



1

2 **Technical Design Report:**
3 **A High-Granularity Timing Detector for the**
4 **ATLAS Phase-II Upgrade**

5

Draft version: 1.5

Created: 1st January 2020 – 14:35

6 **Prepared by: The ATLAS Collaboration**

7 © 2020 CERN for the benefit of the ATLAS Collaboration.

8 Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

Abstract

The large increase of pile-up interactions is one of the main experimental challenges for the HL-LHC physics program. A powerful new way to mitigate the effects of pile-up is to use high-precision timing information to distinguish between collisions occurring close in space but well-separated in time. A High-Granularity Timing Detector, based on low gain avalanche detector technology, is therefore proposed for the ATLAS Phase-II upgrade. Covering the pseudorapidity region between 2.4 and 4.0, this device will significantly improve the performance in the forward region. Taking into account the typical number of hits per track in the detector, the target average time resolution per track for a minimum-ionising particle is 30 ps at the start of lifetime, increasing to 50 ps at the end of HL-LHC operation. The high-precision timing information greatly improves the track-to-vertex association, leading to a performance similar to that in the central region for both jet and lepton reconstruction, as well as the tagging of heavy-flavour jets. These improvements in object reconstruction performance translate into important sensitivity gains and enhance the reach of the HL-LHC physics program. In addition, the HGTD offers unique capabilities for the online and offline luminosity determination, an important requirement for precision physics measurements.

Contents

29	1 Introduction	1
30	2 Detector Requirements and Layout	3
31	2.1 Beam conditions at the HL-LHC	3
32	2.2 Detector overview and requirements	4
33	2.3 Detector layout and optimisation	6
34	2.4 Radiation hardness	14
35	3 Performance and Physics Benchmarks	17
36	3.1 Simulation	17
37	3.1.1 Detector geometry	17
38	3.1.2 Sensor simulation	19
39	3.1.3 Simulation using silicon hits	25
40	3.2 Detector performance	27
41	3.2.1 Detector response characterisation	27
42	3.2.2 Track-level performance	33
43	3.2.3 Determination of the time of the primary vertex	39
44	3.2.4 Suppression of pile-up jets	42
45	3.2.5 Missing transverse momentum	44
46	3.2.6 Tagging of heavy flavour jets	45
47	3.2.7 Lepton isolation	46
48	3.2.8 Non-collision background	48
49	3.3 Physics	50
50	3.3.1 Strategy for the Application of Object Performance to Physics Analyses	50
51	3.3.2 Impact of the luminosity uncertainty	53
52	3.3.3 Measurement of $\sin^2 \theta_{\text{eff}}$	54
53	3.3.4 Monopole searches	55
54	4 Technical Overview	59
55	4.1 Introduction	59
56	4.2 Detector overview and key requirements	59
57	4.2.1 Expected Radiation levels	60
58	4.2.2 Key requirements	61
59	4.2.3 Read-out bandwidth and trigger	62

60	4.3	Hybrid HGTD module	63
61	4.3.1	Sensors	64
62	4.3.2	Front End ASIC	66
63	4.3.3	Module assembly	68
64	4.4	Module loading on support structure	70
65	4.5	Off detector electronics, calibration and luminosity	71
66	4.5.1	Peripheral Electronics Boards	71
67	4.5.2	t_0 time calibration	73
68	4.5.3	Luminosity	74
69	4.6	Power distribution and detector control system	74
70	4.6.1	HV system	74
71	4.6.2	LV system	74
72	4.6.3	Monitoring and Controls	75
73	4.7	Mechanics, Services and Infrastructure	75
74	4.8	Assembly, Installation and Commissioning	76
75	4.9	Next steps towards construction	77
76	5	Sensors	79
77	5.1	Sensor parameters and requirements	79
78	5.2	Low Gain Avalanche Detectors	79
79	5.2.1	Overview	79
80	5.2.2	LGAD productions	82
81	5.3	Radiation damage and irradiations	83
82	5.4	Sensor tests: methodology and experimental techniques	85
83	5.5	LGAD performance before and after irradiation	87
84	5.5.1	Electrical characterisation: I-V and C-V	87
85	5.5.2	Operating bias Voltage and self-triggering	91
86	5.5.3	Collected charge and gain	93
87	5.5.4	Efficiency	94
88	5.5.5	Time resolution	95
89	5.5.6	Uniformity, inter-pad gap and edge region	96
90	5.5.7	Long term and stability tests	99
91	5.5.8	Leakage current and power after irradiation	102
92	5.6	Operational aspects and bias voltage evolution in HGTD	103
93	5.7	Summary of status quo sensor design	105
94	5.8	Roadmap for future sensor productions and activities	106
95	6	Front-end Electronics	109
96	6.1	General requirements	110
97	6.1.1	Data transmission bandwidth requirements	112
98	6.2	ASIC architecture	112
99	6.2.1	Channel architecture	113

100	6.2.2	Readout architecture	115
101	6.3	Single-channel readout electronics	117
102	6.3.1	Preamplifier	117
103	6.3.2	Discriminator and time walk correction	119
104	6.3.3	TDC	120
105	6.3.4	Command pulser for calibration	122
106	6.3.5	Hit processor	122
107	6.3.6	Luminosity processing unit	124
108	6.4	ASIC End of Column logic and digital blocks	125
109	6.4.1	Matrix readout process	125
110	6.4.2	Slow control	127
111	6.4.3	Clock generator and phase shifter	128
112	6.4.4	Clock distribution	128
113	6.5	Radiation tolerance	129
114	6.6	ASIC Power distribution and grounding	130
115	6.7	ASIC prototype measurements	131
116	6.7.1	Test bench performance	132
117	6.7.2	Test beam performance	134
118	6.7.3	Irradiation tests	135
119	6.8	Monitoring	135
120	6.8.1	Temperature monitoring	135
121	6.8.2	Supply voltages monitoring	136
122	6.8.3	Complete monitoring system	137
123	6.9	Roadmap towards production	138
124	7	Module Assembly and Loading	141
125	7.1	Introduction	141
126	7.2	The bare module	141
127	7.2.1	Bare module assembly	142
128	7.2.2	First bare module prototypes: process and results	143
129	7.2.3	Hybridization specifications	146
130	7.2.4	Quality assurance / quality control	148
131	7.2.5	Production hybridization strategy	148
132	7.3	Module design and assembly	149
133	7.3.1	General description	149
134	7.3.2	Voltage distribution and signal readout: flex cables	150
135	7.3.3	Gluing and wire-bonding	154
136	7.3.4	Assembly specifications	158
137	7.3.5	Quality assurance / quality control	159
138	7.3.6	Production strategy for module assembly	159
139	7.4	Module loading	160
140	7.4.1	General description	160

141	7.4.2	Support units and detector units	160
142	7.4.3	Gluing studies	162
143	7.4.4	Procedure for assembly and qualification	163
144	7.4.5	Detector unit assembly strategy	164
145	7.5	Thermal calculation	165
146	8	Power Distribution, Control, and Safety Systems	171
147	8.1	High voltage	171
148	8.2	Low voltage	172
149	8.3	Grounding/shielding	173
150	9	Peripheral Electronics	175
151	9.1	Requirements	177
152	9.1.1	Data transfer	177
153	9.1.2	Physical limitations	177
154	9.1.3	Radiation tolerance	178
155	9.2	Data transfer	178
156	9.2.1	LpGBT	178
157	9.2.2	Optical links	181
158	9.3	DC-DC converters	181
159	9.4	Control and monitoring	182
160	9.5	Connectors	183
161	9.6	Peripheral boards	183
162	9.6.1	Layout considerations	183
163	9.6.2	Layout	184
164	9.7	Power dissipation	185
165	10	DAQ and Luminosity Measurement	187
166	10.1	DAQ interface	187
167	10.1.1	Off-detector electronics	187
168	10.1.2	Calibration and timing	189
169	10.2	Timing correction	190
170	10.2.1	Sources of clock jitter	191
171	10.2.2	Timing correction procedure	192
172	10.3	Luminosity	195
173	10.3.1	HGTD as a luminometer	196
174	10.3.2	Linearity of the luminosity determination	197
175	10.3.3	Noise and afterglow subtraction	197
176	10.3.4	Statistical precision of the luminosity determination	199
177	10.3.5	Systematic uncertainties affecting the luminosity determination	200
178	10.3.6	Occupancy readout at 40 MHz	202
179	10.3.7	Luminosity back-end electronics	205

180	10.3.8 Per-event luminosity information stored in the ATLAS raw data	206
181	10.3.9 Operation in non-Stable Beams conditions	207
182	10.3.10 Minimum-bias trigger at Level-0	207
183	10.4 Detector Control System	208
184	10.4.1 High Voltage	208
185	10.4.2 Low Voltage	209
186	10.4.3 Temperatures	209
187	10.4.4 Pressure and humidity	209
188	10.4.5 Configuring	210
189	10.4.6 DCS software	210
190	10.4.7 Interlock system	211
191	10.4.8 External systems	211
192	11 Detector Mechanics	213
193	11.1 Engineering design overview	213
194	11.2 CO ₂ cooling system	215
195	11.2.1 Requirements	215
196	11.2.2 Cooling design	217
197	11.2.3 Cooling plant demonstrator	220
198	11.3 Moderator	221
199	11.4 Hermetic vessel	222
200	11.4.1 Requirements	222
201	11.4.2 Front cover and heaters	223
202	11.4.3 Back cover and interface with LAr cryostat	224
203	11.4.4 Inner ring design	225
204	11.4.5 Outer ring design	226
205	11.5 Local supports and cooling disks	228
206	11.5.1 Geometry and design	228
207	11.5.2 Thermal performance	229
208	11.6 Detector overall layout	229
209	12 Detector Infrastructure	233
210	12.1 Services	233
211	12.1.1 Specifications	233
212	12.1.2 Services layout	235
213	12.1.3 Patch panels in PP-EC area	237
214	12.1.4 Services routing on the calorimeter front wall	238
215	12.1.5 Services connection to outer ring and inside the vessel	240
216	12.1.6 Services installation	241

217	13 Detector Assembly, Installation, and Commissioning	243
218	13.1 Assembly and commissioning on surface	243
219	13.1.1 Half disks instrumentation	243
220	13.1.2 Detector assembly on hermetic vessel	244
221	13.1.3 Quality Assurance after assembly	244
222	13.2 Installation in the cavern and commissioning	245
223	13.2.1 Access scenarios	245
224	13.2.2 Maintenance, radiation environment, and radio protection	246
225	13.2.3 Transport to the cavern and lowering	251
226	13.2.4 Detector Installation and Commissioning	252
227	14 Demonstrator	253
228	14.1 Introduction	253
229	14.2 Heater demonstrator	253
230	14.2.1 Mechanical structure	256
231	14.3 Peripheral and back-end electronics, data acquisition	257
232	14.3.1 Peripheral electronics demonstrator	257
233	14.3.2 DAQ demonstrator	258
234	14.3.3 HGTD module	259
235	14.4 Full demonstrator	260
236	14.5 Demonstrator tests	260
237	14.6 Schedule and organisation	260
238	15 Project Organization, Costs, and Schedule	263
239	15.1 Organization and management	263
240	15.1.1 Upgrade organisation in ATLAS	263
241	15.1.2 HGTD organisation	264
242	15.1.3 Technical milestones	267
243	15.1.4 Deliverables and WBS	269
244	15.2 Schedule and production schedule milestones	269
245	15.3 Costs	274
246	15.4 Risk management	277
247	A Expected Energy Spectra	283
248	B Technical Drawings	285
249	Bibliography	292

250

1 Introduction

251 The high-luminosity (HL) Phase-II of the Large Hadron Collider (LHC) at CERN is scheduled
252 to start in 2026. This HL-LHC will deliver an integrated luminosity of up to 4000 fb^{-1} over
253 the subsequent decade. The instantaneous luminosity of the HL-LHC will reach up to
254 $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, a large increase from the $2.1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ obtained during LHC Run-2.
255 Two extended periods without physics operation are anticipated prior to Phase-II, Long
256 Shutdown 2 (LS2) in 2019–2020 and Long Shutdown 3 (LS3) from 2024 until mid 2026.
257 During LS3, extensive upgrades to the ATLAS Experiment will be installed to cope with the
258 higher luminosities and add new capabilities.

259 This report describes the technical design of a High Granularity Timing Detector (HGTD)
260 for the ATLAS Phase-II upgrade, an upgrade that will augment the Phase-II silicon-based
261 Inner Tracker, the ITk [1], in the forward region, with the capability to measure charged-
262 particle trajectories in time as well as space. The target average time resolution per track
263 for a minimum-ionising particle is $\approx 30 \text{ ps}$ at the beginning of the HL-LHC, increasing to
264 50 ps per track at the end of the HL-LHC. The HGTD will also provide a precise, real-time
265 luminosity measurement for ATLAS's Phase-II physics programme. The goals for the design
266 are detailed in Chap. 2, and Chap. 3 provides projections for how the detector will improve
267 ATLAS object reconstruction and physics.

268 The technical design of the HGTD is summarized in Chap. 4. The HGTD will consist of
269 many silicon-based low-gain avalanche detectors (LGADs), placed in front of the end-cap
270 and forward calorimeters at $2.4 < |\eta| < 4.0$ and arranged such that a track traverses two
271 or three sensors. Chap. 5 describes the LGAD sensors and their expected performance,
272 based on measurements of prototype devices that include irradiation at the levels expected
273 at the HL-LHC. Chap. 6 describes the front-end electronics, a low-noise, radiation-hard
274 custom ASIC called the ALTIROC, and the performance of the analog front end. Chap. 7
275 discusses the hybridization of the LGAD and ALTIROC into modules of a single LGAD
276 sensor bump-bonded to two ALTIROC chips, their assembly into disks and staves, and their
277 connection via a flex cable to peripheral electronics boards at the outer radii of the disk
278 geometry. Chap. 8 describes the powering and control of the detector. Chap. 9 describes
279 the function and layout of the peripheral electronics boards, and Chap. 10 summarizes the
280 connection of the detector to the ATLAS data acquisition system, the real-time intercalibration
281 of the arrival time within the readout path, and the 40 MHz readout of highly-granular hit
282 multiplicity data for real-time luminosity measurement. Chap. 11 provides the engineering
283 design for cooling the LGADs and front-end electronics, Chap. 12 the mechanical design

284 of the overall detector, the necessary services and routing, and Sec. 13.1 the assembly and
285 commissioning of the detector. Chap. 14 describes a set of intermediate prototypes that will
286 integrate elements of the full design during the remaining R&D period, in order to validate
287 key aspects of the design. Finally, Chap. 15 documents the organisation of the project to
288 deliver and commission the detector for the start of the HL-LHC in 2026.

2 Detector Requirements and Layout

2.1 Beam conditions at the HL-LHC

Pile-up is one of the main challenges at the HL-LHC. The exact beam-spot characteristics of the HL-LHC have not yet been determined. In the nominal operation scheme [2], an average of 200 simultaneous pp interactions ($\langle\mu\rangle = 200$) will occur within the same bunch crossing interval. A major challenge for the ITk is to efficiently reconstruct charged particles created in the primary interactions and assign them to the correct production vertices in this high pile-up environment. The luminous region will have an estimated Gaussian spread of 30 to 60 mm along the beam axis (z direction¹.) The width in time could range from 175 to 260 ps. The case considered in this report is the *nominal* scenario, with Gaussian spreads of 45 mm in along the beam axis and 175 ps in time.

The spatial pile-up line density, i.e. the number of collisions per length unit along the beam axis during one bunch crossing, is a key quantity for evaluating the performance of ATLAS with and without the HGTD. For an average of 200 collisions per bunch crossing, denoted $\langle\mu\rangle = 200$, an average pile-up density of 1.8 collisions/mm is expected. This average masks the effect of the local variations illustrated in Fig. 2.1. The local pile-up vertex density is calculated by computing the average number of interactions per unit length in a window of ± 3 mm around the signal vertex for $\langle\mu\rangle = 200$. This is large enough to avoid quantisation effects and small enough to probe the tails of the distribution. The curve for $\langle\mu\rangle = 30$ is obtained by scaling, which effectively increases the window size.

Fig. 2.1 shows the pile-up densities for $\langle\mu\rangle = 30$ and $\langle\mu\rangle = 200$ for the same beam spot size. The most probable local pile-up density for this scenario is around 1.44 collisions/mm.

Timing information can supplement the tracker measurement of the longitudinal impact parameter, z_0 , improving how often tracks are assigned to the correct vertices and mitigating the impact of a high vertex density. To illustrate this, an example is presented in Fig. 2.2, which shows a single $t\bar{t}$ Hard Scatter (HS) vertex along pile-up vertices in the $z-t$ plane at $\langle\mu\rangle = 200$. The simulated and reconstructed vertices are overlaid. The vertex and track

¹ The ATLAS experiment uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the centre of the LHC ring, and the y -axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudo-rapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$.

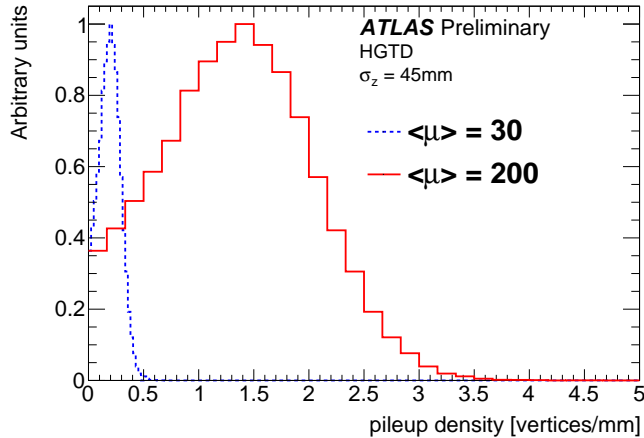


Figure 2.1: Local pile-up vertex densities for different values of $\langle\mu\rangle$.

316 reconstruction used in this event are further described in Sec. 3.2.2. The impact parameter
 317 resolution in the forward region is limited by multiple scattering. A minimal p_T cut of 0.9
 318 GeV is applied for tracks at all η to reject soft forward tracks. Each track is required to have
 319 at least 3 pixel clusters. Tracks are required to have $\sigma(d_0) < 0.3$ mm and $\sigma(z_0) < 0.5$ mm in
 320 order to ensure good precision. This cut effectively removes low- p_T forward tracks because
 321 of their limited resolution.

322 The tracker sees the event as a one-dimensional projection on the z axis, where a large number
 323 of tracks from vertices occurring at different times but close in space lead to ambiguities
 324 in the track-to-vertex association. This happens when the distance between vertices is of
 325 the same order or smaller than the resolution of the longitudinal impact parameter of the
 326 track, which happens more often for tracks in the forward region. The timing information
 327 reduces the density of vertices which are considered for a given track. The figure shows the
 328 reconstructed vertices are also spread in time, due to the association of tracks with timing
 329 information of the HGTD. The time of HS vertex may be reconstructed using tracks with hits
 330 in the HGTD, as discussed Sec. 3.2.3, allows for more effective separation of the HS vertex
 331 from pile-up vertices surrounding it in the z direction.

332 2.2 Detector overview and requirements

333 The HGTD has been designed for operation with $\langle\mu\rangle = 200$ and a total integrated luminosity
 334 of 4000 fb^{-1} . Taking into account the space constraints of the existing ATLAS Experiment,
 335 including the more advanced planning for the tracker upgrade when R&D on the HGTD
 336 began, the HGTD will be located in the gap region between the barrel and the end-cap
 337 calorimeters, at a distance in z of approximately ± 3.5 m from the nominal interaction point.

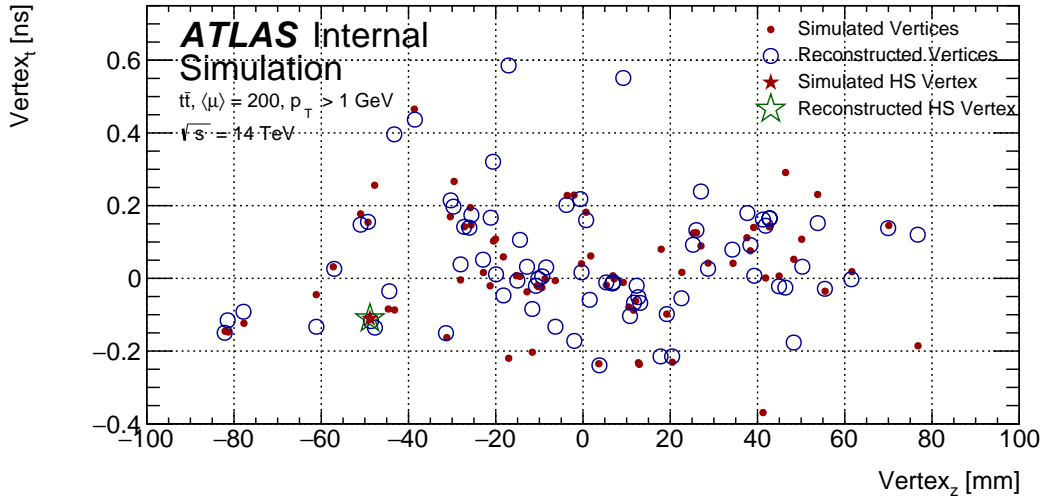


Figure 2.2: Visualisation of the primary vertices in an event in the z - t plane, showing the simulated Hard Scatter (HS) with pile-up interactions superimposed for $\langle \mu \rangle = 200$. The reconstructed vertices (blue circles) are overlaid along with the reconstructed HS vertex (green star).

338 This region lies outside the ITk volume and in front of the end-cap and forward calorimeters
 339 in the volume currently occupied by the Minimum-Bias Trigger Scintillators (MBTS), which
 340 will be removed. The position of the two vessels for the HGTD within the ATLAS detector is
 341 shown in Fig. 2.3.

342 The envelope of the detector vessel has a radial extent of 110 to 1000 mm. The envelope in z ,
 343 including the moderator, supports, and front and rear vessel covers, is 125 mm. A 50 mm-
 344 thick moderator is placed behind the HGTD to reduce the back-scattered neutrons created by
 345 the end-cap/forward calorimeters, protecting both the ITk and the HGTD. A silicon-based
 346 timing detector technology is preferred due to the space limitations. The sensors must be
 347 thin and configurable in arrays. In close collaboration with RD50 [3] and manufacturers, an
 348 extensive R&D program is progressing quickly towards sensors that provide the required
 349 timing resolution in harsh radiation environments. Low Gain Avalanche Detector (LGAD) [4]
 350 pads of $1.3 \text{ mm} \times 1.3 \text{ mm}$ with an active thickness of $50 \mu\text{m}$ fulfil these requirements. This
 351 pad size ensures occupancies below 10% at the highest expected levels of pile-up, small
 352 dead areas between pads, and low sensor capacitance, which is important for the time
 353 resolution.

354 A custom ASIC (ALTIROC), which will be bump-bonded to the sensors, is being developed
 355 to meet the requirements on time resolution and radiation hardness. The ASIC will also
 356 provide functionality to count the number of hits registered in the sensor and transmit
 357 this at 40 MHz to allow unbiased, bunch-by-bunch measurements of the luminosity and
 358 the implementation of a minimum-bias trigger. After optimising the layout for timing
 359 performance and cost, the detector design described in this document will give an average of

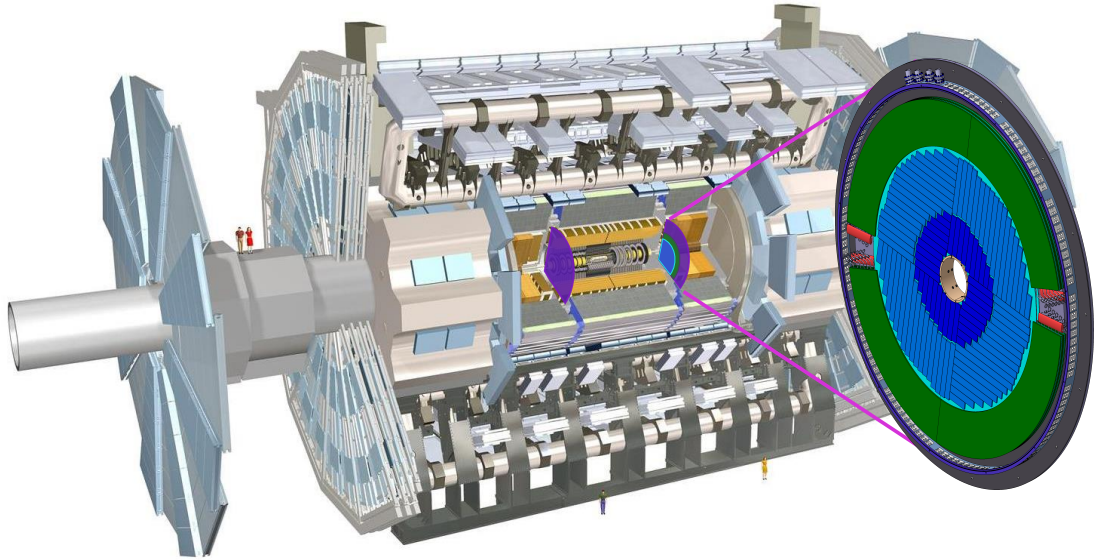


Figure 2.3: *TO BE UPDATED* Position of the HGTD within the ATLAS Detector. The HGTD acceptance is defined as the surface covered by the HGTD between a radius of 120 mm and 640 mm at a position of $z = \pm 3.5$ m along the beamline, on both sides of the detector.

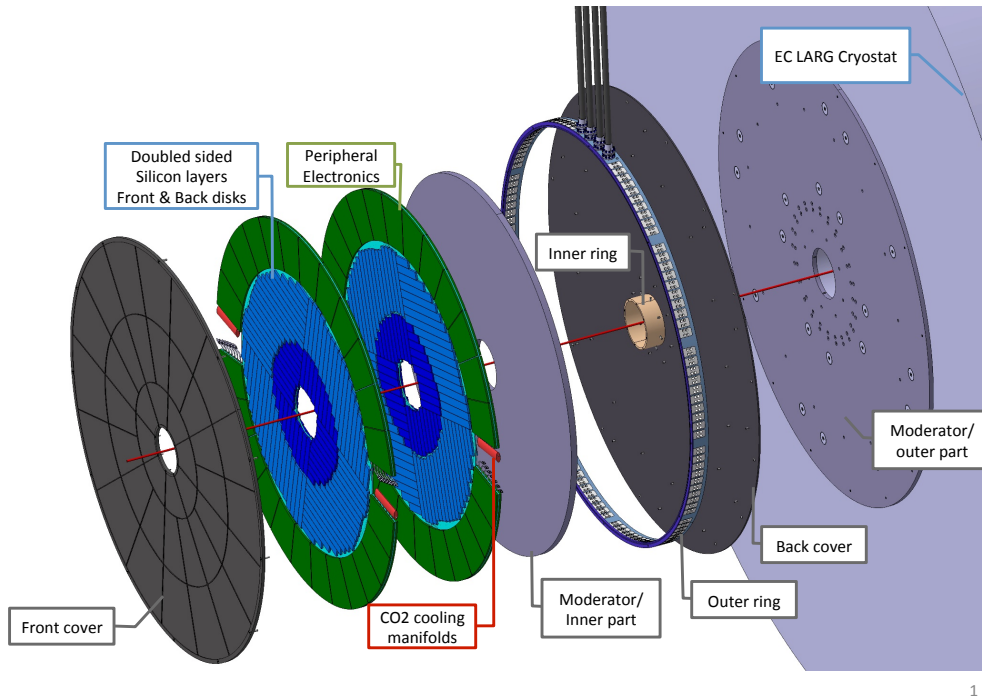
360 2.1, 2.5 and 2.7 hits per track for the regions $230 \text{ mm} > r > 120 \text{ mm}$, $470 \text{ mm} > r > 230 \text{ mm}$
 361 and $640 \text{ mm} > r > 470 \text{ mm}$. A description of the detector layout optimisation is presented
 362 in Sec. 2.3. It covers the pseudo-rapidity range $2.4 < |\eta| < 4.0$.

363 Each HGTD end-cap is the integration of one hermetic vessel, two instrumented double-
 364 sided layers (mounted in two cooling/support disks), and two moderator pieces placed
 365 inside and outside the hermetic vessel. Each cooling/support disk is physically separated
 366 in two half circular disks. Furthermore, the layers are rotated in opposite directions with
 367 respect to one another by 18° to 20° in order to maximize the hit efficiency.

368 A global view of the various components of the detector and its main parameters are
 369 shown in Fig. 2.4 and Tab. 2.1. The time resolution parameters have been optimised using
 370 information from the sensor Chap. 5 and front-end electronics Chap. 6 performance from
 371 lab and testbeam measurements.

372 2.3 Detector layout and optimisation

373 The goal of the detector design is to provide the best possible time resolution in order to
 374 effectively suppress the effects of pile-up in the forward region. The ability to associate tracks
 375 to primary vertices depends on the longitudinal impact parameter resolution of the ITk. The
 376 current ITk layout is shown in Fig. 2.5. Fig. 2.6 shows the resolution, σ_{z_0} , of the longitudinal



1

Figure 2.4: *TO BE UPDATED* Global view of the HGTD to be installed on each of two end-cap calorimeters. The various components are shown: hermetic vessel (front and rear covers, inner and outer rings), two instrumented double-sided layers (mounted in two cooling disks), two moderator pieces placed inside and outside the hermetic vessel.

377 track impact parameter, z_0 , from the ITk as a function of η , for $p_T = 1$ GeV and $p_T = 10$ GeV
 378 muons. The resolution is shown for both digital and analogue clustering. Digital clustering
 379 shows a similar performance to analogue clustering for $p_T = 1$ GeV muons and a 10–20 %
 380 degradation in σ_{z_0} for $p_T = 10$ GeV muons. Digital clustering takes the geometrical average
 381 as the centroid of a given pixel cluster, while analogue clustering improves the ability to
 382 determine the position by weighting the centroid by the charge deposited in each of the
 383 pixels contributing to the cluster. In this report, performance studies have been performed
 384 with an ITk layout and simulation [5] including a sensor pitch of $50 \times 50 \mu\text{m}$ and digital
 385 clustering, except where otherwise specified.

386 For good spatial separation of the HL-LHC collision vertices, σ_{z_0} should be significantly better
 387 than the inverse of the average pile-up density, $600 \mu\text{m}$. Fig. 2.6 shows that, in the central
 388 region, σ_{z_0} is well below this limit. In the forward region, however, the resolution exceeds
 389 the limit by a large factor, reaching 3 mm for particles with low transverse momentum (p_T) at
 390 $|\eta| \approx 4$, due to the combination of geometric projection and, as shown in Fig. 2.7, increased
 391 material. As a result, ITk by itself cannot associate such forward tracks to correct vertices in

Pseudo-rapidity coverage	$2.4 < \eta < 4.0$
Thickness in z	75 mm (+50 mm moderator)
Position of active layers in z	$z = \pm 3.5$ m
Weight per endcap	350 kg
Radial extension:	
Total	$110 \text{ mm} < r < 1000 \text{ mm}$
Active area	$120 \text{ mm} < r < 640 \text{ mm}$
Pad size	$1.3 \text{ mm} \times 1.3 \text{ mm}$
Active sensor thickness	50 μm
Number of channels	3.6 M
Active area	6.4 m ²
Module size	30 x 15 pads (4 cm x 2 cm)
Modules	8032
Collected charge per hit	> 4.0 fC
Average number of hits per track	
$2.4 < \eta < 2.7$ (640 mm $> r >$ 470 mm)	≈ 2.1
$2.7 < \eta < 3.5$ (470 mm $> r >$ 230 mm)	≈ 2.5
$3.5 < \eta < 4.0$ (230 mm $> r >$ 120 mm)	≈ 2.7
Average time resolution per hit (start and end of operational lifetime)	
$2.4 < \eta < 4.0$	≈ 35 ps (start) ≈ 65 ps (end)
Average time resolution per track (start and end of operational lifetime)	≈ 30 ps (start) ≈ 50 ps (end)

Table 2.1: Main parameters of the HGTD.

392 an unambiguous way.

393 In addition to the spatial distribution of collisions, pile-up collisions will also be distributed
394 in time, with a Gaussian width in the range 175 to 260 ps. The current baseline working
395 point is 175 ps, which is used in the simulation of the HGTD performance studies. HGTD is
396 designed to provide a time-of-arrival measurement for incident tracks with a resolution of
397 30 ps at the beginning of the HL-LHC, degrading to 50 ps at the end of the HL-LHC. After
398 determining the vertex position using the ITk, this complementary time measurement is
399 significantly more precise than the spread of the beamspot in time.

400 With the combination of ITk and HGTD measurements, ATLAS can view a portion of the
401 event in space and time, extending the pile-up rejection capabilities of the ITk to the extent
402 of its acceptance. This is one of the main motivations for the HGTD.

The main contributions to the time resolution of a detector element are:

$$\sigma_{\text{total}}^2 = \sigma_{\text{L}}^2 + \sigma_{\text{elec}}^2 + \sigma_{\text{clock}}^2 \quad (2.1)$$

403 where σ_{L}^2 are Landau fluctuations in the deposited charge as the charged particle traverses
404 the sensor, σ_{elec}^2 represents the contributions from the readout electronics, and σ_{clock}^2 is the
405 clock contribution. Beam tests and sensor simulations show that thinner silicon sensors
406 reduce the contribution from Landau fluctuations. With a 50 μm thick LGAD sensor, this
407 contribution amounts to approximately 25 ps. This is further discussed in Chap. 5. With fast

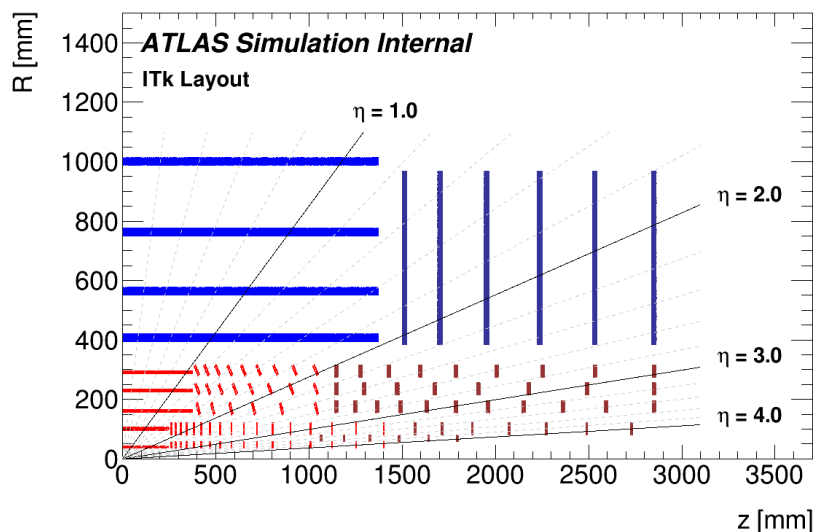


Figure 2.5: Schematic layout of the ITk for the HL-LHC phase of ATLAS. The active elements of the barrel and end-cap ITk Strip detector are shown in blue, for the ITk Pixel detector the sensors are shown in red for the barrel layers and in dark red for the end-cap rings. Here only one quadrant and only active detector elements are shown.

408 detector signals and a high signal-to-noise ratio, the contribution from the electronics can be
 409 kept to approximately 25 ps. This is achievable only if applying corrections for the time walk
 410 induced by different signal amplitudes, using small bins in the time-to-digital conversion
 411 and applying precise in-situ inter-calibration. The details of the design of the readout
 412 electronics to achieve this are described in Chap. 6. The clock contribution is required to be
 413 below 15 ps; its distribution is discussed in more detail in Chap. 10.

414 For simplicity, the same pad size, 1.3 mm × 1.3 mm is used for the entire HGTD. This pad
 415 size balances several characteristics. For smaller pad sizes, both electronic noise and physics
 416 occupancy are smaller, while the number of channels to be instrumented and the cumulative
 417 area of inter-pad dead zones are larger. The size was chosen to give a maximum occupancy at
 418 the lowest instrumented radius, 120 mm, of less than 10%. This also ensures a low double-hit
 419 probability.

420 The layout of modules in each endcap was defined by maximising the coverage and min-
 421 imising the effect of non-instrumented regions due to mechanical tolerances. In the second
 422 step the spacing between modules was defined.

423 The readout rows are sets of modules whose flex cables are guided together towards larger
 424 radii to the peripheral on-detector electronics. Their disposition for the first and last layer is
 425 shown as rectangles in Fig. 2.8. The active width of a module is 39 mm which limits how well
 426 the area near the circular opening at 120 mm can be covered. For $r > 150$ mm, the coverage

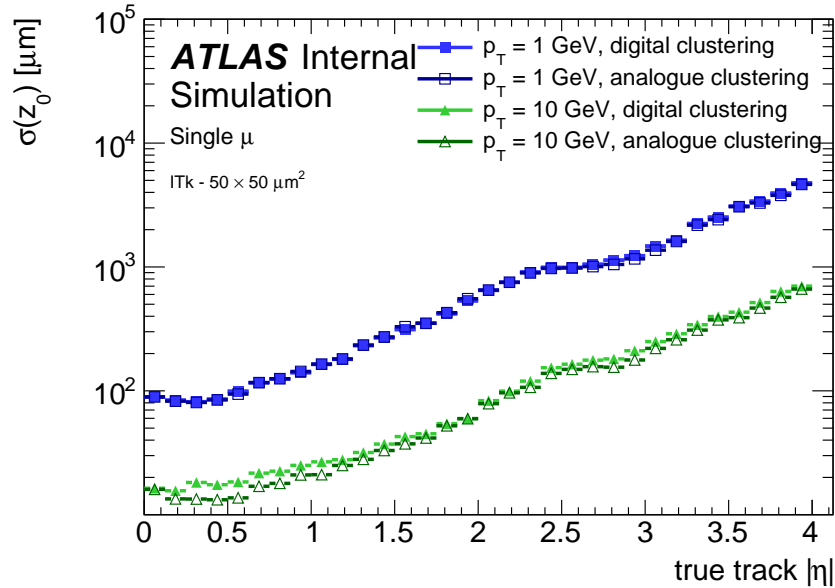


Figure 2.6: Resolution of the longitudinal track impact parameter, z_0 , as a function of η for muons of $p_T = 1$ GeV and $p_T = 10$ GeV. Additionally, the figure highlights the difference in resolution whether analogue or digital clustering is used. Studies for this report use digital clustering, which have roughly a 10 % degradation in z_0 resolution for $p_T = 10$ GeV muons.

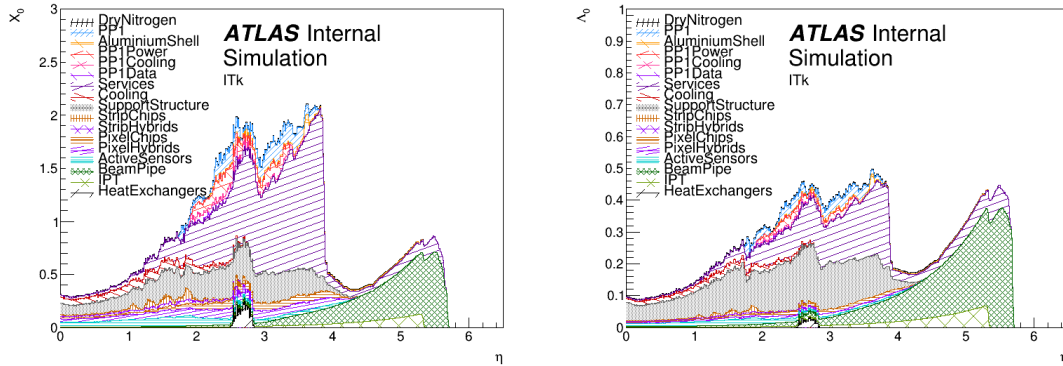


Figure 2.7: Radiation length X_0 (left) and nuclear interaction length λ_0 (right) as a function of pseudo-rapidity η , broken down by type of material for the ITk Layout [5] and beam pipe. *Preliminary, to be updated in new release*

427 is complete. The maximum length of the readout rows is limited by the manufacturing
 428 capabilities for the flexible circuits used for the data transmission. The non-instrumented
 429 zone is 0.5 mm between two readout rows for each row to account for mechanical tolerances.
 430 The effective width of a readout row is therefore 41 mm. These constraints lead to the helix
 431 structure shown in Fig. 2.8. A particle transiting the detector should encounter multiple

432 layers of LGAD sensors as it enters (encountering a “first layer”) and exits (after encountering
 433 a “last layer”). Fig. 2.8(a) shows the geometry of the first layer and Fig. 2.8(b) shows the
 434 geometry of the last layer. The first and last layer are arranged to mirror the geometry of
 435 one another. Therefore the non-instrumented zones of the two disks will not overlap, except
 436 in the case of four readout rows per quadrant. Additionally, each of the layers is rotated in
 437 opposite directions by 18° to 20° as shown in Fig. 2.8(c) for 20° . Any angle of rotation beyond
 438 10° results in similar performance in terms of the number of hits and dead regions. The
 439 baseline angle is chosen largely due to detector services considerations, which are further
 440 discussed in Chap. 12 and Chap. 13. Along with optimising the coverage, the rotation
 441 frees sufficient room at 640 mm to install the cooling equipment between the peripheral
 442 electronics.

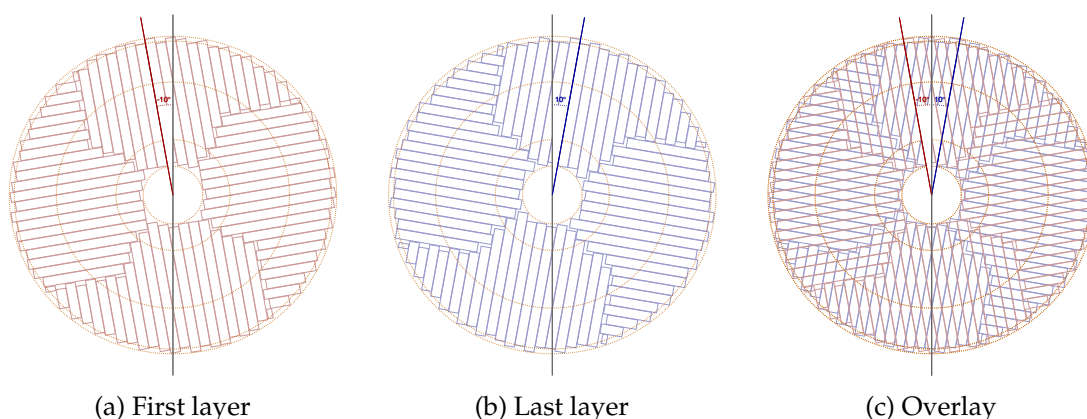


Figure 2.8: The orientation of the readout rows for the first and last layer encountered by a particle, separately and with the overlay of both. Each layer is rotated in alternating directions by 20° .

443 The geometry of the detector has been optimised to approximate a flat timing resolution
 444 as a function of η . Due to radiation damage, the timing resolution of the detector will
 445 be degraded as the integrated luminosity delivered by the LHC increases. This radiation
 446 depends strongly on r , with higher radiation closer to the beam axis. The radiation levels
 447 expected for the full lifetime of the HL-LHC, including safety factors, are discussed in Sec. 2.4.
 448 The geometry of the HGTD is designed such that at $r < 230$ mm on average about 2.7 hits
 449 are obtained for a charged particle, for $230 \text{ mm} < r < 470$ mm on average about 2.5 hits are
 450 obtained, whereas at $r > 470$ mm on average about 2.1 hits are expected to be associated to
 451 a track.

452 Each layer of the HGTD is double-sided, i.e., the modules with sensors and on-detector
 453 electronics are mounted on the front and back sides of a common cooling disk. As illustrated
 454 in Fig. 2.9, the modules on the two sides of a disk are arranged to overlap so that the
 455 number of hits exceeds the number of disks. A study using full simulation was performed
 456 to determine the optimal overlap between modules in $r < 230$ mm to achieve the required
 457 timing resolution via the average number of hits given the expected time resolution of the

458 pads. The maximal overlap is limited by the need for sufficient space between the modules to
 459 allow the read out of the data. For $r > 470$ mm, an overlap of 20%, for $470 \text{ mm} > r > 230$ mm
 460 an overlap of 54% and for $r < 230$ mm an overlap of 70% was the result of the optimisation.
 461 The HGTD acceptance is defined as the surface covered by the HGTD between a radius
 462 of 120 mm and 640 mm. The number of hits as a function of radius and trasvenrse plane
 463 position is shown in Fig. 2.10. The relative fraction of tracks as a function of hits per track for
 464 each ring can be found in Fig. 2.11.

