0.8$  jets (which corresponds to at least four b-tagged subjets, at  
 1294 least two in each  $R = 0.8$  jet). In addition, the leading and sub leading jets must have  $p_T \geq 200$  GeV in  
 1295 order for the event to be accepted. From equation 3.4, a Higgs boson with  $p_T = 200$  GeV leads to a pair  
 1296 of b quarks with  $\Delta R \sim \frac{2m}{p_T} = 1.25$ . Therefore, considering jet with  $R = 0.8$  (diameter equal to 1.6) and  
 1297  $p_T \geq 200$  is a reasonable choice. For a more complete description about the jet radius see appendix A.  
 1298 Due to the b-tagging efficiency formulas (that are described in detail in section 6.2.1), there is a natural  
 1299 cutoff at  $|\eta| = 4$  so we do not place any additional cut in  $\eta$ . From now on, these cuts are referred to as  
 1300 pre-selection cuts.

1301 After the pre-selection cuts we can plot all the relevant distributions that allow us to characterize the  
 1302 signal and background events. These are shown in the next two sections. In addition, these distributions  
 1303 provide a first insight into the discriminating power of the kinematic and substructure variables that are  
 1304 explored.

1305 **6.1.2 Signal characterization**

1306 The cross section for di-Higgs inclusive production from  $pp$  collisions at 100 TeV is 1.2 pb [52]. This value  
1307 is computed at NLO accuracy. The dominant process is gluon-gluon fusion that has been extensively  
1308 discussed in section 2.1.1. The signal process comprehends the di-Higgs inclusive production and the  
1309  $hh \rightarrow b\bar{b}b\bar{b}$  subsequent decays. The total process cross section is then given by  $\sigma(pp \rightarrow hh) \times BR(hh \rightarrow$   
1310  $b\bar{b}b\bar{b})$  and it amounts to approximately 0.4 pb [NOT VERY COHERENT WITH MG VALUES - CHECK].  
1311 For an integrated luminosity of  $3000\text{fb}^{-1}$  of proton collisions at  $\sqrt{s} = 100$  TeV we expect X events.

1312 The Higgs candidates are reconstructed using large R jets that can be directly measured using  
1313 information from the calorimeters and tracking systems. Figure 6.3 shows the  $\Delta\phi$  between the two  
1314 leading jets (Higgs candidates),  $\Delta\phi(hh)$ . For a large fraction of the signal events (red curve)  $\Delta\phi(hh) \sim \pi$   
1315 which indicates that the jets are produced back-to-back in the transverse plane. There is also a small  
1316 fraction of events for which the jets have  $\Delta\phi(hh) \sim 1$ . For these events there must be at least a third jet  
1317 that balances the momentum of the jets pair.

1318 Each wide jet is expected to be consistent with having two subjets associated with the two b quarks  
1319 from the  $h \rightarrow b\bar{b}$  decay. In the Higgs rest frame, the two subjets are produced back-to-back conserving  
1320 momentum. In the laboratory frame, the  $\Delta R$  between the subjets depends on the momentum of the  
1321 Higgs boson, with larger  $\Delta R$  corresponding to Higgs bosons with a lower momentum. This correla-  
1322 tion between the  $\Delta R$  between the subjets of the leading Higgs candidate ( $\Delta R(b\bar{b})$ ) and its momentum  
1323 ( $p_T(h_1)$ ) is shown in figure 6.3.

1324 The  $p_T$  distributions of the leading and sub leading jets (Higgs candidates) are shown in figure 6.7  
1325 on the left and right, respectively. The histograms are normalized to unit area in order to allow for shape  
1326 comparison between them. For the leading jet the  $p_T$  spectrum peaks at approximately 350 GeV and  
1327 has a long tail for large values of  $p_T$ , in particular, longer than any of the backgrounds. For the sub  
1328 leading jet the spectrum is softer, as expected, but the tail of the distribution is still longer than for any of  
1329 the backgrounds. These distributions show that the signal process produces jets with a larger transverse  
1330 momentum than the jets produce by the background processes which indicates that a boosted analysis  
1331 might perform well.

1332 The  $\eta$  distribution of the leading jet is shown in figure 6.8 on the left. This distribution is limited to  
1333  $|\eta| < 4$  because the b-tagging efficiency goes to zero for  $|\eta| > 4$  (see section 6.2.1 for more details). The  
1334 shape is approximately gaussian which is due to the detector acceptance [EXPLAIN BETTER ?]. The  $\eta$   
1335 distribution of the sub leading jet is very similar to this one so we abstain from showing it here.

1336 The softdrop mass distribution of the leading jet is shown in figure 6.8 on the right. For the signal (red  
1337 curve) there is a clear peak at approximately 120 GeV which corresponds to the SM Higgs boson mass  
1338 peak. The peak is quite broad which means the mass resolution in this channel is not very good. On the  
1339 one hand, this is because we are using large-R jets to reconstruct the Higgs candidates. These objects  
1340 have a worse mass resolution than other cleaner objects such as photons or electrons. On the other  
1341 hand, the cuts applied before plotting these distributions are very loose. As we place additional cut the  
1342 mass peak can be made slightly narrower. Another interesting feature of the signal's mass spectrum is  
1343 the existence of a peak close to zero. The explanation is the following. Some jets do not contain both b

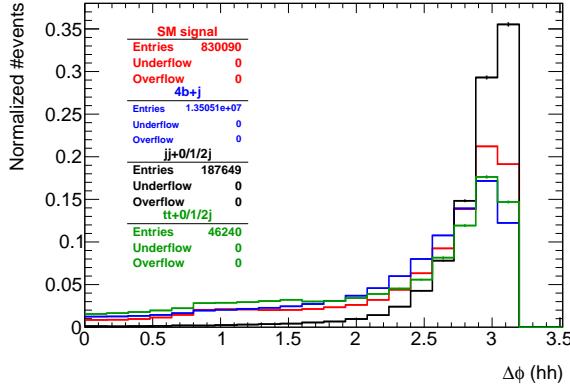


Figure 6.3:  $\Delta\phi$  between the b quarks coming from the decay of the leading Higgs candidate,  $h_1$ .

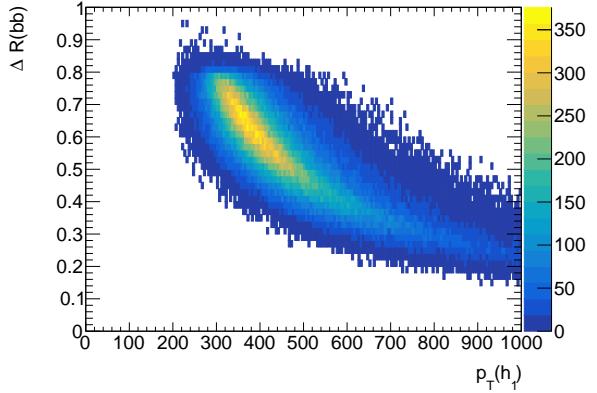


Figure 6.4:  $\Delta\phi$  between the Higgs candidates.

1344 quarks from the Higgs decay such that when applying the mass drop procedure only one b quark with a  
 1345 mass of approximately 5 GeV remains inside the jet creating the peak at lower masses. A more detailed  
 1346 discussion, as well as a plot that supports this interpretation, can be found in appendix C.

1347 In addition to the basic kinematic distributions that we just described there are a multitude of other  
 1348 variables one can look at. Here we show the  $\Delta\eta$  between the Higgs pairs, figure 6.9 on the left, and the  
 1349  $\tau_{21}$  variable for the leading Higgs candidate, figure 6.9 on the right. The first distribution shows that the  
 1350 Higgs pair tends to have a  $\Delta\eta$  very close to zero which indicates that the pair is highly boosted in the  
 1351 longitudinal plane [??]. The  $\tau_{21}$  distribution shows that for the signal this variable takes values close to  
 1352 zero which means the jet is consistent with having two subjets.

### 1353 6.1.3 Backgrounds

1354 The relevant backgrounds for this analysis are QCD (light?) multijet production,  $t\bar{t}$  and  $b\bar{b}b\bar{b}$ . Although  
 1355 the  $b\bar{b}b\bar{b}$  background is a particular case of a QCD multijet production process we consider it separately  
 1356 because it constitutes the irreducible background and therefore it will have a higher efficiency in the  
 1357 analysis. The cross sections for these processes are several orders of magnitude larger than the cross  
 1358 section for the signal, as Table 5.1 shows. In addition, in the case of the  $t\bar{t}$  background, the event  
 1359 topology is expected to be similar to the signal because it also consists in the production of a pair of  
 1360 particles with the same mass.