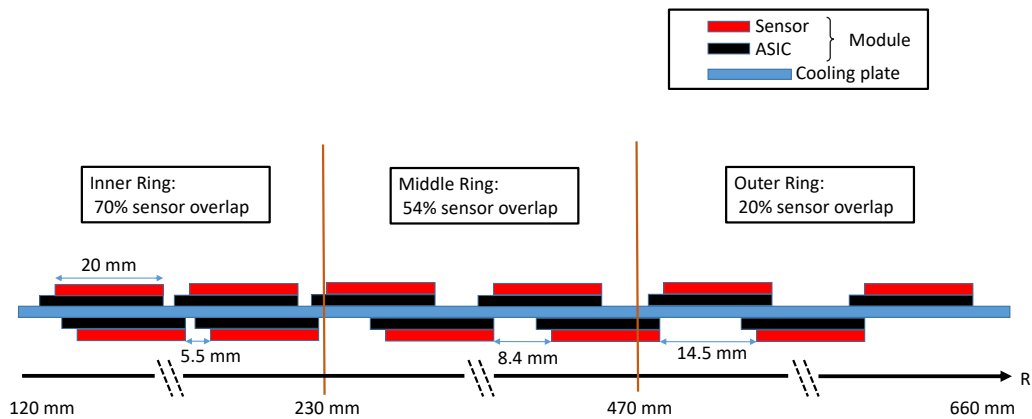


Figure 2.9: The schematic drawing shows the overlap between the modules on the front and back of a cooling disk. There is a sensor overlap of 20% for $r > 470$ mm, 54% for $470 \text{ mm} > r > 230$ mm and 70 % for $r < 230$ mm.

465 The material for the HGTD is highlighted in Fig. 2.12, which includes the material for the
 466 moderator located behind the active sensor area of the HGTD.

467 Beyond pile-up mitigation, HGTD can play a key role in the HL-LHC physics programme
 468 as a luminometer. An accurate luminosity determination will be a critical input for corner-
 469 stone precision measurements. The luminosity uncertainty can be a limiting factor to many
 470 precision cross-section measurements, including achieving $\mathcal{O}(1\%)$ accuracy on certain meas-
 471 urements of Higgs boson production and couplings. It is therefore important to be able to
 472 determine the luminosity as accurately as in Phase-I, which will be a challenge with the
 473 harsh environment at the HL-LHC. For the technologies used traditionally for luminometers,
 474 the increased pile-up leads to increased detector occupancies, posing serious problems. The
 475 HGTD provides unique and pile-up-robust capabilities for measuring the luminosity at the
 476 HL-LHC.

477 Taking advantage of the high granularity of the detector, the luminosity can be measured

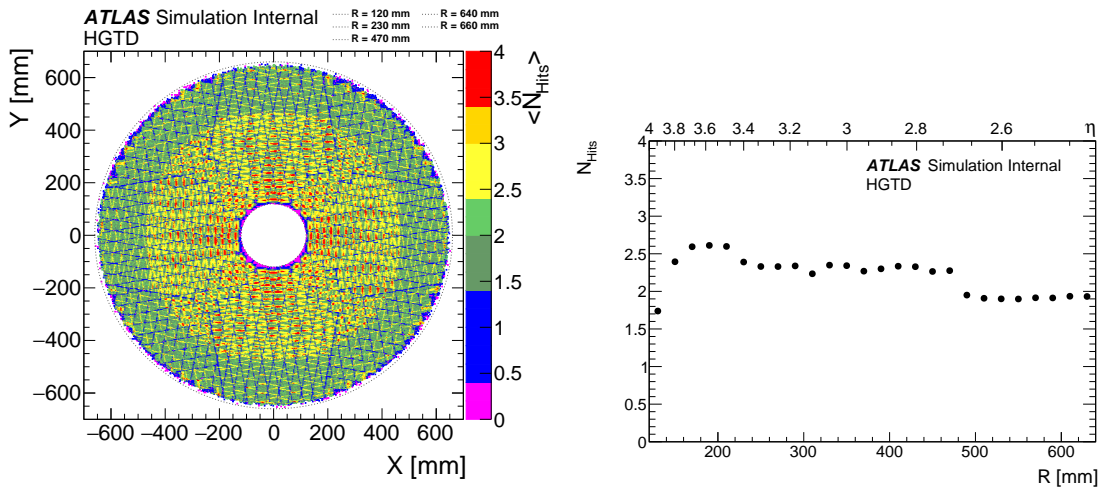


Figure 2.10: Hit multiplicity as function of X,Y and R.

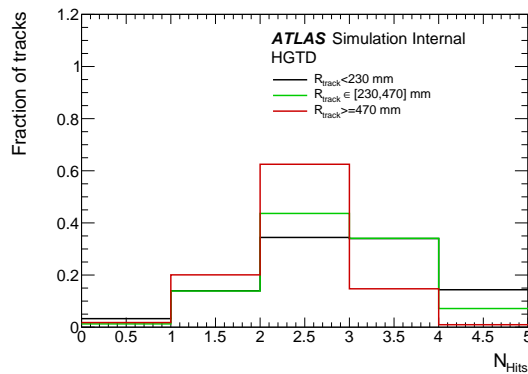


Figure 2.11: Fraction of tracks as a function of number of Hits in the track for tracks in the inner, middle and outer ring.

478 by counting the mean number of hits in the detector, a quantity linearly proportional to the
 479 average number of interactions per bunch crossing. The counting will be done over two
 480 time windows, one centred at the bunch crossing and with a width of 3.125 ns, the other
 481 with both width and relative position tunable with a step of 3.125 ns. The application of
 482 these capabilities and their implementation are further discussed in in Sec. 3.3.2, Chap. 6,
 483 and Chap. 10.

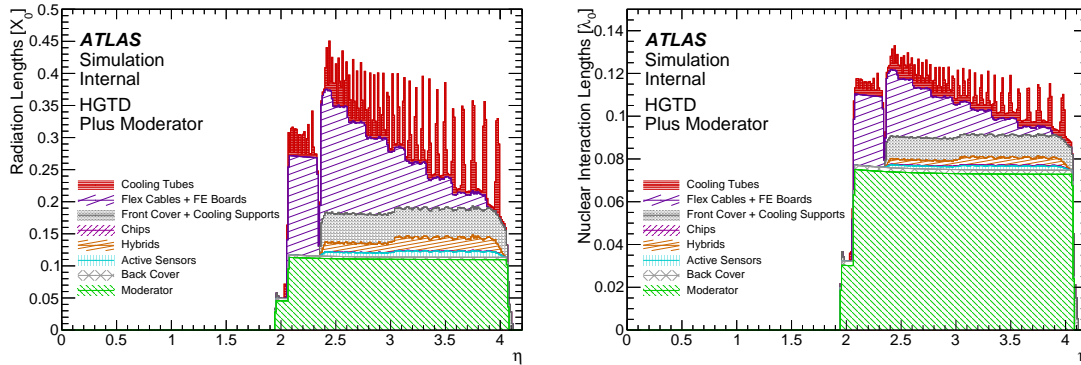
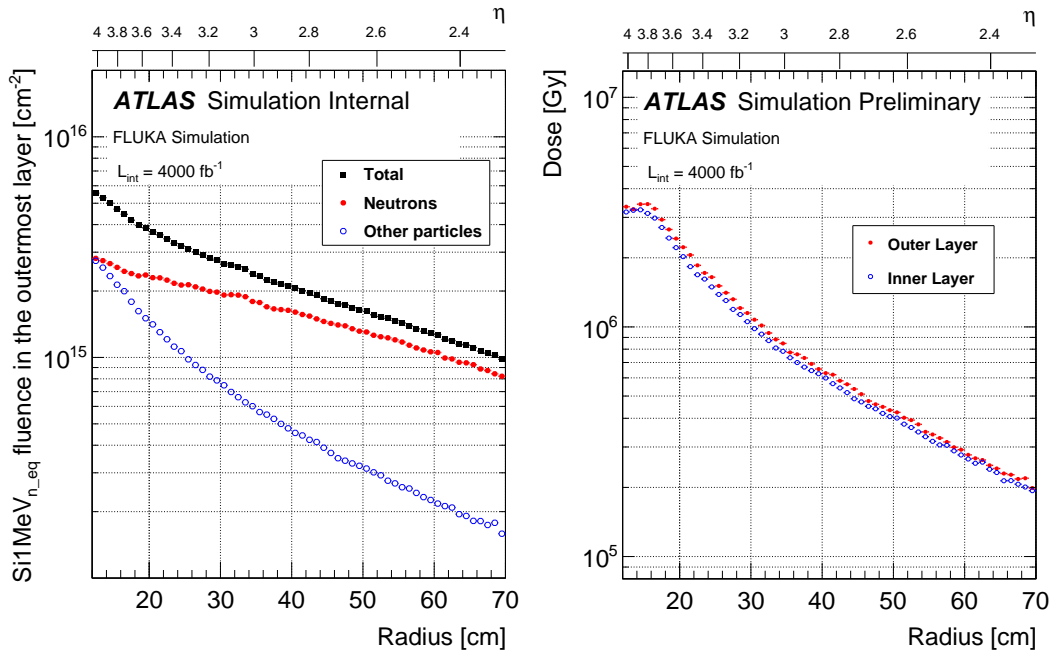


Figure 2.12: Radiation length X_0 (left) and nuclear interaction length λ_0 (right) as a function of pseudo-rapidity η , broken down by type of material for the HGTD. The moderator is included as it is within the hermetic vessel, although it is situated completely behind the active area of the HGTD. The baseline cooling pipes will be made with Titanium, which will reduce the peaks amplitude by almost a factor of 2.

484 2.4 Radiation hardness

485 One of the most important parameters of the HGTD will be radiation hardness of the sensors
 486 and electronics. Given that the HGTD will be installed with a pseudo-rapidity coverage
 487 of $2.4 < |\eta| < 4.0$, it is crucial that the detector can withstand the lifetime of the HL-LHC
 488 running. At the end of the HL-LHC (4000 fb^{-1}), the maximum nominal neutron-equivalent
 489 fluence at a radius of 120 mm, should reach $5.6 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ and the total ionising dose
 490 (TID) will be about 3.3 MGy, as shown in Fig. 2.13. To account for uncertainties in the
 491 simulation, a safety factor of 1.5 is applied to both numbers. An additional factor of 1.5 is
 492 applied to the TID due to uncertainties in the behaviour of the electronics after irradiation.
 493 This leads to a total safety factor of 1.5 for the sensors that are most sensitive to the particle
 494 fluence, and 2.25 for the electronics which are more sensitive to the TID. After applying
 495 these, the detector would need to withstand $8.3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ and 7.5 MGy.

496 This amount of radiation damage to lowest-radius ($r < 230 \text{ mm}$) sensors and electronics
 497 suggests that this innermost part of the detector should be replaced after each 1000 fb^{-1} and
 498 the sensors and electronics within $470 \text{ mm} > r > 230 \text{ mm}$ should be replaced at half lifetime
 499 (2000 fb^{-1}) of data-taking during the HL-LHC program. The plan is therefore to replace the
 500 sensors and ASICs located at a radius up to about $r > 230 \text{ mm}$ three times and the sensors
 501 and electronics located within $470 \text{ mm} > r > 230 \text{ mm}$ once. This corresponds to about 52%
 502 of the sensors and ASICs. Consequently, in the proposed 3 rings layout the maximal TID
 503 and fluence, using the Fluka estimations of September 2019, does not exceed 2 MGy and
 504 $2.5 \times 10^{15} \text{ neq/cm}^2$. In the inner ring the total Si 1MeV neq has a similar contribution from
 505 neutrons and charged particles while in the middle and outer rings the dominant effect
 506 comes from neutrons, as seen in Fig. 2.13. The exact radial transition between the three rings



(a) Nominal Si1MeV_{n,eq} fluence for HL-LHC. (b) Nominal ionising dose for HL-LHC.

Figure 2.13: Expected nominal Si1MeV_{n,eq} fluence and ionising dose as functions of the radius in the outermost sensor layer of the HGTD for 4000 fb⁻¹, i.e. before including safety factors. The contribution from charged hadrons is included in 'Others'. These estimations used Fluka simulations using ATLAS Fluka geometry 3.1Q7 (from December 2019).

507 will be tuned for the final detector layout, once the FLUKA simulations will be updated
 508 with the final ITk layout, and the radiation hardness of the final sensors and ASICs are
 509 re-evaluated.

510 More details can be found in Chap. 5 to Chap. 6. The maximum fluence and total ionising
 511 dose as a function of the radial position including the replacement of the rings can be found
 512 in Fig. 4.2. The expected proton, neutron, and pion energy spectra in the HGTD front and
 513 rear layer after 4000 fb⁻¹ are shown in Fig. A.1, Fig. A.2, and Fig. A.3.

514 **3 Performance and Physics Benchmarks**

515 The precision time measurement capability of the HGTD enhances the performance for hard
516 scatter jet tagging, missing transverse momentum E_T^{miss} , tagging b -jets, and lepton isolation
517 in the forward region. Additionally the timing structure of non collision background (NCB)
518 is analysed. The impact of these improvements on the sensitivity of a few selected physics
519 analyses is presented. The HGTD timing capabilities will provide a luminosity measurement,
520 a discussion on the impact of the precision of this measurement as well as the systematic
521 errors are also included.

522 **3.1 Simulation**

523 The full simulation of the HGTD is performed using a software release dedicated to the
524 ATLAS Upgrade program. The production of simulated samples follows the same steps
525 as regular ATLAS simulation based on the Run 2 offline software chain: event generation,
526 detector simulation, digitisation of simulated energy deposits into the actual detector read-
527 out data format, and event reconstruction starting from the digitised data.

528 The simulation of the HGTD was implemented as close as possible to the layout foreseen.
529 In the direction of the beam axis the HGTD implementation starts at 3420 mm. Behind
530 the HGTD before the cryostat wall two moderators with a total thickness of 50 mm are
531 implemented.

532 **3.1.1 Detector geometry**

533 The GEANT4 toolkit [6] is used to simulate the ATLAS detector. The simulation uses
534 dedicated `GeoModel` packages to implement the detector geometry which is converted, via
535 dedicated tools, to GEANT4 volumes. Particles are propagated through this geometry and
536 the various physics processes, caused by their interaction with the detector material, are
537 simulated. In sensitive detector elements, processes ranging from energies of a few eV, such
538 as the ionisation in gases, up to TeV energies are simulated to provide a detector-response
539 model as realistic as possible. The simulation step produces hits, i.e. small steps in the
540 material with a starting point, an end point and the amount of energy deposited by the
541 particle.

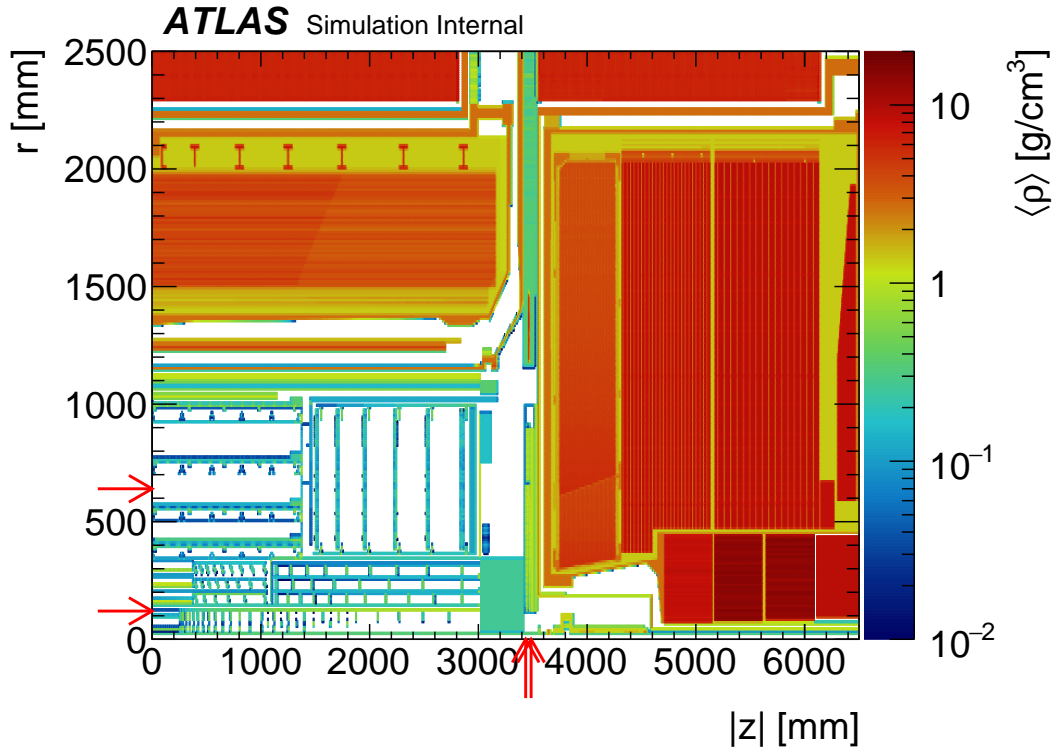


Figure 3.1: The material density is shown as function of z and radius. The extent of the active part of the HGTD in z and r direction, 500 mm and 520 mm respectively, is indicated by the arrows. **TODO:** This figure needs to be updated.

542 In Fig. 3.1, the material density of the upgraded ATLAS detector is shown as a function of
 543 the beam axis and the radius. The active part of the HGTD extends from 120 mm to 640 mm
 544 in radius.

545 There are four individual layers per endcap, the silicon sensors are mounted on both sides
 546 of the two cooling ensembles. The simulation of the HGTD has been extended to implement
 547 the peripheral electronics at radii greater than 640 mm. As an illustration a simulated event
 548 is shown in Fig. 3.2.

549 The simulation of the HGTD includes front and back covers and heaters of the HGTD. The
 550 simulation of the cooling plates, the CO₂ cooling loops as well as the support plates for the
 551 modules is also implemented. Volumes mimicking the peripheral electronics have also been
 552 implemented in the simulation.

553 The modules are the ASICs plus the sensors. These are simulated as boxes of size 22 mm ×
 554 40 mm with silicon sensors of size 20.5 mm × 40 mm perpendicular to the direction of the
 555 beam axis. The modules are slightly larger than the sensors in one dimension to account for

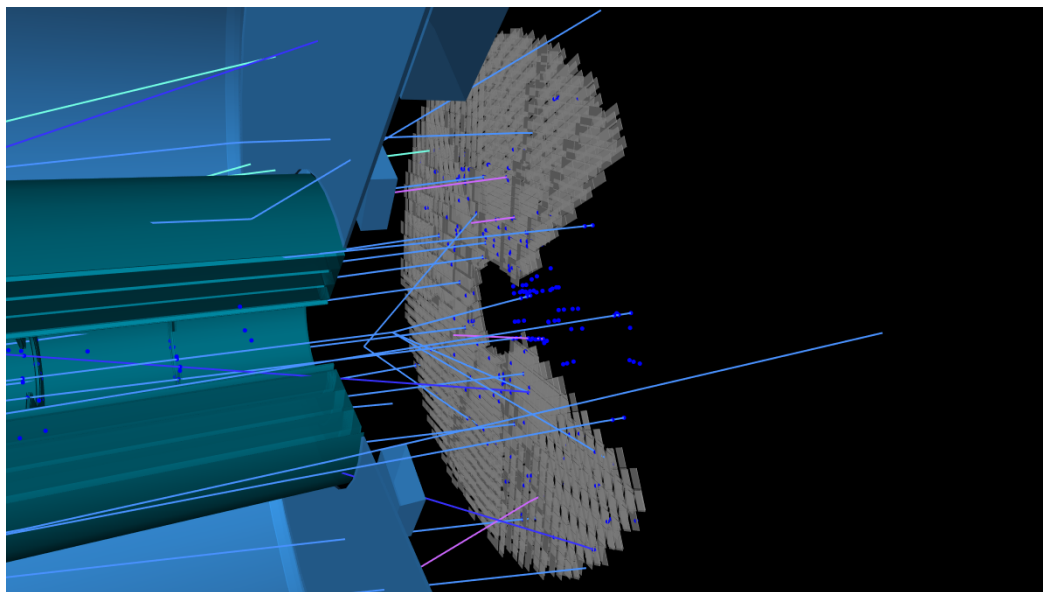


Figure 3.2: The simulation of an event in the HGTD is shown.

556 the wirebonding of the ASIC to the silicon sensor. The total thickness of the silicon sensor is
557 $250\ \mu\text{m}$ with an active thickness of $50\ \mu\text{m}$ a passive thickness of $200\ \mu\text{m}$, corresponding to
558 the LGAD sensor design. In total, 3992 modules are present in each end-cap. **TODO: Check**
559 **these numbers - do we use two-ring or three-ring layout here?**

560 The Flex PCBs connecting the ASICs to the peripheral electronics beyond 640 mm have also
561 been implemented in the simulation. This leads to an increase of the material distribution of
562 the HGTD as the radius increases.

563 3.1.2 Sensor simulation

564 The baseline of $1.3\ \text{mm} \times 1.3\ \text{mm}$ is used for the padsize. Two sources of inefficiency are
565 taken into account at simulation time. The guard ring of 0.5 mm and the inter-pad dead
566 zone of $50\ \mu\text{m}$ are implemented. The active area is 79% of the total silicon area. The different
567 zones of the sensors are illustrated in Fig. 3.3.

568 The digitization and clustering steps are implemented using silicon pixel hits and contains
569 functionality for associating truth information to the simulated hits. The digitization step
570 generates a detector signal with a pulse shape extracted from beam tests of LGAD sensors,
571 adds expected electronic noise and is capable of describing the timing performance expected
572 at various points during the HL-LHC program. The expected timing resolution depends
573 on the radiation the detector has been exposed to, and this is implemented through with
574 increasing integrated luminosity.

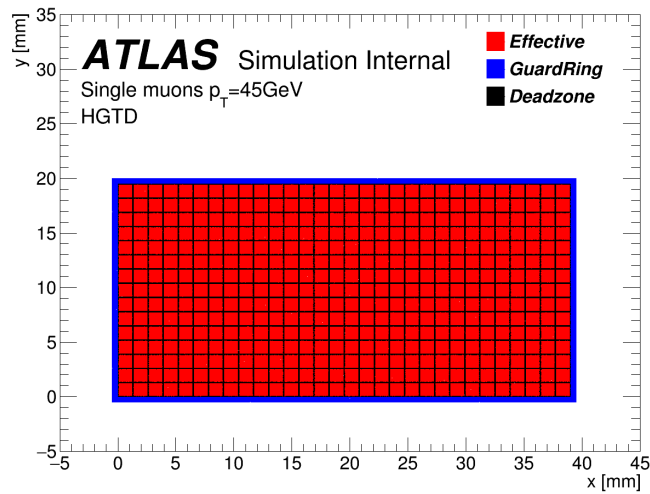


Figure 3.3: The simulation of the HGTD sensor is shown separated into the active area, the guard ring and the inter-pad deadzones using the simulation of single muons.

575 The active area of each pad is associated with a unique identifier. In the digitization step
 576 for $\langle \mu \rangle = 200$, the hits from the different interactions are summed in energy if they are in
 577 the same 5 ps time bin (TODO: Check to what extent this is true for SiHits). To allow for
 578 maximum flexibility at the analysis stage, the hits are then copied down to the format used
 579 for the analysis. The timing information is stored after subtracting a global offset of 11.6 ns,
 580 corresponding to the time of flight from the nominal interaction point to the center of the
 581 HGTD.

582 In Fig. 3.4 the hits in the HGTD are shown in the transverse plane for the two cooling plates
 583 with modules on the front and back for positive z . The position of the modules can be
 584 identified. The displacement of modules mounted on the back of a cooling plate with respect
 585 to those mounted on the front of each cooling plate is shown. The transition around 320
 586 but outofdate from an overlap of 80% to 20% is visible as decreased density of sensors. The
 587 mirror symmetry between the first and second cooling disk as well as the rotation of 15°
 588 each of the cooling disks is observed.

589 The simulation provides the energy deposit in the sensitive layer of the HGTD as single
 590 energy deposit for each particle traversing it. The simulation of the non-uniform distribution
 591 of the charges in the sensitive volume as well as the effect of the electronics chain (time walk,
 592 jitter) are taken into account at analysis level. For each hit a pulse is simulated to compute
 593 the time and energy in each pad. Data derived from the 2016 HGTD test beam were used to
 594 derive the pulse shape. A convolution of a Gaussian with a Landau distribution was found
 595 to give the best description of the pulse shape. The non-uniform energy deposit is modeled
 596 via the width of the Gaussian. The signal time is defined on its leading edge, therefore the

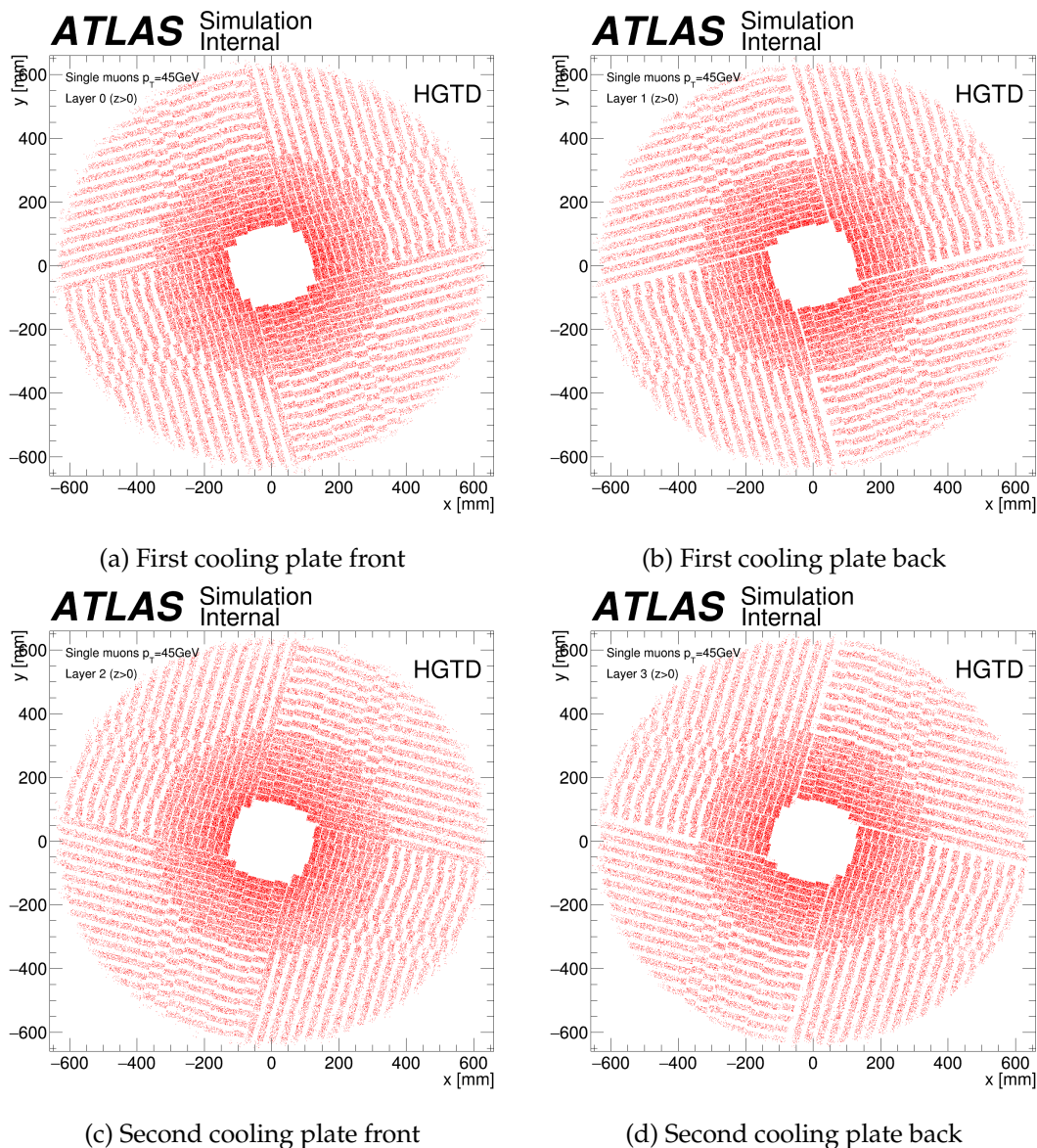


Figure 3.4: The positions of simulated hits in the transverse plane for each of the four layers separately. Only hits in active parts of the detector elements are shown. The position of the modules, the mirror symmetry between the first and second cooling plate as well as the rotation are clearly visible. **TODO:** These plots need to be updated to be made with SiHits, or just replaced with the figures showing the placement of the modules.

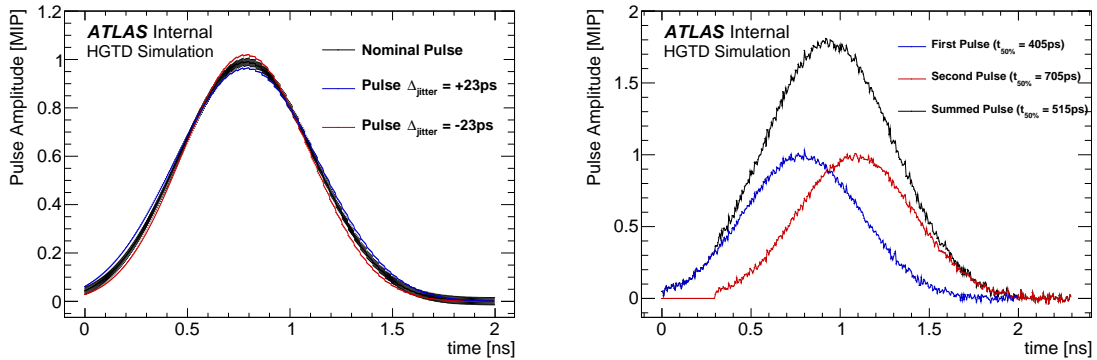


Figure 3.5: The simulated shape in time of the signal in a pad of the HGTD is shown for (a) a normal hit and (b) a double hit, separated by 300 ps.

597 variation models adequately the induced timing uncertainty. Fig. 3.5 shows the nominal
 598 shape and the effect of two hits in the same pad, separated by 300 ps.

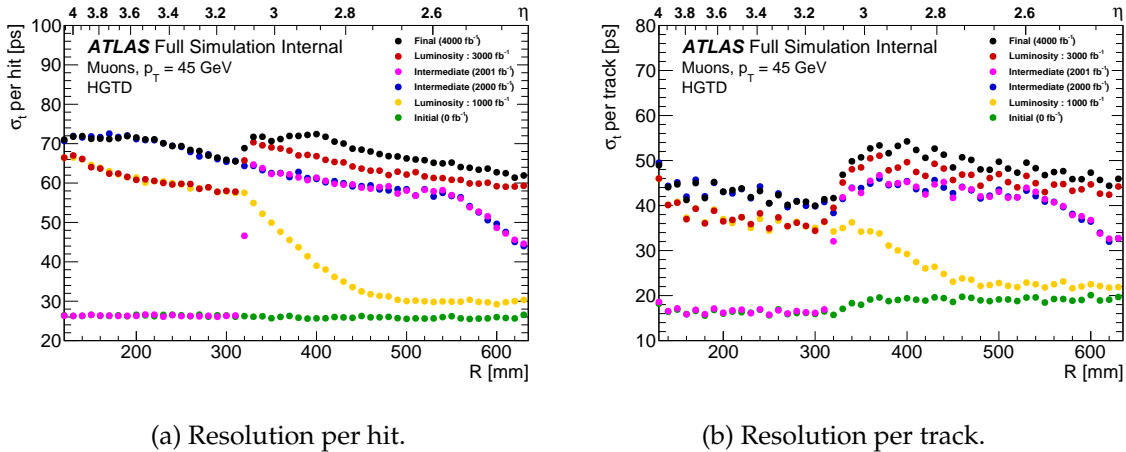
599 For each hit, a pulse is simulated with 200 points of a step size of 5 ps where the width of
 600 Gaussian contribution is driven by the desired timing resolution of the sensor. The maximal
 601 amplitude of the pulse is the deposited energy. The time corresponding to the first point of
 602 the pulse is chosen to be the time of the hit. Additionally a Gaussian noise with of 1.5% of
 603 the energy of a MIP (0.2 keV) is added to the amplitude in each time bin.

604 For each pad, the pulses are then summed together. A pseudo constant fraction discriminator
 605 (CFD) algorithm defines the time as the time of the first point with an energy above 50% of
 606 the maximum amplitude. Therefore the time of a pad is offset by 0.405 ns.

607 The contribution of electronic noise to the timing resolution is taken into account as a function
 608 of the position of the sensor and the accumulated integrated luminosity with a Gaussian
 609 smearing. The dose received by the sensor as a function of its radius was computed using
 610 FLUKA, then data from test bench measurements of sensors define the corresponding gain
 611 for the sensor. The gain is transformed into the timing resolution using measurements with
 612 ALTIROC0. This procedure results in a Gaussian smearing of minimum 10 ps and maximum
 613 60 ps.

614 Timing Resolution Scenarios

615 Four timing performance scenarios are defined: *Initial*, two *Intermediate* scenarios and
 616 *Final* corresponding to integrated luminosities of 0 fb^{-1} , 2000 fb^{-1} , 2001 fb^{-1} and 4000 fb^{-1}
 617 which are shown in Fig. 3.6. The scenarios correspond to the performance expected at
 618 the beginning, after half of the expected integrated luminosity, after half of the expected



(a) Resolution per hit.

(b) Resolution per track.

Figure 3.6: The HGTD timing resolution is shown as function of the radius for four timing scenarios. The sensor resolution and the contribution from the electronics are considered, added in quadrature.

619 integrated luminosity after replacement of the inner part of the HGTD and at the end of the
 620 HL-LHC data taking. Two additional scenarios for 1000 fb^{-1} and 4000 fb^{-1} are also shown to
 621 illustrate the change of the resolution as the integrated luminosity increases. Maintaining
 622 this performance as the clock is distributed across the detector will require intercalibrating
 623 the reference t_0 , discussed in Sec. 10.2.

624 As the integrated luminosity increases, the damage to the sensors and electronics caused
 625 by radiation will deteriorate their timing resolution inducing a dependence on the radial
 626 distance from the beam axis. The replacement of the inner part ($r < 320$ but outofdate)
 627 of the HGTD after half of the HL-LHC programme leads to identical performance of the
 628 timing resolution in this region for the *Intermediate* for 2000 fb^{-1} and *Final* scenarios. In
 629 the simulation the replacement has been implemented only as function of the radius so
 630 only pads at a radius below 320 but outofdate are replaced. This is conservative as the
 631 entire module will be replaced which will lead to an improved performance above 320 but
 632 outofdate. Fig. 3.6 has bins of 10 mm so that after the replacement the transition region at
 633 320 but outofdate has contributions from both regions. The timing resolution for a hit at a
 634 radius of 120 mm is expected to be degraded to the order of 70 ps after $2.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$,
 635 including the replacement.

636 The resulting timing resolution for tracks is shown in Fig. 3.6(b) for the four scenarios as
 637 well as for the two additional scenarios of 1000 fb^{-1} and 4000 fb^{-1} . At radii lower than 320
 638 but outofdate ($|\eta| \approx 3.1$ but outofdate!), the increased number of hits compensates the effects
 639 of radiation damage such that the resulting timing resolution is fairly independent of the
 640 radius. The per-hit resolution was implemented in full simulation. The reconstructed timing
 641 resolution for tracks shown here was determined using the single-muon events by comparing
 642 the reconstructed hit or track time to the expected time from the truth information. The hits
 643 closest to the extrapolation of the reconstructed muon track within a window of 1.4 mm are

Process	Generator	N events, $\langle\mu\rangle = 0$	N events $\langle\mu\rangle = 200$
Minimum Bias, high- p_T	Pythia8	1000000	
Minimum Bias, low- p_T	Pythia8	100000	
Single π^+ , $p_T = 5$ GeV, flat η [2.3-4.3]		20000	20000
Single π^+ , $p_T = 20$ GeV,		20000	20000
Single π^+ , $p_T = 45$ GeV,		200000	200000
Single π^+ , flat η [2.3-4.1], flat p_T [0.1-5.0] GeV		20000	20000
Single π^0 , flat η [2.3-4.1], flat p_T [0.1-5.0] GeV		20000	20000
Single γ , $p_T = 20$ GeV, flat η [2.3-4.3]		20000	20000
Single γ , $p_T = 45$ GeV, flat η [2.3-3.2]		20000	20000
Single γ , $p_T = 100$ GeV, flat η [2.3-3.2]		50000	50000
Single μ , $p_T = 45$ GeV, flat θ		40000	40000
Single μ , $p_T = 45$ GeV, flat η [2.3-3.2]		30000	30000
Single μ , $p_T = 45$ GeV, flat η [3.2-4.3]		10000	10000
Single electron, $p_T = 45$ GeV, flat η [2.3-4.3]		40000	40000
Single electron, $p_T = 20$ GeV, flat η [2.3-4.3]		20000	20000
NCB beam–gas, oxygen		40000	
NCB beam–gas, carbon		40000	
NCB Beam–gas, hydrogen		40000	
$Z \rightarrow ee$	Powheg+Pythia8	10000	10000
$Z \rightarrow \tau\tau$	Powheg+Pythia8	40000	40000
$t\bar{t}$	Powheg+Pythia8	100000	100000
VBF $H \rightarrow ZZ \rightarrow 4\nu$	PowhegPythia8	50000	50000
Dijet production, X GeV $\hat{p}_T < Y$ GeV	Pythia8	1000 000	1000000
Dijet production, X GeV $\hat{p}_T < Y$ GeV	Pythia8	1000 000	1000000
Dijet production, X GeV $\hat{p}_T < Y$ GeV	Pythia8	1000 000	1000000

Table 3.1: Monte Carlo events simulated, digitized and reconstructed with Athena releases AtlasProduction-20.20.14.4-6 using the Step 3.1 ITK geometry (geometry tag ATLAS-P2-ITK-17-04-02).

644 used in this study. The distribution is dominantly Gaussian with negligible tails.

645 Production

646 The simulation, digitisation and reconstruction was implemented in the ATLAS upgrade
647 software releases 20.20.14.1 and 20.20.14.2 deployed on the grid. Samples of single particles,
648 electrons, muons and pions as well as non–collision background (NCB) and selected physics
649 processes such as $t\bar{t}$, VBF $H \rightarrow Z(\nu\nu)Z(\nu\nu)$ and monopoles were produced using the ATLAS
650 production system. Pythia8 [7] was used together with Powheg [8–10] for most of the
651 samples. In the simulation step the beamspot was simulated with the spread in z and t .
652 Samples with $\langle\mu\rangle = 0$ as well as $\langle\mu\rangle = 200$ were processed. A summary of the samples
653 is shown in Tab. 3.1. For the minimum bias (inelastic collisions in the underlying event)
654 samples, a signal set of single neutrinos was used at $\langle\mu\rangle = 200$.

655 3.1.3 Simulation using silicon hits

656 Digitisation

657 The Geant4 simulated hits are processed in a digitisation step in order to emulate the detector
658 electronics output. The HGTD LGADs are simulated adapting the ATLAS ITk pixel offline
659 software. They are described as planar n-in-p pixel sensors with electron carriers. The
660 channel efficiency is simulated as perfect for hits above threshold, also without simulated
661 defects. As described in Sec. 3.1.2, the HGTD sensors are implemented to have pixels with
662 a size of $1.3 \text{ mm} \times 1.3 \text{ mm}$, a 0.5 mm thick guard ring and inter-pixel dead zones of $50 \mu\text{m}$.
663 The active thickness of the silicon sensors is $50 \mu\text{m}$.

664 During digitisation, the energy deposited for each Geant4 step in the active silicon volume
665 is used to evaluate the free charge and the drift time to the readout surface accordingly to
666 the sensor thickness, carries mobility, depletion and bias voltage and Lorentz shift. The
667 front-end electronics in-time threshold is set to 600 electrons, with an intrinsic standard
668 deviation of 40 electrons added in quadrature with a noise standard deviation of 75 electrons.
669 Given the characteristics of the sensors, the capacitive coupling to nearby pixels is considered
670 negligible as studied in LGAD beam tests and discussed in Chap. 5 and cross-talk effects
671 are not simulated.

672 The time measurement associated to each energy deposition in the sensor is obtained
673 smearing the time of the Geant4 interaction in the silicon with a sensor time resolution of
674 30 ps . At this point, the algorithm estimates the total charge released in each pad and checks
675 if it is above threshold. Pads above threshold are labelled as fired.

676 The time structure of the energy depositions in the active area of HGTD sensors in the first
677 layer is shown in Fig. 3.7 for the Initial timing scenario. The time distribution is obtained
678 subtracting for each deposit the expected time of arrival accordingly to the position of the
679 sensor and assuming a straight line trajectory from the center of the detector. For all layers,
680 the deposits originating from primary and secondary particles are in time for the bulk of the
681 distribution. However secondaries create also a pronounced tail in the timing distribution.
682 In the following, only fired pads characterised by energy depositions in $[-1.0, +1.0] \text{ ns}$
683 time window around the expected time of arrival are taken into account. The first energy
684 deposition determines the time measurement to be associated to each of them.

685 Clustering

686 The first step of event reconstruction is the formation of clusters from the individual channels
687 of the HGTD. For silicon-based detectors, like the ITk pixel and strip detectors, this is a local
688 pattern recognition step where adjacent readout channels are grouped together in clusters,

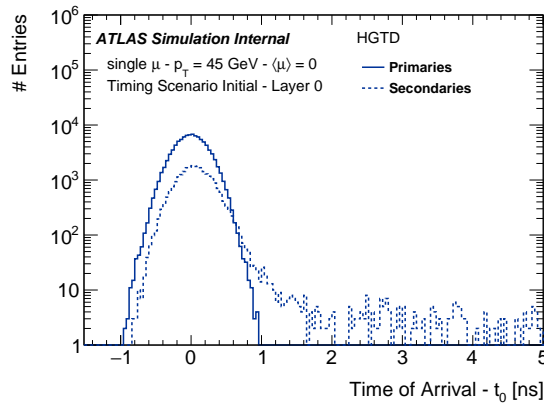


Figure 3.7: The distribution of the time of the hits with respect to the truth time is shown for hits originating from primary and secondary particles from a single muon sample without pile-up.

689 which represent single position measurement. This is done using a connected component
690 analysis which is chosen to be based on eight-cell connectivity [ref.].

691 For the HGTD three different clustering approaches have been studied in order to profit
692 of the time measurement associated to the individual channels and the bigger sensor size,
693 compared to the pad size:

- 694 • Geometric clustering: as in the ITk pixel reconstruction, this approach groups adjacent
695 fired pads to form clusters neglecting the time measurement associated to them. The
696 local position of the cluster is calculated by taking a simple center of gravity of all the
697 pads in the cluster.
- 698 • Geometric clustering with time filtering: the time measurement is used to cluster
699 together adjacent fired pads. These are grouped if the time difference of the considered
700 channels is smaller than 30 ps. The local position of the cluster is calculated as in the
701 previous case, while the average of the time measurements to the pads in the cluster
702 determines the time of the cluster.
- 703 • Single-pad clustering: each fired pad is converted into a cluster. The centre and the
704 time measurement of the fired pad define the local position and the time of the cluster,
705 respectively.