1361 The assumption that QCD multijet production and  $t\bar{t}$  two main backgrounds is corroborated by the  
 1362 ATLAS di-Higgs search performed in the same channel where these backgrounds are found to be the  
 1363 dominant ones and that all other sources of backgrounds, including processes involving Higgs bosons,  
 1364 are found to be negligible [54]. In appendix B we discuss and evaluate the importance of some back-  
 1365 grounds that include Higgs bosons to our analysis.

1366 Figure 6.5 shows examples of LO Feynman diagrams that contribute to  $b\bar{b}b\bar{b}+j$  (left), three (middle)  
 1367 and four (right) light jets production. The  $b\bar{b}b\bar{b}$  background is generated with an extra light jet at generator  
 1368 level. This jet boosts the four b quarks and has a minimum  $p_T$  of 200 GeV. This increases the probability  
 1369 of two b quarks being reconstructed as single large R jet therefore emulating the signal's final state

signature. 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 originated 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.

#### 1376 CROSS SECTION FOR THESE PROCESSES.

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 6.7. In addition, for this background the  $\tau_{21}$  variable takes values close to one, as it is shown in figure 6.9 on the right. This indicates that the jets are not consistent with two subjets.

Figure 6.6 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 (represented as  $t\bar{t}+0/1/2\ j$ ), at generator level. The extra jets can originated from light and b quarks and from gluons. A jet matching procedure is implemented in Pythia in order to avoid double counting. The top quark has a very short life time, predicted to be  $5 \times 10^{-25}$  s, such that it decays before it can hadronize. This sample is inclusive in the top quark decay modes. However, the most favoured decay of the top quark is  $t \rightarrow Wb$  with a branching ratio close to 96% [82]. Therefore,  $t\bar{t}$  events will, most of the times, result in the  $WWbb$  final state. We do not specify any decay mode for the  $W$  such that the sample is also inclusive in the  $W$  decay modes. The  $W$  decays to hadrons ( $W^+ \rightarrow q\bar{q}$ ) and leptons ( $W^+ \rightarrow l^+\nu$ ) with  $BR(W^+ \rightarrow q\bar{q}) \sim 68\%$  and  $BR(W^+ \rightarrow l^+\nu) \sim 10\%$ . If one (or both)  $W$  bosons decays to hadrons then there will be additional jets in the final state. These can be b jets or can be misidentified as such. If both  $W$  bosons decay to leptons there will still be at least two b jets in the final state, coming from the  $t\bar{t}$  decay.

#### 1394 CROSS SECTION.

It is interesting to note that in the softdrop mass plot for the leading Higgs candidate (figure 6.8 on the right) the  $t\bar{t}$  background (green line) shows a small peak around 170 GeV. In this region, all the decay products of the top quark are contained inside the  $R = 0.8$  jet.

## 1398 6.2 Analysis strategy

- 
- 1399
- 1400
- 1401 - Event topology: two boosted jets each corresponding to a Higgs
  - 1402 - Physics objects: particle flow anti-kT R=0.8 jets (discussion about jet radius in appendix)
  - 1403 - Selection criteria
  - 1404 - Substructure variables
  - 1405 - Optimization (efficiency plots, correlations, MVA...)
-

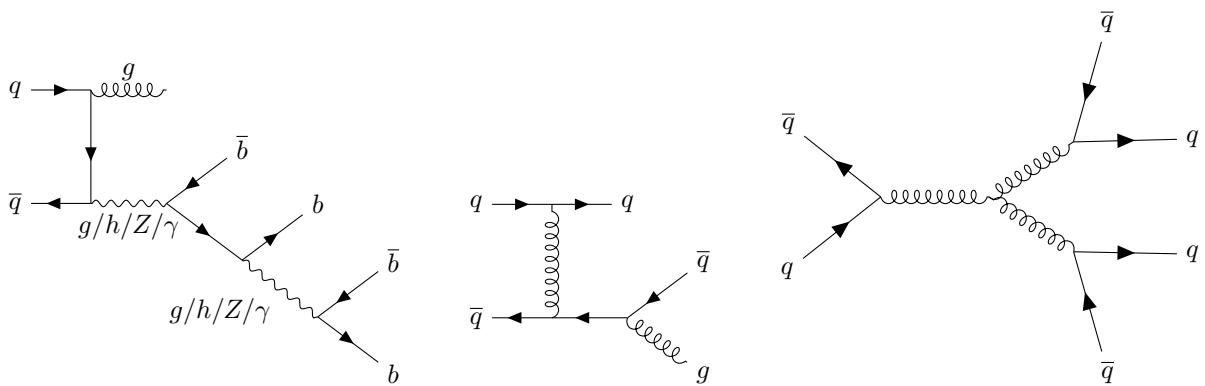


Figure 6.5: Example of diagrams that contribute to the QCD multijet background: five final state jets, four of which are b-jets (left), three final state jets (middle) and four final state jets (right). Here,  $q$  stands for a light quark/jet.

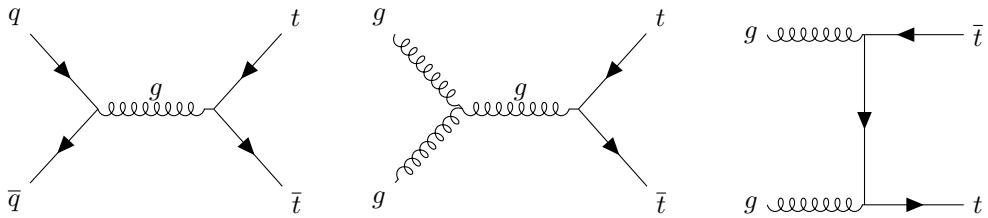


Figure 6.6: Dominant diagrams of  $pp \rightarrow t\bar{t}$  at LO.

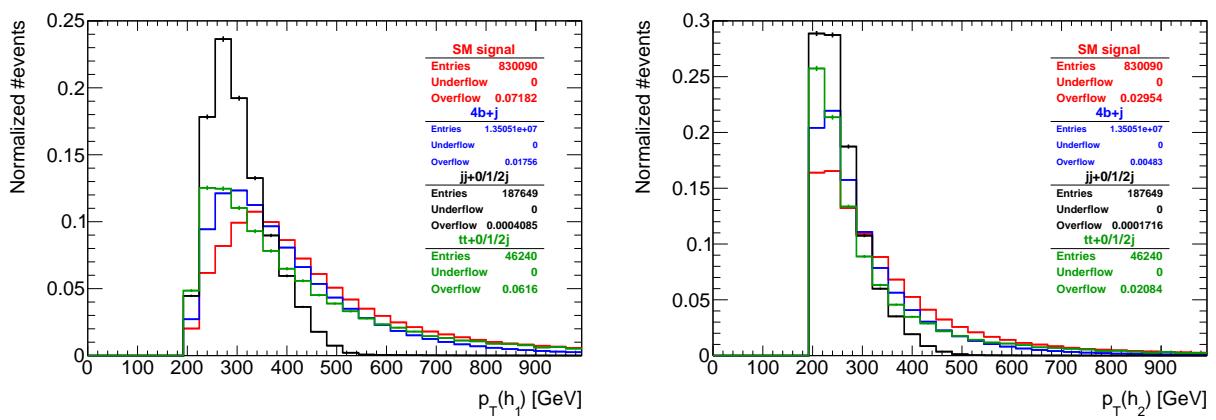


Figure 6.7:  $p_T$  distributions for the leading (left) and sub leading (right) Higgs candidates. The signal is the SM  $hh \rightarrow b\bar{b}$  process. The histograms are normalized to unit area.

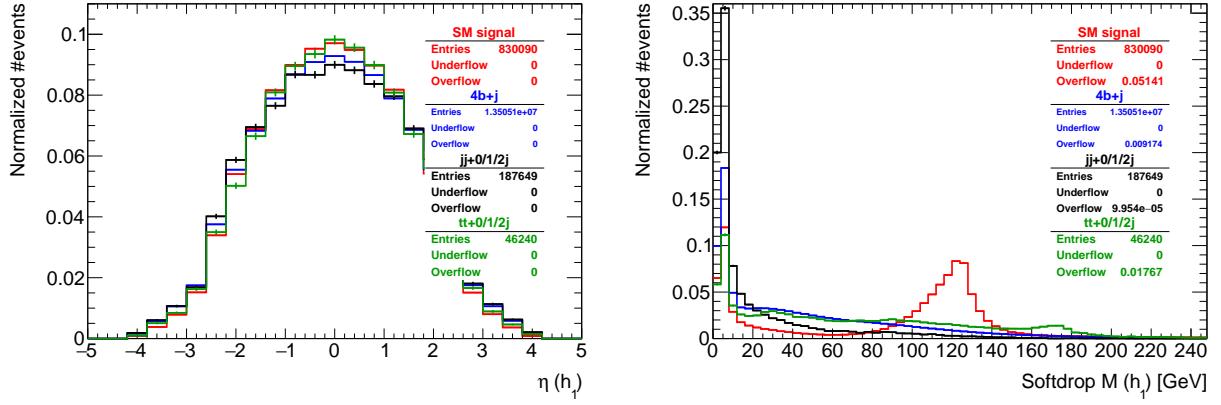


Figure 6.8:  $\eta$  distribution for the leading Higgs candidate (left) and softdrop mass distribution for the leading Higgs candidate (right). The signal is the SM  $hh \rightarrow b\bar{b}$  process. The histograms are normalized to unit area.