706 While the single-pad clustering is always providing clusters with dimension 1 along both
707 local coordinates axis, the other two approaches can give reconstructed cluster sizes that
708 differ from the truth ones. Fig. 3.8 shows the reconstructed cluster size in $local_x$ and $local_y$ ¹
709 using the geometric clustering algorithm with and without time filtering for $t\bar{t}$ Monte Carlo
710 events with 200 pile-up. The reconstructed cluster size is compared to the truth width of

¹ The $local_x$ and $local_y$ coordinates represent the two coordinates along the sensor grid. $local_x$ is in the $R\phi$ plane perpendicular to the beam line while $local_y$ points radially in R .

711 clusters reconstructed with the geometric clustering with time filtering for $t\bar{t}$ events without
 712 pile-up. The large cluster size at higher pseudo-rapidity provided by the geometric clustering
 713 algorithm is due to the higher sensor occupancy. When time filtering is required, smaller
 714 clusters are obtained with an average size close to one pad in both $local_x$ and $local_y$. Fig.
 715 3.9 shows the probability of merging contributions originated by multiple particles in the
 716 same cluster with the geometric clustering algorithm with and without time filtering in $t\bar{t}$
 717 events with 200 pile-up. In order to correctly associate clusters to tracks and provide good
 718 timing measurements, one wants that the rate of merging multiple contributions into one
 719 clusters stays as low as possible. Therefore, taking into account the results described above
 720 and shown in Fig. 3.8 and Fig. 3.9, single-pad clusters have been considered the optimal
 721 collection as input to the tracking reconstruction.

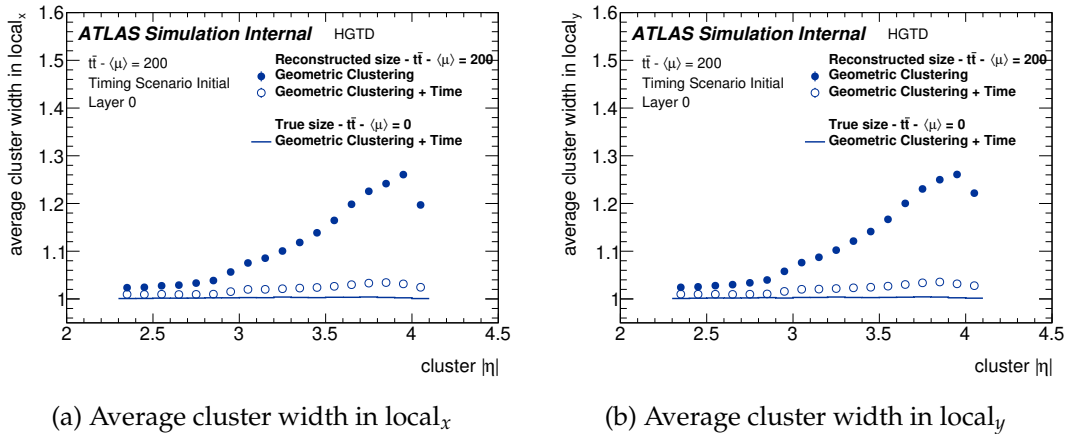


Figure 3.8: Average cluster width in $local_x$ (a) and $local_y$ (b) for clusters obtained with the geometric clustering algorithm with and without time filtering in $t\bar{t}$ events with 200 pile-up. The reconstructed cluster sizes are compared to the truth widths of clusters reconstructed with the geometric clustering with time filtering for $t\bar{t}$ events without pile-up.

722 3.2 Detector performance

723 As a first step the signatures in the detector are studied. This is followed by the discussion
 724 of the track level performance and the performance gains for individual objects.

725 3.2.1 Detector response characterisation

726 The HGTD will be installed at a distance of 3420 mm in z from the nominal interaction point.
 727 Particles produced at the interaction point will traverse the ITk and the material in front of
 728 the HGTD sensors. The HGTD provides signals at four distinct positions in z each with less

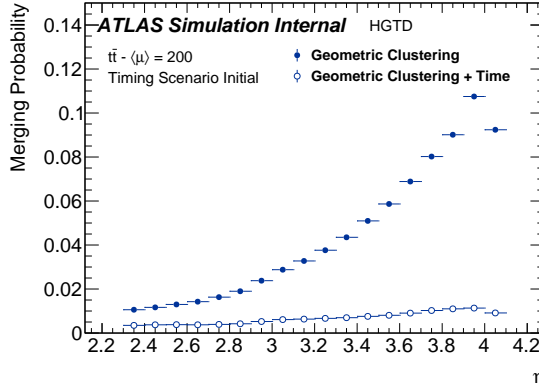


Figure 3.9: Probability of merging contributions originated by multiple particles in the same cluster with the geometric clustering algorithm with and without time filtering in $t\bar{t}$ events with 200 pile-up.

729 than 100% coverage of the transverse plane. Samples of single muons, pions and minimum
 730 bias events with $\langle\mu\rangle = 200$ have been simulated in order to calculate the expected number
 731 of hits per track as function of the incident particle. The effect of the material upstream and
 732 in the HGTD on incident electromagnetically and hadronically interacting particles as well
 733 as the percentage of pads with signal at high pileup was studied.

	$r < 320 \text{ mm}$ ($ \eta > 3.1$ but out of date!)	$r > 320 \text{ mm}$ ($ \eta < 3.1$ but out of date!)
$N_{\text{hits}} \geq 2$	88%	72%
$N_{\text{hits}} = 0$	1.6%	2.8%
$\langle N_{\text{hits}} \rangle$	2.7	1.9

Table 3.2: The percentage of tracks with at least two hits, the percentage of tracks escaping undetected and the average number of hits for muons with a p_T of 45 GeV are shown for the baseline detector layout. The values include the effect of non-instrumented zones, inter-pad dead-zones and guard ring as well as the track-matching efficiency.

734 The hit efficiency is studied using single-muon events with $p_T = 45 \text{ GeV}$. The distribution
 735 of the muons is flat the polar and azimuthal angles. The extrapolation of the tracks to the
 736 HGTD was performed using the last measured point of the track in the ITk. Only hits within
 737 1.4 mm of the result of the extrapolation are accepted. This is only slightly larger than the
 738 pad size. This ensures that the pad extrapolated to and a neighboring pad are candidates for
 739 the track matching. For only 5% of the hits the difference between the extrapolated position
 740 and the position of the closest hit is larger than the acceptance criteria. At most one hit was
 741 accepted for each sensor plane, so that the maximum number of hits is four. Fig. 3.10 shows
 742 the average number of hits per muon in the transverse plane. The results are summarised in
 743 Tab. 3.2. 79% of the extrapolated muon tracks have at least two matched hits in the HGTD.

744 For at most 2.4% of the muons, no hits are registered within the acceptance window around
 745 extrapolated position. The effect of uninstrumented zones and inactive areas on the sensor
 746 such as the guard ring and the inter-pad dead-zones are taken into account.

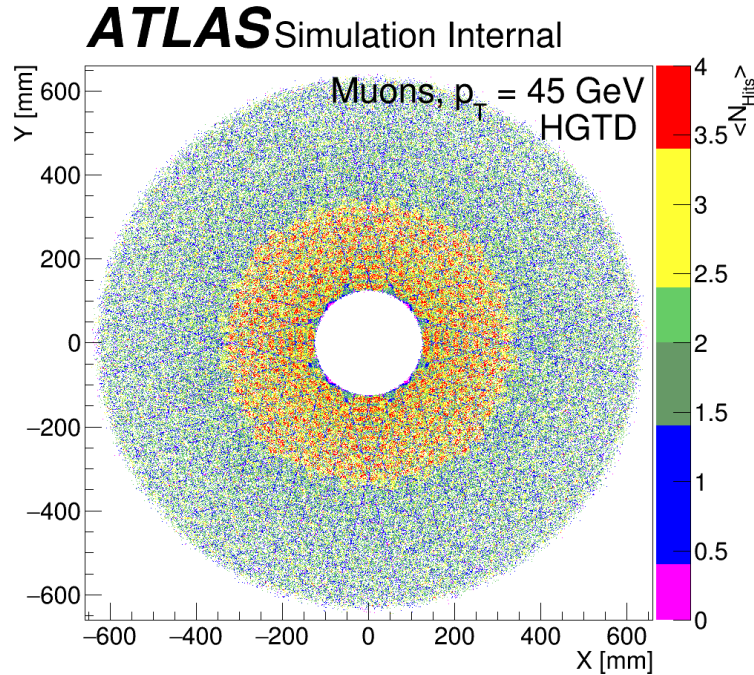


Figure 3.10: The average number of hits as a function of the position in the HGTD is shown for the baseline layout. The overlap is 80% at $r < 320$ but outdated and 20% at larger radii.

747 The average number of hits is shown as a function of the radial distance from the beam axis
 748 in Fig. 3.11. The overlap of 20% between the modules at $r > 320$ but outdated leads to an
 749 average number of hits of 1.9. The overlap 80% for $r < 320$ but outdated results in an average
 750 hit multiplicity of 2.7 in this region in agreement with the requirements listed in Tab. 2.1.

751 In Fig. 3.12 the two dimensional distribution shows the number of hits as function of $|\eta|$. The
 752 profile of the histogram shows that the number of hits exceeds the number of hits expected
 753 by the study of the detector optimization with muons. The muons, as shown in Fig. 3.12(a),
 754 only interact electromagnetically leading to a low multiplicity with an approximately Gaussian
 755 distribution. For pions, generated with flat distribution in transverse momentum between
 756 1 GeV and 5 GeV, Fig. 3.12(b), the hit multiplicity has long tails due to hadronic interactions.
 757 When restricting the hits to those within 1.4 mm of the extrapolated hit position, a similar
 758 behavior in η dependence and magnitude is observed for muons and pions. This shows
 759 that interactions in front of the HGTD are likely, creating showers, as shown by the profile
 760 histogram in Fig. 3.12(b) for a distance of 3.5 mm, leading to a number of hits larger than
 761 expected for a single hit per sensor layer.

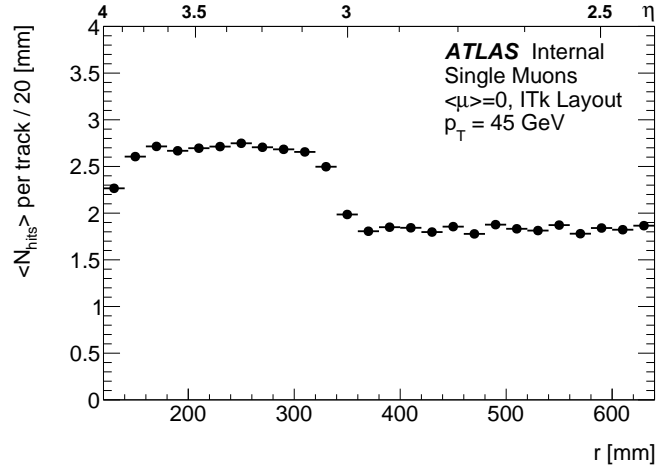
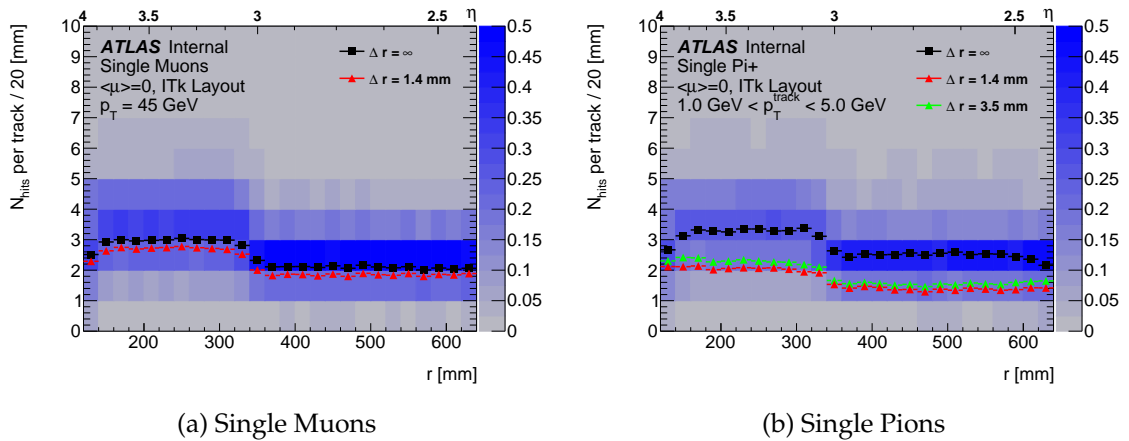


Figure 3.11: The average hit multiplicity as a function of the radius (and pseudo-rapidity) is shown for the baseline HGTD layout of 80/20% overlap below /above 320 but outdated.



(a) Single Muons

(b) Single Pions

Figure 3.12: The distribution of the number of hits as function of radius (and $|\eta|$) is shown as a 2D histogram with its profile (black). Each $|\eta|$ bin is normalised separately. For the profile histograms (red, green) the hits are required to be within a distance of the expected hit position.

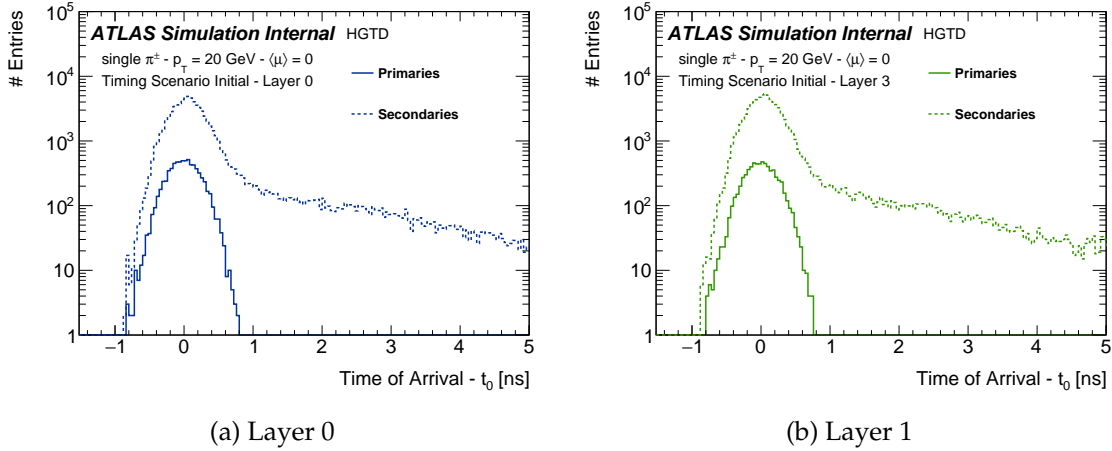


Figure 3.13: The distribution of the time of the hits with respect to the true time of arrival is shown for hits originating from primary and secondary particles.

762 The time structure of the hits in the first and last layer is shown in Fig. 3.13 for the *Initial*
 763 timing scenario. The structure is similar in all layers. The hits originating from primary and
 764 secondary particles are in time for the bulk of the hits. However the secondaries create a
 765 pronounced tail in the timing distribution.

766 The probability to have a hit in a pad decreases as function of the distance from the beam
 767 axis. The fixed pad size of the detector has to ensure a maximal occupancy of less than
 768 about 10% at the lowest instrumented radius of 120 mm. This ensures a low double-hit
 769 probability.

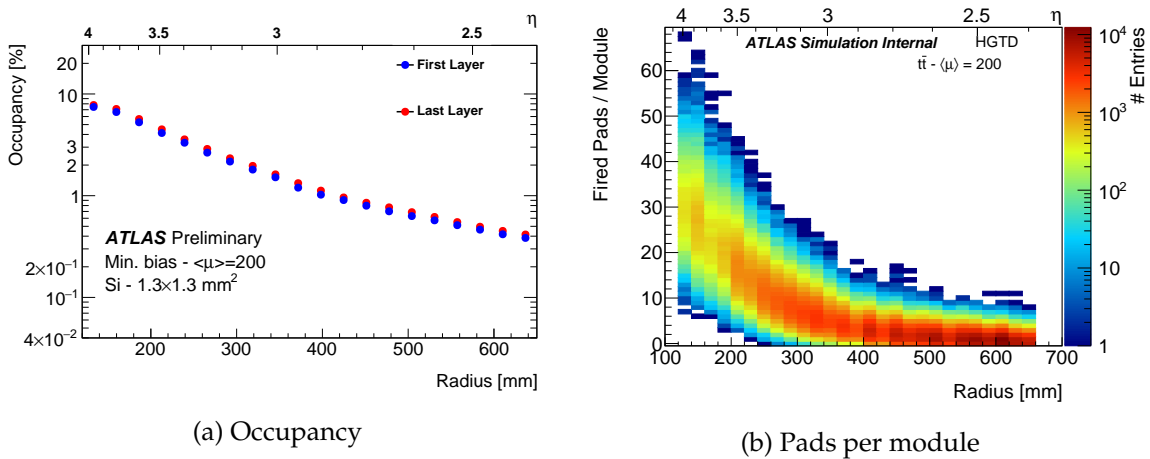


Figure 3.14: The occupancy 3.14(a) and the number of fired pads 3.14(b) per module are shown as a function of the radius for a pad size 1.3 mm × 1.3 mm at a pile-up of $\langle\mu\rangle = 200$.

770 In Fig. 3.14(a) the occupancy expected for a pile-up of $\langle\mu\rangle = 200$ is shown, defined as

771 the percentage of pads of the HGTD registering a hit, for the HGTD baseline pad size of
772 $1.3 \text{ mm} \times 1.3 \text{ mm}$. As expected, the occupancy decreases as a function of radius. A slight
773 increase is observed when moving outwards from the innermost to the outermost layer,
774 primarily due to the increased probability of initiating showers due to hadronic interactions
775 as more material is traversed. At the smallest radius the occupancy is 8%, fulfilling the
776 requirements.

777 The occupancy in Fig. 3.14(a) is calculated from a single bunch crossing at $\langle \mu \rangle = 200$.
778 The contribution from preceding bunch crossings has been estimated by calculating the
779 occupancy for the tails in the timing distribution, i.e., taking into account only hits more
780 than 2 ns later than the average time of flight. A more accurate modeling of hits arriving
781 late was obtained using FLUKA, additionally taking into account the induced nuclear
782 radioactivity. The time measurement is active only during 5 ns out of 25 ns. Given this, the
783 relative contribution to the occupancy is estimated to be of the order of a percent.

784 In Fig. 3.14(b) the average number of pads in a module with signal is shown as function of
785 the radius for a sample of top pair production at $\langle \mu \rangle = 200$. The variation of the number of
786 pads with signal in a module has to be taken into account in the calculation of the bandwidth
787 for the data transfer to the peripheral electronics.

788 Handling the Geant4 truth information it is also possible to study the types of the particle
789 firing the pads. The breakdown of the different contributions is shown as function of the
790 radius in In Fig. 3.15(a). In top pair production events with 200 overlaid pile-up interactions,
791 pile-up particles and secondaries from showers in the upstream detector material dominate
792 the occupancy. If a primary particle and another particle deposit energy in the same pad,
793 the signal of the primary particle can be masked in the LGAD sensor if the other arrives
794 earlier (shadowing). It is therefore important to evaluate the amount of primaries masked by
795 the early arrive of other particles. Fig In Fig. 3.15(b) shows the percentage of pads fired by
796 secondaries and pile-up particles shadowing a primary particle with respect to the number
797 of pads where at least one contribution from a primary particle occurs in $[-1.0, +1.0]$ ns. For
798 events with 200 overlaid pile-up interactions, the percentage of shadowed pads is 4.5% at
799 low radius, where the occupancy is maximal, decreasing to 1% at larger radius. Performing
800 the same analysis for $\langle \mu \rangle = 0$ shows that the level of 1% is due to particles originating from
801 the same primary interaction and characterised by a time of arrival compatible within the
802 timing resolution.

803 In Fig. 3.15(a) the breakdown of the origin of the hits detected in the HGTD within a time
804 window of total width 2 ns centered on the time of the primary is shown as function of
805 the radius. In a $t\bar{t}$ sample at $\langle \mu \rangle = 200$ secondaries and pileup dominate the occupancy. If
806 a primary particle and a pileup particle deposit energy in the same pad, the signal of the
807 primary particle can be deformed. In Fig. 3.15(b) the percentage of pads fired by secondaries
808 particles and pileup shadowing a primary particle with respect to the number of pads in
809 which a primary particle has deposited energy within the 2 ns time window around the
810 expected time of arrival. The distribution is shown as a function of the radius. For $t\bar{t}$ events at

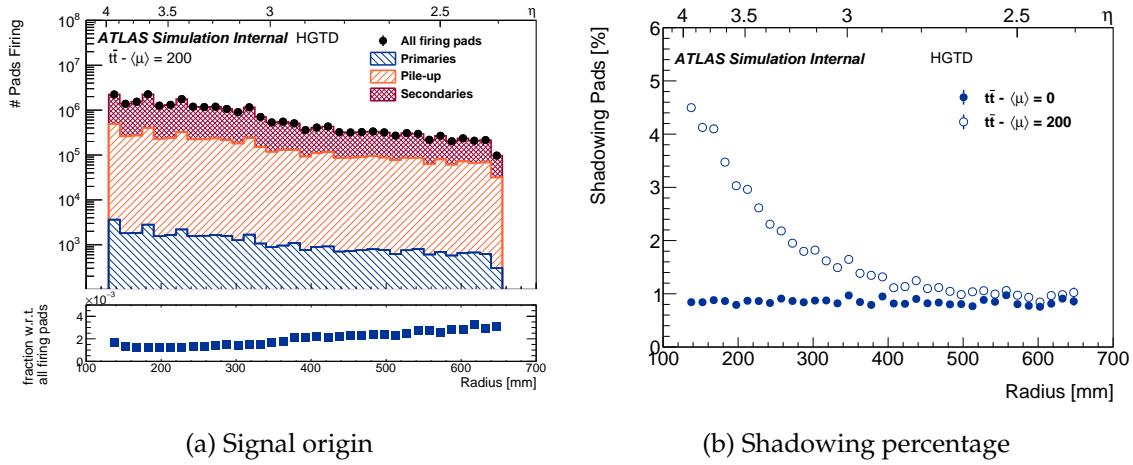


Figure 3.15: The origin of the hits 3.15(a) detected in the HGTD and the percentage of shadowed pads 3.15(b) are shown as function of the radius.

811 $\langle\mu\rangle = 200$, the maximum is 4.5% at low radius, where the occupancy is maximal, decreasing
 812 to 1% at larger radius. Performing the same analysis for $\langle\mu\rangle = 0$ shows that the level of 1%
 813 is due to secondaries and primaries arriving with a time of arrival compatible within the
 814 timing resolution. In these cases the shadowing effect does not bias the time measurement.

815 3.2.2 Track-level performance

816 Track-to-vertex association

817

818 The precise assignment of tracks to primary vertices (track-to-vertex association) is one of
 819 the key elements to mitigate the effects of pile-up on the full suite of event reconstruction
 820 algorithms at hadron colliders. Jet reconstruction and calibration, pile-up mitigation for jets,
 821 b -tagging, lepton isolation, and jet substructure measurements rely strongly on the correct
 822 assignment of tracks to primary vertices and jets.

823 A track is associated to a vertex if its origin is geometrically compatible in z with the vertex
 824 position. The compatibility is determined by the resolution on the track z_0 impact parameter
 825 such that

$$\frac{|z_0 - z_{\text{vertex}}|}{\sigma_{z_0}} < 2.5, \quad (3.1)$$

826 where σ_{z_0} is the per-track resolution on the longitudinal impact parameter and depends
 827 primarily on the track η and p_T .

828 The reliability of the track-to-vertex association depends on the value of σ_{z_0} relative to the
 829 average pile-up density $\langle\rho(z)\rangle$. The average number of interactions within a window of $|z_0 -$
 830 $z_{\text{vertex}}| = 2.5\sigma_{z_0}$ is given by $N = 2\langle\rho(z)\rangle 2.5\sigma_{z_0}$. This means that in order to unambiguously
 831 associate tracks to vertices based on Equation. (3.1), N has to be smaller than 1, or $\sigma_{z_0} <$
 832 $1/5\langle\rho(z)\rangle$. For a track with z_0 at the origin, where the average density peaks, this value is
 833 approximately 130 μm . If σ_{z_0} is larger than this value, the association of tracks to vertices
 834 becomes ambiguous because the same track may be compatible with multiple nearby vertices.
 835 It is important to note that this argument applies to prompt tracks such as those produced in
 836 light-quark and gluon jets, or prompt leptons. The association of displaced tracks from the
 837 decay of B/D hadrons to primary vertices requires the use of a larger z window, which will
 838 lead to greater pile-up contamination.

839 While the longitudinal impact parameter resolution is relatively constant and small for
 840 $|\eta| < 1.5$, it grows rapidly with pseudo-rapidity, reaching several millimetres for $|\eta| \gtrsim 2.5$.
 841 The resolution is further degraded for low p_T tracks due to multiple scattering effects. The η
 842 dependence of the impact parameter resolution is mostly determined by the geometry of the
 843 inner detector. As η increases, tracks become more collinear to the beam line.

844 Based on Fig. 2.6, a 1 GeV track with $|\eta| = 3$ has a z_0 resolution of approximately 1 mm,
 845 leading to a $\pm 2.5\sigma$ window of 5 mm in z for the vertex position. With a most probable
 846 average pile-up vertex density of 1.8 vertices/mm at $z = 0$, this means that, on average, a
 847 forward track can be compatible with up to about 9 near-by vertices on average. Or, in other
 848 words, track-to-vertex association will suffer significantly from pile-up contamination.

849 Track Extrapolation to the HGTD

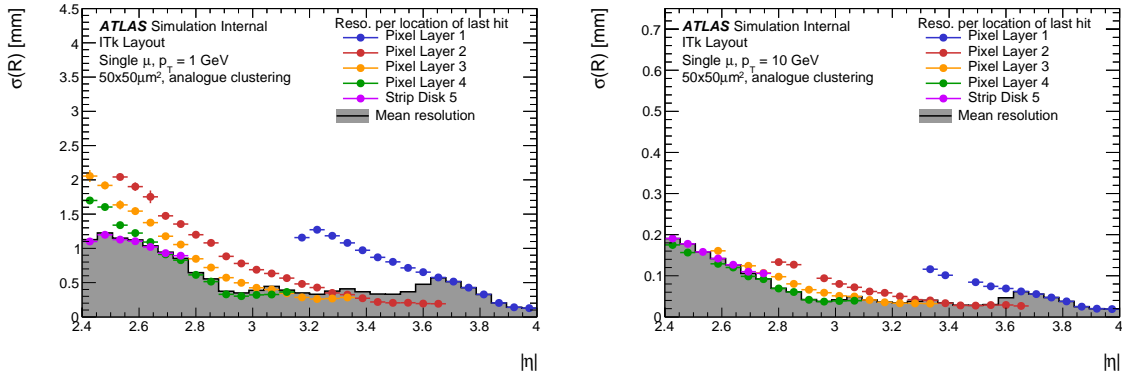
850

851 In order to associate the timing information provided by the HGTD to the correct track, the
 852 track has to be extrapolated from the ITk to the HGTD. The precision of the extrapolation
 853 is affected by the material in the ITk and between the ITk and the HGTD. In Fig. 3.16 the
 854 precision of the extrapolation as a function of η to the HGTD surface is shown for r and
 855 $r \times \phi$. Single muons with transverse momenta of 1 GeV and 10 GeV were analysed. The
 856 extrapolation is performed from the last hit in the ITk associated to the track. For the majority
 857 of tracks $p_T > 1$ GeV, the precision of the extrapolation is better than the pad size used in
 858 the HGTD (1.3 mm \times 1.3 mm).

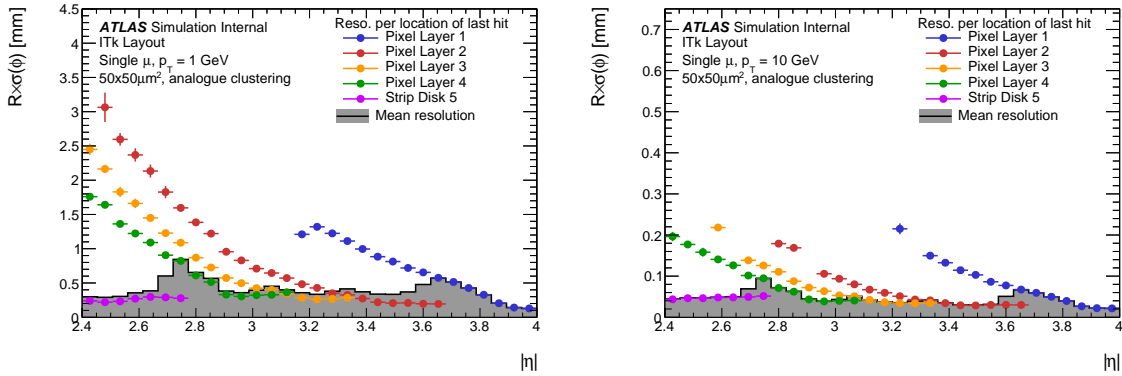
859 Track Extension

860

861 Track candidates reconstructed in the ITk are extended to the HGTD using a progressive
 862 Kalman filter within the HGTD acceptance. The starting point for the extension is the ITk
 863 layer closest to the HGTD within which a measurement has contributed to the track. From
 864 here, the tracks are extrapolated to each layer of the HGTD. In each layer, HGTD clusters
 865 found in a 5 cm \times 5 cm window around the extrapolated crossing location are evaluated for
 866 compatibility with the track by attempting to add them to the track in a forward filtering



(a) Extrapolation resolution in radius for $p_T = 1$ GeV (b) Extrapolation resolution in radius for $p_T = 10$ GeV



(c) Extrapolation resolution in radius $\times \phi$ for $p_T = 1$ GeV (d) Extrapolation resolution in radius $\times \phi$ for $p_T = 10$ GeV

Figure 3.16: The extrapolation resolution in radius r and in the product $r \times \phi$ for tracks with $p_T = 1$ GeV and $p_T = 10$ GeV. The resolution is plotted as a function of η for the extrapolation of the track from the last hit in the ITk. The resolution is similar to the size of a single pad in the HGTD.

867 step. The cluster with the lowest χ^2 resulting from this procedure is considered a a valid
 868 extension of the track into the given layer if the reduced χ^2 value is less than five. If no
 869 cluster satisfying this condition is found, no extension is registered for this layer and the
 870 procedure repeated in the next layer. In case of a successful extension, the updated track
 871 including the newly-associated HGTD cluster is used when extrapolating to further layers
 872 of the HGTD, replacing at each step the track information from the last measurement as a
 873 starting point of the extrapolation.

874 The inclusive extension efficiency is defined as the fraction of tracks that have a valid
 875 extension, considering one or two measurements in the HGTD that were produced by the
 876 truth particle matched to the track as shown in Fig. 3.17(a). In the sample shown, $t\bar{t}$ with
 877 $\langle \mu \rangle = 200$, the dense environment is a challenge to assign the correct HGTD hits to the

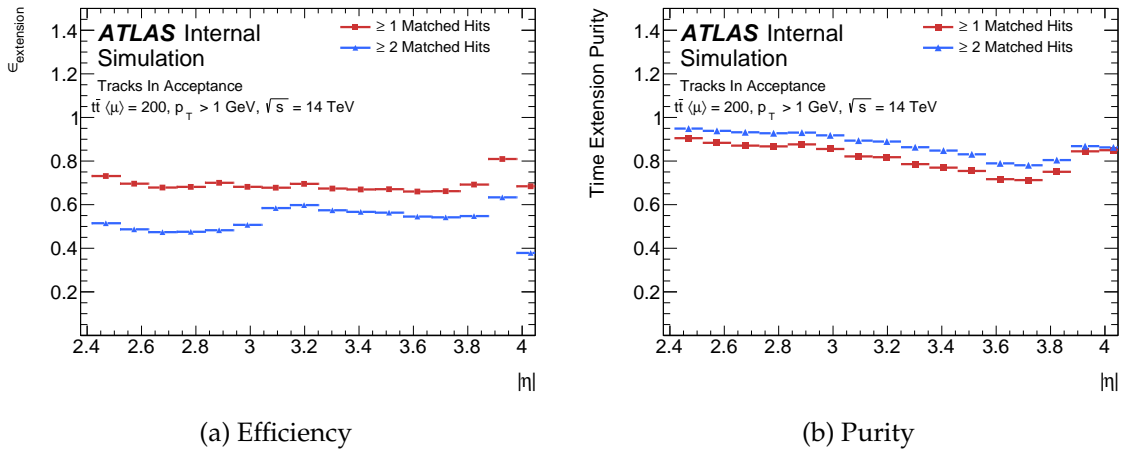


Figure 3.17: The performance of the track extension to the HGTD is shown as function of η for $t\bar{t}$ with $\langle\mu\rangle = 200$.

878 tracks. To estimate the number of wrongly assigned measurements on each extension, the
 879 purity, shown in Fig. 3.17(b), is defined as the fraction of correct assigned measurements
 880 over the total measurements on the track. For each track each hit associated to the track is
 881 classified according to its origin using the Monte Carlo truth link. The purity obtained in $t\bar{t}$
 882 events for $\mu = 200$ is between 80% and 90%.

883 Track Timing Association

884

885 Associating a timing measurement to tracks reconstructed with the ITk using the HGTD hits
 886 was studied using single-pion samples ($\langle\mu\rangle = 0$, generated with a flat distribution in η and
 887 ϕ) and a physics sample with VBF-produced $H \rightarrow Z(\nu\nu)Z(\nu\nu)$ at $\langle\mu\rangle = 200$, both with full
 888 simulation of the HGTD.

889 The association of time information to a track depends on the precision of the track extra-
 890 polation. A track traverses the material in the ITk and the material between the ITk and
 891 the HGTD. Therefore the association is performed by extrapolating the tracks to the HGTD
 892 using the last measured point in the ITk as this method leads to a smaller error on the
 893 extrapolated position in the HGTD with respect to the extrapolation from the perigee. Only
 894 tracks that are reconstructed with a p_T greater the 1 GeV are extrapolated. Furthermore the
 895 extrapolation has to be within the acceptance of the HGTD, i.e., in radius between 120 mm
 896 and 640 mm. The timing scenario *Initial* has been used for the study.

897 The single muon sample at $\langle\mu\rangle = 0$ is used in a first step to check the performance of the
 898 time association as function of the number of hits. In Fig. 3.18(a), the difference between
 899 the reconstructed and the expected time is shown. The distributions are Gaussian with a

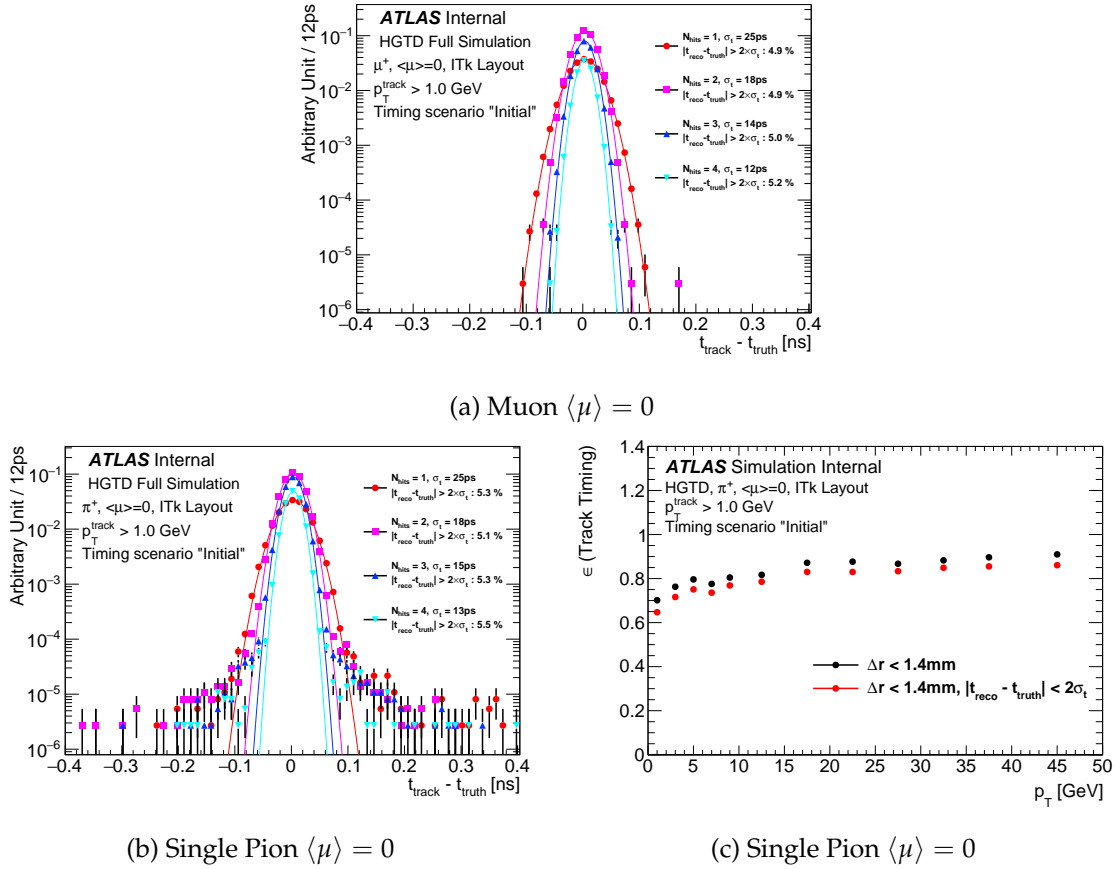


Figure 3.18: The reconstructed timing resolution and efficiency are shown for different samples at $\langle \mu \rangle = 0$.

900 resolution compatible with the simulated resolution, shown in Fig. 3.6(a), divided by the
 901 square root of the number of associated hits.

902 In the analysis of the single pion sample, only primary tracks are accepted by using the
 903 associated truth particle. If several hits are found in the pads that are within 1.4 mm of the
 904 track extrapolation, only the closest one is used. Fig. 3.18(b) shows the difference of the
 905 reconstructed track time and the expected time. The Gaussian core follows the expected
 906 scaling with the number of hits, as for the muons. The tails are increased with respect to
 907 the muons and the distribution is slightly asymmetric as late secondaries (see Fig. 3.13)
 908 contribute to the reconstructed time.

909 The time association efficiency is shown Fig. 3.18(c) as a function of the reconstructed
 910 transverse momentum. Tracks with $|\eta| < 2.4$ or $|\eta| > 4.0$ are not considered in order to
 911 avoid border effects. If, instead of selecting only primary tracks the reconstructed track with
 912 the highest p_T is used, the efficiency to reconstruct the correct timing for the high p_T pions is

913 unchanged.

914 At the highest transverse momentum, the efficiency is 90%. At the lowest momentum,
 915 the efficiency is 70%. The window size of the time association of 1.4 mm is in the region
 916 where the efficiency increases only slowly for high p_T . For low p_T , the same efficiency
 917 can be reached only for a window of 30 mm. The multiple scattering effect in the material
 918 between the last measured point in the ITk and the HGTD is more important for low p_T . A
 919 window size 15 times larger (20 mm) would be necessary to reach an efficiency of 97% for
 920 high p_T pions. For low p_T pions, the same efficiency is attainable only for a window of size
 921 200 mm.

922 While the geometrical acceptance is optimised to compensate for inactive zones in one layer
 923 with an active zone in another layer, the difference between track extrapolation and hit in
 924 different layers is strongly correlated. As an example, the correlation of the first layer with
 925 the last layer for the difference between hit and extrapolated position is more than 70%. The
 926 second set of points in Fig. 3.18(c) is the efficiency to associate a time within $2\sigma_{t_{\text{reco}}}$ of the
 927 expected truth time. The ratio between the two efficiencies is about 0.95, independent of η
 928 and p_T .

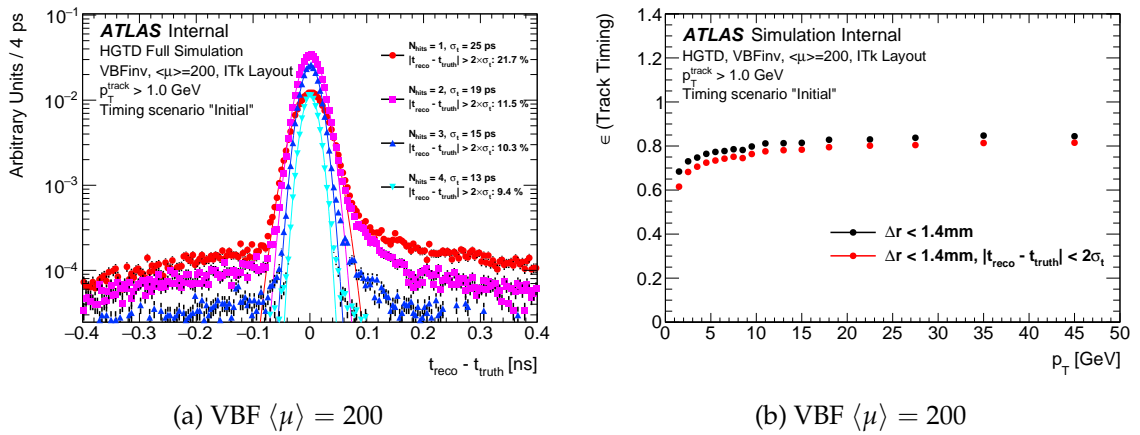


Figure 3.19: The reconstructed time resolution and efficiencies for associating time to tracks is shown for pions in a VBF sample.

929 With a pile-up of $\langle\mu\rangle = 200$, the analysis of the track timing association is complicated by
 930 the presence of pileup. Time measurements in the HGTD have to be associated to tracks
 931 from the hard scatter while keeping the association of pile-up hits to a minimum. Therefore
 932 additional criteria are applied in the time association algorithm. If only one hit is found in
 933 the acceptance window, the time of the hit is kept as track time. If multiple hit candidates
 934 are identified, the hits are additionally required to be compatible in time with each other, i.e.,
 935 within 2.5σ of the hit timing resolution around the time of the hit in the layer closest to the
 936 interaction point.

937 The least squares method is used to choose from all possible combination candidates.

$$S = \sum_i (\vec{x}_{\text{hit}_i} - \vec{x}_{\text{extrapolation}_i})^2 \quad (3.2)$$

938 \vec{x}_{hit_i} is the hit candidate in the layer i , $\vec{x}_{\text{extrapolation}_i}$ is the position given by the track extrapolation in the same layer. The hit combination with the lowest value of S is chosen, and the
 939 mean time of the collected hits is assigned to the track.
 940

941 The result of the algorithm is shown in Fig. 3.19(a). The difference between the reconstructed
 942 time and the expected time is well-described by a Gaussian core. The resolution improves
 943 as in the $\langle \mu \rangle = 0$ case with the square root of the number of hits. The tails are slightly
 944 asymmetric due to secondaries contributing to the reconstructed time. Since the algorithm
 945 can pickup up hits which are due to pileup, the tails increase. These tails decrease as function
 946 of the number of hits as the additional requirements on the timing consistency of the chosen
 947 hits are stronger than, e.g., in the case of one hit where no constraints can be applied.

948 The efficiency for correctly assigning a time to a track in a high-pile-up environment is
 949 shown in Fig. 3.19(b) as a function of track p_T . The contribution of tracks with only one hit
 950 associated in the HGTD to the total efficiency is about 18%. The dependence of the efficiency
 951 with p_T is similar to that obtained with single-pion samples, reaching a plateau of 83%. The
 952 efficiency increase due to pileup hits being associated with the track is compensated by the
 953 additional association criteria. In total the efficiency is lower. Requiring a reconstructed time
 954 compatible with the expected time within two standard deviations reduces the efficiency by
 955 about 5%.

956 This study is a first step in the development of the association of the track information
 957 with timing information in the HGTD. The main inefficiency, of not finding a hit within
 958 the window, has been identified as interactions in the material in front of the HGTD. More
 959 sophisticated pattern recognition algorithms are being studied to improve this efficiency.

960 3.2.3 Determination of the time of the primary vertex

961 Once time measurements are associated to reconstructed tracks a time can be associated to
 962 the reconstructed vertices. For a precise knowledge of the time of the hard scatter vertex it is
 963 essential to remove tracks that are out of time with respect to the vertex.