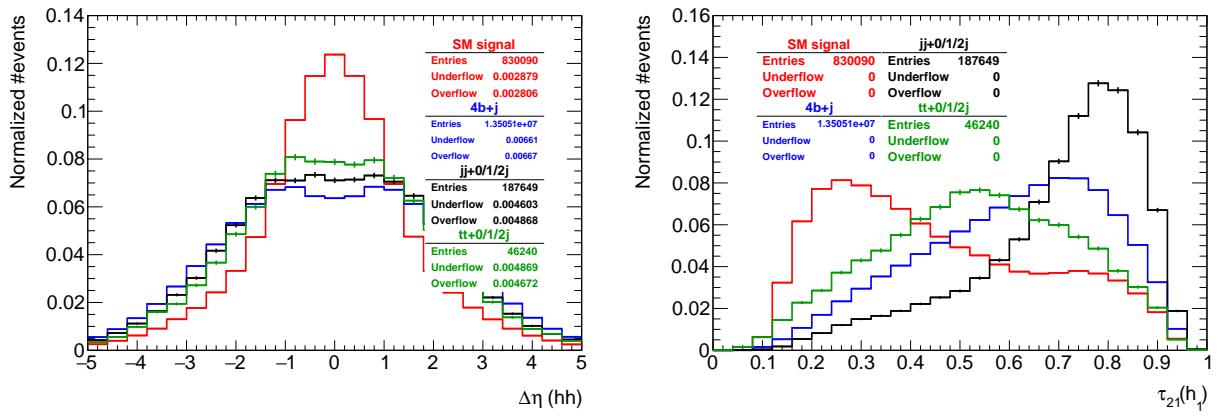


Figure 6.9:  $\text{oi}$

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

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

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

The events are reconstructed using particle flow (or eflow) or pure calorimeter jets (we explored both approaches) with  $R = 0.8$ , 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/strategy that is described in section 6.2.1.

### 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  $\Delta R$  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,  $\Delta R(\text{subjet}, \text{parton})$ . 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  $\Delta R(\text{subjet}, \text{parton}) < 0.3$ , as indicated by the plot in Figure 6.10. In this plot there is a peak for  $\Delta R < 0.3$  that corresponds to the If the subjet is matched to at least a b quark, we b-tag the subjet with a given probability. If the subjet is not matched to any b quark but it is matched to at least one c quark we apply a c mistag rate. If the subjet is not matched to any b or c quark we apply a light mistag rate. The b-tag probability and mistag rates were obtained from the Delphes FCC-hh card. They depend on the momentum of the jet and on its  $\eta$  coordinate. They are summarized in Table 6.1. Note that a jet cannot be b-tagged if  $|\eta| > 4$  or if its momentum is smaller than 10 GeV or larger than 15000 GeV.

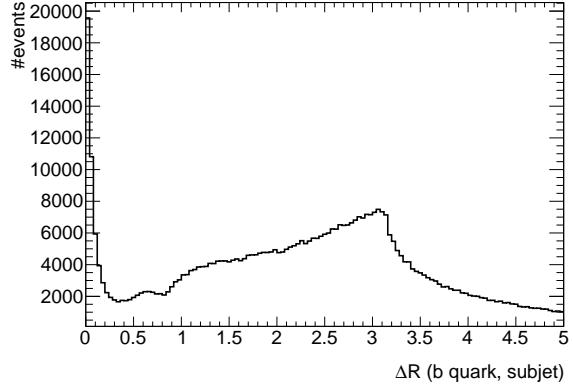


Figure 6.10: Minimum  $\Delta R$  between b quarks and subjets of the  $R = 0.8$  jets.

If the subjet is matched to at least a b quark, we b-tag the subjet with a given probability. If the subjet is not matched to any b quark but it is matched to at least one c quark we apply a c mistag rate. If the subjet is not matched to any b or c quark we apply a light mistag rate. The b-tag probability and mistag rates were obtained from the Delphes FCC-hh card. They depend on the momentum of the jet and on its  $\eta$  coordinate. They are summarized in Table 6.1. Note that a jet cannot be b-tagged if  $|\eta| > 4$  or if its momentum is smaller than 10 GeV or larger than 15000 GeV.

1437 **6.2.2 Baseline analysis**

1438 The analysis described in this section and on the ones that follow were developed using the sample  
1439 simulated with the default FCC-hh detector implementation. The analysis was then applied to the sam-  
1440 ples generated using the different detector configurations. This allows for a straightforward comparison  
1441 between 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 momentums of the leading and sub leading Higgs candidates,  $p_T(h_1)$  and  $p_T(h_2)$ , and of the Higgs pair,  $p_T(hh)$ :

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

1442 These cuts follow from both distributions shown in figure 6.11 and guarantee that we are choosing events  
1443 for which the Higgs candidates are sufficiently boosted. In addition, a high threshold for the  $p_T$  of the  
1444 jets suppresses  $t\bar{t}$  events because the decay products of the top quark are reconstructed in a single jet.  
1445 In a real analysis, these cuts would also be dictated by the trigger threshold.

From figure 6.12, on the left, we then apply a cut on the  $\tau_{21}$  of the leading Higgs candidate to select jets that are more compatible with a "2-prong" structure:

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

The cut on this variable works as an Higgs tagging method. Therefore we apply the same cut on the  $\tau_{21}$  of the sub leading Higgs candidate although that does not necessarily follow from the distribution on the right of figure 6.12. Then, from the distribution on the right of figure 6.13, we place a cut on the second Fox-Wolfram momentum of the leading Higgs candidate,  $H_2(h_1)$ :

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

1446 This substructure variable is particularly interesting because it helps supress the  $t\bar{t}$  background.

Finally, based on the distribution shown in figure 6.13 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:

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

1447 Using this analysis, we achieve a significance,  $S/\sqrt{B}$ , of approximately 6 for an integrated luminosity  
1448 of  $30 \text{ ab}^{-1}$ .

1449 **6.2.3 Optimization**

1450 We explored different methods to optimize the baseline analysis in order to try to increase the achieved  
1451 significance.

1452 TO INCLUDE:

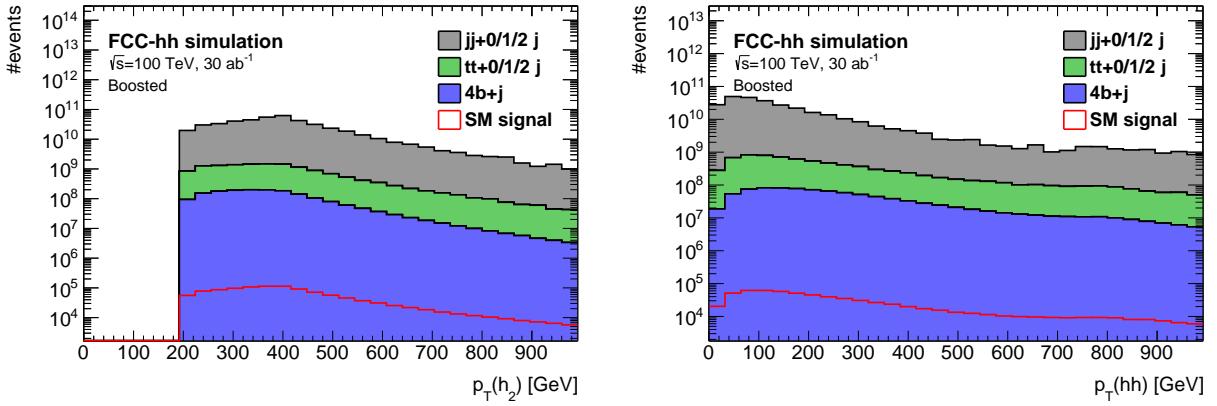


Figure 6.11:  $p_T$  distributions for the sub leading Higgs candidate (left) and for the Higgs pair (right). The histograms are normalized to  $\mathcal{L} = 30 \text{ ab}^{-1}$ .

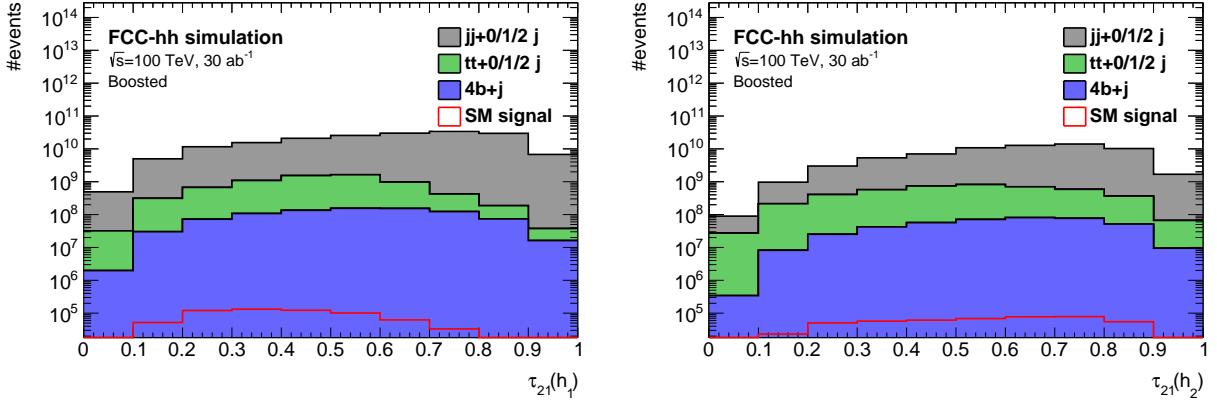


Figure 6.12: Distributions of the  $\tau_{21}$  variable for the leading (left) and sub leading (right) Higgs candidates. The histograms are normalized to  $\mathcal{L} = 30 \text{ ab}^{-1}$ .

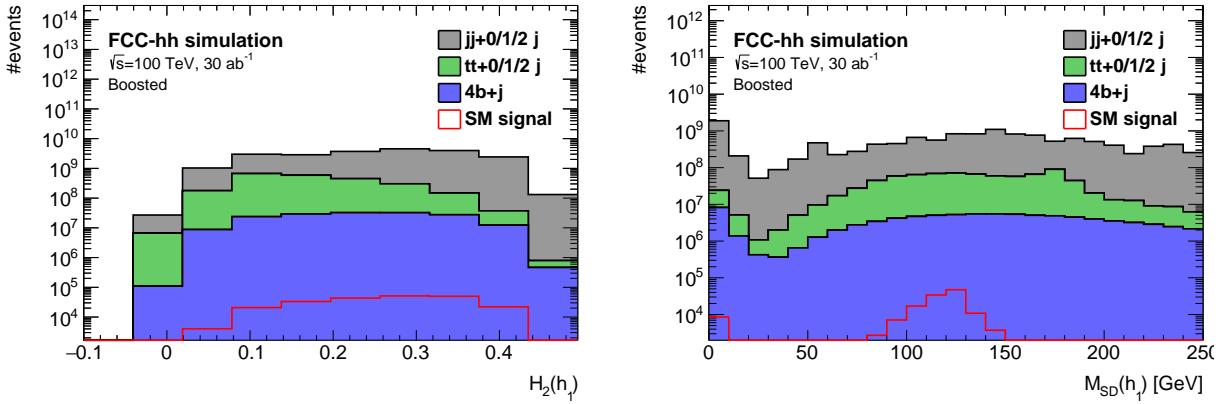


Figure 6.13:  $H_2$  variable for the leading Higgs candidate (left) and softdrop mass distribution for the leading Higgs candidate. The histograms are normalized to  $\mathcal{L} = 30 \text{ ab}^{-1}$ .

- 1453 - Mention study on correlations ?
- 1454 - Optimization of cut-based analysis based on  $S/\sqrt{B}$  plots
- 1455 - MVA analysis ?



1456 **Chapter 7**

1457 **Results**

1458 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  
1459 benchmark luminosities,  $3 \text{ ab}^{-1}$  and  $30 \text{ ab}^{-1}$  (section 7.1). The statistical analysis used to extract the  
1460 signal strength and to set limits on the Higgs boson triple coupling is also discussed. In section 7.2 we  
1461 show how the significance of the analysis varies as a function of the granularity of the HCAL and/or the  
1462 detector configuration. We also compare the results obtained using particle flow and pure calorimeter  
1463 jets.

1464 **7.1 Di-Higgs discovery potential at the FCC-hh**

1465 **7.1.1 Statistical analysis**

1466 **7.2 Hadronic calorimeter granularity studies for future colliders**

1467 **7.3 Benchmarking against ATLAS analysis at  $\sqrt{s} = 13 \text{ TeV}$**

<b>Selection</b>	SM	2HDM	4b+j	jj+0/1/2 j	t̄t+0/1/2 j
Gen level	100 (3.46e7)	100 (5.55e7)	100 (4.90e10)	100 (5.47e14)	100 (2.25e12)
N(b-tags) $\geq 4$	92.5	91.8	75.8	3.963	53.5
$p_T(j_1, j_2) \geq 200 \text{ GeV}$	16.6 (5.75e6)	30.6 (1.70e7)	17.8 (8.73e9)	0.74 (4.06e12)	1.06 (2.38e10)
<b>Pre-selection cuts</b>					
$p_T(j_1) \geq 400 \text{ GeV}$	8.623	18.2	7.0	0.18	0.446
$p_T(j_2) \geq 350 \text{ GeV}$	5.7	11.2	3.9	0.121	0.263
$p_T(j_1 + j_2) \geq 100 \text{ GeV}$	4.648	8.16	3.3	0.07	0.223
$\tau_{21}(j_1, j_2) < 0.55$	1.7	3.3	0.54	0.005	0.069
$FW2(j_1) > 0.2$	1.28	2.7	0.32	0.003	0.020
$(100 < M_{SD}(j_1, j_2) < 135) \text{ GeV}$	0.422 (1.46e5)	1.09 (6.07e5)	0.0136 (6.66e6)	0.00008 (4.38e8)	0.00078 (1.75e7)

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

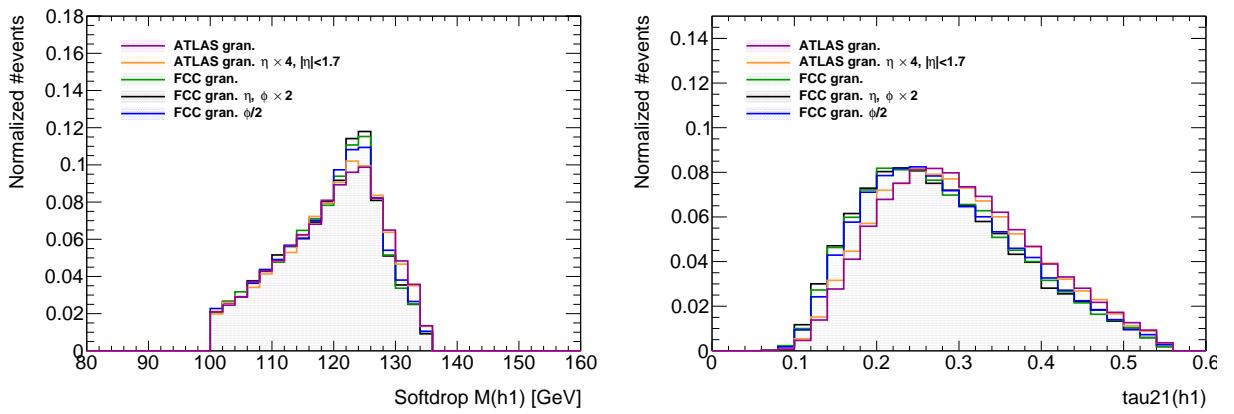


Figure 7.1: Leading Higgs candidate softdrop mass (left) and  $\tau_{21}$  (right) plots for the different detector configurations for the SM signal sample. The plots contain all the signal events that passed all the cuts of the baseline analysis. For the plot on the left note that the x axis range is from 80 GeV to 160 GeV in order to make the differences between the histograms more clear.