964 First a study was performed only using the tracking information. For each reconstructed ver-
 965 tex, tracks with associated times differing by less than 50 ps are clustered together iteratively.
 966 If several clusters are reconstructed for a vertex, the one with the highest $\sum p_T^2$ is assigned to
 967 the vertex. Fig. 3.20 shows the reconstructed vertex time as function of the true vertex time.
 968 Spurious measurements off-diagonal are due to splitting and merging effects and may be
 969 improved by including the time measurements in the vertex finding and fitting procedure.

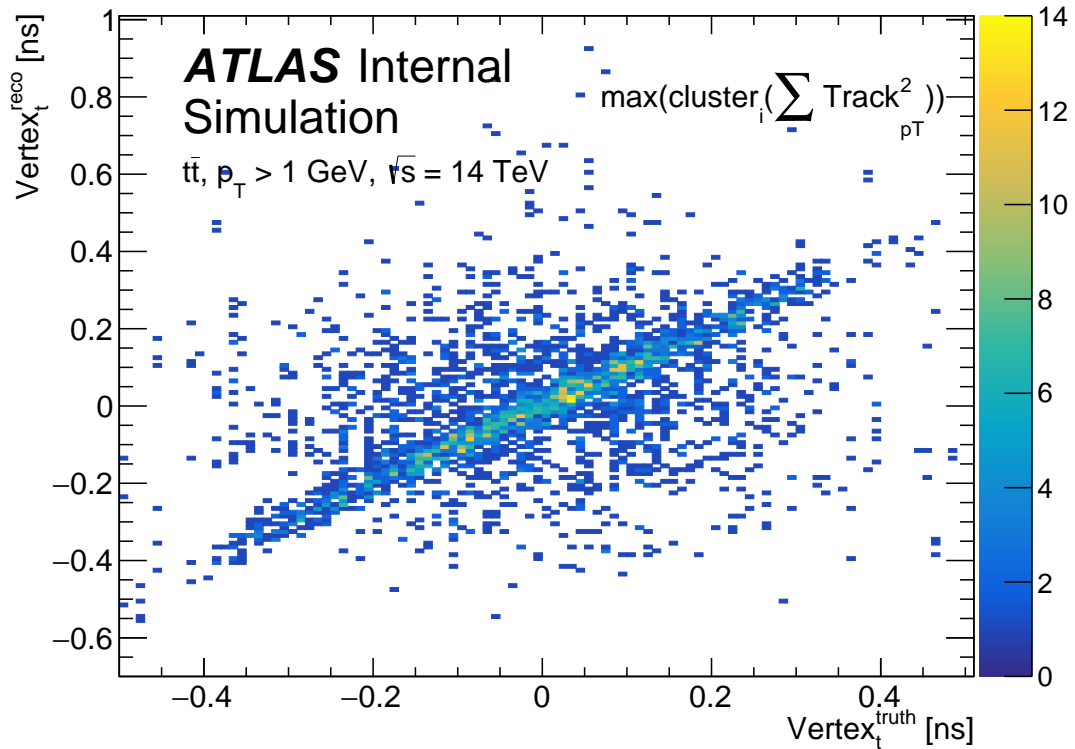


Figure 3.20: The reconstructed vertex time is shown as function of the true vertex time.

970 Next the time vertex efficiency was studied for tracks associated to a jet in order to determine
 971 the time and the resolution of the hard scatter vertex. The restriction of the tracks to these
 972 tracks from all possible tracks improves the determination of the primary t_0 . Studies have
 973 shown that the pile-up track contamination of hard scatter truth matched jets represent
 974 on average, depending on η , less than 30% of the tracks associated to a jet. The track-
 975 time clustering takes advantage of this, searching for the largest cluster in the set of times
 976 associated to the tracks in the jet in order to reject the pileup.

977 With this method, each associated track of the jet is assigned a time, unless no hits fulfilling
 978 the requirements are found. Further, the individual track-times are clustered iteratively,
 979 starting with those closest to each other in units of their resolution. The resolution used in
 980 the algorithm was determined from the Gaussian core of the track time resolution studies.
 981 These times are merged, and a resolution weighted time average calculated as well the
 982 resolution of this merged time.

$$t_{\text{merged}} = \frac{\sigma_2^2 t_1 + \sigma_1^2 t_2}{\sigma_1^2 + \sigma_2^2} \quad (3.3)$$

$$\sigma_{\text{merged}} = \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (3.4)$$

983 More time values are added to this average, repeating the update of time and resolution,
 984 until the next value is outside three times the uncertainty σ_{merged} .

985 To define a time for the jet, the biggest time cluster, i.e., the t_{merged} with the largest number
 986 of t_i , is chosen. The only additional condition on this cluster is to have at least two entries,
 987 i.e., at least two tracks. If there is a second cluster with the same number of entries, none is
 988 selected and the time for this jet is not defined.

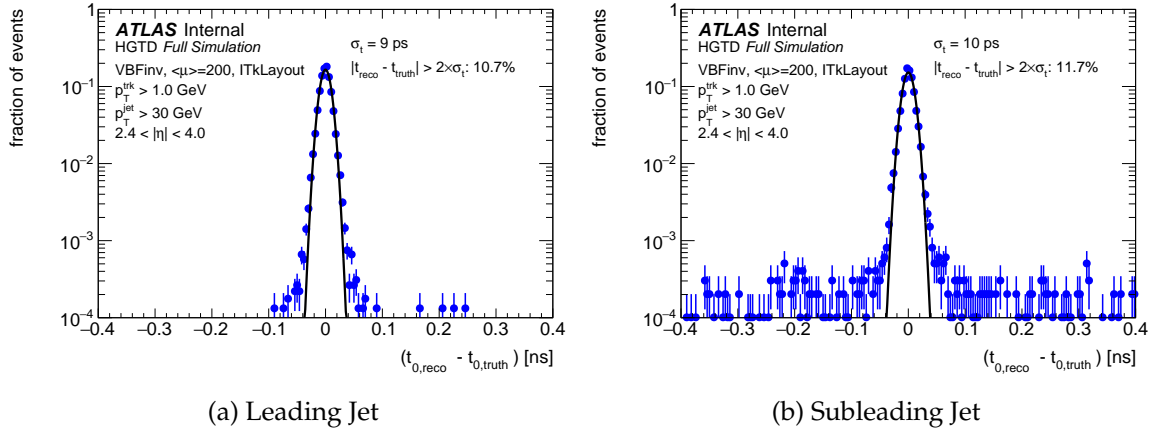


Figure 3.21: The precision of the t_0 determination using the HGTD is shown.

989 To assign a time to the hard scatter event, t_0 , usage of the time of the leading or sub-leading
 990 jet was studied on a sample of vector boson fusion Higgs production, followed by an
 991 invisible decay of the Higgs boson for $\langle \mu \rangle = 200$. In order to separate the performance of
 992 the algorithm from the effect of vertexing, the truth primary vertex is required to be within
 993 0.5 mm of the reconstructed primary vertex, reducing the sample by 37%. The leading or
 994 subleading jet is required to fall into the acceptance of the HGTD which is the case in 59% of
 995 the remaining events. Applying the algorithm on the leading jet leads to an efficiency of the
 996 t_0 determination of 85%.

997 The resolution of the algorithm is shown in Fig. 3.21(a) for the leading jet. The Gaussian core
 998 of the distribution has a width of 9 ps. Only 11% of the reconstructed times are outside a 2σ
 999 window of the central value.

1000 The subleading jet can be used for t_0 determination in case the leading one is in the central
 1001 region. The algorithm applied to subleading jet leads to an efficiency of 60%, where 12% of

1002 the reconstructed times are outside a 2σ window of the central value as shown in Fig. 3.21(b).
 1003 The precision of the timing reconstruction of 11 ps is similar to the precision using the leading
 1004 jet.

1005 In total, excluding requirement to reconstruct the correct vertex in space within 0.5 mm, in
 1006 43% of the events a t_0 can be reconstructed. Requiring this reconstructed time to be within
 1007 2σ of the expected vertex time reduces the number of 38%.

1008 3.2.4 Suppression of pile-up jets

1009 Pile-up is one of the most difficult challenges for object identification under HL-LHC condi-
 1010 tions. Particles produced in pile-up interactions can contaminate the jets of interest coming
 1011 from the hard-scatter vertex, thereby reducing the accuracy of the jet energy determination.
 1012 Pile-up interactions can also produce additional jets which do not originate from the primary
 1013 hard-scatter interaction. These pile-up jets can be produced as the result of a hard QCD
 1014 process from a pile-up vertex, or by random combinations of particles from multiple vertices.
 1015 At low jet p_T , the latter mechanism is dominant, whereas at high jet p_T , the majority of
 1016 pile-up jets are QCD jets.

1017 Pile-up jets can reduce the precision of Standard Model measurements and the sensitivity to
 1018 discover new physics. For example, additional jets can increase the amount of background
 1019 events passing a selection, as well as reduce the efficacy of kinematic variables or discrimin-
 1020 ants to separate signals from backgrounds. Hence, the efficient identification and rejection
 1021 of pile-up jets is essential to enhance the physics potential of the HL-LHC.

The key element to suppress pile-up in jets is the accurate association of jets with tracks and primary vertices. A simple but powerful discriminant is the R_{p_T} jet variable, defined as the scalar sum of the p_T of all tracks that are inside the jet cone and originate from the hard-scatter vertex PV_0 , divided by the fully calibrated jet p_T , i.e.

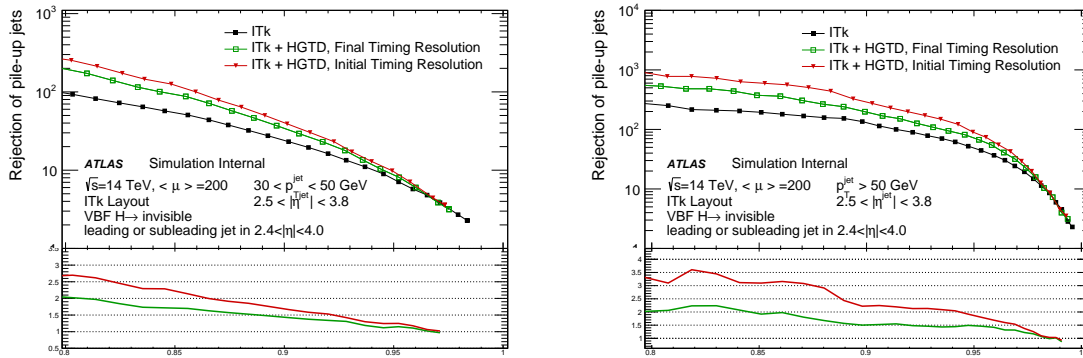
$$R_{p_T} = \frac{\sum p_T^{\text{trk}}(PV_0)}{p_T^{\text{jet}}}.$$

1022 The tracks used to calculate R_{p_T} fulfil the quality requirements defined in Ref. [11] and are
 1023 required to have $p_T > 1$ GeV. The tracks used in the R_{p_T} calculation are required to satisfy
 1024 Eq. (3.1).

1025 Hard-scatter and pile-up jets for simulated events are defined by their matching to truth jets,
 1026 which are reconstructed from stable and interacting final state particles coming from the
 1027 hard interaction. The matching criteria are defined in Ref. [12]. Reconstructed hard-scatter
 1028 jets are required to be within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ of a truth jet with $p_T > 10$ GeV.
 1029 The pile-up jets must be at least $\Delta R > 0.6$ away from any truth jet with $p_T > 4$ GeV. The

1030 performance has been studied using a mixture of full reconstruction (for tracks and jets) and
 1031 fast simulation (for the HGTD).

1032 At moderate levels of pile-up, where track impact parameter measurements can be used to
 1033 assign tracks to vertices with relatively little ambiguity, small values of R_{p_T} correspond to jets
 1034 which have a small fraction of charged-particle p_T originating from the hard-scatter vertex
 1035 PV_0 . These jets are therefore likely to be pile-up jets. However, at high pile-up conditions,
 1036 and particularly in the forward region, the power of this discriminant is reduced. The effect
 1037 can be mitigated by including timing information from the HGTD, removing tracks outside
 1038 a $2\sigma_t$ window around the time of the hard-scatter vertex. The main impact of the HGTD in
 1039 this study is to remove stochastic pileup jets.



(a) Jets with $30 \text{ GeV} < p_T < 50 \text{ GeV}$.

(b) Jets with $p_T > 50 \text{ GeV}$.

Figure 3.22: Pile-up jet rejection as a function of hard-scatter jet efficiency in the $2.4 < |\eta| < 4.0$ region, for the ITk-only and combined ITk + HGTD scenarios with different time resolutions.

1040 Samples of VBF Higgs production with invisible Higgs decays with $\langle \mu \rangle = 200$ were used in
 1041 this study. The events were required to have the leading or subleading jet reconstructed in
 1042 the HGTD acceptance. This ensures consistency with the algorithm to determine the time of
 1043 the primary vertex.

1044 Fig. 3.22 shows the rejection, i.e., the inverse of the mis-tag efficiency, of pile-up jets as a
 1045 function of the efficiency for selecting hard-scatter jets using the R_{p_T} discriminant for jets
 1046 with low and high p_T in dijet events with $\langle \mu \rangle = 200$ without and with the HGTD for the
 1047 different timing resolution scenarios. A significant improvement in performance of a factor
 1048 of 1.9 and 1.5 higher pile-up rejection for jets at an efficiency of 88% is achieved with the use
 1049 of timing information in the *Initial* and *Final* timing scenarios. For a fixed pileup rejection of
 1050 50 the absolute improvement of the hard scatter efficiency is 3% for the *Initial* and 2% for the
 1051 *Final* scenario. Both the time association efficiency and the t_0 determination precision were
 1052 taken into account with smearing functions.

1053 **3.2.5 Missing transverse momentum**

1054 The missing transverse momentum, denoted E_T^{miss} , is computed as the negative vector
 1055 momentum sum of high- p_T physics objects in the event, plus a soft-term component, from
 1056 particles which do not constitute high- p_T objects. The soft term is calculated using tracks
 1057 associated to the hard-scatter vertex [13].

1058 In this study, E_T^{miss} is reconstructed using selected muons, with $p_T > 2.5 \text{ GeV}$ or $p > 4 \text{ GeV}$,
 1059 and electrons with $p_T > 10 \text{ GeV}$. Tracks are required to satisfy the criteria described in
 1060 Sec. 3.2.4. Jets with $p_T > 20 \text{ GeV}$ are selected with a pileup jet rejection of 50 using the R_{p_T}
 1061 as discriminant discussed in the previous section. The association of tracks to the primary
 1062 vertex, optimised for the jet selection, has been done according to Eq. (3.1). Selected tracks
 1063 originating from the hard-scatter vertex, with $p_T^{\text{trk}} > 1 \text{ GeV}$ and $|\eta^{\text{trk}}| < 4$, not associated
 1064 to muons, electrons and jets, are used to reconstruct the soft-term component of the E_T^{miss} .
 1065 The HGTD timing information is used for the soft-term by requiring the track time to be
 1066 compatible with the reconstructed vertex time within two standard deviations. The distance
 1067 of the reconstructed vertex to the truth vertex is required to be less than 0.1 mm. 57% of
 1068 the events remain after this requirement. Events where a jet overlaps with an electron are
 1069 rejected.

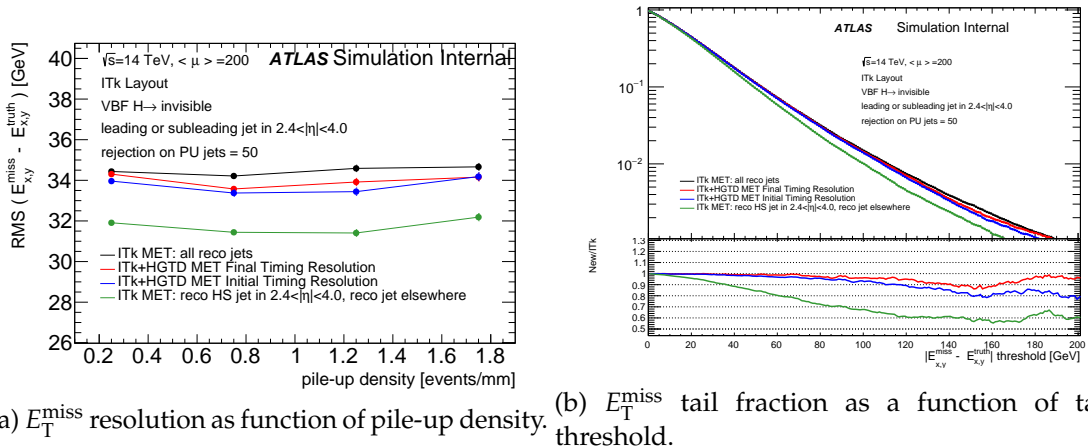


Figure 3.23: Three scenarios are shown for the E_T^{miss} : ITk scenario, ITk and HGTD as well as hard scatter truth jets in the HGTD acceptance. For 3.23(b) the ratio of the two curves is shown in the lower panel.

1070 Fig. 3.23(a) shows the resolution of the x - and y -components of the E_T^{miss} as a function of
 1071 the pile-up density in the event for samples of VBF Higgs production with invisible Higgs
 1072 decays for $\langle \mu \rangle = 200$. The results using the ITk alone are only improved slightly using
 1073 the timing capabilities of the HGTD in spite of the gain shown in Sec. 3.2.4. While the jet
 1074 performance gain is defined per object, E_T^{miss} is an event level quantity. For a fixed pileup

1075 rejection of 50, the hard scatter jet efficiency improves by about 3% for the timing scenario
 1076 *Final*. In VBF Higgs production 57% of the events, normalized relative to the requirement
 1077 on the maximal distance between the reconstructed and true vertex, have at least one hard
 1078 scatter jet reconstructed in the HGTD acceptance. Therefore only 1 event out of ≈ 60 will
 1079 have a jet configuration that is modified. Therefore the improvement on the jet performance
 1080 is diluted in the E_T^{miss} analysis.

1081 In many analysis, especially new physics searches, it is crucial to minimize the E_T^{miss} tails. In
 1082 Fig. 3.23(b) the fraction of events above a $E_T^{\text{miss}} - E_T^{\text{miss,truth}}$ threshold is shown as a function of
 1083 the $E_T^{\text{miss}} - E_T^{\text{miss,truth}}$ threshold for a fixed pile-up jet rejection of 50 in VBF Higgs production
 1084 samples with $\langle \mu \rangle = 200$. A lower fraction of events passing high $E_T^{\text{miss}} - E_T^{\text{miss,truth}}$ thresholds
 1085 corresponds to a better E_T^{miss} reconstruction. The ideal scenario using the generated quantities
 1086 instead of the reconstructed ones in the acceptance of the HGTD shows the maximal potential
 1087 for improvement. When adding the HGTD, for a threshold of 160 GeV, the improvement is
 1088 10% for the timing scenario *Final* and 20% for *Initial*. The decrease of the improvement for
 1089 high values of the threshold is a statistical fluctuation.

1090 3.2.6 Tagging of heavy flavour jets

1091 The efficient identification of b -jets and high rejection of light-quark jets is of central import-
 1092 ance in the HL-LHC physics program. Tagging b -jets is particularly sensitive to pile-up-track
 1093 contamination. This is due to the fact that b -tagging algorithms consider tracks with large
 1094 impact parameters (in both the transverse and longitudinal directions) from the decay of
 1095 displaced vertices. With a larger z_0 window, tracks from nearby pile-up interactions are
 1096 more likely to be selected, leading to an increased rate of misidentified light-quark jets.

1097 Using simulated $t\bar{t}$ events at $\langle \mu \rangle = 200$, the impact of the HGTD on the performance of
 1098 b -tagging algorithms is studied for forward jets ($|\eta| > 2.4$). Fig. 3.24 shows the light-jet
 1099 rejection versus b -tagging efficiency for the IP3D+SV1 b -tagging algorithm. The addition
 1100 of the HGTD removes pile-up tracks from the track selection. As a result, the performance
 1101 of the b -tagger is significantly improved. For a b -tagging efficiency of 70% and 85%, the
 1102 corresponding light-jet rejection for MV1 is increased by factors between 1.3 and 1.4. These
 1103 factors could be greater for processes where more b -jets are expected in the forward region.
 1104 The performance is shown for the ITk-only scenario as well as three scenarios with HGTD
 1105 timing performance representing different stages of the HL-LHC program. It can be seen
 1106 that all timing scenarios yield significant improvements in the performance, even in the *Final*
 1107 scenario. Importantly, significant improvements are observed also after the full radiation
 1108 damage expected during HL-LHC operation.

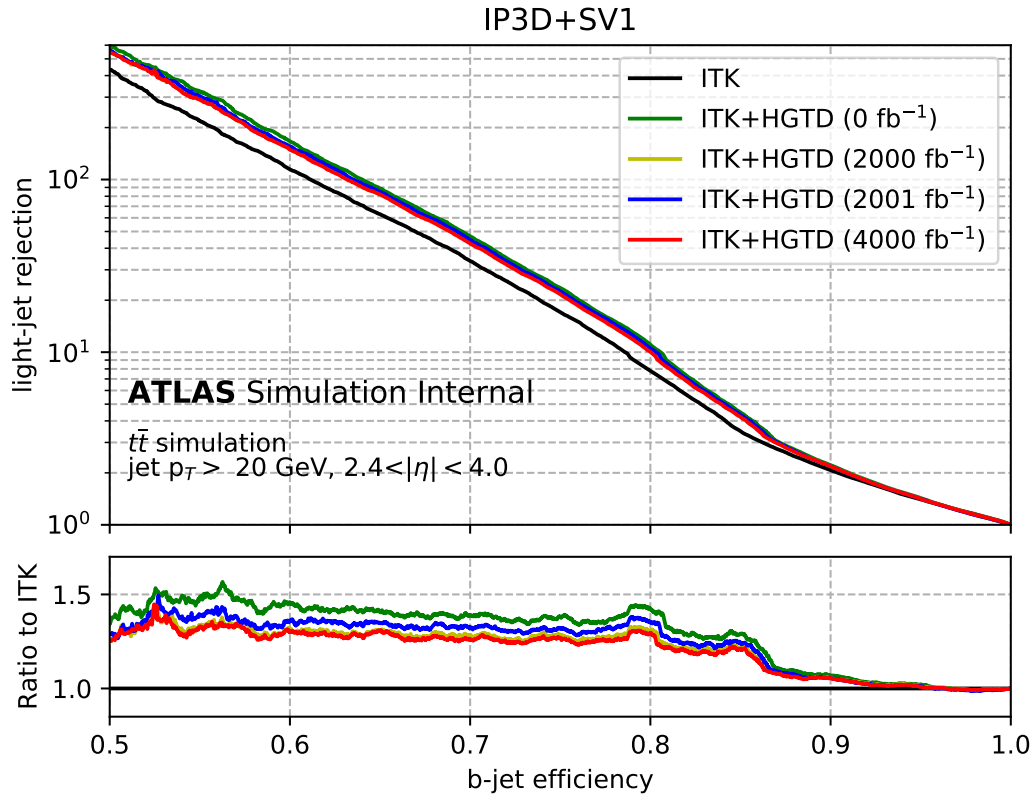


Figure 3.24: Light-jet rejection versus b -tagging efficiency for the IP3D+SV1 tagger and The study uses $t\bar{t}$ events at $\langle \mu \rangle = 200$ and shows the achieved performance for different time resolution scenarios. The ratio plots at the bottom show the relative performance achieved with the HGTD with respect to the ITk-only scenario.

1109 3.2.7 Lepton isolation

1110 In this section, studies of how the electron isolation efficiency is improved with HGTD are
 1111 presented based on the full simulation of the HGTD. The HGTD can be used to assign a time
 1112 to leptons in the forward region. This information can be exploited to reject tracks which
 1113 come from other interactions but are spatially close to the energy deposits in the calorimeter
 1114 and/or the track associated to the lepton. The timing information can reject additional tracks
 1115 from interactions close in z , according to Eq. (3.1) but separated in time from the hard-scatter
 1116 vertex. The isolation efficiency is defined as the probability that no track with $p_T > 1$ GeV is
 1117 reconstructed within $\Delta R < 0.2$ of the electron track.

1118 The physics process used in this study is Z boson production followed by a decay either
 1119 to electrons or tau leptons. Only electronic decays of the tau lepton were used. Forward

1120 electrons with a $p_T > 20$ GeV passing the standard ATLAS medium cuts were selected,
 1121 keeping only those matched with a truth electron with a $p_T > 20$ GeV in a cone of $\Delta R < 0.2$.
 1122 In order to study a reasonable p_T range the cut was reduced to 10 GeV for the tau sample. The
 1123 electron track is defined as the track closest to the electron cluster having a ratio of transverse
 1124 track momentum to transverse cluster energy greater than 0.1. The correct electron track
 1125 is selected in 95% of the cases. The tracks are extrapolated to the HGTD surfaces using
 1126 the last measured point in the ITk. The closest HGTD hit in a window of 3.5 mm between
 1127 the extrapolated position and the hit position is associated to the track. The window is
 1128 larger for electrons than for pions and muons to account for Bremsstrahlung in the material
 1129 between the last ITk measurement and the first sensitive layer of the HGTD. A time is
 1130 then reconstructed in the HGTD for the electron track as well as for all other tracks with
 1131 $p_T > 1$ GeV which are within $\Delta R < 0.2$ of the electron track. These time are compared with
 1132 the time of the electron track. If the time difference between the two is larger than twice the
 1133 quadratic sum of the timing resolution of both tracks the track is discarded.

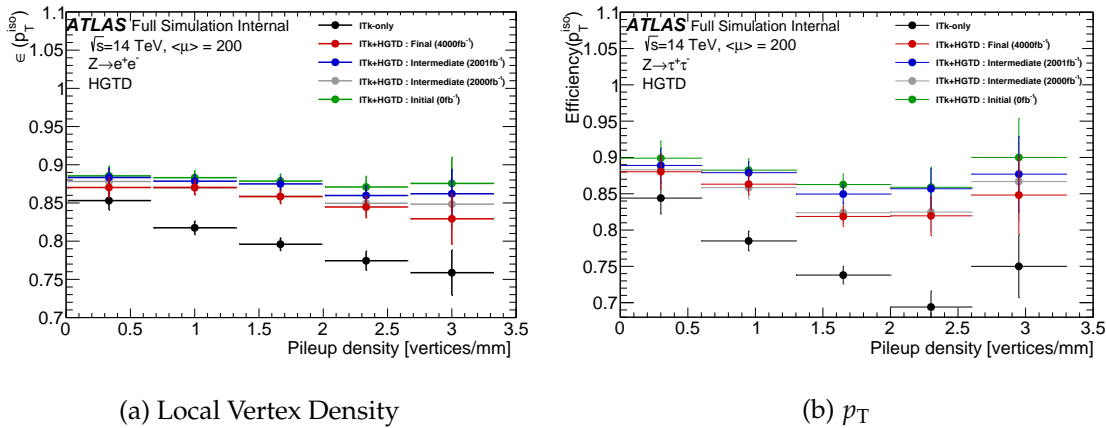


Figure 3.25: The selection efficiency is shown for the electron isolation criteria using the ITk and ITk + HGTD for different timing resolutions.

1134 The isolation efficiency is shown in Fig. 3.25 for the ITk-only scenario and four HGTD
 1135 timing resolution scenarios. For Fig. 3.25(a) the samples with the electronic decay of the
 1136 Z boson was used. While the efficiency drops strongly with the increase of the pile-up
 1137 density when using only the ITk, the addition of the HGTD timing information reduces this
 1138 drop, keeping an efficiency above 85% even at high pile-up density, i.e. with up to three
 1139 additional vertices around the hard-scatter vertex. For a local pile-up density of the order
 1140 of 1.6 vertices/mm the electron isolation efficiency is improved by about 9% for the *Initial*
 1141 and by about 6% for the *Final* scenario. Even in the *Final* timing scenario, the resolution is
 1142 sufficient to achieve an isolation efficiency essentially independent of the pile-up density at
 1143 the end of the HL-LHC.

1144 As an example for a local pileup density of 1.6 vertices/mm the inefficiency of the lepton
 1145 isolation is due to either tracks of hadrons or tracks originating from electromagnetic interac-

1146 tions of the electron (shower). In the ITk only scenario 15% of the electrons are classified as
1147 non-isolated due to tracks truth matched to a hadron and 6% due to tracks truth matched to
1148 a lepton. The HGTD timing information reduces the hadronic inefficiency by 53% (relative)
1149 by removing the pileup track. For the leptonic inefficiency the relative reduction is only 9%
1150 since the extra tracks originate from interactions of the primary electron. They are therefore
1151 in time with the signal electron.

1152 In Fig. 3.25(b) the p_T dependence of the isolation efficiency is shown. As the direct decay
1153 of the Z boson to electrons leads to a peak at 45 GeV, the use of the decay through a tau
1154 lepton allows to probe more easily a broader range of p_T . The improvement is about 10-15%
1155 integrated over the local vertex density, essentially independent of η . The isolation efficiency
1156 as function of the local vertex density for these events is about the same as the direct decay
1157 isolation efficiency.

1158 The timing resolution of the electron is 16 ps for the *Initial* timing scenario with only 9% of the
1159 electron candidates outside a 2σ window centered on the true time. For the timing scenario
1160 *Final* the performance is 42 ps due to the effect of irradiation on the timing resolution on the
1161 sensor and electronics also with 8% outside the 2σ window. Less than 1% of the electron
1162 tracks in the 2 sigma window have only pileup hits associated to them. The performance in
1163 the forward region reaches a level similar to that in the central region. The improvement of
1164 the performance depends only modestly on the timing scenario.

1165 3.2.8 Non-collision background

1166 The ATLAS detector signals from the products of proton-proton collisions, as well as from
1167 non-collision backgrounds (NCB). These consist of beam halo and beam-gas backgrounds,
1168 cosmic rays and detector noise. Beam halo, elastic and inelastic beam-gas are the main
1169 processes creating non-collision backgrounds reaching the detector. Beam halo is the small
1170 fraction of particles surrounding the dense beam core. The collimator system of the LHC
1171 is designed to stop off-momentum and off-position particles. However leakage from the
1172 collimation insertions and from the tertiary collimators near ATLAS allow a fraction of
1173 beam halo background to enter the detector. Beam-gas background is related to the pressure
1174 in the beam pipe; protons colliding with the residual gas molecules in the vacuum create
1175 background particles. NCB events can be created in the vicinity of ATLAS or travel in the
1176 beam pipe or parallel to the beam line over large distances.

1177 As a result of the unavoidable nature of these backgrounds, their implications on physics
1178 studies and detector occupancy of the non-collision background must be understood. This
1179 study investigates the beam halo background in the HGTD. Events are selected from the
1180 FLUKA beam halo simulation [14] (using HL-LHC optics version 1.0) after interacting on
1181 the TCTs located approximately 150 m upstream of ATLAS [15]. The timing analysis of the
1182 HGTD hits provides information on NCB events as well as nominal collision events.

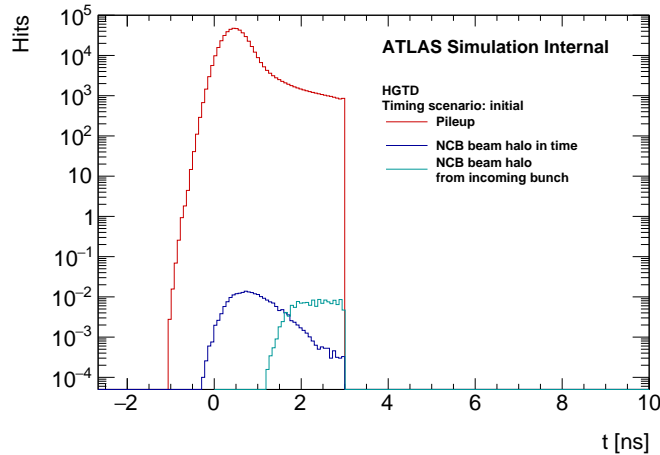


Figure 3.26: The time distribution, normalized per bunch, is shown for minimum bias events with $\langle \mu \rangle = 200$ (red), the background halo samples of the colliding bunch (blue) and the incoming next colliding bunch (cyan). The distribution is cut at 3 ns to mimic the timing window of the TDC in the electronics.

1183 The time distributions for different sources of detector hits are shown in Fig. 3.26. The time t
 1184 is offset by 11.6 ns with respect to the nominal collision time as explained in Sec. 3.1.2. As
 1185 a way to mimic the timing window of the readout electronics TDC, a cut on events with
 1186 $t > 3$ ns is applied. The arrival time of the signal correlated with the nominal collision in the
 1187 HGTD is at approximately 0.405 ns. That signal is coincident with the in-time non-collision
 1188 background of the HGTD. The NCB events associated with the next bunch coming from the
 1189 direction opposite to the interaction point reach the HGTD, on average, 1.6 ns after the time
 1190 of arrival of nominal collision hits.

1191 Only signals in the HGTD for negative z are shown. As the NCB is symmetrical, the
 1192 timing structure will be similar for positive z . Non-collision background and especially
 1193 beam halo events are expected to increase during the HL-LHC operation compared to
 1194 Run-2 [16], nevertheless the NCB events in this study are normalised according to Run-2
 1195 BCM at 1 Hz, 10^{11} p. The NCB events are different from pile-up not only by time of arrival
 1196 but other parameters as well, for example the NCB hits increase with radius, from 0 to
 1197 320 but outofdate where the overlap of the sensors changes and the hit count drops and
 1198 then increases again up to the maximum radius of the detector, while pile-up decreases as
 1199 function of the radius.

1200 The timing structure of the NCB is not identical to the pile-up time distribution. The impact
 1201 of the HGTD on the study of NCB will depend on the magnitude of its expected increase.
 1202 The HGTD will most likely only be able to contribute to the study of NCB in the startup of
 1203 the HL-LHC before its nominal luminosity is reached.

1204 3.3 Physics

1205 *Results on VBF and tH production are in preparation and are foreseen to be part of the*
1206 *second ATLAS circulation.*

1207 3.3.1 Strategy for the Application of Object Performance to Physics Analyses

1208 The improvements in the performance of forward leptons, b -tagging, jets, E_T^{miss} , as well as
1209 the determination of the vertex time translate into improvements in the physics potential of
1210 ATLAS.

1211 The uncertainty on the integrated luminosity will be one of the largest, and in many cases
1212 dominant, uncertainties in Higgs physics and many other precision measurements at the
1213 HL-LHC program. One of the most direct ways in which HGTD can significantly enhance
1214 the ATLAS physics program at the LHC is by providing an additional measurement of the
1215 luminosity to reduce the uncertainty on the integrated luminosity. Sec. 3.3.2 discusses some
1216 concrete examples in the context of Higgs cross section measurements.

1217 HGTD can also be utilized to trigger on highly ionizing particles in the forward region,
1218 enhancing the discovery reach to new particles such as magnetic monopoles. This is the
1219 subject of Sec. 3.3.4.

1220 A key application of improved forward lepton isolation is the measurement of the weak
1221 mixing angle, discussed in Sec. 3.3.3. In the case of lepton isolation, the lepton track provides
1222 the reference t_0 of the event, so there is no need for a global t_0 reconstruction to take
1223 advantage of the track time measurements to suppress pile-up tracks in the isolation cone as
1224 described in Sec. 3.2.

1225 Improved b -tagging in the forward region can be exploited in physics analysis with forward
1226 b -quarks. b -tagging does not have a large reliance on the t_0 determination with the exception
1227 of the case of single-track jets. In this case, a single time measurement does not provide a
1228 useful handle as the the jet vertex time algorithm requires at least two tracks to be clustered
1229 in time. Relaxing the requirement to one track would not improve the situation as the t_0
1230 would be identical to the track time, so the track will pass the cuts. To account for this case,
1231 we conservatively assume that HGTD cannot be used for single track jets.

1232 In tH production the final state will consist of mostly one single forward b -jet. If additionally
1233 this jet has only one track associated to it, a global event t_0 is needed. A global t_0 reconstruc-
1234 tion algorithm is difficult because of the lack of an additional object for the t_0 determination.
1235 Ideally, the hard-scatter vertex time t_0 would be determined based on tracks not used in
1236 the target jet. Such algorithm has not yet been developed, and will be investigated in the
1237 future.

1238 Additionally, it is possible to consider a different approach for the use of timing information
 1239 in the context of b -tagging. The idea is similar in concept to the case of lepton isolation. In
 1240 the self-tagging method, tracks within the b -jet are first split into sub-jets according to their
 1241 times. For example, a candidate b -jet containing two tracks with times t_1 and t_2 (assuming
 1242 both tracks were assigned the correct time information) may be considered as a single 2-track
 1243 jet if both times are compatible with each other, or two one-track subjets otherwise. Note
 1244 that the splitting of the jet into subjets is done without any knowledge about the global
 1245 vertex t_0 . It is only based on local (relative) time information. In a second step, the b -tagging
 1246 algorithm is applied to each sub-jet separately and the new b -tag weight will be the largest
 1247 weight from each of the subjets. For the simplest example of a mistag jet consisting of
 1248 a one hard-scatter and one pile-up track (associated to the same spatial vertex) in which
 1249 the two tracks form a fake secondary vertex, the self-tagging approach would result into
 1250 two one-track jets with no secondary vertex information. The development and study of
 1251 the self-tagging approach for b -tagging will be developed as a future next step. Such an
 1252 approach could yield similar results as of having the knowledge of the hard-scatter t_0 . The
 1253 b -tagging performance obtained with truth-based t_0 determination will be used to estimate
 1254 the gains in physics sensitivity.

1255 Improvements in the suppression of pile-up jets and E_T^{miss} can have a direct application
 1256 in VBF/VBS physics analyses. However, this is a case where the knowledge of the global
 1257 vertex t_0 is more important than for all previous objects considered and more care is needed
 1258 to be able to translate these performance improvements into VBF/VBS physics sensitivity
 1259 gains with HGTD. In a typical VBF/VBS event selection, two jets above 30 GeV and large
 1260 invariant mass m_{jj} are required. The large invariant mass requirement results in at least
 1261 one forward jet almost always present in the final state. This leads to final states that can
 1262 contain one or two forward ($|\eta| > 2.5$) tag jets. For Standard Model VBS processes after
 1263 typical m_{jj} and $\Delta\eta(j, j)$ selections, about 80% of the time there is one central and one forward
 1264 jet. For heavy mass objects, such as searches for heavy Higgs bosons, the proportion of
 1265 forward-forward jets increases. Backgrounds to VBS/VBF topologies can arise from single or
 1266 di-boson processes $V(V)+1$ jet ($V = W/Z$) with an extra jet from a pile-up (merged) vertex.
 1267 The two main background topologies to VBF/VBS analyses are hence central-forward (case 1)
 1268 and forward-forward (case 2), where one of the two jets is from a pile-up interactions.

1269 In case 1 there is only one jet within the HGTD acceptance. This is an example where
 1270 t_0 is required. Given a measurement of the vertex t_0 of the jet, it is possible, to improve
 1271 the R_{p_T} calculation of the forward jet, or to check the Δt between the jet time and t_0 if an
 1272 independent determination of t_0 is available. Using the t_0 algorithm described in Sec. 3.2.3,
 1273 if the pile-up forward jet is stochastic, i.e. made of tracks from various different interactions,
 1274 the t_0 algorithm will remove pileup tracks from hard scatter jets thus improving R_{p_T} . If
 1275 it is a QCD pile-up jet, then all its tracks will have the same t_0 time, so that jet is likely to
 1276 pass. The relative fraction of QCD vs. stochastic pile-up jets is very final state and topology
 1277 dependent so it is not simple to estimate the gain in pile-up jet suppression when there is
 1278 one forward and one central jet in the final state. In order to be able to suppress QCD pile-up

1279 jets in central-forward VBF/VBS topologies, a new t_0 reconstruction algorithm based on
1280 tracks outside jets will be required. This will be investigated in the future.

1281 Case 2 consists of one hard-scatter jet and another pile-up jet. Since in this case both jets
1282 are within HGTD acceptance, HGTD can be used to assign a time to each jet and check for
1283 consistency. Similar to the discussion about b -tagging and lepton isolation, the knowledge
1284 of the global t_0 of the event is not required to identify this topology. Only relative time
1285 information is enough to determine if the event is signal and should pass (both jets belong to
1286 the same interaction) or if the event consists of two jets from two different interactions that
1287 have similar z vertex position but different time and should be rejected. Cases where a third
1288 forward jet is present are more complex and would require special consideration. However,
1289 this case is expected to be less important.

1290 Therefore the impact of HGTD to the suppression of pile-up jets depends on the event
1291 topology and selection cuts, particularly the m_{jj} cut that determines the fraction of forward-
1292 forward jet events, and the specifics of the t_0 reconstruction algorithm. In the most conservat-
1293 ive case, using the existing t_0 algorithm, HGTD can improve the suppression of background
1294 events with a QCD or stochastic pile-up jet in the forward-forward topology and with a
1295 stochastic pile-up jet in the central-forward topology. More advanced t_0 methods, yet to be
1296 developed, may allow to resolve the central-forward QCD pile-up case.

1297 There is another potential HGTD improvement in VBF/VBS analyses related to jets that
1298 does not rely on the knowledge of t_0 . This is the case of improvements in the jet energy
1299 resolution through improved particle-flow. Tracks within the jet that are not consistent in
1300 time with the rest of the jet can be rejected as pile-up tracks. The contamination of pile-up
1301 tracks in the forward region might be one of the main limitations of particle flow algorithms
1302 in the forward region since it will not be possible to correctly tag and remove pile-up energy
1303 contributions to the jet only on the basis of the longitudinal impact parameter, without the
1304 additional time information. The study of particle flow algorithms using timing information
1305 is outside the scope of this TDR due its complexity, but a clear area of potential improvement
1306 for HGTD.

1307 Improvement in the reduction of E_T^{miss} tails can impact searches for new physics and precision
1308 measurements with neutrinos in the final state. Signatures with large E_T^{miss} comprise a very
1309 broad class of events at the LHC. The impact of reduced transverse missing energy tails on
1310 specific physics analyses has not yet been studied but will be pursued as a next step.

1311 The next sections provide detailed studies showing the expected impact of HGTD in some
1312 selected physics analyses from improved luminosity, lepton isolation, and triggering on
1313 highly ionizing particles. Many other physics analysis are being pursued.

1314 3.3.2 Impact of the luminosity uncertainty

1315 The uncertainty on the measurement of the integrated luminosity affects the majority of
 1316 physics analyses at the LHC. It is especially relevant for precision measurements, for which
 1317 the total uncertainty is dominated by systematic effects. An example of such an effect, that
 1318 is the dominant source of uncertainty for some measurements, is the knowledge of the
 1319 integrated luminosity.

1320 The luminosity has been measured in Run 1 and Run 2 of the LHC using several detectors.
 1321 Uncertainties for the nominal proton-proton collision data are currently 2.1%, 2.2%, 2.4%,
 1322 and 2.0% for data from 2015, 2016, 2017 and 2018, respectively. For the combined 2015-
 1323 2018 dataset, taking into account correlated and uncorrelated effects between the different
 1324 years, an uncertainty of 1.7% is obtained. These uncertainties are derived following a
 1325 methodology similar to that detailed in Ref. [17], and using the LUCID-2 detector for the
 1326 baseline luminosity measurements [18], from calibration of the luminosity scale using x - y
 1327 beam-separation scans (van der Meer scans).

1328 In Ref. [19], the Higgs boson analyses performed during Run-2 have been extrapolated to
 1329 the HL-LHC dataset. The performance of the analyses have been updated, including the
 1330 expected changes to the uncertainties affecting the analyses, the increase of the collision
 1331 energy and the increase in the integrated luminosity. An ambitious uncertainty of 1% is
 1332 assumed for the integrated luminosity, as to not dominate all other sources of uncertainty, in-
 1333 stead of a more realistic 2% uncertainty. Since the Higgs boson analyses cannot constrain the
 1334 uncertainty on the luminosity, it is straightforward to compare any value for the luminosity
 1335 uncertainty to the magnitude of the other uncertainties affecting these analyses.