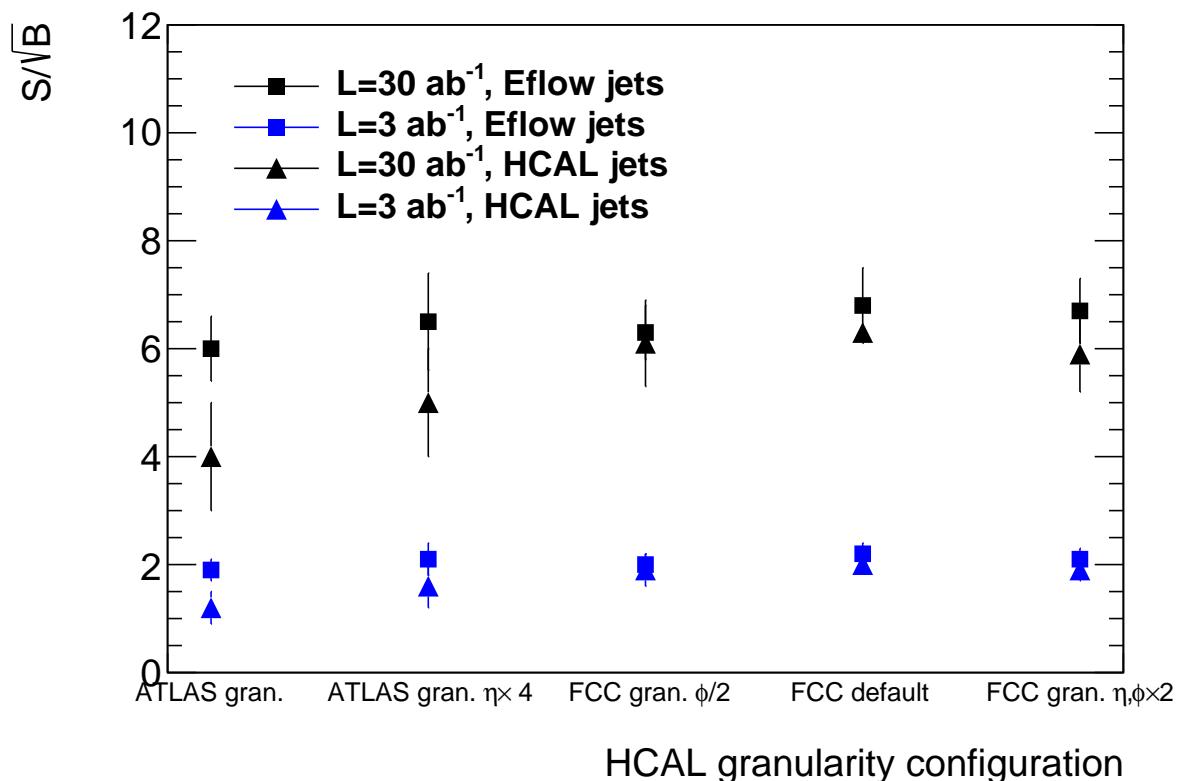


Figure 7.2: Significance ( $S/\sqrt{B}$ ) as a function of the detector configuration for  $\mathcal{L} = 30 \text{ ab}^{-1}$  (black) and  $\mathcal{L} = 3 \text{ ab}^{-1}$  (blue) and for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles) for the SM signal sample.

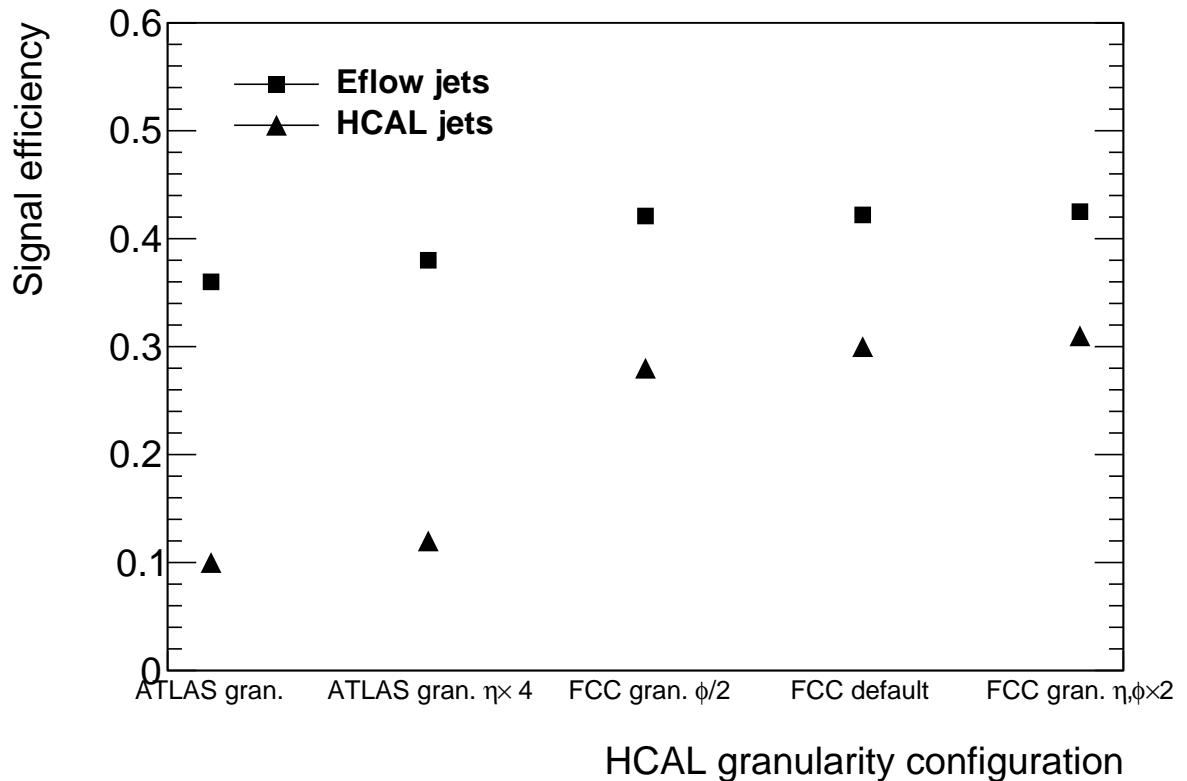


Figure 7.3: Signal efficiency as a function of the detector configuration for particle flow jets (squares) and pure calorimeter (HCAL) jets (triangles).

1468

# Chapter 8

1469

## Conclusions

1470

### BOOSTED H TO BB OBSERVATION AT THE FCC-HH

- 1471 - Using a simple cut based analysis we achieved a  $S/\sqrt{B}$  of approximately 6 for  $\mathcal{L} = 30 \text{ ab}^{-1}$ . This is  
1472 above the observation threshold ( $5\sigma$ ). This result indicates that using the full dataset collected during  
1473 the operation of the FCC-hh could lead to the observation of di-Higgs production, using the final state  
1474 with four b quarks.
- 1475 - The sensitivity to the Higgs triple coupling is ...
- 1476 - For  $\mathcal{L} = 3 \text{ ab}^{-1}$  the achieved significance is approximately 2 which is still below the evidence threshold  
1477 ( $3\sigma$ ).

1478

1479

### GRANULARITY STUDIES - SM

- 1480 - Based only on the plots of the Higgs invariant mass and of some substructure variables for the sig-  
1481 nal sample we can see a clear difference between the different granularity configurations. In particular,  
1482 the mass resolution increases slightly and for the tau\_21 we see a shift to the left. These plots work as  
1483 good safety checks and highlight the impact of the different granularity configurations. However they are  
1484 based only on the signal sample and therefore do not provide any information regarding the change in  
1485 significance.
- 1486 - Using Eflow (particle flow) jets there is not a big change in  $S/\sqrt{B}$  as we change the detector configu-  
1487 rations. The largest difference is of X %.
- 1488 - Based on this observation we redid the analysis using pure calorimeter jets (reconstructed using only  
1489 information from the HCAL). The difference in  $S/\sqrt{B}$  over the detector configurations tested increases.  
1490 The largest difference is of Y %.
- 1491 - Eflow jets use information from both the tracking and the hadronic calorimeter (in the case of hadronic  
1492 jets). Based on the fact that  $S/\sqrt{B}$  varies very little when using eflow jets and that its variation increases  
1493 if we use calorimeter jets we conclude that the resolution of the tracking system is so high that it domi-  
1494 nates the jet reconstruction.
- 1495 - Nonetheless the signal efficiency (computed as the ratio between the number of signal events after all  
1496 the analysis cuts and the total number of events) increases as the granularity of the detector increases

<sup>1497</sup> when using eflow and calorimeter jets. This increment is more accentuated in the case of calorimeter  
<sup>1498</sup> jets where the maximum difference is Z %. For Eflow jets the maximum variation is W %.