1336 Tab. 3.3 lists the largest sources of uncertainty affecting three important Higgs boson cross
 1337 section measurements; gluon-fusion (ggH) production of Higgs bosons with decays to $\gamma\gamma$
 1338 and ZZ^* , and combined gluon-fusion and vector boson fusion (VBF) production of Higgs
 1339 bosons with decay to $\tau\tau$. For all these measurements, an uncertainty of 2% on the integrated
 1340 luminosity would be the single largest source of uncertainty on the results.

Analysis channel	Largest uncertainty	$\Delta\sigma/\sigma_{\text{SM}}$
Cross section for $ggH(\rightarrow \gamma\gamma)$	Photon isolation efficiency	1.9%
Cross section for $ggH(\rightarrow ZZ^*)$	Electron eff. reco. total	1.5%
Cross section for $ggH + \text{VBF}, H \rightarrow \tau\tau$	QCD scale $ggH, p_{\text{T}}^H \geq 120 \text{ GeV}$	1.7%

Table 3.3: List of dominant uncertainties (excluding the uncertainty on the integrated luminosity) affecting various expected Higgs boson cross section results at the HL-LHC. An uncertainty on the luminosity measurement of 2% would be the dominant source of uncertainty for all these measurements.

1341 The above considerations illustrate the importance of a precise luminosity measurement
 1342 for the Higgs boson physics program at the HL-LHC. The same concerns apply to any
 1343 measurement of processes with similar, or larger, cross sections compared to the Higgs boson.

1344 These include important processes such as W and Z boson production, and measurements
 1345 of single and pair production of top quarks. As will be described in Chap. 6 and Chap. 10,
 1346 the HGTD will have the capability to measure the luminosity at the HL-LHC. Chap. 10 also
 1347 discuss the main sources of uncertainty affecting the luminosity determination, many which
 1348 will take operational experience with the HGTD to provide numerical estimates for.

1349 3.3.3 Measurement of $\sin^2 \theta_{\text{eff}}$

1350 In the Standard Model (SM), the Z boson couplings differ for left- and right-handed fermions
 1351 due to the mixing between the neutral states associated to the $U(1)$ and $SU(2)$ gauge groups.
 1352 The difference leads to an asymmetry in the angular distribution of positively and negatively
 1353 charged leptons produced in Z boson decays and depends on the weak mixing angle, $\sin^2 \theta_{\text{eff}}$
 1354 [20].

Experimentally, this asymmetry can be expressed as simply as

$$A_{\text{FB}} = \frac{N(\cos \theta^* > 0) - N(\cos \theta^* < 0)}{N(\cos \theta^* > 0) + N(\cos \theta^* < 0)},$$

1355 where θ^* is the angle between the negative lepton and the quark in the Collins-Soper
 1356 frame [21] of the dilepton system. This asymmetry is enhanced by Z/γ^* interference and
 1357 exhibits significant dependence on the dilepton mass.

1358 The weak mixing angle is one of the fundamental parameters of the SM. Several measure-
 1359 ments of $\sin^2 \theta_{\text{eff}}$ have been made at previous and current colliders, and the current world
 1360 average is dominated by the combination of measurements at LEP and at SLD, which gives
 1361 $\sin^2 \theta_{\text{eff}} = 0.231530 \pm 16 \times 10^{-5}$. However, the two most precise measurements differ by
 1362 over 3σ [20].

1363 At HL-LHC, the best sensitivity to $\sin^2 \theta_{\text{eff}}$ is at high Z rapidity when at least one lepton is
 1364 present in the forward region [22]. Only Z bosons decaying to electrons are considered in
 1365 this analysis since this final state provides the best experimental precision within the largest
 1366 acceptance.

1367 The fiducial acceptance of $Z/\gamma^* \rightarrow ee$ events is split into three independent channels de-
 1368 pending on the electron $|\eta|$: CC, CF, FF when C represents electron reconstructed in the
 1369 central region ($|\eta| < 2.47$) and F represents electron reconstructed in the forward region
 1370 ($2.5 < |\eta| < 4.2$). Both electrons are required to have $p_T > 25$ GeV. The invariant mass of the
 1371 electron pair is required to be loosely consistent with the Z boson mass, $60 < m_{\ell\ell} < 200$ GeV,
 1372 and the events are further categorised in 10 equal-size bins in absolute dilepton rapidity up
 1373 to $|y_{ee}| = 4.0$.

1374 The contribution of jets misidentified as electrons is suppressed using a tight electron iden-
 1375 tification and a track isolation requirement. In the forward region, the timing information

1376 provided by the HGTD is used to improve the electron isolation by rejecting additional
 1377 tracks from interactions close in space, but separated in time from the hard-scatter vertex.
 1378 The purity of the candidate sample is determined with simulation, and is found to be greater
 1379 than 99% in the CC channel, between 90 and 98% in the CF, and between 60 and 90% in the
 1380 FF channel. The signal significance with HGTD is up to 20% higher with respect to the case
 1381 of ITk only in the CF channel.

1382 A_{FB} is calculated from the selected electron pairs, and unfolded to correct for detector effects
 1383 and migrations in $m_{\ell\ell}$ and $|y_{ee}|$ bins. In the CF and FF channels migrations in the $m_{\ell\ell}$ are up
 1384 to 50 and 60% respectively. Various sources of uncertainty are considered. Those associated
 1385 with background are mostly relevant in CF and FF channel and are estimated to be 5% on
 1386 the background yield and considered uncorrelated for each $m_{\ell\ell}$ and $|y_{ee}|$ bin.

1387 Significant uncertainties arise from knowledge of the momentum scale and resolution
 1388 for the electrons. Following Reference [23] a systematic of 0.5% (0.7%) is considered to
 1389 account for possible non-linearity in the energy scale of electron reconstructed in the central
 1390 (forward) region with $E_T < 55$ GeV and up to 1.5% (2.1%) for central (forward) electron with
 1391 $E_T > 100$ GeV.

1392 The expected sensitivity to particle level A_{FB} as a function of m_{ee} is shown in green in
 1393 Fig. 3.27 for each channel for chosen rapidity bin. As expected the larger asymmetry is
 1394 observed in the CF channel. The extraction of $\sin^2 \theta_{\text{eff}}$ is done by minimising the χ^2 value
 1395 between particle-level A_{FB} distributions with different weak mixing angle hypotheses, at
 1396 LO in QCD, with NNLO CT14 parton distribution function (PDF). As shown in Fig. 3.27,
 1397 the imperfect knowledge of the PDF results in sizeable uncertainties on A_{FB} , in particular
 1398 in regions where the absolute values of the asymmetry is large, i.e. at high and low $m_{\ell\ell}$.
 1399 On the contrary, near the Z boson mass peak, the effect of varying $\sin^2 \theta_{\text{eff}}$ is maximal,
 1400 while being significantly smaller at high and low masses. Thus, in this projection a global
 1401 fit is performed where $\sin^2 \theta_{\text{eff}}$ is extracted while constraining at the same time the PDF
 1402 uncertainties [22]. With this analysis, the expected sensitivity of the extraction of $\sin^2 \theta_{\text{eff}}$
 1403 are respectively 25×10^{-5} , 21×10^{-5} and 40×10^{-5} for the CC, CF and FF channel. The
 1404 uncertainty of the results is dominated by the currently limited knowledge of the PDFs. If
 1405 looking purely at the experimental uncertainties, including the HGTD in the ATLAS forward
 1406 region brings a 13% improvement on the $\sin^2 \theta_{\text{eff}}$ sensitivity in the CF channel. Combining
 1407 the three channels together the expected sensitivity reaches a precision of $\Delta \sin^2 \theta_{\text{eff}} =$
 1408 $18 \times 10^{-5} \pm 16 \times 10^{-5}$ (PDF) $\pm 9 \times 10^{-5}$ (exp.) which exceeds the precision achieved in all
 1409 previous single-experiment results so far.

1410 3.3.4 Monopole searches

1411 Magnetic monopoles are elementary particles with a single pole magnetic charge, unlike
 1412 dipoles that have both a north and south magnetic pole. These particles were first theorized

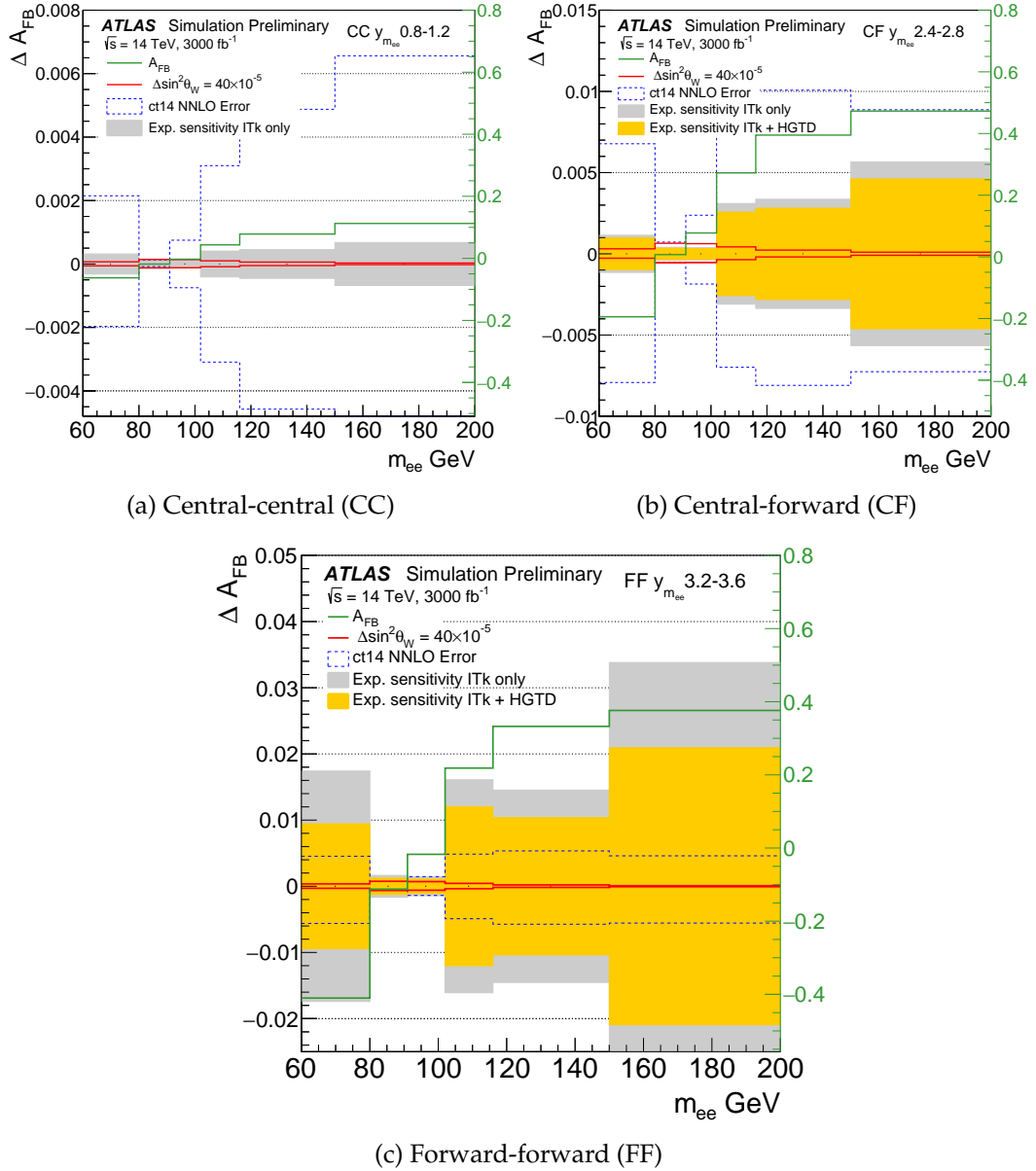


Figure 3.27: Distribution of ΔA_{FB} as a function of mass for the CC, CF and FF channels. The filled bands correspond to the experimental sensitivity with and without the HGTD. The solid red lines correspond to a variations of $\sin^2 \theta_{\text{eff}}$ corresponding to 40×10^{-5} . The dashed blue lines illustrate the total error from CT14 NNLO PDF. Overlaid green line shows the particle-level A_{FB} distribution.

1413 by physicist Paul Dirac in 1931 [24]. Until today no direct evidence of the existence of
 1414 magnetic monopoles was found. A consequence of the existence of monopoles is the perfect
 1415 symmetrization of Maxwell equations. With the existence of magnetic charge the divergence
 1416 term of the magnetic field would no longer be 0 and the flux of magnetic particles would
 1417 give an additional term to the rotor of the electric field. Monopoles can be either scalar
 1418 bosons (Spin-0) or fermions (spin-1/2). The Bethe-Bloch formula for magnetic particles is
 1419 different from the one for electrically charged particle. Electrons and protons have high
 1420 ionization power at low energy and then stabilize to MIPs at high energy due to the $1/\beta^2$
 1421 term. For magnetic monopoles the ionization at high energy grows as a function of β^2 ,
 1422 deviation significantly from the MIP signature.

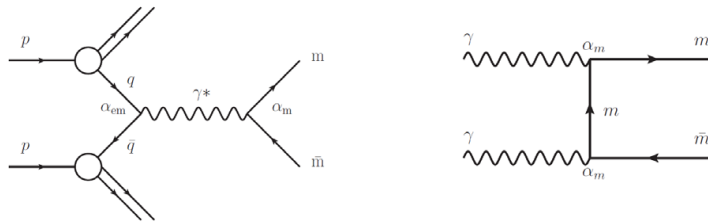


Figure 3.28: Production process of Monopole in LHC for Drell-Yan (left) and photon fusion (right)

1423 Magnetic monopoles were also predicted as part of grand unified theories as described
 1424 in [25]. Several new electroweak models predicted monopoles with a mass that could be
 1425 accessible at LHC [26, 27] thus allowing the detection in LHC experiments. Such monopoles
 1426 would be generated at LHC with Drell-Yan or photon fusion processes as shown in the
 1427 Feynman diagrams in Fig. 3.28. However due to the large coupling of monopoles with
 1428 photons the production cross section cannot be reliably calculated.

1429 According to these models magnetic monopoles act as long lived and highly ionizing
 1430 particles. A single Dirac (magnetic) charge corresponds to around 60 electric elementary
 1431 charges. The increased ionizing power results in an energy deposit 100 times the energy
 1432 deposit of a MIP in a silicon detector such as the HGTD. Past searches for monopoles in
 1433 the ATLAS detector have not turned up evidence for monopoles [28–30]. These searches
 1434 however only cover the barrel region of the detector and rely on a low-threshold EM trigger
 1435 that will likely be scaled up in HL-LHC.

1436 Monopoles were simulated as single particles with $\langle\mu\rangle = 0$. The samples were simulated
 1437 with a special setting, the 4DL package [4DL] in full simulation. Samples with Monopoles of
 1438 Dirac charge 1 and 2 were produced. Higher values of the Dirac charge leads to a high level
 1439 of interaction in the ITk that the particles do not reach the HGTD.

1440 The energy distribution of the hits in HGTD for a monopole event is shown in Fig. 3.29(a).
 1441 Several low-energy hits and one or two hits with a very high-energy deposit are recorded.
 1442 The distribution of the hit energy for monopoles with Dirac charge 1 and 2, of minimum
 1443 bias samples and pions is shown in Fig. 3.29(b). The high-energy deposit is clearly separated

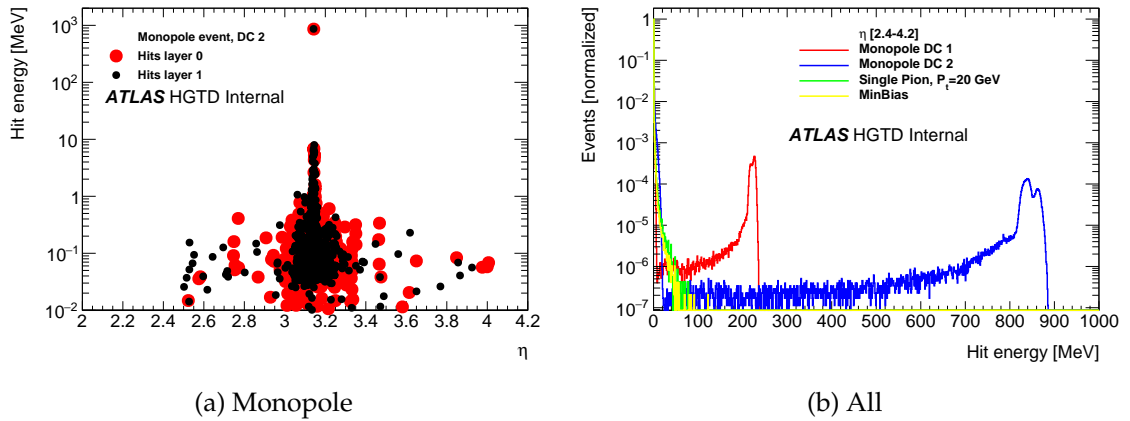


Figure 3.29: The energy deposit vs η) for an event of single Monopole with Dirac charge 2 as well as the energy distribution for minimum bias, pions and monopoles with Dirac Charge 1,2 are shown.

1444 from the other types of particles. Thus a high-energy single hit would give a clear and
 1445 unique signature for an interacting magnetic monopole.

1446 This distinctive signature can also be exploited by HGTD electronics at trigger level. A
 1447 special bit can be added in the luminosity processing electronics to flag high-energy deposits
 1448 in a single pad. This way is possible to recognize candidate Monopole events at trigger level
 1449 in the end-cap region of the detector and make them available for offline analysis.

1450 4 Technical Overview

1451 4.1 Introduction

1452 This chapter summarizes the most important aspects of the design of the HGTD. The main
1453 requirements that drives the design and the proposed technical solutions are discussed.
1454 Measurements from the on-going R&D program are presented, especially on sensors and
1455 electronics, that demonstrate the achieved performance. More detailed descriptions of all
1456 these items will be presented in subsequent chapters, including the next steps towards the
1457 construction.

1458 4.2 Detector overview and key requirements

1459 The detector has been designed for an operation with 200 proton-proton collisions per bunch
1460 crossing and a total integrated luminosity of 4000 fb^{-1} . The HGTD will be located in the
1461 gap region between the barrel and the end-cap calorimeters, at a distance of approximately
1462 $\pm 3.5 \text{ m}$ from the interaction point. Fig. 4.1 shows a transverse view of the detector, without
1463 the front cover of the vessel, where the front layer of the first double-sided active layer (in
1464 blue) and the peripheral electronics boards location (in green) can be seen. The envelope of
1465 the detector vessel has a radial extent of 110 to 1000 mm. The envelope in z , including the
1466 moderator, supports and front and rear vessel covers is 125 mm. This includes the moderator
1467 that is placed behind the HGTD with a total thickness of 50 mm, to reduce the back-scattered
1468 neutrons created by the end-cap/forward calorimeters, protecting both the ITk and the
1469 HGTD. Each end-cap is made of one hermetic vessel, two instrumented double-sided layers
1470 (mounted in two cooling/support disks), and two moderator pieces placed inside and
1471 outside the hermetic vessel. The total detector weight per end-cap is approximately 350 kg,
1472 of which 150 kg comes from the moderator, weight equally shared between the moderator
1473 located inside and outside the vessel.

1474 The front vessel cover and each cooling/support disk are physically separated in two half
1475 circular disks to allow the opening of the detector in presence of the beam pipe.

1476 The active detector element is made of Low Gain Avalanche Silicon Detectors (LGADs)
1477 read-out by dedicated front-end electronics ASICs (ALTIROC). It covers the pseudo-rapidity
1478 range $2.4 < |\eta| < 4.0$ ($120 \text{ mm} < R < 640 \text{ mm}$). The active area is divided in three rings

1479 (inner, middle and outer ring). The inner ring covering the region $3.5 < |\eta| < 4.0$ (120 mm
 1480 $< R < 230$ mm) is equipped with modules mounted on the front and back sides of a given
 1481 cooling plate, with 70 % overlap along the readout row direction, in order to provide on
 1482 average 2.7 hits per track in the most irradiated and higher occupancy region.

1483 The middle ring covering the region $2.7 < |\eta| < 3.5$ (230 mm $< R < 470$ mm) is equipped
 1484 with modules overlapping 54 % providing on average 2.5 hits per track.

1485 The outer ring covers the region $2.4 < |\eta| < 2.7$ (470 mm $< R < 640$ mm) is equipped with
 1486 modules overlapping only 20 % providing on average 2.1 hits per track.

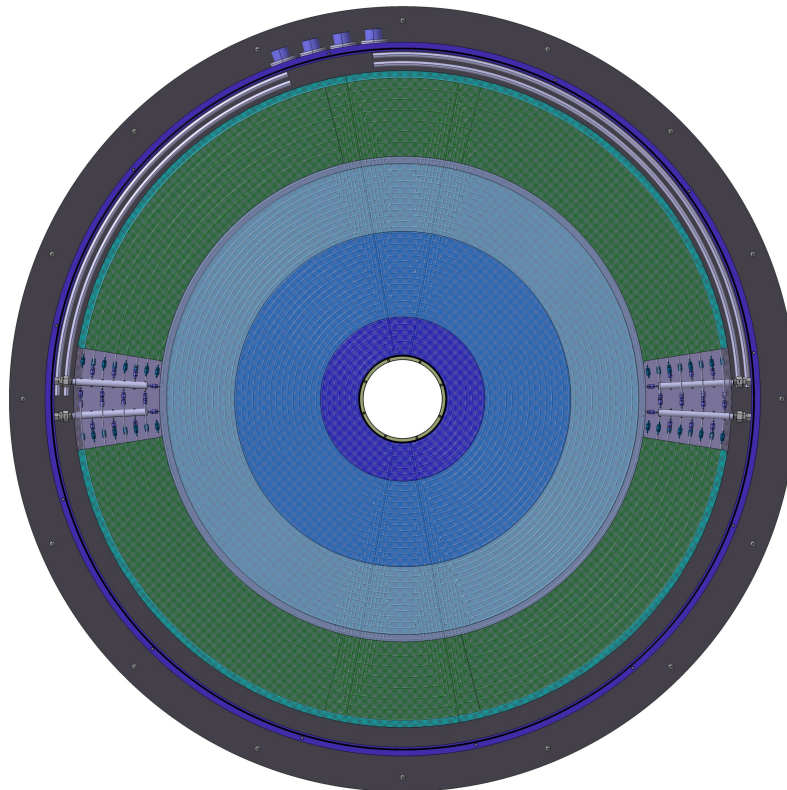


Figure 4.1: Transverse view of the detector. The blue active region is shown in 3 blue tones indicating the 3 rings with different sensors density. The green areas indicate the region of the peripheral electronics boards with some open space for the CO₂ cooling manifold in gray. TODO: REPLACE FIGURE BY ONE WITH THE READOUT ROW LINES, PROVIDED BY ABOUD.

1487 4.2.1 Expected Radiation levels

1488 As discussed in Chap. 2, the radiation levels in the forward region exceed the radiation
 1489 hardness of both the sensors and the front-end electronics, especially at low radius. In

1490 order to mitigate the radiation levels and fulfil the detector requirements during the full life
 1491 time of the HL-LHC the plan is to replace the most inner ring after each 1000 fb^{-1} (3 times
 1492 in total) and the middle ring at 2000 fb^{-1} (once), during long shutdowns. Fig. 4.2(a) and
 1493 Fig. 4.2(b) show respectively the maximum expected neutron-equivalent fluence and TID as
 1494 a function of the detector radius. A factor of 1.5 was included to account for uncertainties in
 1495 the simulation. An additional factor of 1.5 was applied to the total ionising dose (TID) to
 1496 account for low dose rate effects on the ASICs. In the proposed 3 rings layout the maximal
 1497 TID and fluence, using the Fluka estimations of September 2019, does not exceed 2 MGy
 1498 and $2.5 \times 10^{15} \text{ neq/cm}^2$. In the inner ring the total Si 1MeV neq has a similar contribution
 1499 from neutrons and charged particles while in the middle and outer rings the dominant effect
 1500 comes from neutrons, as seen in Fig. 2.13. The exact radial transition between the three rings
 1501 will be tuned for the final detector layout, once the FLUKA simulations will be updated
 1502 with the final ITk layout, and the radiation hardness of the final sensors and ASICs are
 1503 re-evaluated.

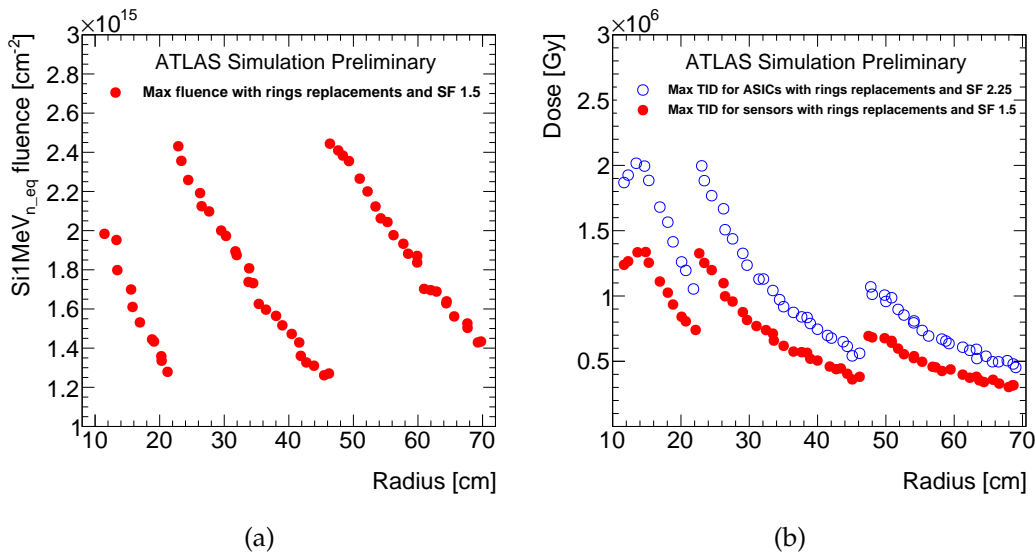


Figure 4.2: Expected Si1MeV_{n,eq} radiation levels in HGTD, using Fluka simulations, as a function of the radius considering a replacement of the inner ring every 1000 fb^{-1} and the middle ring replaced at 2000 fb^{-1} . These curves included a factor of 1.5 to account for simulation uncertainty. An additional factor of 1.5 is applied to the TID to account for low dose rate effects on the electronics, leading to a SF = 2.25.

1504 4.2.2 Key requirements

1505 A high intrinsic single hit efficiency is essential for through the lifetime of the HGTD. This
 1506 puts stringent constraints on the smallest charge delivered by the sensor and the lowest

1507 achievable threshold of the electronics discriminator. Charge as low as 4 fC should be
 1508 detected with a signal over noise (S/N) larger than 7, while keeping a low rate of fake hits
 1509 induced by the electronics noise ($< 0.1\%$). As measured with testbeam, in these conditions
 1510 an efficiency larger than 95% can be obtained.

1511 The target time resolution per track, combining multiple hits, is from 30 ps at the start of
 1512 lifetime to 50 ps after 4000 fb^{-1} . To achieve this performance, the time resolution per hit
 1513 should be about 35 ps at the start of lifetime and 65 ps at the end of lifetime over the full
 1514 surface of the detector.

The main contributions to the time resolution of a hit are given by:

$$\sigma_{\text{hit}}^2 = \sigma_{\text{Landau}}^2 + \sigma_{\text{elec}}^2 + \sigma_{\text{clock}}^2 \quad (4.1)$$

- 1515 • where σ_{Landau} is the time resolution induced by the Landau fluctuations in the depos-
 1516 ited charge as the charged particle traverses the sensor
- 1517 • σ_{elec} is the contribution from the electronics read-out (jitter and time-walk). It is
 1518 required to be about 25 ps for a MIP with an LGAD gain of 20 (corresponding to a
 1519 charge of 10 fC) at the start of the HL-LHC, and at most 60 ps after 4000 fb^{-1} for a
 1520 charge of 4 fC. The TDC contribution is expected to be negligible.
- 1521 • σ_{clock} is the non-deterministic jitter contribution from the clock distribution, to be
 1522 smaller than 15 ps after calibration.

1523 In addition, the detector should be sensitive to hits in the same pad that come from consec-
 1524 utive bunch crossings and should provide the sum of the number of hits per ASIC for each
 1525 bunch crossing. The latter is used in the luminosity measurement.

1526 4.2.3 Read-out bandwidth and trigger

1527 With the baseline ATLAS architecture, the ATLAS detector is read-out with a single Level 0 (L0)
 1528 trigger running at an maximum rate of 1 MHz, with a maximum latency of $10 \mu\text{s}$ [31]. The
 1529 time information of the HGTD hit cells will be read out on reception of this L0 trigger signal.
 1530 In the evolved scheme considered by ATLAS, called L0–L1, the HGTD will be read-out on
 1531 the reception of a L1 trigger signal with a maximum frequency of 800 kHz and a maximum
 1532 latency of $35 \mu\text{s}$. The time information from each ASIC is read-out by only one data line to
 1533 the lpGBT. Therefore the maximal bandwidth is limited to 1.28 Gbit s^{-1} . The luminosity data
 1534 is transmitted at each bunch crossing to dedicated lpGBTs, requiring a 640 Mbit s^{-1} e-link
 1535 bandwidth.

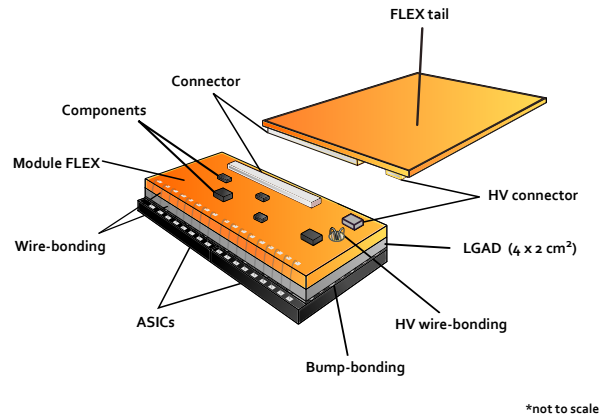


Figure 4.3: View of an HGTD hybrid module equipped with its read-out flex cable tail. The bare module, glued on the flex cable, is made of a $4 \times 2 \text{ cm}^2$ sensor with two bump bonded ASICs. The signal lines of the ASIC are wire bonded on one side of the cable, while the bias voltage of the sensor is provided to the back-side of the sensor through a hole in the cable.

1536 4.3 Hybrid HGTD module

1537 Fig. 4.3 shows a view of an hybrid module made of two parts: an LGAD sensor and two
 1538 ASICs, called a bare module, and a flexible printed circuit board (flex cable) made of two
 1539 pieces, 1 small flex board permanently glued to the bare module and a long flex tail whose
 1540 length, not exceeding 60 cm, depends on the module position in the detector. The sensor and
 1541 the ASICs are connected through a flip-chip bump bonding process called hybridization. All
 1542 connections between the ASIC and the peripheral electronics are routed through the flex
 1543 cable. The bare module is glued in the back side of the sensor to the flex module small piece,
 1544 and all the signals are wire bonded between the ASIC and the flex cable and for the high
 1545 voltage between the sensor and and the flex.

1546 The characteristics of the bare modules are:

- 1547 • The size of the bare modules, all equal, is approximately $2 \times 4 \text{ cm}^2$ and each bare
 1548 module contains 450 pads (15x30). Its size has been defined to optimize the coverage
 1549 at the inner radius a and to provide a good yield for the hybridization process. The
 1550 nominal total amount of bare modules is 8032.
- 1551 • The size of the pad, $1.3 \times 1.3 \text{ mm}^2$, results from a compromise between smaller pads,
 1552 leading to lower occupancy and smaller capacitance and thus low electronics jitter,
 1553 and larger pads, which provide better geometric coverage with large fill factors and
 1554 less power dissipation from the ASIC.

- The sensor is connected to two ASICs, each of them reading a matrix of 225 (15x15) pads.
The size of the ASIC is about $20 \times 22 \text{ mm}^2$.

The status of the R&D of key components is discussed briefly below, with more technical details in subsequent chapters.

4.3.1 Sensors

The sensors are based on LGAD technology, pioneered 5 years ago by the Centro Nacional de Microelectrónica (CNM) Barcelona in close collaboration with the RD50 collaboration.

LGADs are n-on-p silicon detectors containing an extra highly-doped p-layer below the n-p junction to create a high field which causes internal amplification as displayed in Fig. 4.4(a). When a charged particle crosses the detector, an initial current is created from the drift of the electrons and holes in the silicon. When the electrons reach the amplification region, new electron/hole pairs are created and the holes drift towards the p^+ region and generate a large current resulting as the gain of the LGAD. This current, much larger than in a standard diode, is the key ingredient to get an excellent time resolution for energy deposited by MIP particles. The expected current for different irradiation levels (therefore different gains) are presented in Fig. 4.4(b). For large gain, the rise time is about 500 ps and the signal duration is approximately 1 ns. For large fluences, the charge is smaller and the rise time and the signal duration are shorter.

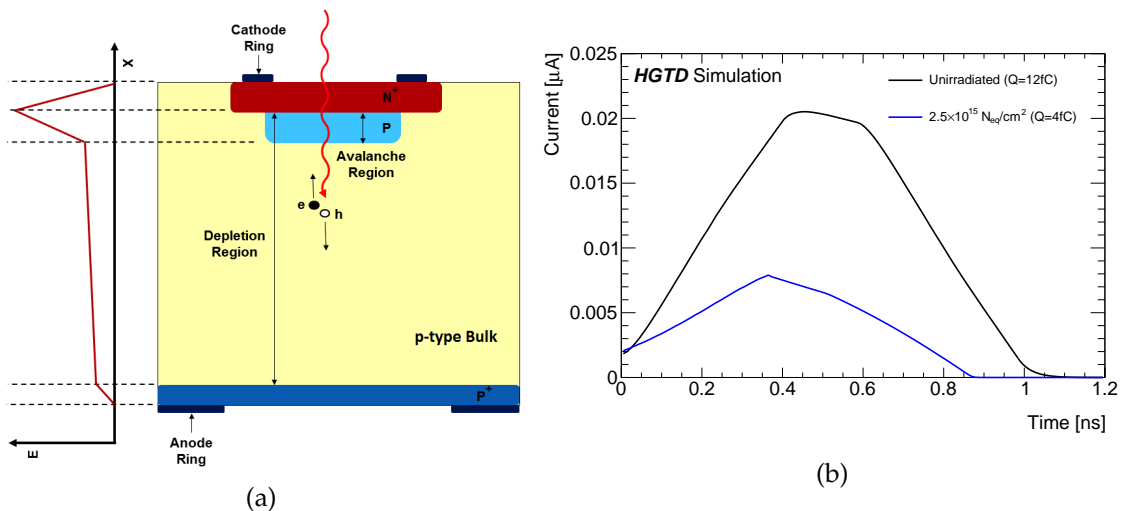


Figure 4.4: Cross section of an LGAD ((a)) and simulated signal current in LGADs at start and after full integrated neutron fluence ((b)).

After amplification in the gain layer, the height of the LGAD signal is proportional to the gain, M , but is independent of the detector thickness. On the other hand, the slope dV/dt

1575 depends on the thickness of the sensor, favouring thin sensors since the electronics jitter
 1576 scales as the inverse of the slope. However, the jitter depends also linearly on the detector
 1577 pad capacitance, therefore limiting the potential use of very thin sensors. Consequently, the
 1578 optimal thinness relies strongly on the performance of the read-out ASIC. The baseline active
 1579 thickness has been chosen to be $50\ \mu\text{m}$ while the total thickness is $250\ \mu\text{m}$. The pad size is
 1580 $1.3 \times 1.3\ \text{mm}^2$ which resulted from an optimization discussed in the previous section.

1581 Over the last years LGAD sensors are been produced by CNM/Spain, HPK/Japan, FBK/Italy
 1582 and recently NDL/China with different doping levels, active thickness, pad size, or inter-
 1583 pad gaps. These detectors have been exposed to protons, neutrons and Xrays up to the
 1584 expected maximum radiation levels (including the safety factors) and intensively charac-
 1585 terized in laboratory (with probe station, β source, laser) or in beam tests (at CERN, DESY,
 1586 FERMILAB).

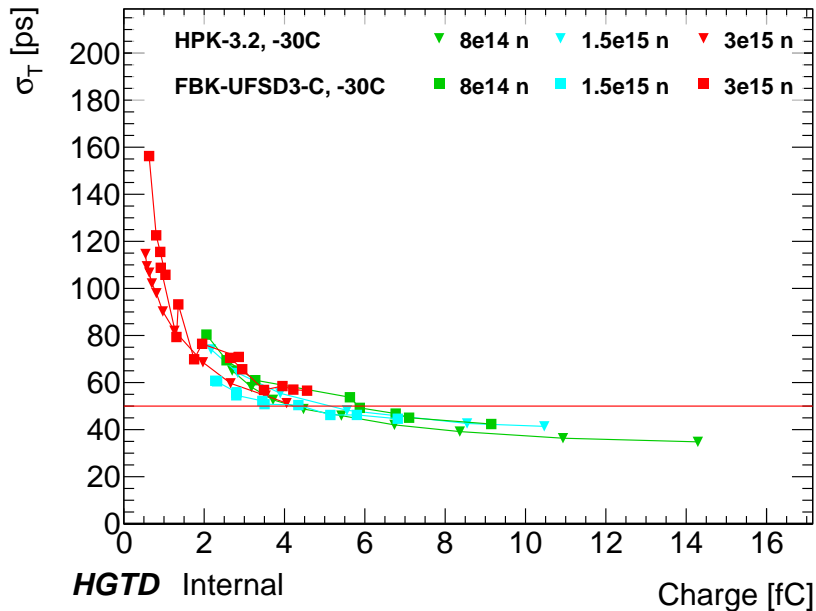


Figure 4.5: Time resolution as function of the collected charge for neutron irradiated LGADs from different producers (HPK, FBK) with a $50\ \mu\text{m}$ active thickness. These measurements have been made at $-30\ ^\circ\text{C}$, in the laboratory with a β source using a custom electronics read-out board (not the ALTIROC).

1587 Under irradiation, the expected decrease of the charge yield can be mitigated by increasing
 1588 the bias voltage (up to $700\ \text{V}$) and operating at low temperature ($-30\ ^\circ\text{C}$). Fig. 4.5 summarizes
 1589 results obtained in the laboratory, with dedicated electronics, for sensors from different
 1590 producers exposed to neutron fluence up to $2.5 \times 10^{15}\ \text{n}_{\text{eq}}\ \text{cm}^{-2}$. A charge of $4\ \text{fC}$ can be
 1591 reached up to a fluence of $2.5 \times 10^{15}\ \text{n}_{\text{eq}}\ \text{cm}^{-2}$, providing a time resolution of about $50\ \text{ps}$.

1592 The performance of sensors from all manufacturers is similar, even if before irradiation the
1593 optimal bias voltage might be different because the doping concentration is different. With a
1594 minimal charge of about 4 fC and a discriminator threshold of about 2 fC, a hit efficiency of
1595 95% is expected. For the largest fluence, the Boron doping has mostly migrated from the
1596 avalanche region and the remaining small gain comes from the bulk diode, due only to the
1597 large high bias voltage applied. The time resolution in this domain is fully dominated by the
1598 electronics jitter, thus the ASIC performance at low charge

1599 Intense R&D is still ongoing to improve the radiation hardness with deep narrow doping
1600 implantation, different doping (Ga instead of B), C implantation. Depending on the results
1601 of these studies, discussed in detail in Chap. 5, the exact radius of the inner and middle rings
1602 might be tuned. The working points for the bias voltage need to be adjusted with respect to
1603 the radiation flux: ongoing studies of the breakdown voltage should define safe criteria to
1604 operate the detector.

1605 Already, many single pads (> 1000), small arrays of 2×2 and 5×5 pads from various
1606 companies have been measured in the laboratory and test beams, showing an excellent
1607 yield. The first matrices of 15×15 pads, delivered by HPK, have also been characterized.
1608 They show an excellent uniformity both for the operating bias voltage and the low leakage
1609 current.

1610 Following almost four years of R&D activities, shared in part with CMS timing detector
1611 and RD50, a first set of criteria for the parameters of the final sensor design have been
1612 established, constituting our baseline. These include: 50 μm active thickness, narrow and
1613 deep doping profile, 70 μm inter-pad gap, and a 300 μm slim edge distance with two guards
1614 rings. However, some of the parameters will need to be further validated up to the Final
1615 Design Review, scheduled for 2021.

1616 4.3.2 Front End ASIC

1617 As discussed previously in combination with the LGAD sensor, the Front End ASIC is
1618 challenging. Taking into account the expected TID radiation levels and needed low jitter, the
1619 technology to be used is CMOS TSMC 130 nm. The global architecture of the ASIC, called
1620 ALTIROC, is similar to the ASICs developed for pixel detectors but with a significantly
1621 reduced number of channels and a quite different single pixel Front End for the the time
1622 measurement. Fig. 4.6 presents the general architecture with a matrix of 225 channels
1623 organized along columns for the read-out and with common digital electronics at the
1624 bottom.

1625 The analog Front End electronics of each channel is the most critical element to reach low
1626 jitter. The sensor signal is amplified using a voltage preamplifier. Taking into account
1627 the non-negligible duration of the LGAD signal (approximately 1 ns), a preamplifier with
1628 about a 1 GHz bandwidth, is enough. The preamplifier is followed by a fast discriminator.

1629 The leading edge of the output (Time Of Arrival, or TOA) provides the start of a Time to
1630 Digital Converter (TDC) using a Vernier delay line configuration. The stop is given by
1631 the clock. This start-stop structure minimizes the power dissipation when hits are absent.
1632 The quantisation step is 20 ps, which contributes little to the expected time resolution. The
1633 TOA measurements are restricted to a 2.5 ns window centered on the bunch crossing. The
1634 expected time dispersion of the hits has a r.m.s of 300 ps so that such a window contains all
1635 the hits of the collisions if centered with about 100 ps accuracy with a phase shifter. The
1636 falling edge of the output provides the start of a second TDC, with 40 ps quantisation step,
1637 in order to measure the Time Over Threshold (TOT), which may be used as an estimate of
1638 the signal amplitude. The TOT information is used offline to correct the TOA for the time
1639 walk effect. After correction, the residual variations are well within ± 10 ps. The digital Front
1640 End is used to store the time data up to the reception of a trigger and buffers the data in
1641 order to be read by the End Of Column cells. This buffer is needed to cope with event to
1642 event fluctuations in the number of hits.

1643 The preamplifier and discriminator performance has been validated using a four-channel
1644 prototype (ALTIROC0), bump-bonded to a sensor of 2×2 pads first and a 25 channel
1645 prototypes (ALTIROC1-V2), bump-bonded to sensors of 5×5 pads including the TDC and
1646 SRAM. The complete analog Front End, adding the TDCs and a SRAM (ALTIROC1), was
1647 tested in the laboratory with the ASIC wire bonded to a specific board (see Fig. 4.7). This
1648 figure also shows preliminary results using the TDC for the jitter and the TOA variation as a
1649 function of an injected calibration charge. With the ASIC alone and an input capacitance of 4
1650 pF, the threshold can be as low as 2 fC, allowing a measurement of an input charge of 4 fC
1651 with a calibration input signal. The jitter for an injected charge of 10 fC (4 fC) is about 20 ps
1652 (60 ps). The variation of the TOA versus the input charge, about 500 ps, is compatible with
1653 the preamplifier bandwidth. This time walk effect needs to be corrected. While the TOT can
1654 be used on testbench with the ASIC alone to correct it, some couplings are distorting the TOT
1655 distribution when bump bonded to sensor preventing to apply this correction. Investigations
1656 are on-going to understand the origin of these couplings. Preliminary measurement with
1657 beam show that a time resolution of 40 ps can be reached with non irradiated sensors

1658 The common digital electronics must satisfy a wide variety of requirements. It first retrieves
1659 the time information of the matched hits and the luminosity hits sum computed in the End
1660 of Column. The luminosity hits are summed in two different windows, a 3.125 ns window
1661 centered on the bunch crossing and a second one with a larger size adjustable by slow control.
1662 In a second step, it formats these data, and provides them to the serializer, which transfers
1663 the data on the e-link to the lpGBT. The speed of the serializer can be selected through
1664 slow control at 320 Mbit s^{-1} , 640 Mbit s^{-1} or 1.28 Gbit s^{-1} , in order to maximize the use of
1665 the bandwidth. A control unit receives the fast commands from the lpGBT (clock, BCID,
1666 L01/L1,...) and through I²C the slow control parameters. A phase-locked loop (PLL) and a
1667 phase shifter are used to clean the jitter of the clock and adjust the clocks with a 100 ps step.
1668 This allows the time and luminosity windows to be centered on the bunch crossing clock for

1669 each individual ASIC. Finally, monitoring blocks are included to measure the temperature
 1670 and the leakage current.

1671 The next major ASIC iteration, ALTIROC2, will integrate all the functionality of the final
 1672 ASIC and will have it's final size. Triply redundant registers will mitigate against SEE and
 1673 will be implemented for all controls and signals registers but not for the read-out data. The
 1674 first iteration should be submitted in 2020 and a second iteration 1 year later. The Final
 1675 Design Review is planned early 2022.

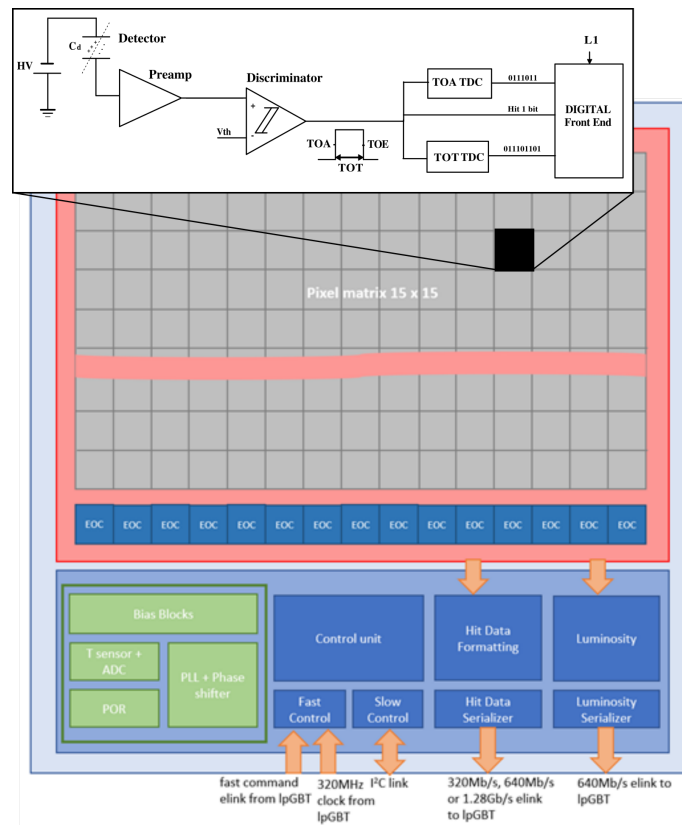


Figure 4.6: Global architecture of the ALTIROC ASIC. The schematic of one Front End electronics channel is displayed on top of the channels matrix, with the preamplifier followed by a discriminator, two TDCs, and a digital front end block.

1676 **4.3.3 Module assembly**

1677 After having qualified separately the sensor and the ASIC at the wafer level, they will be
 1678 connected through a flip-chip bump bonding process. Under Bump Metal (UBM) will be
 1679 deposited on the sensor wafer before dicing. The next step of the hybridisation consists in

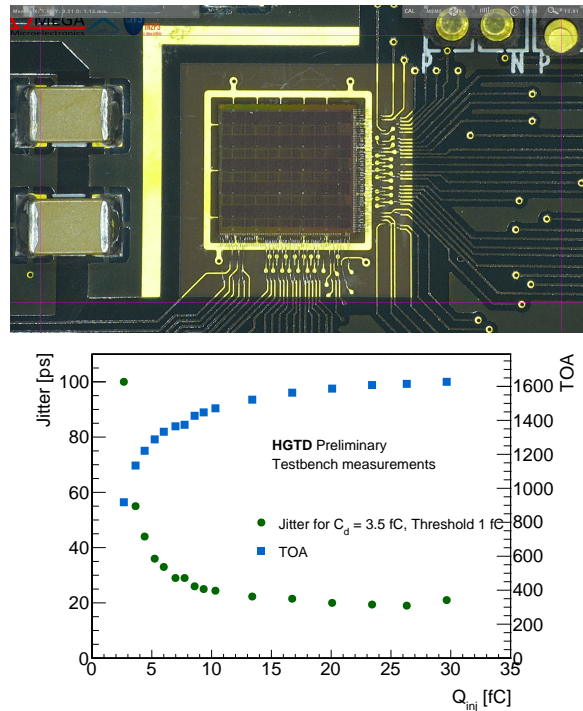


Figure 4.7: Picture of a ALTIROC1 die with 5×5 channels (left) and preliminary measurements of the average time and jitter as a function of the injected charge using calibration injection with one channel of ALTIROC1 (right). PLOT WITH ALTIROC1-V2/TDC IF POSSIBLE AND 2FC THERSHOLD . TO BE DONE BY NIKOLA

1680 the flip-chipping during which the sensor and ASIC are aligned, heated, and compressed,
 1681 so that each solder bump melts and provides the electrical contact between the sensor pad
 1682 and the channel readout. With the large pad size of $1.3 \text{ mm} \times 1.3 \text{ mm}$, solder bump as large
 1683 as $90 \mu\text{m}$ can be used, making the process standard for a few companies, contrary to the
 1684 hybridization of the ATLAS ITk pixel detector. The bump bonding of the prototypes has been
 1685 done in Collaborating institutes and also in Industry with ALTIROC1 and 5×5 channels
 1686 sensors (both doing Under Bump MetaLlization and flip-chip). Satisfactory performance
 1687 results have been obtained, both for connectivity and mechanical stress. The aim is to start
 1688 the qualification of their processes in view of future production in Industry. The final design
 1689 review of the bump-bonding process is planned at the end of 2022.

1690 AS shown in Fig. 4.3 the bare module is glued to a small flex cable pcb, on which the ASIC
 1691 signals are wire bonded. This small flex is connected to the the long flex cable tail though a
 1692 connector in order to transmit the signals to the peripheral electronics boards.

1693 Taking into account the space constraints, the flex tail is a two layers design with a maximum
 1694 thickness of $220 \mu\text{m}$. The longest readout row contains 19 modules. As displayed in Fig. 4.8,

1695 each flex transfers four type of signals:

- 1696 • the data to be read out (time information or luminosity) on two differential-pair e-links
1697 per ASIC. The speed of the data transmission varies from 1.28 Gbit s^{-1} for the inner
1698 radius modules, for the longest flex ($L \sim 60 \text{ cm}$) to 320 Mbit s^{-1} for the outer radius,
1699 for the shortest flex ($L \sim 15 \text{ cm}$). For the luminosity, the speed is 640 Mbit s^{-1} .
- 1700 • the fast commands from the lpGBT (clock, L0/L1 trigger, BCID and configuration
1701 parameters) and the slow control parameters through I²C.
- 1702 • the ASIC power supplies (1.2 V), setting a strong constraint on the flex plane resistance
1703 to minimize the voltage drop and the power dissipation ($< 200 \text{ m}\Omega$) Digital and analog
1704 lines are separated.
- 1705 • the bias voltage for the sensor (up to 800 V requiring excellent insulation).

1706 The first flex cable prototypes, still made of a single piece and longer than required, have
1707 been manufactured in two companies and at CERN PCB workshop. When normalised to
1708 the maximum expected length of $L \sim 60 \text{ cm}$ it satisfies the data transmission, bias voltage
1709 insulation and resistance requirements.

1710 The R&D is still on going on the flex design to ensure it satisfies the strong thickness
1711 constraint along Z and to identify/develop reliable dense and thin mini-connectors for
1712 the connection between the module flex and the flex tail (in the module region) and the
1713 connection to the peripheral electronics board. A few companies have been contacted for
1714 these specific R&D and the final design review should take place in 2022.

1715 The bare module will be glued to the flex small pcb board piece. Some tests of the glue to be
1716 used are ongoing in close collaboration with the ITk Pixel community, as the requirements
1717 are similar. To exercise the module assembly, real size modules made in a first step of heaters
1718 will be mounted in spring 2020 and with real modules (in 2021-2022). This activity will be
1719 done in the framework of the demonstrator activity (detailed in Chap. 14).

1720 4.4 Module loading on support structure

1721 The modules are loaded on an intermediate support structure where the modules are inserted
1722 and glued in pre-defined holes, insuring the exact position of each module and the alignment
1723 along the x and y readout row directions, as displayed in Fig. 4.9. These structures are later
1724 screwed to the cooling plates, using a thermal conductive grease to insure the direct contact
1725 of the modules with the Aluminium cooling plate. These support structures are separated in
1726 3 radial regions to allow a fast replacement of the rings planned to take place at surface in
1727 the long shutdowns.

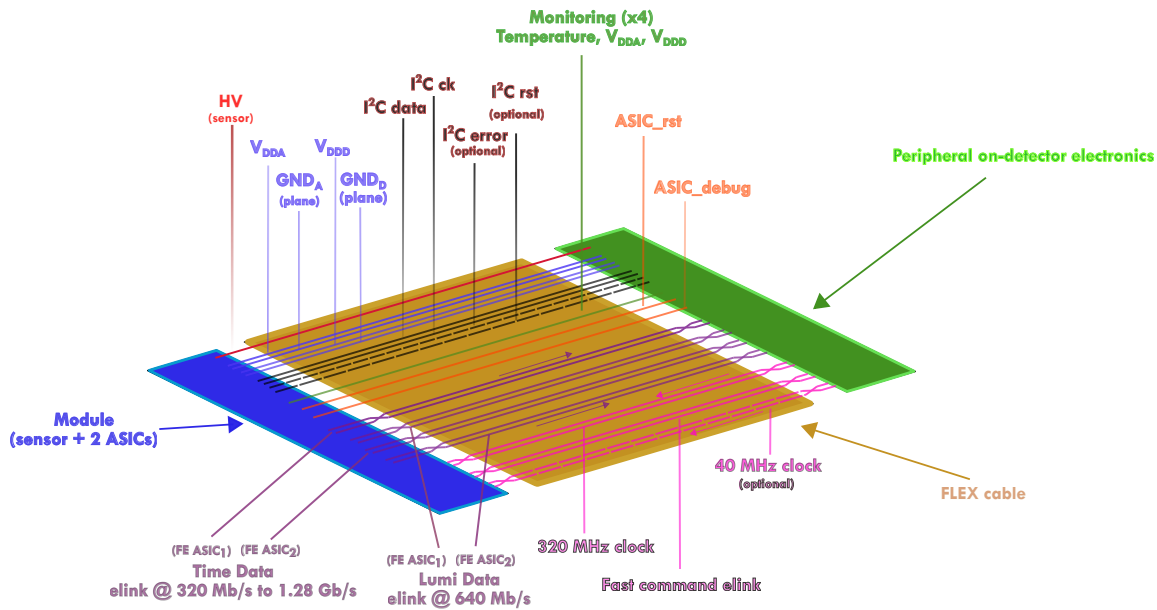


Figure 4.8: Signal transmitted from the ASICs to the peripheral electronics. Each ASIC has a dedicated e-link for luminosity and time data transmission while the other signals are common to both ASICs. The HV line is connected to the sensor.

1728 **4.5 Off detector electronics, calibration and luminosity**

1729 Fig. 4.10 shows the data path from the front-end ASIC to the off-detector backend. Different
 1730 data and control signals from the flex cable are connected to the Peripheral Electronics
 1731 Boards (PEB), where the electrical signals are encoded and transmitted via optical link (at
 1732 $10.24 \text{ Gbit s}^{-1}$) to the off-detector electronics located in USA15.

1733 The off-detector electronics consist of Front End Link eXchange (FELIX) system and Data
 1734 Handler and will be described in section Sec. 10.1.1. The main data stream is read out at
 1735 L0 trigger rate and is meant for ATLAS event process. The luminosity stream is read out
 1736 by dedicated FELIX boards which sum the number of hits over a large enough region to
 1737 provide an accurate online luminosity measurement, the use of 16 regions per layer is under
 1738 study.

1739 **4.5.1 Peripheral Electronics Boards**

1740 NEED TO UPDATE THIS SECTION.....

1741 The PEBs are still at an early design stage and will use components already developed by
 1742 CERN for the LHC upgrades. A single PEB will group together up to four read-out rows in

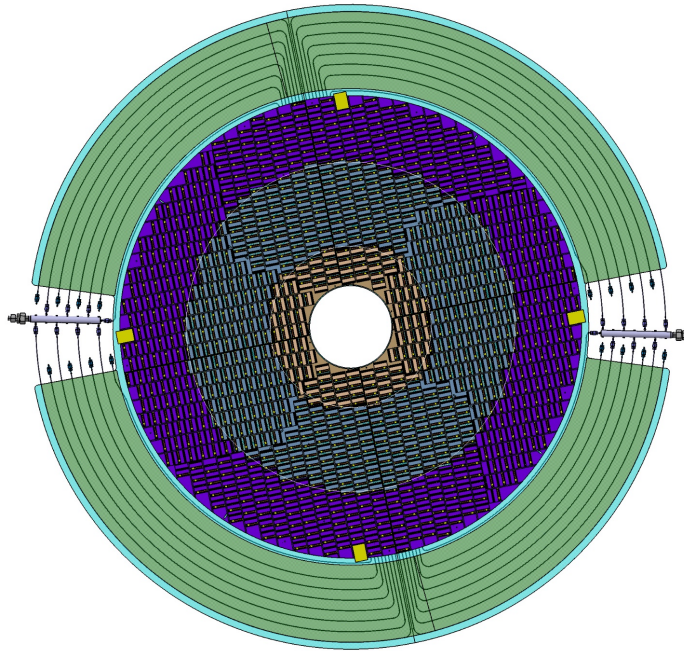


Figure 4.9: View of the modules mounted in the intermediate support structures. The module supports are separated in three radial rings (shown in different colours), and are screwed to the cooling/support plate. The cooling pipes circulating at the outer radius (in green) will serve as pre-heaters and will cool the peripheral electronics boards, to be located in this region.

1743 order to optimize the numbers of components. The expected component densities are high,
 1744 with the PCB carrying large numbers of signals with different properties. Typically, a PEB
 1745 will contain:

- 1746 • 3 to 8 lpGBTs both for the timing data and the luminosity data and the equivalent
 1747 numbers of optical links.
- 1748 • 6 to 8 VTRx optical receivers and transmitters
- 1749 • 25 to 35 DC/DC converters, still to be optimized, depending on their performance.
 1750 Both the BPOL12V and BPOL2.5V are expected to be used.
- 1751 • 3 to 5 multiplexers developed in TSMC 130 nm as input of the monitoring signals to
 1752 the lpGBT ADC.

1753 The placement of these components needs a careful PCB design layout, to fit inside the
 1754 allowed envelope dimensions (in z and r). R&D is still needed to develop flex connectors
 1755 but also HV connectors. A first functional prototype of these PEBs is expected in 2020.

1756

1757 4.5.2 t_0 time calibration

1758 The t_0 knowledge of each individual channel (3.59 millions of channels) is crucial to achieve
 1759 the expected time resolution. The irreducible and non deterministic clock contribution to
 1760 the the resolution is expected to be around 10 ps, coming mainly from the lpGBT clock jitter
 1761 and the additional contribution from the flex cable and ASIC. However this performance
 1762 assumes that all channels are ideally in time with the bunch crossing clock. The use of HGTD
 1763 for the physics strongly relies on the relative comparison of the time of different channels
 1764 within an event. Consequently the geometrical (time static) inter-calibration of all channels
 1765 t_0 is the most crucial while global time drifts over large regions will have smaller impact on
 1766 the performance.

1767 Geometric and time static effects can be corrected with the calibration injection signals in the
 1768 ASIC (different flex cable length, systematic difference between channels in ASIC due the
 1769 imperfect clock tree distribution, etc.) or computed (geometrical time of flight). Calibration
 1770 sequence between LHC fills will be used to monitored these calibration constants.

1771 The variation with time of the 40 MHz phase, therefore of the t_0 is a correction to be determ-
 1772 ined in-situ using the data. Such low frequency clock phase variations can arise in the HGTD,
 1773 for instance with temperature variations at module level from the CO₂ cooling, variations
 1774 from one lpGBT to another (serving a few modules), or from the known day/night effect
 1775 of the LHC clock, probably common to an entire HGTD endcap. The calibration procedure
 1776 will consist in measuring the inclusive average time of each channel with the data triggered
 1777 at 1 MHz. Depending on the time period of these effects, and on the affected component
 1778 and area (ASIC, module, group of ASIC of same lpGBT, PEB board), they may be calibrated
 1779 with a good accuracy. For instance, at the ASIC level, a preliminary study shows that by
 1780 computing the t_0 online, a 20 ps (50 ps) contribution can be reached at low (high) radius
 1781 for periodic effects with a time period beyond 20 ms. The final calibration will need an
 1782 additional offline calibration combining the information from many calibration windows.

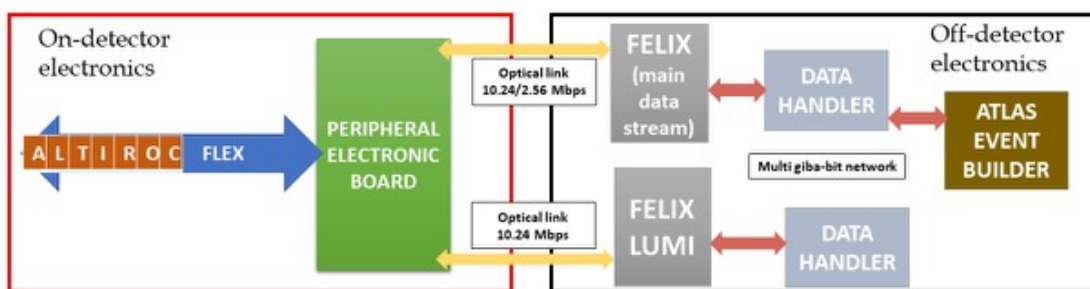


Figure 4.10: Data transmission paths for the main stream and the luminosity stream.

1783 4.5.3 Luminosity

1784 NEED TO FILL TEXT.....

1785 4.6 Power distribution and detector control system

1786 4.6.1 HV system

1787 A schematic layout of the high voltage system is show in Fig. 8.1 and detailed in Chap. 8.
1788 Each of the 8032 modules require a bias voltage in a range from approximately 300 V, at
1789 the start of the HL-LHC, up to 750 V, after the detector has been exposed to the expected
1790 maximum irradiation levels of $2.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, as detailed in Chap. 5. The irradiation
1791 of each module will strongly depend on its radial position in the detector, seen in Fig. 4.2,
1792 with an expected maximum variation of < 15% inside each module. The ultimate goal is
1793 to use individual adjustable voltages for each module, to allow for optimal operation. As
1794 a compromise of costs and performance, at least at the start of the HL-LHC, two to three
1795 modules (instead of one) will be connected to one HV channel in the outer region of each
1796 ring that have similar radial position.

1797 This merging can be done either in USA15 services cavern, where the HV power supplies
1798 will be located or in the patch panels region , exact location still to be decided. All HV cables
1799 will be routed from the beginning to allow in a later stage, to feed one module from one
1800 HV channel. Consequently HV power supplies that deliver up to 800 V and up to 6 mA
1801 current are needed in order to feed simultaneously two modules. The power supplies will
1802 be based on commercial multi-channel rack mounted units. Monitoring the leakage current
1803 and the TOT as an indicator of the collected charge will give a good estimate of the sensor
1804 gain evolution during data taking, allowing to perform the necessary HV adjustments.

1805 4.6.2 LV system

1806 A schematic layout of the LV voltage system is shown in Fig. 8.2 and detailed in Chap. 8. The
1807 low voltages needed by the front-end and peripheral electronics will be able to deliver almost
1808 20 kW and will be provided in a three stage system. Bulk power supplies, located in USA15,
1809 will provide 300 V DC current to DC-DC converters to be placed in the patch panel areas(PP-
1810 EC), located around the end-cap calorimeter outer radius surface, and accessible during
1811 technical stops and shutdowns. The second-stage multi-channel DC-DC units convert 300 V
1812 to 10 V which is distributed to radiation hard DC-DC converters, located on the peripheral
1813 electronics boards. The last stage converts the power to the front-end electronics (ASICs)
1814 and the peripheral electronics boards providing mainly 1.2 V DC power but also 2.5 V for

1815 the optical receivers/transmitters . The converters of the peripheral boards are based on the
1816 bpol12V, being developed by CERN for the HL-LHC upgrades.

1817 4.6.3 Monitoring and Controls

1818 A Detector Control System (DCS) will be implemented to control and monitor the various
1819 detector parameters: the power (HV, LV) supplied to the detector; the temperatures of the
1820 modules and of the peripheral electronics, the cooling system and the pressure of the N₂. It
1821 provides the tools to monitor the operational parameters of the detector, to bring the detector
1822 into any desired operational state, and to signal any abnormal behaviour by allowing for
1823 manual and automatic actions. More details are given in Chap. 8.

1824 4.7 Mechanics, Services and Infrastructure

1825 The detector mechanics and services were designed taking into account the severe constraints
1826 of space to accommodate the detector and the services that need to be routed in the gap
1827 between the barrel and end-cap calorimeters, sharing the space with ITk and the Tile
1828 calorimeter crack counters. The use of light structures were prioritized to minimize the
1829 amount of material in front of the active layers and minimize the potential increase in the
1830 radiation levels, leading to a total detector weight per end-cap of approximately 350 kg
1831 (275 kg without the external moderator).

1832 The hermetic vessel provides a robust support structure to the detector disks in a cold and dry
1833 volume, with radial dimensions of $100 \text{ mm} < r < 1000 \text{ mm}$. It has four main components:
1834 the front and back covers, the inner ring and the outer ring (which will hold all the service
1835 feedthroughs), as illustrated in Fig. 2.4. The front cover is divided in two half disks to allow
1836 it's manipulation in the presence of the beam pipe. It consists of a honeycomb core placed
1837 between two thin carbon fibre reinforced panels to reduce deflection. The thickness of front
1838 and rear covers are 13 mm and 7 mm respectively. To avoid condensation in the external
1839 face of the HGTD vessel during operation, heaters will be placed on the external face of the
1840 front cover, insuring a minimal temperature of 20 °C outside the HGTD vessel. An air gap of
1841 3 mm will be kept between the HGTD detector and the end-cap LAr calorimeter.

1842 Each double-sided layer (two per end-cap) is divided in two half circular disks of 30 kg each
1843 with 120 mm inner radius and 980 mm outer radius. Their shape allow for a completion, in
1844 case of delays, in the ATLAS detector even when the beam pipe is in place, provided that the
1845 back vessel covers and moderator are installed in LS3 when the beam pipe is not in place.
1846 The detector concept should facilitate rapid and safe removal of the detector to the surface
1847 in the high radiation environment. This operation is envisaged at each long shutdown of the
1848 HL-LHC for the replacement of the innermost or middle rings. The rotation of the two disk

1849 layers inside the vessel by 20° with respect to each other, as seen in Fig. 11.15, allows for a
1850 better integration of the cooling pipes inside the vessel while minimising the regions with
1851 zero hits resulting from the dead zones between the staves and imperfect coverage in the
1852 inner most radius.

1853 The expected maximum power consumption of the detector, to operate at -35°C and to
1854 reach the required performance, amounts to 39.3 kW in total (19.7 kW per end-cap); details
1855 of the various components are summarized in Tab. 11.1. An evaporative CO_2 cooling system
1856 of (50 kW will be used and part of its infrastructure and cooling spare unit will be shared
1857 with ITk.

1858 The evaluation of the amount of services required to operate the detector, summarized in
1859 Tab. 12.1, and respective routing design was subject to a careful evaluation and optimisation.
1860 This is due to the limited space in the detector vessel outer ring allocated to the services
1861 feedthroughs, limited space in the barrel-end-cap calorimeter gap region and, last but not
1862 least, in the ATLAS flexible chains that allow maintaining part of the services connected
1863 during the end-cap calorimeters opening and closure. The detector services routing is shown
1864 in Fig. 4.11 for the calorimeter extended barrel face. The cables, exiting in four layers in the
1865 feedthroughs region, will pass to one layer at $r > 1.3$ m to fit within an envelope of 17 mm.
1866 At the outer radius of the calorimeter, services are routed in various layers in z but narrow
1867 slots in ϕ to pass in between the Tile fingers, a space also shared with ITk services. The
1868 exception will be a dedicated slot in ϕ , on the top of the calorimeter, to be given from the ITk
1869 original envelope, to route the four CO_2 cooling pipes of 50 mm diameter maximum each
1870 pipe. The priority for services installation in flexible chains, still to be confirmed, will be
1871 given to optical fibres, cooling pipes, interlock and cooling temperature sensor cables. The
1872 other services need to go through fixed cable trays and should be disconnected before the
1873 extended barrel calorimeters are moved for maintenance of the ATLAS detector. For that
1874 purpose the patch panels (PP-EC) will be organised on the end-cap Tile calorimeter outer
1875 surface in accessible places. The patch panel boxes will be also used for re-mapping the
1876 cables to match connectors on the detector.

1877 4.8 Assembly, Installation and Commissioning

1878 The final assembly of the detector and quality assurance, e.g. mounting the modules support
1879 frames and peripheral electronics boards into the half circular disks, connecting each flex
1880 cable to the respective peripheral electronics boards, and global certification, should take
1881 place at CERN with the participation of several collaborating Institutes. After the assembly,
1882 the detector will be transported to the pit. Each end-cap, HGTD A and HGTD C, will be
1883 lowered on side A and side C respectively and lowered directly from the surface to the
1884 minivans. The final installation of the detector should take approximately 1 month per
1885 end-cap and is planned for April 2026 (HGTD A) and January 2027 (HGTD C).

1886 Dedicated tools are needed for assembly, lowering, and final installation of the detector.
1887 These tools are still at a conceptual stage and where possible will use synergies with already
1888 developed tools for other sub-detectors.

1889 The overall commissioning will start immediately after the connectivity of the services to
1890 the detector. The access to the detector components during the commissioning should be
1891 possible until approximately May 2026, close to the expected end-cap calorimeters closure.
1892 This will leave at least 6 months of intense commissioning while access is still possible. Both
1893 the installation and commissioning of HGTD will be done with the participation of several
1894 collaborating Institutes.

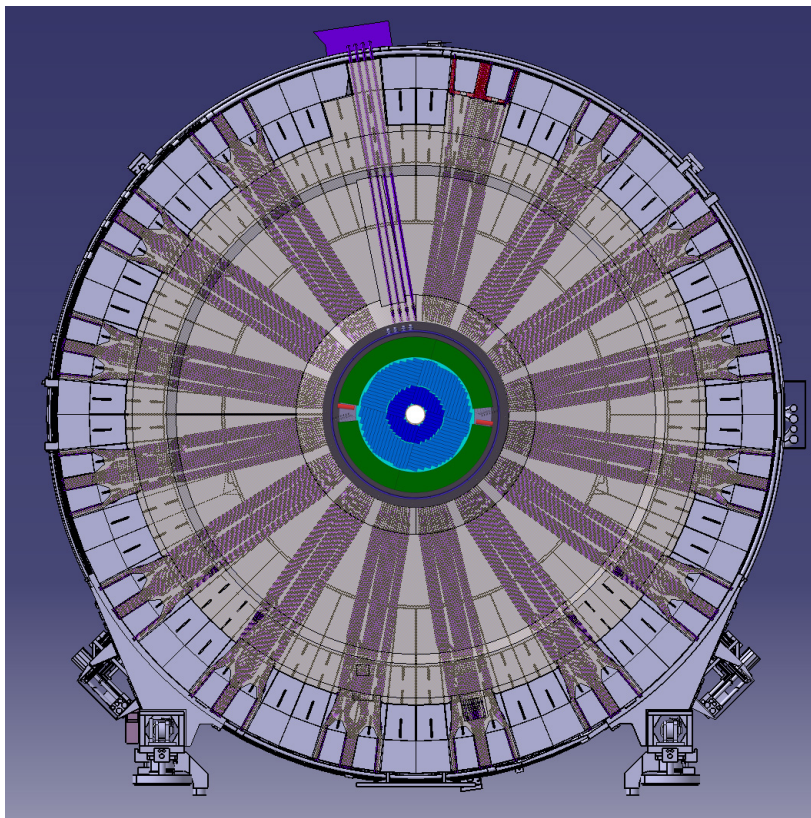


Figure 4.11: Transverse view of HGTD with services routed along the end-cap calorimeter face.
UPDATE THIS FIGURE WITH LATEST VERSION BY VLADIMIR/SERGEY

1895 **4.9 Next steps towards construction**

1896 NEED TO FILL TEXT.....

1897 5 Sensors

1898 5.1 Sensor parameters and requirements

1899 The HGTD sensor parameters and requirements are summarized in Tab. 5.1. The sensors are
1900 intended to provide a fast signal in response to charged particles for a time resolution per
1901 hit of 40 ps at the start and 70–85 ps at the end of lifetime (combined performance with the
1902 electronics and other contributions). The charge should be at least 4 fC and the hit efficiency
1903 at least 95%. The granularity should be 1.3 mm × 1.3 mm and the physical thickness below
1904 300 μm. The sensor should be of total active size of 39 mm × 19.5 mm with 30 × 15 pads and
1905 bump-bonded to two readout chips (ALTIROC) of 15 × 15 pads. The inactive edge around
1906 the sensor should be less than 500 μm. In the baseline scenario, discussed in Chap. 4, the
1907 innermost part of the detector ($r < 230$ mm) should be replaced after each 1000fb^{-1} and the
1908 middle ring within $470\text{ mm} > r > 230$ mm should be replaced at half lifetime (2000fb^{-1})
1909 of data-taking during the HL-LHC program. The sensors are then required to sustain a
1910 1 MeV-neutron equivalent particle fluence of maximally $2.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ and a TID of
1911 2.0 MGy, including safety factors.

1912 The leakage current should be less than 5 μA per pad, the applied bias voltage less than 800 V
1913 and the power density less than 100 mW/cm² at an operation temperature of maximally
1914 −30 °C on-sensor. The technology chosen for the HGTD sensors is Silicon Low Gain Ava-
1915 lanche Detectors (LGAD) with a baseline active thickness of 50 μm. The target gain (charge)
1916 is 20 (10 fC) at the start and at least 8 (4 fC) at the end of lifetime.

1917 5.2 Low Gain Avalanche Detectors

1918 5.2.1 Overview

1919 LGADs are segmented planar Silicon detectors with internal gain as sketched in Fig. 5.1.
1920 The gain depends on the doping dose of the multiplication layer as seen in Fig. 5.2 and
1921 diminishes with radiation fluence as shown in Sec. 5.5.3. They have been pioneered by the
1922 Centro Nacional de Microelectronica (CNM) Barcelona [4] and developed during the last 5
1923 years within the CERN-RD50 community [3] also in collaboration with other LGAD vendors
1924 as Hamamatsu Photonics (HPK, Japan) and Fondazione Bruno Kessler (FBK, Italy). An

Technology	Silicon Low Gain Avalanche Detector (LGAD)
Time resolution $2.4 < \eta < 2.7$	≈ 40 ps (start); ≈ 70 ps (end of lifetime)
Time resolution $2.7 < \eta < 3.5$	≈ 40 ps (start); ≈ 70 ps (end of lifetime)
Time resolution $3.5 < \eta < 4.0$	≈ 40 ps (start); ≈ 70 ps (end of lifetime)
Gain	≈ 20 (start); > 8 (end of lifetime)
Minimal charge	4 fC
Hit efficiency	$>95\%$
Granularity	1.3 mm \times 1.3 mm
Physical thickness	<300 μm
Active thickness	50 μm
Active size	39 mm \times 19.5 mm (30 \times 15 pads)
Inactive edge	<500 μm
Radiation tolerance	2.5×10^{15} n _{eq} cm ⁻² , 2.0 MGy
Maximum operation temperature on-sensor	-30 °C
Maximum leakage current per pad	5 μA
Maximum bias voltage	800 V
Maximum power density	100 mW/cm ²

Table 5.1: Sensor parameters and requirements.

1925 introduction of the technology is given in Chap. 4. Additional background and details are
 1926 given in Reference [32].

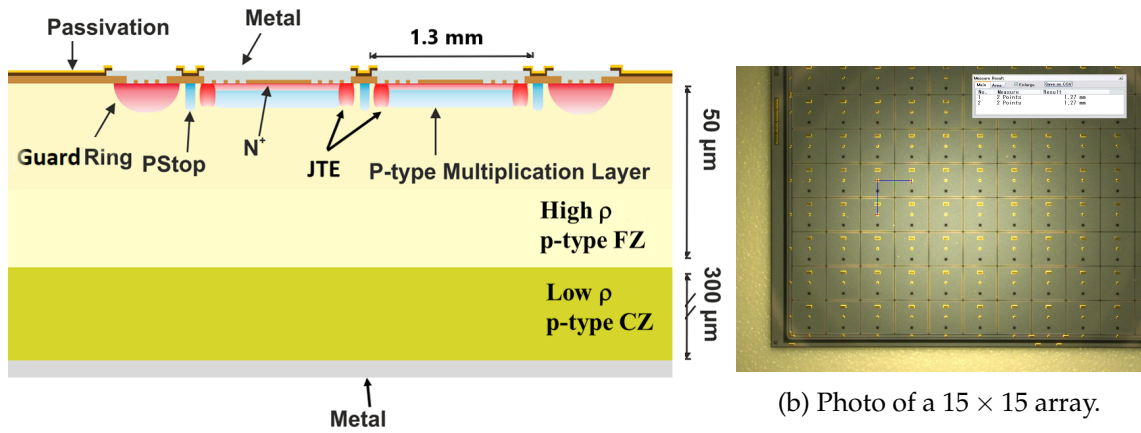


Figure 5.1: (a) Cross section of a 2 \times 2 array including a JTE around each sub-pad (SiSi wafer, CNM design) [33]. (b) Microscope photo of an HPK-3.1 15 \times 15 array.

1927 Three major effects determine the time resolution: time walk from amplitude variations,
 1928 jitter from electronic noise and “Landau fluctuation” from charge deposition uniformities
 1929 along the particle path. Time walk and noise jitter depend on the type of readout electronics
 1930 chosen. Both depend inversely on the signal slope (voltage slope at the output of the

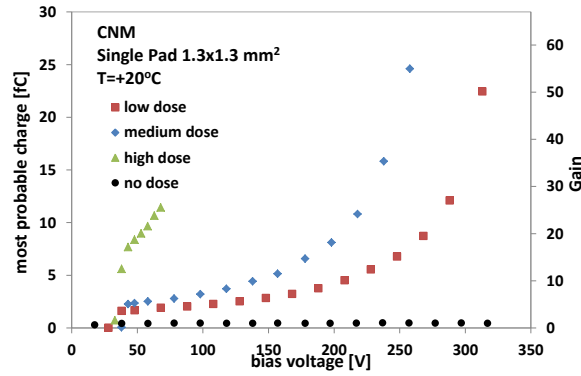


Figure 5.2

Figure 5.3: Gain and charge as a function of bias voltage for CNM LGADs with different doping concentration of the multiplication layer.

1931 amplifier) dV/dt :

$$\sigma_{\text{TimeWalk}} = \left[\frac{V_{\text{th}}}{S} \right]_{\text{RMS}}, \quad \sigma_{\text{Jitter}} = \frac{N}{(dV/dt)} \simeq \frac{t_{\text{rise}}}{(S/N)}, \quad (5.1)$$

1932 where S refers to the signal which is proportional to the gain, N to the noise, t_{rise} to the
 1933 rise time and V_{th} to the threshold voltage. It can be seen that the lowest noise jitter and
 1934 time walk are achieved with sensors with high signal-to-noise ratio (S/N) and small rise
 1935 time, i.e. with thin sensors and large gain. Time walk can usually be corrected for by a
 1936 large extent using time reconstruction algorithms such as constant-fraction discrimination
 1937 (CFD) or amplitude or time-over-threshold (ToT) corrections. The third effect called “Landau
 1938 fluctuation” is due to the non-uniform charge deposition along the particle path leading to
 1939 time-of-arrival fluctuations. It is a contribution depending on the thickness of the sensor
 1940 (thin is beneficial) and the setting of the threshold. Adding the three contributions in
 1941 quadrature yields the overall time resolution. After time-walk correction, the noise jitter is
 1942 the dominating contribution for low S/N and the Landau term takes over for high S/N .

1943 An example for a measured LGAD time resolution is shown in Fig. 4.5 as a function of
 1944 gain, with the time walk corrected using CFD. As expected from Eq. (5.1) the resolution
 1945 improves with increasing gain due to the reduced noise jitter, but then levels off to the
 1946 Landau fluctuation of about 30 ps for 50 μm thickness.

1947 This observation feeds into the plan to operate LGADs at a gain of about 20 before irra-
 1948 diation and as close as possible to that value after irradiation given restrictions from the
 1949 leakage current, the breakdown voltage, and the noise, including the excess noise from the
 1950 multiplication process. The gain target of 20 was chosen since the time resolution fulfills
 1951 already the HGTD requirement of 40 ps per hit at the start of operation (see Tab. 2.1) and

1952 is improving only slowly when going to higher gains as seen in Fig. 4.5. Moreover, the
1953 maximum achievable gain reduces after irradiation, hence an optimisation of the detector to
1954 higher gains before irradiation would only benefit a short period at the start of operation.
1955 At high fluences, operation at charges down to 4 fC corresponding to a gain of 8 becomes
1956 necessary (see Sec. 5.5.3).

1957 The field in the Silicon bulk (i.e. no-gain) region should be high enough ($1 \times 10^4 \text{ V cm}^{-1}$) to
1958 saturate the drift velocity of about $100 \mu\text{m ns}^{-1}$ for a reduced rise time.

1959 An LGAD active thickness of $50 \mu\text{m}$ has been adopted as the best compromise between
1960 capacitance and deposited charge (favouring a large thickness) and signal slope and Landau
1961 fluctuations (favouring a small thickness). LGADs of $30 \mu\text{m}$ active thickness have been
1962 studied as an option in the past and showed a better sensor-only performance before
1963 irradiation, but were discarded due to the higher capacitance and higher power dissipation
1964 at similar performance after irradiation compared to $50 \mu\text{m}$. Such small active thicknesses
1965 are usually achieved by different techniques that all use a thin active high resistivity layer
1966 on top of a thicker insensitive Silicon substrate of low resistivity, such as Silicon-on-Insulator
1967 (SOI), Silicon-Silicon Wafer Bonding (SiSi) or epitaxial (Epi) wafer techniques.

1968 Fig. 5.1(a) shows the cross section of a 2×2 LGAD array. Each pad consists of the p-type
1969 multiplication layer underneath the n^+ implantation, surrounded by a Junction Termination
1970 Extension (JTE). The JTE is an n^+ implantation that is deeper than the one of the central
1971 pad. It controls the electric field at the edges to avoid early breakdown, but also leads to an
1972 inter-pad gap with no or reduced gain and hence worse time resolution and hit efficiency in
1973 this region. The complete sensor is surrounded by a guard ring (GR). Fig. 5.1(b) shows a
1974 photo of an HPK 15×15 array.

1975 As a dopant for the p-type multiplication layer, Boron (B) is typically used. Additional Car-
1976 bon (C) implantation or the substitution of B by Gallium (Ga) are investigated as candidates
1977 for improved radiation hardness.

1978 5.2.2 LGAD productions

1979 At present, LGADs have been produced in six manufacturing sites, shown in Tab. 5.2
1980 along with their production capabilities: Hamamatsu Photonics (HPK), Japan; CNM, Spain;
1981 Fondazione Bruno Kessler (FBK), Italy; Micron, UK; Brookhaven National Lab (BNL), USA;
1982 and NDL, China. Further vendors are interested in LGAD productions.

1983 There are plans to use LGADs in three experiments at the HL-LHC (ATLAS, CMS, LHCb).
1984 There has been fruitful collaboration and coordination between ATLAS-HGTD and CMS-
1985 ETL [34] with respect to simulations, design, manufacturing and testing.

1986 The design and production of LGADs for HGTD had two distinct phases: an early R&D
1987 phase of about 6 years with much of the activities carried out within the RD50 collaboration

where the basic parameters were investigated and the suitability of LGADs for large scale application has been determined. The different manufacturers tended to concentrate on different parameters (like multiplication layer doping profile and dose, variation of the types of dopant, thickness). In general, the LGAD sensors produced by different manufacturers appear to perform similarly, with the exception of the leakage current before irradiation, and the bias voltage reach after irradiation.

In the second phase into which the collaboration is now entering the focus will be geared towards the production of sensors for HGTD specific application once the sensor requirements are better understood, and thus the options are reduced. For example, the decision to fix early on the pitch of the pads in the detector arrays to 1.3 mm provided a needed stable ground so that the development of other parts of the detector (electronics, modules, mechanical layout) could proceed. At this point, the need to investigate issues of manufacturing (yield, uniformity, large arrays, fill-factor, under-bump-metalization (UBM¹), etc.) and operations (bias voltage, power, reliability, breakdown) have become more important.

Manufacturer	Wafer Size [inch]	Thickness [μm]	Ga Implant	C Implant	Array 5 \times 5	Array 15 \times 15	Array 30 \times 15	UBM
CNM	4-6	30 - 300	x	x	x	(x)	(x)	
FBK	6	60 - 300	x	x	x			
HPK	6	20 - 80			x	x	(x)	x
BNL	4	50						
Micron	4	100 - 300						
NDL	6	33			x	x		

Table 5.2: LGAD manufacturers and production capabilities achieved to-date. Crosses in brackets (x) are for ongoing runs.

The results in the following have been mainly obtained from the LGAD types shown in Tab. 5.3. These runs include LGAD sensors of HGTD geometry. Many more runs not mentioned here have been studied in addition for R&D purposes. Typically in a run there are sensors of varied nominal inter-pad gaps (IP) or slim edges (SE).

5.3 Radiation damage and irradiations

As explained in Sec. 2.4, the detector has to withstand a 1 MeV neutron equivalent particle fluence of maximally $2.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, assuming one replacement of the inner part after half of the total integrated luminosity of 4000 fb^{-1} . In the innermost region, the radiation field is roughly equal for neutrons and charged hadrons (Fig. 2.13), but the contribution by charged hadrons decreases steeply with radius, so that the field is dominated by neutrons in the outer regions due to backscatter from the calorimeters. The energy spectrum of protons and pions roughly peak between 50 MeV and 10 GeV, whereas the neutron spectrum peaks

¹ UBM is part of the hybridisation process as explained in Sec. 7.2.1.

Manufacturer	Name	Thickness [μm]	Gain layer dopant	C implant	Gain layer depth [μm]	Gain layer depletion [V]
HPK	HPK-3.1	50	Boron	No	1.6	40
HPK	HPK-3.2	50	Boron	No	2.2	55
FBK	FBK-UFSD3-C	60	Boron	Yes	0.6	20
CNM	CNM-AIDA1/2	50	Boron	No	1.0	45
NDL	NDL-BV60	33	Boron	No	1.0	20
Manufacturer	Name	Full depletion [V]	V_{BD} -30 °C [V]	Nominal IP [μm]	Nominal SE [μm]	Max. Array Size
HPK	HPK-3.1	50	200	30→95	200→500	15 × 15
HPK	HPK-3.2	65	70	30→95	200→500	15 × 15
FBK	FBK-UFSD3-C	25	170	37	200→500	5 × 5
CNM	CNM-AIDA1/2	50	220/90	37→57	200→500	5 × 5
NDL	NDL-BV60	35	70	55	450	15 × 15

Table 5.3: Design, geometrical and electrical properties of LGAD types.

2014 at about 1 MeV, but has large contributions over a large range from 0.1 eV to 100 MeV (see
2015 appendix A).