<sup>1499</sup>

<sup>1500</sup>

<sup>1501</sup> GRANULARITY STUDIES - BSM

<sup>1502</sup> - In addition to the SM production of Higgs pairs we also analyzed two benchmark models in which new  
<sup>1503</sup> heavy particles couple to the SM Higgs boson through the s-channel diagram.

# 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  $s^{**}(1/2) = 540\text{-GeV}$ . *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>.

- 1530 [11] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Phys. Rev.*, 109:1015–1017,  
 1531 Feb 1958. doi: 10.1103/PhysRev.109.1015. URL <https://link.aps.org/doi/10.1103/PhysRev.109.1015>.
- 1532
- 1533 [12] S. M. Bilenky and J. Hosek. GLASHOW-WEINBERG-SALAM THEORY OF ELECTROWEAK IN-  
 1534 TERATIONS AND THE NEUTRAL CURRENTS. *Phys. Rept.*, 90:73–157, 1982. doi: 10.1016/0370-1573(82)90016-3.
- 1535
- 1536 [13] M. Mangano. *Physics at the FCC-hh, a 100 TeV pp collider*. CERN Yellow Reports: Monographs.  
 1537 CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2270978>.
- 1538 [14] T. Plehn, M. Spira, and P. M. Zerwas. Pair production of neutral Higgs particles in gluon-  
 1539 gluon collisions. *Nucl. Phys.*, B479:46–64, 1996. doi: 10.1016/0550-3213(96)00418-X,10.1016/0550-3213(98)00406-4. [Erratum: Nucl. Phys.B531,655(1998)].
- 1540
- 1541 [15] S. Borowka, N. Greiner, G. Heinrich, S. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke. Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence. *Phys. Rev. Lett.*, 117(1):012001, 2016. doi: 10.1103/PhysRevLett.117.079901,10.1103/PhysRevLett.117.012001. [Erratum: Phys. Rev. Lett.117,no.7,079901(2016)].
- 1542
- 1543
- 1544
- 1545 [16] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke. Full top quark mass dependence in Higgs boson pair production at NLO. *JHEP*, 10:107, 2016. doi: 10.1007/JHEP10(2016)107.
- 1546
- 1547
- 1548 [17] H. C. van de Hulst, E. Raimond, and H. van Woerden. Rotation and density distribution of the Andromeda nebula derived from observations of the 21-cm line. 14:1, Nov. 1957.
- 1549
- 1550 [18] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567,  
 1551 1998. doi: 10.1103/PhysRevLett.81.1562.
- 1552
- 1553 [19] Q. R. Ahmad et al. Measurement of the rate of  $\nu_e + d \rightarrow p + p + e^-$  interactions produced by  
 $^8B$  solar neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, 2001. doi:  
 1554 10.1103/PhysRevLett.87.071301.
- 1555
- 1556 [20] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current  
 1557 interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002. doi:  
 10.1103/PhysRevLett.89.011301.
- 1558
- 1559 [21] B. Gripaios. Lectures on Physics Beyond the Standard Model. 2015.
- 1560
- 1561 [22] H. Sun, Y.-J. Zhou, and H. Chen. Constraints on large-extra-dimensions model through 125-  
 GeV Higgs pair production at the LHC. *Eur. Phys. J.*, C72:2011, 2012. doi: 10.1140/epjc/s10052-012-2011-4.
- 1562
- 1563 [23] P. Huang, A. Joglekar, M. Li, and C. E. M. Wagner. Corrections to di-Higgs boson production  
 1564 with light stops and modified Higgs couplings. *Phys. Rev.*, D97(7):075001, 2018. doi: 10.1103/PhysRevD.97.075001.

- 1565 [24] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with  
 1566 the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- 1567 [25] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment  
 1568 at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
- 1569 [26] M. Aaboud et al. Evidence for the  $H \rightarrow b\bar{b}$  decay with the ATLAS detector. 2017.
- 1570 [27] Evidence for the higgs boson decay to a bottom quarkantiquark pair. *Physics Letters B*, 780:501  
 1571 – 532, 2018. ISSN 0370-2693. doi: <https://doi.org/10.1016/j.physletb.2018.02.050>. URL <http://www.sciencedirect.com/science/article/pii/S0370269318301618>.
- 1572 [28] A. M. Sirunyan et al. Observation of the Higgs boson decay to a pair of  $\tau$  leptons with the CMS  
 1573 detector. *Phys. Lett.*, B779:283–316, 2018. doi: 10.1016/j.physletb.2018.02.004.
- 1574 [29] A. M. Sirunyan et al. Observation of  $t\bar{t}H$  production. *Phys. Rev. Lett.*, 120:231801, 2018. doi:  
 1575 [10.1103/PhysRevLett.120.231801](https://doi.org/10.1103/PhysRevLett.120.231801).
- 1576 [30] M. Aaboud et al. Observation of Higgs boson production in association with a top quark pair at the  
 1577 LHC with the ATLAS detector. 2018.
- 1578 [31] Measurement of the Higgs boson mass in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  channels with  
 1579  $\sqrt{s}=13\text{TeV}$   $pp$  collisions using the ATLAS detector. Technical Report ATLAS-CONF-2017-046,  
 1580 CERN, Geneva, Jul 2017. URL <http://cds.cern.ch/record/2273853>.
- 1581 [32] G. Aad et al. Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and  
 1582 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015. doi: 10.1103/  
 1583 [PhysRevLett.114.191803](https://doi.org/10.1103/PhysRevLett.114.191803).
- 1584 [33] M. Aaboud et al. Measurement of the  $W$ -boson mass in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS  
 1585 detector. *Eur. Phys. J.*, C78(2):110, 2018. doi: 10.1140/epjc/s10052-017-5475-4.
- 1586 [34] R. Nisius. Measurements of the top quark mass with the ATLAS detector. *PoS*, EPS-HEP2017:  
 1587 453, 2017. doi: 10.22323/1.314.0453.
- 1588 [35] Measurement of the top quark mass with muon+jets final states in  $pp$  collisions at  $\sqrt{s} = 13$  TeV.  
 1589 Technical Report CMS-PAS-TOP-16-022, CERN, Geneva, 2017. URL [http://cds.cern.ch/](http://cds.cern.ch/record/2255834)  
 1590 record/2255834.
- 1591 [36] CERN. The high-luminosity lhc. <http://cds.cern.ch/record/2114693>, 2015. Accessed  
 1592 2018/01/11.
- 1593 [37] T. Koffas. ATLAS Higgs physics prospects at the high luminosity LHC. *PoS*, ICHEP2016:426, 2017.
- 1594 [38] A. M. Henriques Correia. The ATLAS Tile Calorimeter. Technical Report ATL-TILECAL-PROC-  
 1595 2015-002, CERN, Geneva, Mar 2015. URL <https://cds.cern.ch/record/2004868>.
- 1596 [39] ATLAS liquid argon calorimeter: Technical design report. 1996.