2016 Radiation damage in Silicon mainly results in the change of the effective doping concen-
2017 tration, the introduction of trapping centers that reduce the mean free path of the charge
2018 carrier, and the increase of the leakage current [3]. For LGADs, one of the main effects is
2019 the degradation of gain with fluence at a fixed voltage due to removal of initial acceptors in
2020 the multiplication layer [35, 36], which implies the need to increase the applied bias voltage
2021 after irradiation to at least partly compensate for this.

2022 To study the LGAD performance after irradiation, sensors have been irradiated up to fluences
2023 of $6 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ at various facilities with different particle types and energies that are
2024 representative for the ones expected in HGTD. Tab. 5.4 gives an overview on the facilities,
2025 their parameters and maximum fluences as well as Total Ionising Dose (TID) achieved for
2026 different LGAD types irradiated. The hardness factor is used for the conversion of the actual
2027 particle fluence to the 1 MeV-neutron equivalent fluences, which is used throughout this
2028 document.

2029 First prototypes were irradiated in all facilities except for CYRIC, and it was found that
2030 acceptor removal seems to be faster after irradiations with 200 MeV–23 GeV charged hadrons
2031 than with neutrons [35, 36]. However, CERN PS is in shutdown now until 2021 and cam-
2032 paigns in Los Alamos are still ongoing. Hence results for LGADs with the HGTD geometry
2033 presented here are mostly after irradiations with neutrons at Ljubljana and 70 MeV protons
2034 at CYRIC due to the sites' availability. As will be shown below (see e.g. Sec. 5.5.3), the
2035 performance of samples irradiated at these two sites at similar 1 MeV-neutron equivalent
2036 fluences is similar, in contrast to the earlier results using higher energy charged hadrons.
2037 These studies will be followed up by irradiations with higher energy charged hadrons
2038 at PSI and Los Alamos when these facilities become available again in 2020. Also mixed
2039 neutron-proton irradiations are planned.

2040 It should be noted that irradiations at CYRIC with 70 MeV protons led to a maximum TID of
 2041 4.0 MGy, close to the HGTD requirement of 2.0 MGy. To study in more detail the effect of
 2042 TID such as changes in the surface conditions, presently there are irradiations with X-rays
 2043 under way at IHEP.

2044 The measurements with irradiated sensors were done after annealing for 80 min at 60 °C, if
 2045 not noted otherwise. Dedicated annealing studies are presented in Sec. 5.5.7.

Facility & Abbreviation	Particle Type	Hardness Factor	TID [MGy] / $10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$	Max. Fluence [$10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$]	Max. TID [MGy]	LGAD Types Irradiated
JSI Ljubljana (<i>n</i>)	$\approx 1 \text{ MeV n}$	0.9	0.01	6	0.06	all
CYRIC (<i>pCY</i>)	70 MeV p	1.5	0.81	5	4.0	HPK-3.1/3.2, NDL FBK-UFS3-C
Los Alamos (<i>pLA</i>)	800 MeV p	0.7	0.43	1	0.4	early prototypes, HPK-3.1
CERN PS (<i>pPS</i>)	23 GeV p	0.6	0.44	6	2.7	early prototypes
PSI (<i>pi</i>)	192 MeV pions	1	0.32	2	0.6	early prototypes

Table 5.4: Irradiation facilities and parameters and maximum achieved fluence and TID, as well as LGAD types irradiated.

2046 5.4 Sensor tests: methodology and experimental techniques

2047 The LGAD sensors have been tested before and after irradiation by various HGTD groups,
 2048 as well as within the RD50 community.

2049 Electrical measurements including capacitance-voltage (C-V) and current-voltage (I-V) char-
 2050 acteristics have been performed in laboratory probe stations. For the probing of large arrays,
 2051 custom-made probe cards for the simultaneous contact of 5×5 pads have been developed.
 2052 For the measurement of larger arrays like the 15×15 single-chip sensor, the probe card
 2053 is applied subsequently to 5×5 sub-blocks. A probe card with 15×15 contacts is under
 2054 development. An alternative is the subsequent probing of one single pad after another on
 2055 a semi-automatic probe station that allows to scan over an arbitrary number of pads in an
 2056 array, while the neighbouring pads and the guard ring are floating.

2057 The dynamic properties of LGADs, such as charge collection, gain and time resolutions,
 2058 have been measured in response to ionising particles, both in the laboratory with ^{90}Sr β
 2059 particles [32, 35–41] and lasers, as well as in beam tests [37, 38, 42]. Beam tests have been
 2060 performed by the HGTD community in more than ten periods between 2016 and 2019 at the
 2061 H6 beam line of the CERN SPS with 40 to 120 GeV pions, at SLAC with 15 GeV electrons, at
 2062 FermiLab with 120 GeV protons, and at DESY with 5 GeV electrons [42]. Data were taken in
 2063 two modes: stand-alone and integrated into a beam telescope that provided track position
 2064 information with about $3 \mu\text{m}$ precision.

2065 Most of the measurements on irradiated sensors were performed at a temperature of $-30\text{ }^{\circ}\text{C}$,
2066 the lowest temperature reachable in standard laboratory climate chambers.

2067 The dynamic measurements in the laboratory and beam tests were all obtained using custom-
2068 made HGTD-specific readout boards with an integrated high bandwidth amplifier with a
2069 gain of about 10, followed by a second commercial 2 GHz amplifier of gain 10, allowing
2070 the recording of the pulse shape of the fast LGAD signals [37, 42] with a high bandwidth
2071 oscilloscope (1–2.5 GHz). The noise was measured as the RMS fluctuation of the base line
2072 of the oscilloscope trace. It typically amounts to 1.6 mV–2.5 mV (roughly corresponding
2073 to a charge of 0.12 fC–0.20 fC) depending on the type and vertical scale of the oscilloscope,
2074 the board type, and the physical location. Measurements at test beam facilities tend to be
2075 noisier than laboratory measurements since machinery and magnets are operated in the
2076 same areas. This reflects the performance of the sensors with discrete electronics optimized
2077 for precision timing. It should be clearly noted that measurements with the ALTIROC are
2078 not part of this section since the chip has not been available yet for large-scale sensor testing.
2079 The measurements presented here will be repeated with the ALTIROC as soon as enough
2080 chips are available. First measurements of the combined sensor-ALTIROC performance on
2081 few bump-bonded hybrid prototypes are presented in Sec. 6.7.2.

2082 Position-sensitive scans using red and infrared laser to deposit charge carriers inside the
2083 sensors have been made at various institutes, using the Transient Current Technique (TCT)
2084 setup.

2085 The gain is extracted by dividing the collected charge in an LGAD device by the charge of
2086 no-gain PIN diodes of the same thickness without multiplication layer (for β s and MIPs of
2087 about 3 ke^- or 0.5 fC for $50\text{ }\mu\text{m}$ thickness).

2088 Time resolutions are typically extracted from the spread of the time-of-arrival difference
2089 between two sensors when a particle passes through both. Either at least two LGADs are used
2090 or LGADs and a fast Cherenkov counter based on quartz bars and a Silicon photo multiplier
2091 (SiPM) with typical time resolution of about 10 ps. If at least three devices are measured
2092 simultaneously, a χ^2 minimisation is used to obtain the time resolution of all devices. In
2093 case only one device under test (DUT) is measured with respect to one reference device of
2094 known resolution, the DUT resolution is obtained by subtracting quadratically the reference
2095 contribution. Timewalk effects are usually corrected for using time reconstruction algorithms
2096 such as the Constant-Fraction Discriminator (CFD), the Zero-Crossing Discriminator (ZCD)
2097 or corrections using the amplitude or Time-Over-Threshold (TOT) of the signal [42].

2098 LGAD behavior such as time resolution and collected charge was simulated using the
2099 software WeightField 2 [43]. The simulation were tuned using laboratory measurements
2100 from different sensor types and are extrapolated to foresee future improvements as seen in
2101 Sec. 5.8. Also, the software TCAD sentaurus [44] was used in aid of the sensor production.

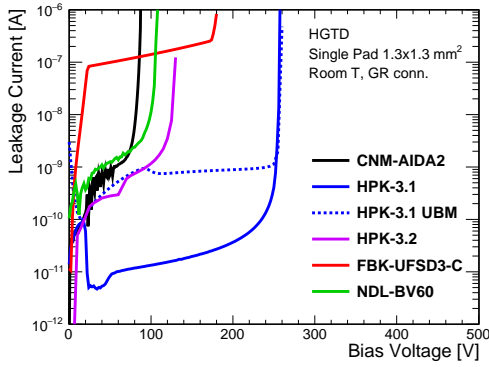
2102 5.5 LGAD performance before and after irradiation

2103 5.5.1 Electrical characterisation: I-V and C-V

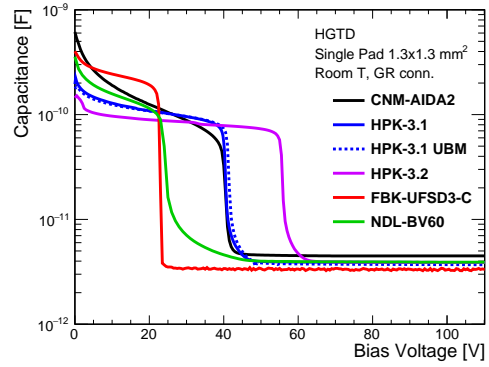
2104 Fig. 5.4(a) and Fig. 5.4(b) show the I-V and C-V curves of $1.3\text{ mm} \times 1.3\text{ mm}$ LGAD pads of
2105 different vendors and runs, measured with the guard ring (GR) connected to ground. Most
2106 of the vendors and runs achieve nA leakage current levels or below before breakdown, well
2107 below the ALTIROC leakage current limit of $5\text{ }\mu\text{A}$ per pad. The addition of the UBM process
2108 at HPK in this prototype run led to an increased leakage current by 2 orders of magnitude
2109 to about 1 nA with respect to wafers without UBM, which is still safe for operation and
2110 expected to improve in future productions. No influence on the C-V behavior was found.
2111 The FBK-UFSD3-C sensors with Carbon exhibits currents of about 100 nA , which are higher
2112 than HPK Boron-only sensors but are still safely below the ALTIROC limit. After irradiation,
2113 the currents of FBK-UFSD3-C become more similar to the other types. The breakdown
2114 voltage increases with decreasing multiplication layer dose.

2115 Also the range of the "foot" of the C-V curve (i.e. the voltage region where C stays at high
2116 values while the multiplication layer is being depleted, starting from the n-p junction at the
2117 front) is an indicator of the multiplication layer dose. Foot values between 20 V and 60 V
2118 indicate substantial gains, as verified below. The depletion of the bulk (indicated by the
2119 sharp fall of the C-V curve) happens rather fast within a few V due to the high resistivity and
2120 the small thickness. The end capacitances of about 3 pF – 4 pF for $1.3\text{ mm} \times 1.3\text{ mm}$ LGAD
2121 pads (measured with a connected guard ring) are consistent with active thicknesses of
2122 $40\text{ }\mu\text{m}$ – $60\text{ }\mu\text{m}$.

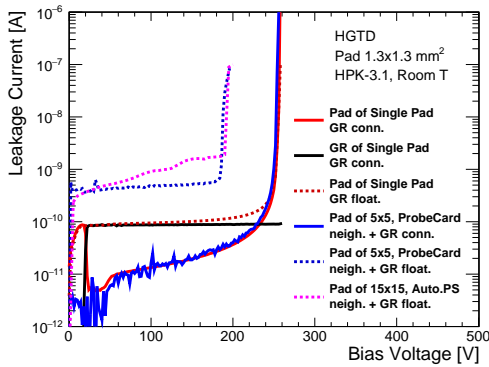
2123 Fig. 5.4(c) shows the I-V curves for HPK-3.1 sensors of the LGAD pad and guard ring (GR)
2124 with either GR connected to ground (as the pad) or floating. For the single pad sensor,
2125 it can be seen that the current through the pad in case of floating guard ring is roughly
2126 the sum of pad and guard ring current in case the guard ring is connected. However, the
2127 breakdown voltage, V_{BD} , where the current increases rapidly, is found not to be affected
2128 by the GR biasing condition for single pads. For a pad in an HPK-3.1 array, the I-V curve
2129 is found to be almost identical to the one of a single pad in case the neighbors and the
2130 guard ring are connected to the same potential, as measured with a 5×5 probe card on a
2131 5×5 array (see Fig. 5.4(c) and Fig. 5.5(a)). However when leaving neighboring pads and
2132 GR floating, the current level is increased by 2 orders of magnitude (presumably due to
2133 punch-through to the neighbors) and V_{BD} is observed to be reduced from about 250 V to
2134 about 190 V , consistently measured with a probe card when connecting only one channel
2135 and an automatic probe station with only one needle (see Fig. 5.4(c)). It should be noted that
2136 this behavior of shifting V_{BD} in case of floating neighbors and GR was not observed for the
2137 5×5 arrays of the CNM-AIDA run. This indicates that it depends on the sensor design and
2138 the exact production process. Probing with an automatic probe station turned out to be a
2139 powerful tool to identify individual faulty pads inside an array and is so far the only method



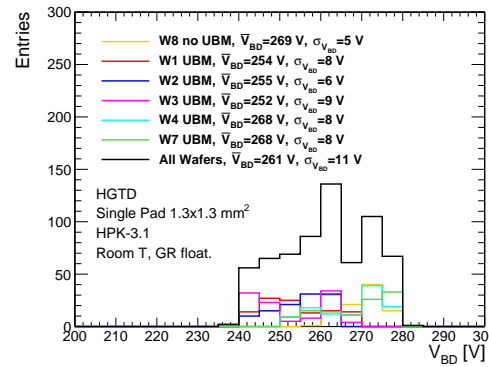
(a) I-V for different runs.



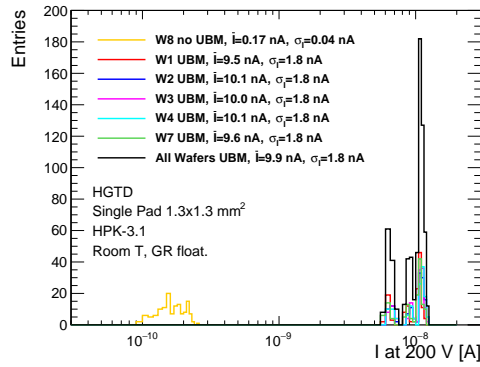
(b) C-V for different runs.



(c) I-V for different types and conditions.



(d) V_{BD} distribution.



(e) Current at 200 V distribution.

Figure 5.4: Measurements of current-voltage I-V (a) and capacitance-voltage C-V (b) characteristics comparing different vendors and runs, as well as device types and biasing conditions (c). (d) and (e) show the distributions of V_{BD} and the current at 200 V for single pads of different wafers of HPK type 3.1 (with and without UBM).

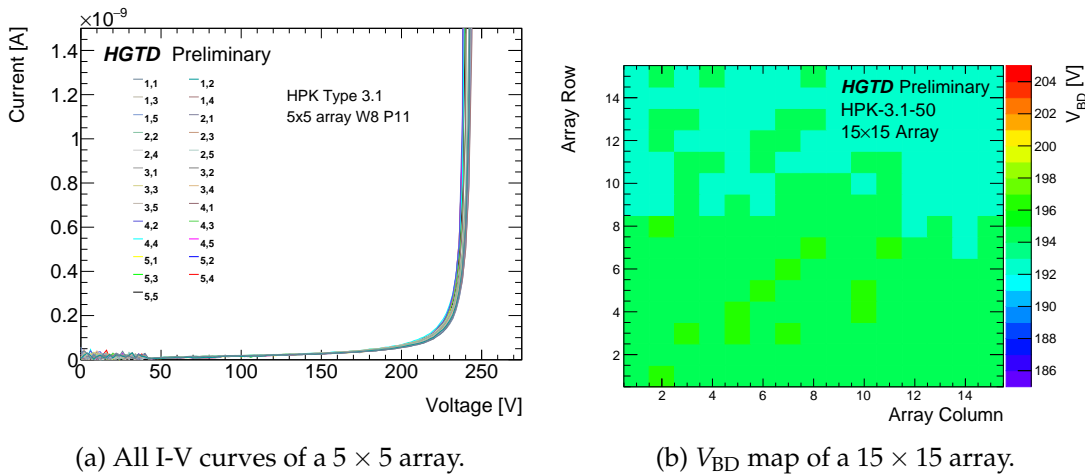


Figure 5.5: (a) I-V measurement of 25 pads from an unirradiated HPK-3.1 5×5 array without UBM measured with a 5×5 probe card at room temperature (all pads and GR grounded). (b) V_{BD} map of a 15×15 HPK-3.1 array without UBM measured with an automatic probe station at room temperature (neighbors and GR floating). **TODO: change with HPK 3.2?**

2140 to probe 15×15 arrays efficiently until the development of a 15×15 probe card is finished.
 2141 In most cases all pads inside an array behave uniformly for HPK-3.1 5×5 and 15×15 arrays
 2142 (see Fig. 5.5). The V_{BD} spread between pads in an array is found to be typically between 1 V
 2143 and 2 V.

2144 With the probe card, also the situation was studied that only one pad in the center of a 5×5
 2145 array was floating, while the other 24 pads and the GR were connected. This was to simulate
 2146 the behavior of a faulty pad that needs to be disconnected to make the sensor operable. The
 2147 floating of only one pad had an influence on the breakdown voltage of all other pads in the
 2148 array by introducing a shift to lower V_{BD} by less than 10 V and producing a more sudden
 2149 and steeper breakdown.

2150 HGTD institutes measured a large number of single pads and arrays from different pro-
 2151 ductions, in particular HPK-3.1/3.2 and CNM-AIDA1. **TODO: Add NDL/FBK?** HPK also
 2152 provides their in-house Quality-Control (QC) results with an automatic probe station (GR
 2153 floating) of each single pad they delivered. The HPK results have been verified by HGTD
 2154 institutes. Fig. 5.4(d) and Fig. 5.4(e) show the corresponding distributions of V_{BD} and the
 2155 current at 200 V for all HPK-3.1 single pads on different wafers, with and without UBM,
 2156 demonstrating a good uniformity. The mean of V_{BD} for all wafers is 261 V with a spread
 2157 of 11 V. The per-wafer spread varies between 5 V and 9 V. No single pad sensor has a V_{BD}
 2158 of less than 235 V or more than 285 V. For the current at 200 V, two distinct distributions
 2159 are found as expected from the results discussed above: one for sensors without UBM with
 2160 a mean of 0.17 nA, and one after applying UBM with a mean of about 10 nA (it should be
 2161 noted again that the GR was floating), the spread is found to be about 20%.

LGAD Type	Sensor Type	Nominal Edge [μm]	Nominal IP gap [μm]	Sensors tested	Pads tested	Fraction of Perfect Sensors [%]	Fraction of Good Pads [%]
HPK-3.1	Single	Sum all	95	648	648	100	100
		500	95	360	360	100	100
		300	95	144	144	100	100
		200	95	144	144	100	100
	2 × 2	Sum all	Sum all	13	52	100	100
		500	30	1	4	100	100
		300–500	50	2	8	100	100
		300–500	70	2	8	100	100
		200–500	95	8	32	100	100
	5 × 5	500	95	19	475	100	100
15 × 15	500	95	27	6075	85.2	99.5	
HPK-3.2	Single	Sum all	Sum all	216	216	100	100
		500	95	120	120	100	100
		300	95	48	48	100	100
		200	95	48	48	100	100
	2 × 2	Sum all	Sum all	26	104	100	100
		500	30	2	8	100	100
		300–500	50	4	16	100	100
		300–500	70	4	16	100	100
		200–500	95	16	64	100	100
	5 × 5	500	95	6	150	100	100
15 × 15	500	95	23	5175	91.3	99.8	
FBK-UFSD3	2 × 2	?	40?	x	x	x	x
	5 × 5	?	40?	x	x	x	x
CNM-AIDA1	Single	500	37	84	84	69	69
		500	47	39	39	95	95
		500	57	42	42	100	100
	5 × 5	500	37	6	150	50	66
		500	47	6	150	83	90
		500	57	6	150	100	100

Table 5.5: Number of tested devices and fraction of good pads and sensors for HPK-3.1/3.2 and CNM-AIDA1 of different sensor types, edge and inter-pad (IP) gap designs. An array of 15 × 15 pads corresponds to the final ALTIROC size and half of the full final sensor area. **TODO: update CNM? Add NDL/FBK?**

2162 Tab. 5.5 shows the fraction of good individual pads in single pads and arrays defined as
 2163 having a breakdown voltage above 90% of the expected one for the respective biasing
 2164 condition of GR and neighbors. Moreover, the fraction of perfect sensors is displayed, which
 2165 are defined by requiring all pads in a sensor to be good. For HPK, the fraction of good
 2166 pads turned out to be 99.7–100%. No dependence on the edge design between 200 μm and
 2167 500 μm edge was found. Only one HPK-3.1 and one HPK-3.2 15×15 HPK array with one
 2168 or two bad pads, respectively, were found, all other sensors were classified as perfect. For
 2169 CNM-AIDA1 the result was found to depend on the inter-pad gap parameter (IP): the largest
 2170 inter-pad gap (IP57) is found to give 100% good sensors and pads, which reduces to about
 2171 70% good pads and 50% perfect sensors for the smallest inter-pad gap (IP37).

2172 5.5.2 Operating bias Voltage and self-triggering

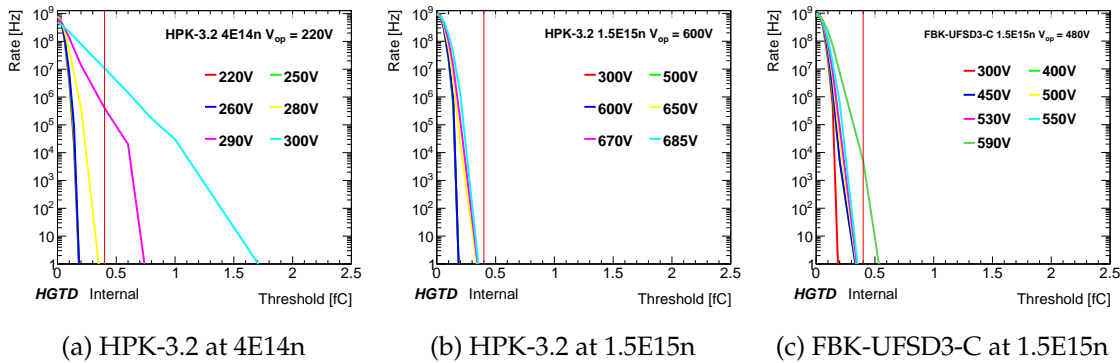


Figure 5.6: Self-trigger rate as a function of collected charge. Fig. 5.6(a): for HPK-3.2 sensor at $4 \times 10^{14} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ of neutron irradiation. Fig. 5.6(b): same sensor after $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ of neutron irradiation. Fig. 5.6(c): for FBK-UFSD3-C sensor at $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ of neutron irradiation. A threshold of 5 mV corresponds to roughly 0.4 fC of collected charge, represented by the red line in the plots. The operating voltage V_{op} for each is written in the legend, as shown no self-triggering is present at that V_{op} .

2173 As mentioned in Sec. 5.4, dynamic measurements in response to particles have been per-
 2174 formed in the laboratory and beam tests on custom-made HGTD-specific readout boards.

2175 The maximum applicable bias voltage plays a crucial role in determining the performance of
 2176 the sensors before and after irradiation, since the gain depends on the bias voltage, and this
 2177 dependence changes with irradiation. It is important to realize that for thin sensors the effect
 2178 of trapping is reduced due to the smaller electrode distances so that the charge collected
 2179 from the bulk before charge multiplication does not change much even after irradiation to
 2180 $1 \times 10^{16} \text{ n}_{\text{eq}} \text{ cm}^{-2}$.

2181 The operating voltage (V_{op}) is defined as a stable and safe operation voltage where the sensor
 2182 has reasonable performance in term of time resolution and gain. To evaluate it, several

2183 aspects are taken into account. At this voltage the sensor can be operated for a prolonged
 2184 period of time and under a constant flux of particles (similar to LHC conditions 40 MHz)
 2185 without the risk of inducing breakdown or electrical arcing between the sensor structures
 2186 (see Sec. 5.5.7). The noise increase should be less than 20% when compared to lower voltages,
 2187 plus the signal to noise ratio must be higher with respect to previous voltages. The maximum
 2188 leakage current allowed is limited to $5 \mu\text{A}$ per pixel and the power less than $100 \text{ mW}/\text{cm}^2$.

2189 Furthermore at this voltage the sensor must not present self-triggering events (events caused
 2190 by discharges not caused by particle hitting the detector) with a rate higher than 1 kHz for a
 2191 trigger threshold of $\pm 5 \text{ mV}$ or collected charge of 0.4 fC . An excessive self-triggering would
 2192 increase the dead time of the HGTD detector hindering its operation and is potentially
 2193 detrimental.

2194 This was studied in detail for HPK-3.2 and FBK sensors (studies for other types are ongoing):
 2195 the self trigger rate increases dramatically if the sensor is operated near the breakdown with
 2196 gain higher than around 30. This statement is valid both for unirradiated and irradiated
 2197 (with neutrons/protons) sensors of HPK-3.2 and FBK as shown in Fig. 5.6. For a neutron
 2198 fluence of $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ even at the highest voltage no self-triggering is observed since
 2199 the gain is low.

2200 Fig. 5.7 shows V_{op} as a function of fluence after neutron irradiation for different LGAD
 2201 types. It can be seen that it increases with fluence up to maximally about 750 V for $50 \mu\text{m}$
 2202 sensors. Sensors of $30 \mu\text{m}$ can only sustain about 500 V maximally (see Sec. 5.5.7). The
 2203 minimum voltage tested is usually chosen as when no clear signal is seen (usually the case
 2204 for irradiated sensors) or the sensor is believed to be not fully depleted (usually the case for
 2205 non irradiated sensors).

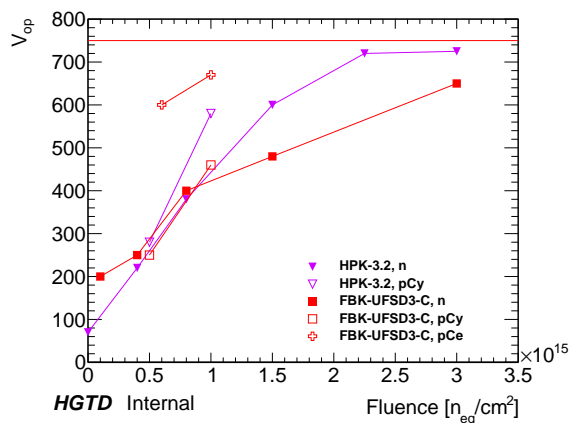


Figure 5.7: V_{op} as a function of fluence after irradiation for different LGAD types.

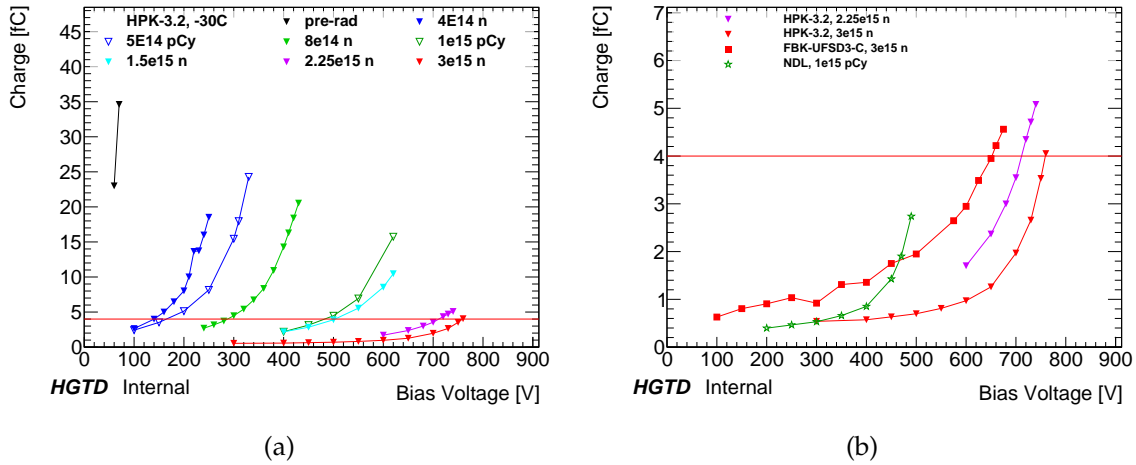


Figure 5.8: Collected charge as a function of bias voltage for different fluences for HPK-3.2 and for all vendors at maximum fluence. The horizontal lines indicate the HGTD lower charge limit of 4 fC at all fluences. Solid markers indicate n irradiation (n), open markers p irradiation at CYRIC (pCy). Measurements were performed at -30°C .

2206 **5.5.3 Collected charge and gain**

2207 Fig. 5.8 shows the collected charge as a function of bias voltage after neutron and proton
 2208 irradiation up to $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ for different LGAD types: HPK-3.2, FBK-UFSD3-C and
 2209 NDL sensors. The charge at V_{op} , as defined in Sec. 5.5.2, and 95% of V_{op} as a function of
 2210 fluence is shown in Fig. 5.9. The sensors are measured up to the maximum safe bias voltage
 2211 V_{max} where there is no risk of sensor breaking and no self triggering is observed.

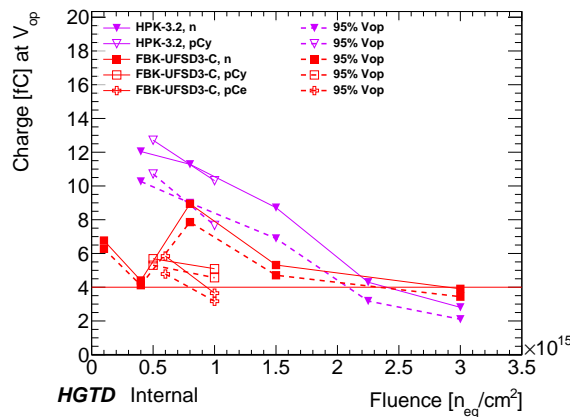


Figure 5.9: The charge at V_{op} and 95% of V_{op} as a function of fluence. Solid markers indicate n irradiation (n), open markers p irradiation at CYRIC (pCy) or CERN (pCe)

2212 It is evident that by going to higher fluences the increase in bias voltage can only partially

2213 compensate for the loss in gain due to the acceptor removal. A charge of 4 fC was found to
 2214 be the lower limit that still satisfies the HGTD science requirements in terms of hit efficiency
 2215 (see Sec. 5.5.4) and time resolution including the ALTIROC jitter (see Sec. 6.7). This level
 2216 is indicated by the horizontal lines. The corresponding “charge bias working point”, i.e.
 2217 the bias voltage needed for 4 fC, as well as the difference V_{diff} between V_{max} and the bias
 2218 working point WP (indicating the amount of bias head room for sensor operation) are shown
 2219 in Fig. 5.9.

2220 The following observations are made for the different types: **TODO: 3.1 is not anymore the**
 2221 **baseline, reorder**

2222 a. Baseline 50 μm sensor (HPK-3.1)

2223 HPK-3.1 can reach the target charge of 2.5 fC up to $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. The bias working
 2224 point is larger than for other types, however with a sufficient head room. At a fluence of
 2225 $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, only 2.0 fC is reached. Measurements after neutron and 70 MeV proton
 2226 irradiation seem to give similar results at similar fluences, although it is difficult to compare
 2227 directly since due to technical reasons not exactly the same fluence points could be taken. In
 2228 the future more studies with protons, also at higher energies, will be carried out.

2229 b. 50 μm sensor with higher doping and deep gain layer (HPK-3.2)

2230 HPK-3.2 sensors have a deeper and higher-dose multiplication layer, which leads to a
 2231 reduced acceptor removal rate. Hence, this type can reach the target charge of 2.5 fC up to
 2232 the higher fluence of $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. The bias working point is less than for HPK-3.1 and
 2233 exhibits a sufficient head room. At a fluence of $6 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, only 2.0 fC is reached.

2234 c. 60 μm sensor with gain layer infused with carbon (FBK-UFSD3-C)

2235 The main contributor to radiation damage in LGADs is the acceptor removal, i.e. the
 2236 reduction of the doping concentration of the gain layer which results in loss of gain. Addition
 2237 of Carbon in the gain layer reduces the acceptor removal. The required bias voltage is thus
 2238 lower than for other types with a large head room. Irradiations up to $6 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$
 2239 are ongoing. Further studies to extend the beneficial effect of carbon to higher fluences are
 2240 envisaged.

2241 5.5.4 Efficiency

2242 The hit efficiency of LGAD sensors on HGTD-specific readout boards was measured in
 2243 HGTD beam tests using an external telescope for reference tracks [42]. Fig. 5.10 shows
 2244 the efficiency as a function of most probable charge collected, compiled from 16 different
 2245 single pad sensors before and after irradiation at different bias voltages. The threshold
 2246 to accept events with a hit was chosen at a measured noise occupancy of 0.1% and 0.01%,
 2247 respectively.

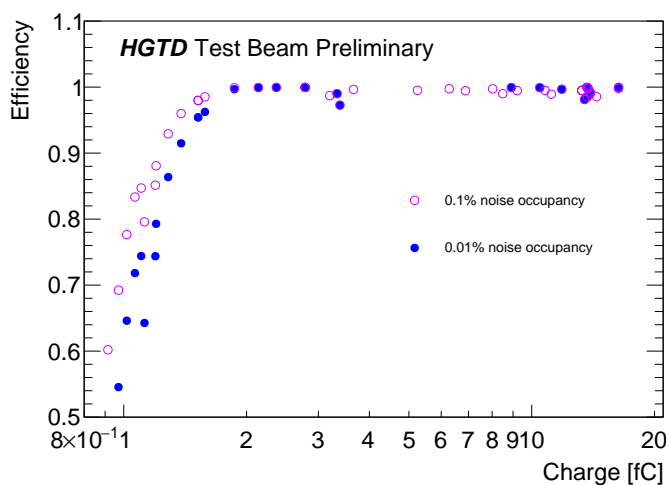


Figure 5.10: Hit efficiency as a function of collected charge at a measured noise occupancy of 0.1% and 0.01%.

2248 It can be seen that a universal curve is obtained, irrespective of fluence, indicating that the
 2249 charge is the main parameters on which the hit efficiency depends, given a certain noise
 2250 occupancy. A hit efficiency above 99% is obtained at the HGTD target working point of
 2251 2.5 fC mentioned in Sec. 5.5.3 for both noise occupancy working points. The measurements
 2252 will be repeated with the ALTIROC electronics once available for large-scale testing.

2253 2D efficiency maps are shown in Sec. 5.5.6 for arrays before and after irradiation.

2254 The cross talk between different pads of a 2×2 array was also measured and found to be
 2255 below 1% before and after irradiation.

2256 5.5.5 Time resolution

2257 The time resolutions of LGAD devices have been extensively studied in various beam
 2258 tests [37, 38, 42] and ^{90}Sr setups [32] on custom-made HGTD-specific readout boards. It
 2259 should be noted again that the measurements on such test boards exhibit noise and jitter
 2260 properties different from the final ALTIROC electronics, for which these measurements
 2261 will be repeated once available for large-scale testing. Hence, the results presented in this
 2262 section are best-case scenarios for the time resolution and demonstrate what is possible
 2263 when not limited by power and size constraints of a readout chip. For the remaining part
 2264 of the TDR the time resolution is not the one presented here with HGTD-specific readout
 2265 boards. Instead, the measured charge of the sensor was taken as an input to the ALTIROC
 2266 time resolution vs. charge function (see Fig. 3.6). This is to have a more realistic estimate of
 2267 the final time resolution with the ALTIROC.

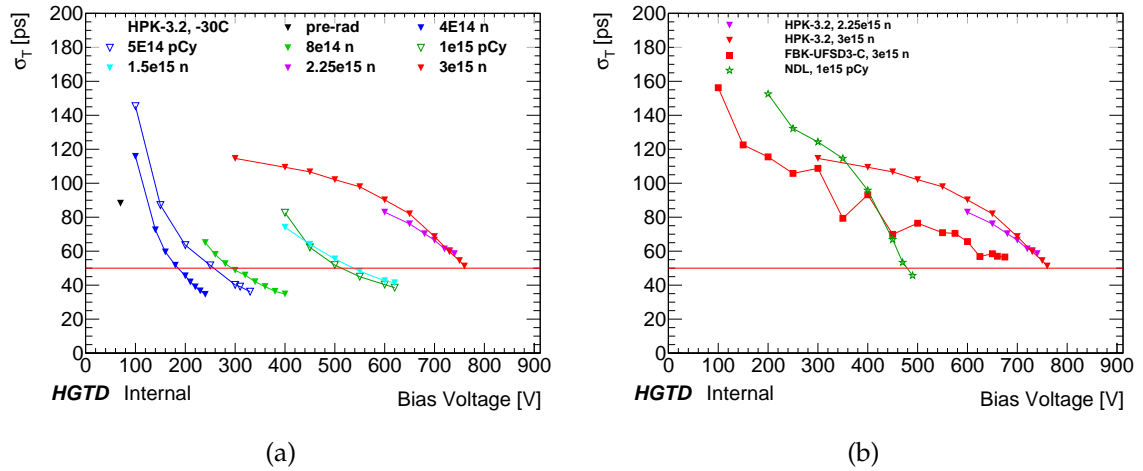


Figure 5.11: Time resolution as a function of bias voltage for different fluences for HPK-3.2 and for all vendors at maximum fluence. Solid markers indicate n irradiation (n), open markers p irradiation at CYRIC (pCy). Measurements were performed at $-30\text{ }^{\circ}\text{C}$.

2268 On custom-made HGTD-specific readout boards, it has been consistently shown that sub-
 2269 30 ps time resolution can be achieved below the breakdown point before irradiation for
 2270 sensors from all vendors with pad widths up to 1.3 mm and up to 5 pF capacitance [32,
 2271 37–42].

2272 The time resolution of HPK-3.2 was measured (up to a maximum voltage Sec. 5.5.2) in
 2273 the β -telescope after irradiation with 1 MeV neutrons at Ljubljana, and 70 MeV protons at
 2274 CYRIC. The results shown in ?? indicate that a resolution of 35 ps and better is achieved up
 2275 to a fluence of $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. At a fluence of $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, the resolution at V_{max}
 2276 is slightly over 50 ps.

2277 The time resolution for FBK-UFSD3-C with Carbon addition is around 40 to 50 ps, generally
 2278 worse than HPK-3.1/3.2-50, until $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. For a fluence of $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, it
 2279 reaches slightly over 50 ps, similarly to HPK 50 μm .

2280 In Fig. 5.12 the time resolution at V_{op} is shown. A time resolution of 60 ps is reached for
 2281 almost all sensors at all fluences with sufficient head room until $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, while
 2282 50 ps is not.

2283 5.5.6 Uniformity, inter-pad gap and edge region

2284 One crucial parameter of HGTD is the sensor fill factor, corresponding to the portion of
 2285 the detector which is able to detect particles efficiently. In the original plans a fill factor of
 2286 90% was foreseen, this would correspond to an inactive region between two pads of around
 2287 70 μm for a pad size of $1.3 \text{ mm}^2 \times 1.3 \text{ mm}^2$. Furthermore, the dead region at the edge of the

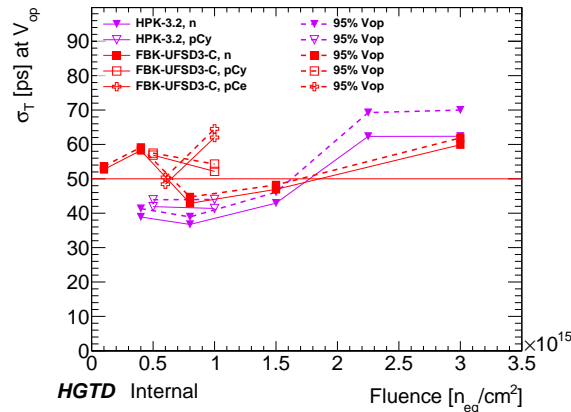


Figure 5.12: The time resolution at V_{op} and 95% of V_{op} as a function of fluence. The WP voltage curves are shown with a continuous line, the dotted line represents the voltage difference V_{diff} between WP and maximum bias voltage (head room). The results for HPK-3.1 include both n and pCy irradiation, the others are for n irradiation.

2288 arrays comprehending the guard ring has to be taken into account for the evaluation of the
2289 dead area.

2290 CNM and HPK provided the HGTD collaboration with multi pad LGAD arrays of different
2291 geometries (2×2 , 3×3 , 5×5 , 15×15) with different inter-pad and edge distances (see
2292 Sec. 5.2.2). The nominal values quoted by the vendor corresponds to distances between
2293 structures in the design of the detector. However it does not reflect perfectly the electric
2294 field configuration of the sensors. For this reason the values of inter-pad region and edge
2295 distances have to be measured in the laboratory with a focused infra-red laser beam or at
2296 test beam facilities.

2297 The sensors were studied at CERN's test beam facility. Thanks to the tracking system it
2298 was possible to evaluate the efficiency and the time resolution (using a SiPM as timing and
2299 efficiency reference) as a function of position. The hit efficiency and time resolution map for
2300 a 2×2 array is shown in Fig. 5.13 before and after irradiation.

2301 In the laboratory the sensors were tested with an infrared laser of 1060 nm wavelength
2302 focused to $10 \mu\text{m}$ – $20 \mu\text{m}$ FWHM. The light was injected through the sensor's rear opening
2303 of the metalization and scanned from one pad to the other. The two profiles of the pulse
2304 maximum are fitted with a step function and the distance between the pads is evaluated.
2305 The measured effective distance between the neighboring pads can be estimated as the
2306 distance where charge collection efficiency drops to 50% on first pad and rises to 50% on the
2307 neighbour (50%-50% point). The inter-pad scans for HPK-3.1 can be seen in Fig. 5.14(a). The
2308 measured values are around $40 \mu\text{m}$ higher than the nominal values quoted by the vendor.
2309 An overview of the measured vs. nominal values for the HPK-3.1/3.2 and CNM-AIDA
2310 sensors can be seen in Fig. 5.14(b). As shown in Tab. 5.6, the lowest measured value per

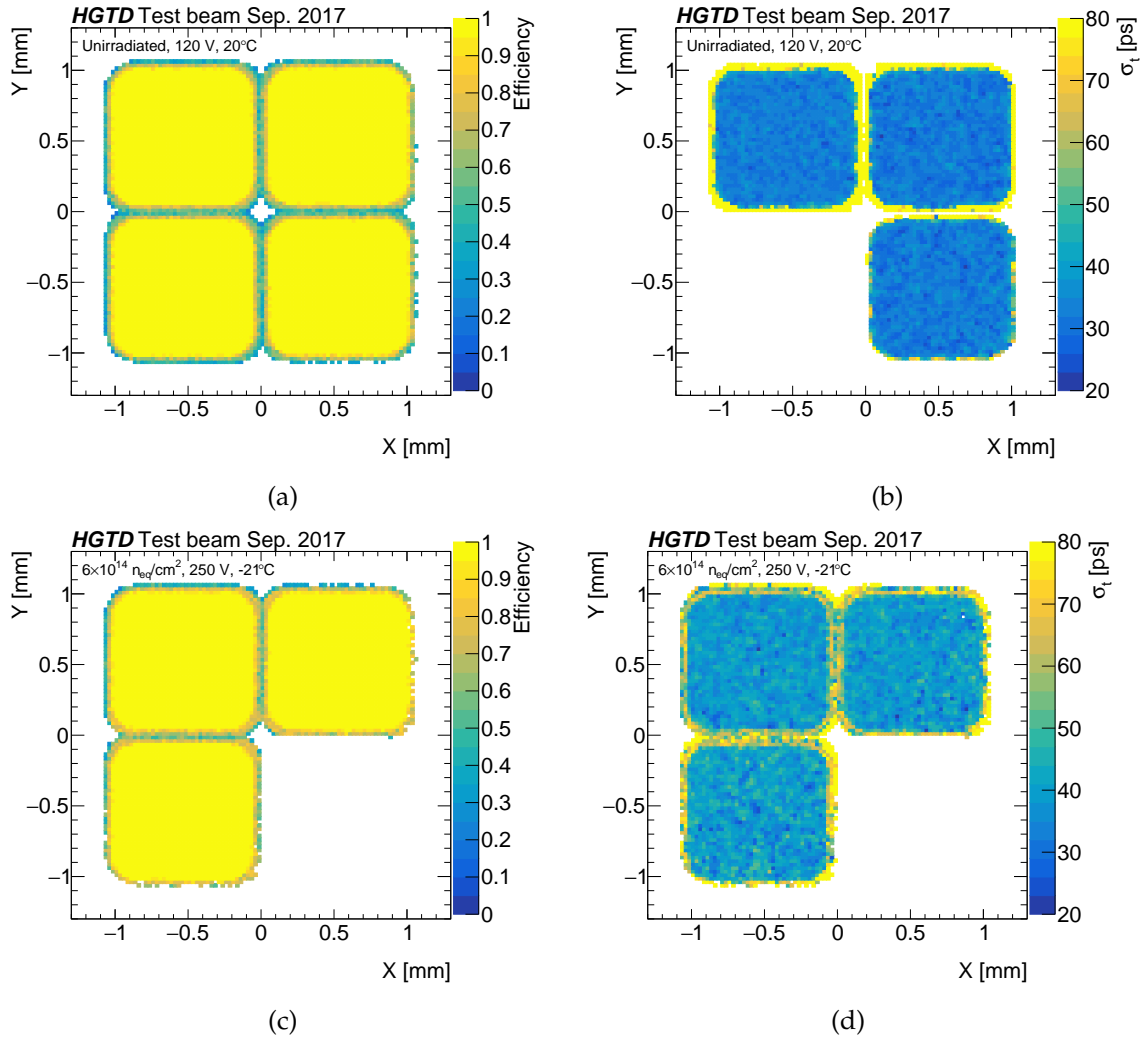


Figure 5.13: 2D maps of efficiency (left) and time resolution (right) before (top) and after n irradiation to $6 \times 10^{14} n_{eq} cm^{-2}$ (bottom) for a 2×2 array from CNM-10478-50 as measured in HGTD beam tests [42]. Sometimes only 3 channels were measured. The efficiency was evaluated at a threshold of 3 times the noise here. A mean efficiency in the pad center of 99% is maintained up to a threshold of 5 times the noise level. The time resolution for this sensor is 39 ps with a spread of 3 ps in the pad center.