- 1598 [40] M. Cacciari, G. P. Salam, and G. Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008. doi:  
1599 10.1088/1126-6708/2008/04/005.
- 1600 [41] Optimisation of the ATLAS  $b$ -tagging performance for the 2016 LHC Run. Technical Report ATL-  
1601 PHYS-PUB-2016-012, CERN, Geneva, Jun 2016. URL <https://cds.cern.ch/record/2160731>.
- 1602 [42] A. Ruiz-Martinez and A. Collaboration. The Run-2 ATLAS Trigger System. Technical Report ATL-  
1603 DAQ-PROC-2016-003, CERN, Geneva, Feb 2016. URL <https://cds.cern.ch/record/2133909>.
- 1604 [43] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*,  
1605 D48:3160–3166, 1993. doi: 10.1103/PhysRevD.48.3160.
- 1606 [44] M. Wobisch and T. Wengler. Hadronization corrections to jet cross-sections in deep inelastic scat-  
1607 tering. In *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany,*  
1608 1998–1999, pages 270–279, 1998. URL [https://inspirehep.net/record/484872/files/arXiv:  
1609 hep-ph\\_9907280.pdf](https://inspirehep.net/record/484872/files/arXiv-hep-ph_9907280.pdf).
- 1610 [45] M. Cacciari, G. P. Salam, and G. Soyez. The Anti- $k(t)$  jet clustering algorithm. *JHEP*, 04:063, 2008.  
1611 doi: 10.1088/1126-6708/2008/04/063.
- 1612 [46] A. Altheimer et al. Boosted objects and jet substructure at the LHC. Report of BOOST2012, held  
1613 at IFIC Valencia, 23rd-27th of July 2012. *Eur. Phys. J.*, C74(3):2792, 2014. doi: 10.1140/epjc/  
1614 s10052-014-2792-8.
- 1615 [47] G. A. et al. Performance of jet substructure techniques for large- $r$  jets in proton-proton collisions  
1616 at  $\sqrt{s} = 7$  tev using the atlas detector. *Journal of High Energy Physics*, 2013(9), Sep 2013. doi:  
1617 10.1007/JHEP09(2013)076. URL [https://doi.org/10.1007/JHEP09\(2013\)076](https://doi.org/10.1007/JHEP09(2013)076).
- 1618 [48] J. Thaler and K. Van Tilburg. Identifying Boosted Objects with N-subjettiness. *JHEP*, 03:015, 2011.  
1619 doi: 10.1007/JHEP03(2011)015.
- 1620 [49] C. Chen. New approach to identifying boosted hadronically-decaying particle using jet substructure  
1621 in its center-of-mass frame. *Phys. Rev.*, D85:034007, 2012. doi: 10.1103/PhysRevD.85.034007.
- 1622 [50] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam. Jet substructure as a new Higgs  
1623 search channel at the LHC. *Phys. Rev. Lett.*, 100:242001, 2008. doi: 10.1103/PhysRevLett.100.  
1624 242001.
- 1625 [51] I. Hinchliffe, A. Kotwal, M. L. Mangano, C. Quigg, and L.-T. Wang. Luminosity goals for a 100-TeV  
1626 pp collider. *Int. J. Mod. Phys.*, A30(23):1544002, 2015. doi: 10.1142/S0217751X15440029.
- 1627 [52] N. Arkani-Hamed, T. Han, M. Mangano, and L.-T. Wang. Physics opportunities of a 100 TeV pro-  
1628 tonproton collider. *Phys. Rept.*, 652:1–49, 2016. doi: 10.1016/j.physrep.2016.07.004.
- 1629 [53] J. Baglio, A. Djouadi, R. Grber, M. M. Mhlleitner, J. Quevillon, and M. Spira. The measurement  
1630 of the Higgs self-coupling at the LHC: theoretical status. *JHEP*, 04:151, 2013. doi: 10.1007/  
1631 JHEP04(2013)151.

- 1632 [54] M. Aaboud et al. Search for pair production of Higgs bosons in the  $b\bar{b}b\bar{b}$  final state using proton–  
 1633 proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *Phys. Rev.*, D94(5):052002, 2016. doi:  
 1634 10.1103/PhysRevD.94.052002.
- 1635 [55] A. M. Sirunyan et al. Search for a massive resonance decaying to a pair of Higgs bosons in the  
 1636 four b quark final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2017.
- 1637 [56] G. Aad et al. Searches for Higgs boson pair production in the  $hh \rightarrow b\bar{b}\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$   
 1638 channels with the ATLAS detector. *Phys. Rev.*, D92:092004, 2015. doi: 10.1103/PhysRevD.92.  
 1639 092004.
- 1640 [57] A. M. Sirunyan et al. Search for Higgs boson pair production in events with two bottom quarks and  
 1641 two tau leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2017.
- 1642 [58] V. Khachatryan et al. Search for two Higgs bosons in final states containing two photons and  
 1643 two bottom quarks in proton-proton collisions at 8 TeV. *Phys. Rev.*, D94(5):052012, 2016. doi:  
 1644 10.1103/PhysRevD.94.052012.
- 1645 [59] G. Aad et al. Search For Higgs Boson Pair Production in the  $\gamma\gamma b\bar{b}$  Final State using  $pp$  Collision  
 1646 Data at  $\sqrt{s} = 8$  TeV from the ATLAS Detector. *Phys. Rev. Lett.*, 114(8):081802, 2015. doi: 10.1103/  
 1647 PhysRevLett.114.081802.
- 1648 [60] A. M. Sirunyan et al. Search for resonant and nonresonant Higgs boson pair production in the  
 1649  $b\bar{b}lnlnu$  final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2017.
- 1650 [61] J. K. Behr, D. Bortoletto, J. A. Frost, N. P. Hartland, C. Issever, and J. Rojo. Boosting Higgs pair  
 1651 production in the  $b\bar{b}b\bar{b}$  final state with multivariate techniques. *Eur. Phys. J.*, C76(7):386, 2016. doi:  
 1652 10.1140/epjc/s10052-016-4215-5.
- 1653 [62] D. Wardrope, E. Jansen, N. Konstantinidis, B. Cooper, R. Falla, and N. Norjoharuddeen. Non-  
 1654 resonant Higgs-pair production in the  $b\bar{b} b\bar{b}$  final state at the LHC. *Eur. Phys. J.*, C75(5):219, 2015.  
 1655 doi: 10.1140/epjc/s10052-015-3439-0.
- 1656 [63] D. E. Ferreira de Lima, A. Papaefstathiou, and M. Spannowsky. Standard model Higgs boson pair  
 1657 production in the  $(b\bar{b})(b\bar{b})$  final state. *JHEP*, 08:030, 2014. doi: 10.1007/JHEP08(2014)030.
- 1658 [64] S. Banerjee, C. Englert, M. L. Mangano, M. Selvaggi, and M. Spannowsky.  $hh +$  jet production at  
 1659 100 TeV. *Eur. Phys. J.*, C78(4):322, 2018. doi: 10.1140/epjc/s10052-018-5788-y.
- 1660 [65] N. P. Hartland. Double Higgs production from  $HH \rightarrow (b\bar{b})(b\bar{b})$  at a 100 TeV hadron collider. *PoS*,  
 1661 DIS2016:115, 2016.
- 1662 [66] S. V. Chekanov, M. Beydler, A. V. Kotwal, L. Gray, S. Sen, N. V. Tran, S. S. Yu, and J. Zuzelski. Initial  
 1663 performance studies of a general-purpose detector for multi-TeV physics at a 100 TeV pp collider.  
 1664 *JINST*, 12(06):P06009, 2017. doi: 10.1088/1748-0221/12/06/P06009.

- 1665 [67] S. Chekanov. Boosted jets in high-granularity calorimeter at a 100 TeV pp collider. BOOST 2017,  
1666 2017.
- 1667 [68] S. Chekanov. Physics requirements for a Hadron Calorimeter for a 100 TeV proton-proton collider.  
1668 FCC week 2015, 2015.
- 1669 [69] S. Chekanov. Simulation of a high-granular hadronic calorimeter for multi-TeV physics. FCC week  
1670 2016, 2016.
- 1671 [70] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer.
- 1672 [71] T. Sjostrand, S. Mrenna, and P. Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys.*  
1673 *Commun.*, 178:852–867, 2008. doi: 10.1016/j.cpc.2008.01.036.
- 1674 [72] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lematre, A. Mertens, and M. Selvaggi.  
1675 DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:  
1676 057, 2014. doi: 10.1007/JHEP02(2014)057.
- 1677 [73] M. Cacciari, G. P. Salam, and G. Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012. doi:  
1678 10.1140/epjc/s10052-012-1896-2.
- 1679 [74] D. Buskulic et al. Performance of the ALEPH detector at LEP. *Nucl. Instrum. Meth.*, A360:481–506,  
1680 1995. doi: 10.1016/0168-9002(95)00138-7.
- 1681 [75] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical  
1682 Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009. URL <http://cds.cern.ch/record/1194487>.
- 1684 [76] FCC study group. Fccsw. <https://github.com/HEP-FCC/FCCSW>, 2018.
- 1685 [77] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing  
1686 applications. *Comput. Phys. Commun.*, 140:45–55, 2001. doi: 10.1016/S0010-4655(01)00254-5.
- 1687 [78] F. Gaede, B. Hegner, and P. Mato. Podio: An event-data-model toolkit for high energy physics  
1688 experiments. *Journal of Physics: Conference Series*, 898(7):072039, 2017. URL <http://stacks.iop.org/1742-6596/898/i=7/a=072039>.
- 1690 [79] FCC-hh study group. Eventproducer. <https://github.com/FCC-hh-framework/EventProducer>,  
1691 2018.
- 1692 [80] N. D. Christensen and C. Duhr. FeynRules - Feynman rules made easy. *Comput. Phys. Commun.*,  
1693 180:1614–1641, 2009. doi: 10.1016/j.cpc.2009.02.018.
- 1694 [81] P. M. Ferreira, R. Santos, M. Sher, and J. P. Silva. 2HDM confronting LHC data. In *Proceedings,*  
1695 *48th Rencontres de Moriond on QCD and High Energy Interactions: La Thuile, Italy, March 9-*  
1696 *16, 2013*, pages 67–70, 2013. URL <https://inspirehep.net/record/1234236/files/arXiv:1305.4587.pdf>.

<sub>1698</sub> [82] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016. doi: 10.1088/  
<sub>1699</sub> 1674-1137/40/10/100001.



1700 **Appendix A**

1701 **Jet radius discussion**

1702 It is important to discuss if the radius of the jet that is used to reconstruct the boosted Higgs boson  
1703 candidates ( $R = 0.8$ ) is the appropriate one. The question arises because in ATLAS boosted objects  
1704 with  $p_T \sim 200$  GeV are reconstructed using  $R = 1.0$  jets. In this work, we work with a CM energy of 100  
1705 TeV and require that the two leading jets have  $p_T > 200$  GeV and use jets with  $R = 0.8$ . It is necessary  
1706 to understand if these jets are large enough to fully reconstruct the Higgs candidates.

1707 As a first approximation we compute the angle between the b quarks produced by the decay of  
1708 a Higgs boson. We assume that the b quarks are massless and that the Higgs moves only in the  
1709 transverse plane (perpendicular to the beam pipe) such that it has no longitudinal momentum and the  
1710 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  
1711 smaller than the jet's diameter (1.6) and therefore the two b quarks can both be contained inside the jet  
1712 and the Higgs boson fully reconstructed.

1713 Another test we can make is to compute the  
1714  $\Delta R$  between the b quarks coming from the lead-  
1715 ing Higgs candidate with  $p_T$  larger than a given  
1716 value using truth level information. In figure A.1  
1717 we show the distribution of  $\Delta R(b, \bar{b})$  for  $p_T$  of  
1718 the leading Higgs candidate larger than 200 GeV  
1719 (solid blue) and 300 GeV (solid black). To obtain  
1720 the dashed histograms we apply the  $p_T$  to both  
1721 Higgs candidates. The integral of the histograms  
1722 between 0 and 1.6 gives an estimate of the frac-  
1723 tion of signal we keep if we apply these  $p_T$  cuts.  
1724 For  $p_T(h_1) > 200(300)$  GeV we get that 93(98) %  
1725 of the signal has b quarks with  $\Delta R < 1.6$  and  
1726 therefore can be fully reconstructed using a jet  
1727 with  $R = 0.8$ .

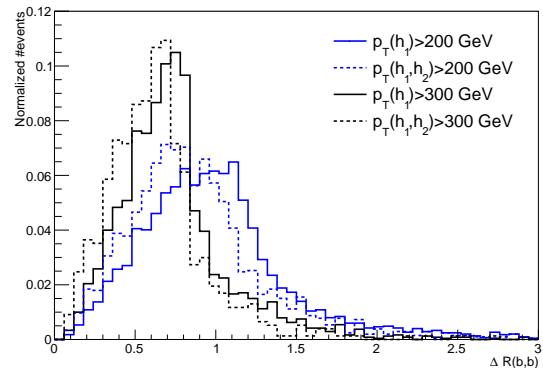


Figure A.1:  $\Delta R(b, \bar{b})$  distributions for  $p_T(h_1) > 200/300$  GeV (solid blue/black) and for  $p_T(h_1, h_2) > 200/300$  GeV (dashed blue/black).

1728 **Appendix B**

1729 **Additional background processes**

1730 In this section we discuss the importance of additional background processes and estimate how they  
1731 influence the analysis. In particular, we investigate backgrounds involving Higgs bosons.

<sub>1732</sub> **Appendix C**

<sub>1733</sub> **Softdrop mass**

<sub>1734</sub> It is noticeable in the softdrop mass spectrum of both Higgs candidates (for signal and backgrounds) the  
<sub>1735</sub> existence an atypical peak close to zero. In this section we explain the origin of this feature.

<sub>1736</sub> We believe that the peak close to zero corresponds to Higgs candidates ( $R = 0.8$  jets) that do not  
<sub>1737</sub> contain both b quarks from the Higgs decay. The plot that support this conclusion is shown in figure  
<sub>1738</sub> C.1. It shows the correlation between the maximum  $\Delta R$  between the Higgs candidate jet and one of  
<sub>1739</sub> the b quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to  $\Delta R = 0.8$ . For  
<sub>1740</sub> small masses ( $< 80$  GeV) the maximum  $\Delta R$  is usually larger than 0.8 which means that at least one  
<sub>1741</sub> of the b quarks is not contained in the jet's cone. When applying the soft drop procedure to these jets,  
<sub>1742</sub> soft radiation is removed and we are left with a single b quark. The mass of b quarks is  $\sim 5$  GeV and  
<sub>1743</sub> therefore we get a peak at this mass.

<sub>1744</sub> In practice, this does not affect our analysis because we place a mass window cut around the Higgs  
<sub>1745</sub> boson mass and therefore the low mass peak is removed. Nonetheless, the study presented here is  
<sub>1746</sub> extremely important because it helps rule out possible malfunctions of the soft drop algorithm and gives  
<sub>1747</sub> us confidence that we understand exactly what is happening.

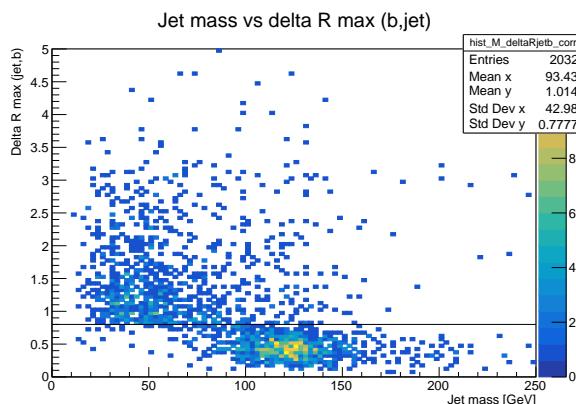


Figure C.1: Correlation between the maximum  $\Delta R$  between the Higgs candidate jet and one of the b quarks (y axis) and the jet's mass (x axis). The horizontal line corresponds to  $\Delta R = 0.8$ .

1748 **Appendix D**

1749 **Samples generation: parameters**

1750 In this appendix we provide we list the values of parameters used in MadGraph (section D.1) and Pythia  
1751 (section D.2) to generate the samples used in this work.

1752 **D.1 MadGraph**

1753 The MadGraph5 level cuts are summarized in table D.1. We show only the most relevant cuts for this  
1754 analysis: the minimum  $p_T$  of light and b quarks,  $p_{T,j}^{\min}$  and  $p_{T,b}^{\min}$ , the maximum pseudorapidity range for  
1755 light and b quarks,  $\eta_j^{\max}$  and  $\eta_b^{\max}$  and the  $\Delta R$  separation between two light quarks,  $\Delta R(jj)$ , two b  
1756 quarks,  $\Delta R(bb)$ , and between a light and b quarks,  $\Delta R(jb)$ . The  $xqcut$  parameter is a measure of the  
1757 required parton separation at Madgraph level. Whenever MadGraph produces two partons,  $i$  and  $j$ ,  
1758 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  
1759 of this expression is smaller than the specified value of  $xqcut$  then we do not generate the event. The  
1760  $bwcutoff$  parameter defines what is considered to be on-shell s-channel resonances. The  $H_T$  variable  
1761 is the scalar sum of the  $p_T$  of all truth level partons, including b quarks.

1762 **D.2 Pythia**

1763 The settings for jet matching can be found in table D.2 under the corresponding samples' columns. We  
1764 perform the jet matching procedure (merge=on) using the MLM matching scheme and the appropriate  
1765 algorithm for a parton level process generated in MadGraph (scheme=1). We do not read the matching  
1766 parameters from the MadGraph file (setMad=off) because this option is not available for these files. The  
1767 size of the cone drawn around the jet's center, the maximum pseudorapidity and the maximum number  
1768 of jets to be matched are given by coneRadius, etaJetMax and nJetMax, respectively. The cone radius  
1769 is set to one. The maximum allowed pseudorapidity of jets is ten which is a much loser cut than the  
1770 acceptance of any current detector. The maximum number of jets is set to four for the jj+0/1/2 j and to  
1771 two for the t̄t+0/1/2 j. The qCut parameter defines the  $k_T$  scale for merging shower products into jets.

<b>MadGraph5</b>	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.

<b>Pythia</b>	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	—	<b>Jet matching:</b> merge=on scheme=1 setMad=off coneRadius=1.0 etaJetMax=10 nJetMax=4 qCut=30	<b>Jet matching:</b> 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.