2311 type (roughly 70, 90, and 110 μm for HPK-3.1, CNM-AIDA and HPK-3.2) correspond to fill
 2312 factors of 90%, 87% and 84%, respectively. HPK-3.2 shows an inter pad gap that is slightly
 2313 larger than HPK-3.1 before irradiation, however after irradiation of $1.5 \times 10^{15} n_{eq} cm^{-2}$ its
 2314 performance is better.

2315 The edge area is evaluated in a similar way by scanning over the edge of the sensor pad
 2316 with a laser. Several types of HPK-3.1 with different edge distances were measured in this

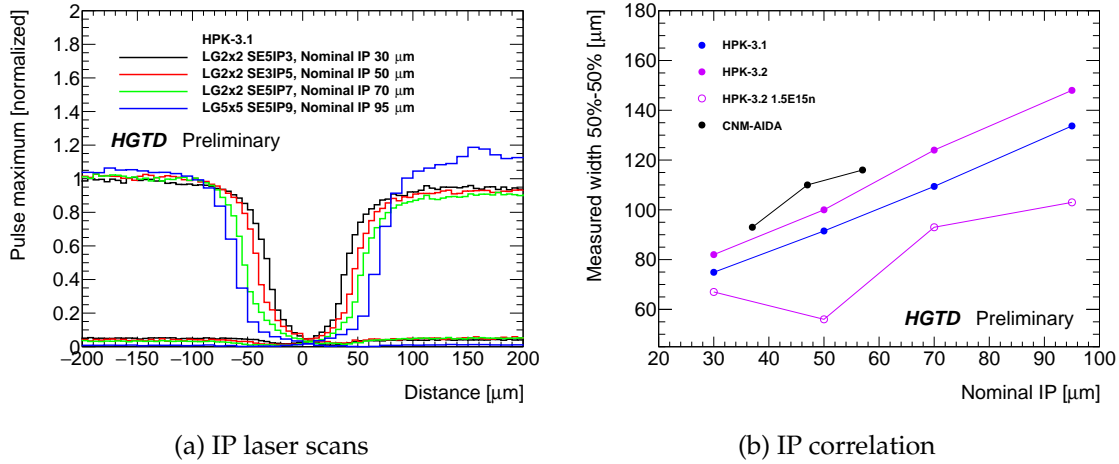


Figure 5.14: Fig. 5.14(a): Inter-Pad distances for several HPK-3.1 sensors. Fig. 5.14(b): Nominal vs. measured inter-pad distances for HPK-3.1, HPK-3.2 (before and after irradiation) and CNM sensors.

Effective IP gap	Fill factor
70 μm	90 %
90 μm	87 %
110 μm	84 %

Table 5.6: Fill factor for different effective (i.e. not nominal) IP gap distances.

2317 way and they all showed a 95%-5% drop from the maximum of around 60 μm showing no
 2318 distortion induced by slimmer edges. Furthermore the guard ring of the sensor was read
 2319 out and the width of it evaluated with the same technique. For an edge distance of 200 μm
 2320 a width of 200 μm was seen. For nominal edges of 300 μm and 500 μm a guard ring width
 2321 of around 350 μm–400 μm was measured, however the sensor with nominal edge of 500 μm
 2322 has an additional smaller guard ring that is left floating and cannot be read out.

2323 So far no change in sensor performance (collected charge, time resolution) or fragility was
 2324 observed for the different IP gaps and edge distances.

2325 5.5.7 Long term and stability tests

2326 Long term and high flux

2327 HGTD sensors were typically tested to evaluate the performance at low rate, with a laborat-
 2328 ory ⁹⁰Sr source, and medium rate, at test beams. Furthermore, they were biased on the scale
 2329 of tens of hours. Nevertheless, during the running of the ATLAS experiment, the sensors
 2330 will be operated continuously for days to weeks in a high particle flux. For this reason, the

2331 resilience of the sensors was tested by applying high voltage for an extended period of time.
2332 To simulate a high flux, an IR laser was pulsed continuously with a frequency of 50 MHz and
2333 the intensity of several MiPs on irradiated HPK-3.1, HPK-3.2 and FBK sensors while biased
2334 up to a voltage of 750V. No change in the behavior of sensors was observed in the timescale
2335 of tens of hours. Tests are ongoing with unirradiated and irradiated sensors. **TODO: Update**
2336 **results when available. Add statement that no change was observed after hours of intense**
2337 **charge injection.**

2338 **Sensor breaking and head room**

2339 It is important to find a safe bias voltage V_{op} at which the sensors can be operated, as
2340 mentioned in Sec. 5.5.2. In addition, the sensor can be protected by operating them at
2341 the minimum bias working point at which the requirements for collected charge and time
2342 resolution are satisfied instead of V_{op} , as shown in Fig. 5.9 and Fig. 5.12.

2343 The difference (V_{diff}) in bias voltage between V_{max} and V_{op} is the “safety head room” and is
2344 usually quite large at low to medium fluences, although it was found that for thin sensors
2345 (below 30 μm) this head room is small and might impact their usefulness. For very high
2346 fluences the “head room” diminishes drastically for all sensors.

2347 During the LGAD R&D phase, these principles were explored with existing sensors listed
2348 in Sec. 5.2.2. As part of the learning curve to define safe operating conditions some of the
2349 sensor were broken during testing. Excluding breaking due to mishandling in the large scale
2350 lab and beam testing campaign, a few general conclusions can be reached for the four sensor
2351 types that were tested in depth.

2352 It was discovered that sensors break when the bias exceeds the critical bias voltage V_{crit} .
2353 Since V_{op} increases with fluence almost all breaking occurred at high fluences. The thickness
2354 of the sensors is an important parameter. It was observed that thin sensors would break
2355 immediately when reaching a certain V_{crit} which depends on the sensor thickness. Sensors
2356 with thickness of 50 μm thickness (like HPK-3.1, HPK-3.2) would break for bias voltages
2357 greater than 750 V. After breaking, a burn mark usually appears in the interface between
2358 pad and guard ring, most of the time at the detector corner where the fields are largest. This
2359 observation motivates future layout studies of the interface of guard ring and multiplication
2360 area. Another study investigates operation of the sensors during temperature and humidity
2361 changes and at different particle rates. These few general observations motivate us to make
2362 the increase of the bias head room as one of the research areas of the next prototype run.

2363 **Annealing**

2364 Most of the measurements with irradiated sensors were done after annealing for 80 min at
2365 60 °C, which roughly simulates the operational conditions in one year of LHC operation

2366 since higher temperature accelerates the annealing (the Arrhenius factor between 60 °C and
 2367 -30 °C is more than 1×10^6 , 80 min simulates hundreds of years at -30 °C, and tens of days
 2368 at room temperature).

2369 A prolonged annealing study was carried out with CNM-10478-50 and HPK-3.1 samples
 2370 with an area of $1.3 \text{ mm}^2 \times 1.3 \text{ mm}^2$ to check the performance in case of unpredicted situ-
 2371 ations where sensors would be exposed to longer times at elevated temperatures or when
 2372 intentional annealing may be used to reduce leakage current and power dissipation.

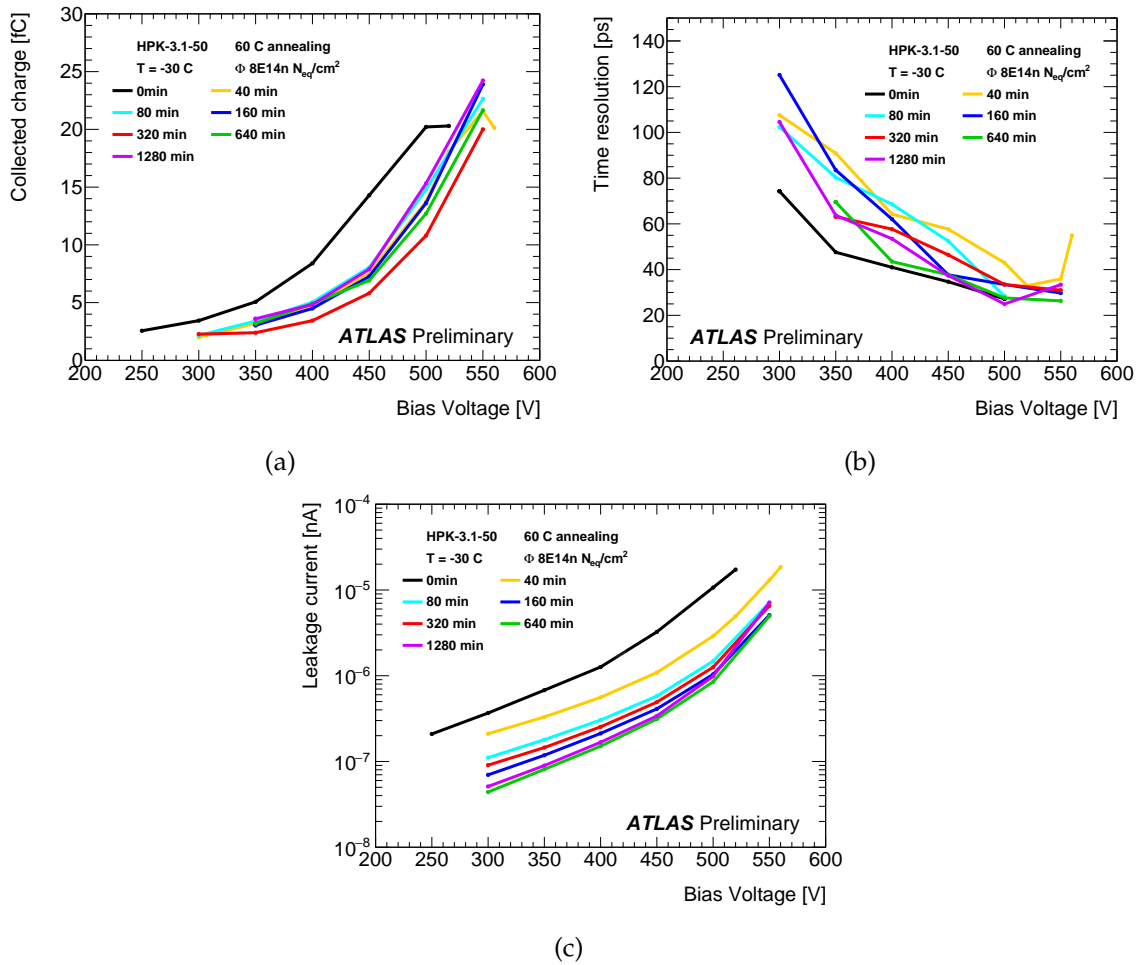


Figure 5.15: Voltage dependence for different annealing times for (a) collected charge, (b) time resolution and (c) leakage current.

2373 The dependence of collected charge on bias voltage for different annealing times is shown
 2374 in Fig. 5.15(a) for HPK samples. It can be seen that the effect of the annealing is limited.
 2375 There seems to be a decrease of initial acceptors in the gain layer with annealing on a time
 2376 scale of tens of minutes, but thereafter the charge stays relatively constant. Even if full

2377 reverse annealing of deep acceptors takes place the applied bias voltages are high enough
 2378 to fully deplete thin detectors and also saturate drift velocity. The effect of annealing on
 2379 time resolution remains limited (10–20 ps maximal spread at high voltages, with an initial
 2380 increase and then decrease again) as shown in Fig. 5.15(b). A much larger beneficial effect of
 2381 annealing can be observed on the leakage current as shown in Fig. 5.15(c).

2382 There were no significant differences in annealing performance observed between the two
 2383 producers HPK and CNM. The annealing studies will be extended further to the whole
 2384 fluence range, different temperatures and producers, so that an accurate running scenario
 2385 can be made.

2386 5.5.8 Leakage current and power after irradiation

2387 In standard Silicon sensors without gain, the leakage current originating from volume
 2388 generation current increases linearly with fluence. However, for LGADs the situation is
 2389 more complex due to the gain and its fluence evolution. The operation in gain mode leads to
 2390 an increase of the leakage current, which is given by the product of the volume generation
 2391 current and the current multiplication factor. As the gain decreases with irradiation and the
 2392 generation current increases, the leakage current does not necessarily increase monotonically
 2393 with fluence. The leakage current in multiplication mode contributes to parallel noise
 2394 linearly, hence it is of high importance to run the sensors at low temperatures since cooling
 2395 decreases the leakage current (roughly by a factor of 2 every 7 °C).

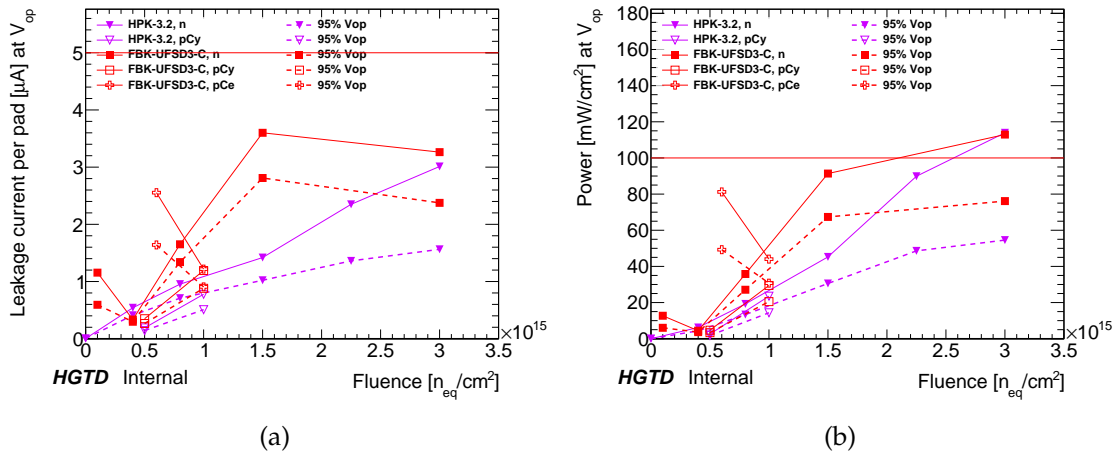


Figure 5.16: (a) Leakage current for single pads at $-30^\circ C$ as a function of bias voltage for HPK-3.2 irradiated with 1 MeV neutrons (solid lines) and 70 MeV protons (dashed lines). The horizontal line represents the ALTIROC maximum acceptable current of 5 μA . (b) Power density as a function of fluence at the operation bias voltage measured V_{op} at $-30^\circ C$ [39, 40]. The horizontal line represents the maximum acceptable power of 100 mW/cm^2 .

2396 The leakage currents for $1.3 \text{ mm}^2 \times 1.3 \text{ mm}^2$ HPK-3.2 single pads for the different fluences
2397 as a function of the bias voltage shown in Fig. 5.16(a) exhibit large increases for increased
2398 bias, partially due to the increased gain. The ALTIROC maximum acceptable current is $5 \mu\text{A}$
2399 (dotted line in Fig. 5.16(a)). HPK-3.1 satisfies this requirement up to the highest fluence
2400 and voltage. From this the power density (power/area) can be derived. The power can be
2401 minimised by operating the sensors at as low temperature and bias voltage as possible. For
2402 the assumed operating temperature (-30°C), Fig. 5.16(b) shows the measured power density
2403 of HPK-3.1 as a function of fluence for V_{op} . It can be seen that there is typically one order of
2404 magnitude difference between these two voltage points. For the 2.5 fC working point, the
2405 power density reaches maximally $30 \text{ mW}/\text{cm}^2$, however, this working point can only be
2406 reached up to $1.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. For other types like HPK-3.2 and FBK-UFSD3-C with a
2407 higher fluence reach, evaluations are still ongoing. The final power dissipation in HGTD will
2408 depend on the sensor type choice as well as the operational scenario as detailed in Sec. 5.6.

2409 5.6 Operational aspects and bias voltage evolution in HGTD

2410 As shown in Sec. 5.5.3, the bias voltage needs to be increased with increasing fluence,
2411 which is a function of radius and integrated luminosity (i.e. period over lifetime) in HGTD.
2412 Monitoring of the leakage current and the TOT as an indicator of collected charge will give a
2413 good estimate of the gain evolution during operation, allowing to perform the necessary
2414 adjustments of the bias voltage.

2415 For a first scenario, it is assumed that the detector is operated at operating bias voltage V_{op}
2416 (see Fig. 5.7). This would give the best possible performance (charge and time resolution),
2417 however, also implies the highest risk and power dissipation (see Fig. 5.16(b)). In the
2418 future, more complex and realistic scenarios will be developed that will be a trade-off
2419 between required performance on the one hand and risk and power constraints on the other
2420 hand (keeping the sensor power at an average level of $30\text{--}60 \text{ mW}/\text{cm}^2$ as mentioned in
2421 Tab. 11.1).

2422 The expected dependence of the fluence on the radius (Fig. 2.13) and the required bias
2423 voltage V_{op} for the increasing fluence permits a prediction of the bias-voltage distribution
2424 along the length of a readout row containing individual modules mounted on support plates
2425 (see Chap. 11). This is shown in Fig. 5.17 for different integrated luminosities for HPK-3.1
2426 sensors. It shows that the ability to connect several nearby modules to the same bias supply
2427 allowing a 10% variation in the bias to modules on one bias supply will be limited.

2428 Note that the exact behavior depends on the sensor type chosen since different sensor types
2429 require different bias voltages for the same performance (see Sec. 5.5.3). Scenarios for other
2430 types with improved radiation hardness like HPK-3.2 and FBK-UFSD3-C will be evaluated
2431 in the future.

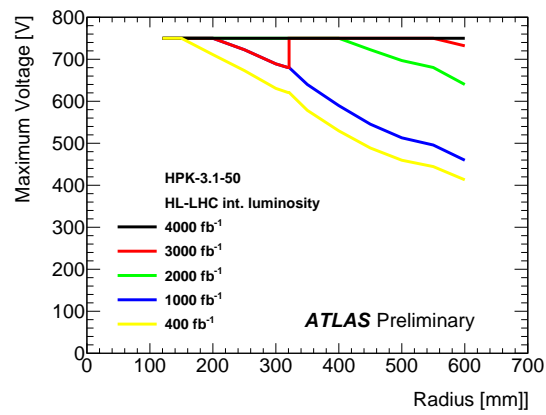


Figure 5.17: Required bias voltage as a function of position along the longest readout row for different integrated luminosities for HPK-3.1 sensors based on V_{max} . The sudden change for the 3000 fb⁻¹ line corresponds to the replacement of the inner ring. (TODO: redo completely with 3-rings, use Vop)

2432 5.7 Summary of status quo sensor design

2433 Through the R&D program of the last few years, which involved three large LGAD suppliers,
 2434 several LGAD designs have been investigated. A recommendation for the final design
 2435 choices will be a “snapshot” taking into account the fact that some of the options need more
 2436 investigation. So the choices will be tilted towards conservative performance and operations,
 2437 making use of the “bias working points” of Fig. 5.9.

- 2438 • Thickness of the high resistivity bulk
 2439 **50 μm** : highest fluence reach, largest head room.
- 2440 • Doping profile
 2441 **Narrow (and deep) shows improved radiation hardness**: however, the performance
 2442 before irradiation is degraded due to the low breakdown voltage. A compromise
 2443 between performance before irradiation and radiation hardness needs to be developed.
- 2444 • Replacing Boron, the dopant in the gain layer, with Gallium
 2445 **No Ga**: advantage not clearly established (although for different doping profiles) after
 2446 neutron irradiations. Indications of small improvement for 23 GeV protons irradiation.
- 2447 • Adding Carbon to the dopant in the gain layer
 2448 **C implantation is a promising candidate that shows improved radiation hardness**
 2449 **up to at least $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$, but is not yet fully understood**: The performance
 2450 at higher fluences needs to be established. Also noise and time resolution need to
 2451 be understood. Moreover, it is not available yet by all vendors. Further studies are
 2452 ongoing.
- 2453 • Inactive distance between pads (inter-pad gap)
 2454 **70 μm** : lowest distance measured, leading to a fill-factor of 90%. With irradiation fill
 2455 factor may improve due to increased operating voltage and relatively larger multiplic-
 2456 ation at the edges of the pads.
- 2457 • Slim edge distance
 2458 **300 μm** : Studies showed same performance as samples with wider edge.
- 2459 • Covering the pads with metal
 2460 **Complete metal cover of pads**: sensors with pads fully covered with metal showed
 2461 better performance in terms of collected charge than sensors with large non-metal
 2462 openings.

2463 With this selection of parameters, the science goals will be reached up to the HGTD target
 2464 fluence of $2.5 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. In general, more studies after high energy charged hadron
 2465 irradiation are needed.

2466 5.8 Roadmap for future sensor productions and activities

2467 During the last year the HGTD collaboration have profited from a open and fruitful collab-
2468 oration with CMS. Our ability to exchange ideas, designs and results, share beam test and
2469 financing of common production runs was a major reason for our rapid progress. Hopefully
2470 this collaboration will continue in the same way.

2471 Two more prototype runs are anticipated, one in 2019 to implement our idea to extend
2472 substantially the fluence reach and to optimize several geometrical layout issues. In 2020,
2473 the final prototype will be produced and tested. A market survey will be conducted in 2020,
2474 based on the understanding of the design issues solved in 2019. Then a pre-production run
2475 will follow in 2021.

2476 During the next prototype runs we will address sensor issues which we discovered during
2477 the recent tests by working with the three manufacturers. One issue is the need to overcome
2478 the loss in gain due to acceptor removal. Another is the need to increase the bias voltage head
2479 room. Another one is the the power consumption. All three issues need to be addressed by
2480 optimizing the presently available LGAD technology. An example is shown in the following,
2481 which leads to reduction in required bias voltage (and hence power) while improving the
2482 gain.

2483 In Sec. 5.7 is mentioned the fact that the radiation hardness of the sensors available at
2484 the moment has not been established up to the HGTD baseline target fluence of $2.5 \times$
2485 $10^{15} n_{\text{eq}} \text{ cm}^{-2}$ assuming one replacement of the inner ring at half lifetime (it should be noted
2486 that scenarios to mitigate the fluence requirements were developed as discussed in Sec. 5.8).
2487 This can be seen in Fig. 5.18, where the collected charge for high fluences is shown for the
2488 strongest candidates. For the fluence of $3 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$ HPK-3.1 sensor does not reach
2489 the required level of 2.5 fC at the highest bias voltages reached, and for $6 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$
2490 the collected charge is even lower. The FBK-UFSD3-C sensor with Carbon implantation
2491 shows better performance at $3 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$, the testing up to $6 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$ will be
2492 done in the near future. Also shown is the collected charge from the HPK-3.2 sensor, which
2493 has a higher doping and deeper implant than the HPK-3.1 sensor (Tab. 5.3); nevertheless, at
2494 $6 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$ the goal of 2.5 fC is still not reached.

2495 The data of the collected charge for both sensors have been used to tune WeightField 2 so
2496 that the simulations match the measured data. Then the collected charge can be simulated
2497 for a sensor which combines both the deep implant of HPK-3.2 with the carbon implant
2498 of FBK-UFSD3-C. The collected charge of the prototype is shown for $3 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$ in
2499 Fig. 5.18(a) and for $6 \times 10^{15} n_{\text{eq}} \text{ cm}^{-2}$ in Fig. 5.18(b). Two important characteristics for such
2500 sensors are evident: much lower operating voltage and much higher collected charge when
2501 compared with the sensors without carbon or with shallow implant.

2502 Further topics of future R&D:

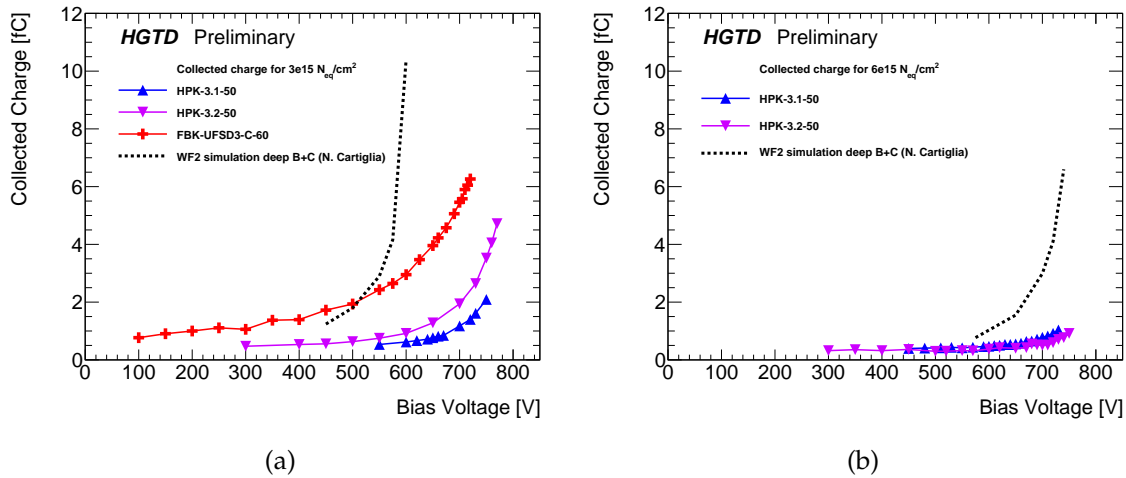


Figure 5.18: Collected charge vs bias voltage for sensors irradiated to $3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ (a) and $6 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ (b), respectively. In the plots are measured data of the existing prototypes and the simulated prospect of the proposed sensors combining deep implantation of the Boron gain layer with carbon implantation.

- 2503 • Reduce the inactive pad distance.
- 2504 • Produce first full-size sensors: 30×15 ($4 \times 2 \text{ cm}^2$).
- 2505 • Establish the robustness of LGADs under stressful operating conditions.
- 2506 • Improve breakdown between guard ring and pad area.

2507 New groups have joined the HGTD sensor R&D effort in the last year and will participate in
 2508 large scale testing both for labs measurements and at beam tests.

2509

6 Front-end Electronics

2510

2511

2512

2513

2514

2515

2516

2517

2518

2519

2520

2521

2522

2523

2524

2525

This chapter describes the required performance, design, and latest prototype testing of the ASIC chip, ALTIROC, that will be bump-bonded to the LGAD sensor. It will have 225 readout channels, thus two ASICs will read out each LGAD. The main challenge in the design of this ASIC is the fact that it needs to have a small enough contribution to the timing resolution, in order to match the excellent performance of the LGAD. As introduced in Sec. 4.2.2, this contribution comes mainly from the time-walk and the jitter. The first one will be addressed by applying a correction based on the fact that the variations in the time-of-arrival (TOA) of the pulse are related to the time-over-threshold (TOT); this is presented in Sec. 6.3.2. The most critical aspect concerning the jitter is the design of the analog front-end electronics, which are composed of a voltage preamplifier followed by a fast discriminator. The measured time-of-arrival and time-over-threshold are digitized using two Time to Digital Converters (TDCs), and stored in a local memory at the channel level. An end-of-column logic is implemented to collect the information for each of the 15 columns (with 15 pads each). The ASIC common digital part is composed of different blocks necessary to generate and align the clocks, receive the slow control commands to configure the ASIC and transmit the digitized data.

2526

2527

2528

2529

2530

2531

Two iterations of this chip have been produced and tested so far: the first, ALTIROC0, integrated four pads in a 2×2 array, with the analog part of the single-channel readout: the preamplifier and the discriminator. The results of the test beam and test bench studies performed on this version of the ASIC can be found in [45]. The second iteration, ALTIROC1, consists of a 5×5 pad matrix, in which the digital components have been added to the single-channel readout.

2532

2533

2534

2535

2536

2537

2538

2539

2540

The requirements imposed by the data taking conditions, the sensor and the targeted performance are presented first in Sec. 6.1. The ASIC architecture is described in Sec. 6.2, first going through the single-channel architecture and then the entire ASIC. Sec. 6.3 describes in detail the design of the single-channel readout electronics, followed by the description of the ASIC common digital part in Sec. 6.4. The radiation tolerance is described in Sec. 6.5 and the power distribution in Sec. 6.6. The performance results obtained so far in testbench and testbeam are described in Sec. 6.7. The description of the monitoring can be found in Sec. 6.8. Lastly, a brief account is given on the future steps towards the completion of the design and testing of the ASIC in Sec. 6.9.

2541 6.1 General requirements

2542 This section presents a brief description of the requirements of the front-end readout electron-
 2543 ics. The requirements of the ASIC can be divided in two types. On one side the considerations
 2544 regarding the operational environment of the ASIC, its powering and electrical connections.
 2545 These requirements are summarized in Tab. 6.1. The second group is concerning the ASIC
 2546 performance, driven by the targeted time resolution. A summary of these requirements is
 2547 presented in Tab. 6.2.

- 2548 • The ASIC will have to withstand high radiation levels and, as in the case of the sensors,
 2549 some ASICs will have to be replaced during the HL-LHC period. The expected
 2550 radiation levels have been presented in Sec. 2.4, considering a 2.25 safety factor for
 2551 the electronics. Thus, the maximal TID is 2.0 MGy (at $r = 120$ mm) and decreases with
 2552 radius. At $r = 320$ but out of date (the edge of the region that will be replaced by half the
 2553 HL-LHC lifetime) it reaches a value of of 2.0 MGy.
- 2554 • Each single-channel readout needs to fit within the sensor pad, with sides of 1.3 mm.
 2555 It will be capable of handling up to 5 μ A leakage current from the sensor.
- 2556 • The target for the electronics is to be able to read out signals from 4 fC up to 50 fC
 2557 throughout the HGTD lifetime.
- 2558 • The electronics jitter for an input charge of about 10 fC is required to be smaller than
 2559 25 ps, i.e smaller than the dispersion induced by the Landau fluctuations on the energy
 2560 deposit which limits the time resolution to 25 ps at large sensor gain. Such charge is
 2561 equivalent to deposit of a MIP in a 50 μ m thick LGAD with a gain of 20. A detector
 2562 capacitance of around 4 pF is considered. The contribution to the time resolution from
 2563 the TDC should be negligible and leads to a 20 ps TDC bin for the TOA measurement
 2564 and a 40 ps TDC bin for the TOT measurement. The time walk should be smaller than
 2565 10 ps over the dynamic range after correction. **TO BE UPDATED**
- 2566 • Because the signal from the sensor will degrade due to the effects of irradiation, it
 2567 should be possible to set the discriminator threshold for small enough values of input
 2568 charge. The minimum threshold should be so that an efficiency above 95% is achieved
 2569 for an input charge of 4 fC (although with a jitter larger than 25 ps). To enable the
 2570 possibility to set such low thresholds, the cross-talk between channels should be kept
 2571 below 5%.
- 2572 • The TOA and TOT information are transferred to the data acquisition system only
 2573 upon L0/L1 trigger reception with latency up to 35 μ s [46], therefore necessitating a
 2574 large size memory. The trigger rate depends on the final scheme adopted. It will be
 2575 1 MHz for a L0 trigger, or 0.8 MHz for a L1 trigger in a L0/L1 scheme with a L0 at
 2576 4 MHz.

- 2577 • The global phase adjustment of the clock should be guaranteed to a precision of 100 ps
2578 in order to properly center the 2.5 ns measuring window at the bunch-crossing.

- 2579 • The ASIC will need to handle the information to perform the luminosity measurement,
2580 computing the number of hits per ASIC on a bunch-by-bunch basis. To limit the
2581 bandwidth required, the information of only a subset of the ASICs is used. The
2582 current proposal is to use the sensors located at $320 \text{ mm} < r < 640 \text{ mm}$, or
2583 equivalently $2.4 < |\eta| < 3.1$. The use of both layers will not provide a
2584 significant increase in coverage with respect to one of the layers, but the redundancy
2585 aids in estimating and reducing the systematic uncertainty on the measured luminosity
2586 and provides contingency in the event of failures in the instrumentation. **TO BE**
2587 **UPDATED**

- 2588 • Finally the ASIC power dissipation should be kept as low as possible, in order to limit
2589 the size required for a single CO₂ cooling unit (for more details on the cooling system
2590 see Sec. 11.2).

TID tolerance	Inner region: 2.0 MGy Outer region: 2.0 MGy
Pad size	$1.3 \times 1.3 \text{ mm}^2$
Voltage	1.2 V
Power dissipation per area (per ASIC)	300 mW cm^{-2} (1.2 W)
e-link driver bandwidth	320 Mbit s^{-1} , 640 Mbit s^{-1} , or 1.28 Gbit s^{-1}
Temperature range	$-40 \text{ }^\circ\text{C}$ to $40 \text{ }^\circ\text{C}$
SEU probability	$< 5\%/\text{hour}$

Table 6.1: Physical, power, environmental and electrical requirements.

Maximum leakage current	$5 \mu\text{A}$
Single pad noise (ENC)	$< 1500 e^- = 0.25 \text{ fC}$
Cross-talk	$< 5\%$
Threshold dispersion after tuning	10%
Maximum jitter	25 ps at 10 fC
TDC contribution	$< 10 \text{ ps}$
Time walk contribution	$< 10 \text{ ps}$
Dynamic range	4 fC–50 fC
TDC conversion time	$< 25 \text{ ns}$
Trigger rate	1 MHz L0 or 0.8 MHz L1
Trigger latency	10 μs L0 or 35 μs L1
Clock phase adjustment	100 ps

Table 6.2: Performance requirements. The values given for the noise, minimum threshold and jitter have been estimated considering a detector capacitance $C_d = 4 \text{ pF}$.

2591 6.1.1 Data transmission bandwidth requirements

2592 The bandwidth of each ASIC strongly depends on the radial region it covers, as shown by
2593 the distribution of the average number of hits per ASIC in Fig. 9.3.

2594 Each module consisting of two ALTIROC ASICs is connected via a flex cable to a Peripheral
2595 Electronics Board (PEB), described in Chap. 9. The PEB transfers digital signals from the flex
2596 cables to optical fibres connected to the back-end DAQ. Flex cables for modules placed at a
2597 radius above **320 mm** also carry two differential e-links with luminosity data. For error-free
2598 data transmission at the bandwidths required by the expected HGTD data volume, the PEB
2599 uses the low-power GigaBit Transmission chip (lpGBT [47]). A dedicated buffer is needed in
2600 each ASIC to average the rate variation and match the best speed of the e-link drivers/lpGBT
2601 transceiver inputs:

- 2602 • The largest average hit rate at small radius does not exceed 20 hits per event, equivalent
2603 to a rate of 500 Mbit s^{-1} (not including header). In the current design a bandwidth of
2604 up to 1.28 Gbit s^{-1} was considered for the innermost radius ASICs (up to $r \simeq 150 \text{ mm}$),
2605 taking into account a considerable safety margin. However if further studies confirm
2606 this, a lower maximum bandwidth could be considered, thus reducing the number of
2607 necessary lpGBTs.
- 2608 • For larger radii, a 320 Mbit s^{-1} bandwidth can be used.
- 2609 • For the luminosity data, the maximal number of hits per ASIC at $r > 320$ but out of date
2610 should be considered **TO BE UPDATED**. This number does not exceed 30. With a
2611 4-bit header in addition to the 12 bits of data for the counts in the larger and smaller
2612 window (see Sec. 6.2.1), a 640 Mbit s^{-1} e-link driver and lpGBT speed is needed.

2613 6.2 ASIC architecture

2614 With an area of $20 \text{ mm} \times 22 \text{ mm}$, the largest part of the chip will be occupied by the channel
2615 matrix: each pad being $1.3 \text{ mm} \times 1.3 \text{ mm}$, arranged in a matrix of 15×15 channels. The
2616 channel matrix will thus have an area of $19.5 \text{ mm} \times 19.5 \text{ mm}$; the additional space is needed
2617 to accommodate the end-of-column logic and the common digital blocks.

2618 This section presents an overall description of the three main structures of the ASIC:

- 2619 • the *single-channel readout* cell, which is repeated 225 times. It integrates the preamplific-
2620 ation, the discrimination and the digitization of the hits as well as the local storage (or
2621 buffering) of the digitized data until an L0/L1 trigger is received.
- 2622 • the *end-of-column logic* (EOC) which performs the readout of the 15 columns and
2623 transfers the data to the trigger data and luminosity process units.

- the *ASIC common digital part* which formats the digitized data before sending it to the peripheral off-detector electronics that will be described in Chap. 9. This stage also contains common cells such as a phase shifter, a Phase Locked Loop (PLL) and a fast command decoder that will be described in Sec. 6.4.

The ASIC has been designed using TSMC ¹ 130 nm technology. Simulations have been performed using the 130 nm TSMC kit provided by CERN. The TSMC 130 nm has been tested up to 400 Mrad and the effects at transistor level are known. The analog part of the ASIC (preamplifier, discriminator and TDC) has been designed to ensure its radiation hardness.

6.2.1 Channel architecture

A conceptual schematic for the single-channel readout is presented in Fig. 6.1. Each readout channel will consist of a preamplifier followed by a discriminator, both of which critical elements for the overall electronics time performance. A detailed characterization of the preamplifier is presented in Sec. 6.3. The time of the pulse will be determined using a discriminator that follows the preamplifier. As a consequence, a time-walk correction needs to be applied in order to account for the dispersion in the TOA due to the different pulse heights. Since the time walk will be measured using a Time Over Threshold architecture (described in Sec. 6.3.2), two TDCs are necessary to digitize the discriminator output. The first is for the digitization over 7 bits of the TOA, which corresponds to the position of the rising edge of the discriminator output. The range used is 2.5 ns, and it will be done with a bin of 20 ps. The second TDC will be used for the digitization over 9 bits of the width of the discriminator output. The bin and range of the TOT-TDC will be 40 ps and 20 ns respectively. Further details on the TDCs are presented in Sec. 6.3.3.

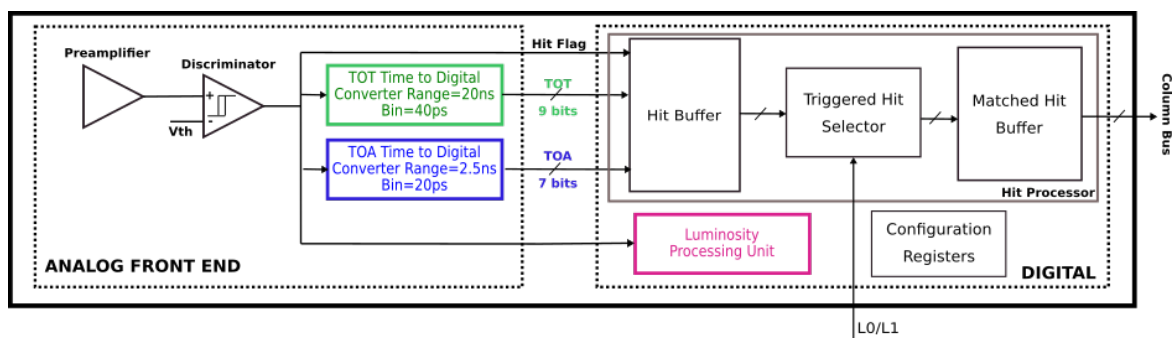


Figure 6.1: Schematic of the single-channel readout electronics. Two main blocks are identified, the analog and the digital part. The input pulse from the sensor enters the preamplifier on the left. The TOA and TOT data are read out by the column bus on the right.

¹ TSMC stands for Taiwan Semiconductor Manufacturing Company. The technology has been qualified up to 4 MGy [48, 49].

2647 The output of the analog read-out is processed by the digital stage providing two different
 2648 measurements: timing and luminosity. The 16 bits of the time measurement data, combined
 2649 with 1 bit for a hit flag, are then stored in a local memory (named *hit buffer*). The content
 2650 of this buffer is processed by a triggered-hit selector circuit on arrival of an L0/L1 trigger
 2651 signal, so this memory should allow latencies of up to 35 μ s. If a trigger signal is received,
 2652 the information is passed on to a secondary buffer named *matched hit buffer*, where it remains
 2653 there until it is retrieved for transmission to the common digital part. These local memories
 2654 are further described in Sec. 6.3.5.

2655 In order to measure the online bunch-by-bunch luminosity, each ASIC will report the sum
 2656 of hits within two different time windows. A schematic drawing of the windows is shown
 2657 in Fig. 6.2. A first 3.125 ns wide window (S1) is centred at the expected arrival time of the
 2658 particles from the collisions. The second window (S2) is adjustable in length and position in
 2659 steps of 3.125 ns, and will count the number of particles arriving before and/or after those
 2660 from the collisions. This side-band will provide valuable information about the background,
 2661 as described in Sec. 10.3.3. The windows are aligned to the expected arrival time of the
 2662 particle, with their length and alignment adjustable via configuration parameters. The
 2663 window generator is a control unit within the logic at the end of each column that contains a
 2664 4-bit counter running at 640 MHz and synchronized to the 40 MHz clock (both provided by
 2665 the phase-shifter further described in Sec. 6.4.3). These parameters will be optimised based
 2666 on operational experience.

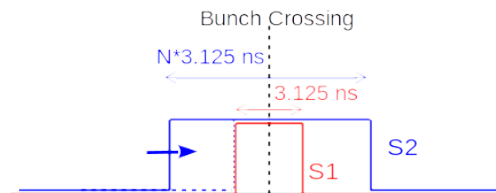


Figure 6.2: Illustration of the time windows used for counting hits for the luminosity data. The smaller window (S1, in red) is 3.125 ns wide and is centred at the bunch crossing time. The width and relative location of the larger window (S2, in blue) can be set in steps of 3.125 ns through the control parameters.

2667 This measurement is done in three steps. For the first step, performed at the single-channel
 2668 level, the output of the discriminator is passed through two programmable windows to
 2669 determine whether the hit happened within them. The way this is done is further described
 2670 in Sec. 6.3.6. Secondly the number of hits per column is computed by the EOC logic, and
 2671 thirdly the data is transferred. These last two steps are described in the next section and in
 2672 Sec. 6.4.

2673 Lastly, there are four 8-bits configuration registers per channel. They are read/written by the
 2674 slow control unit through a Wishbone bus. The configuration registers allow to configure

2675 several features of the TDCs, to enable/disable the discriminator and preamplifier, and to
2676 configure the per-channel threshold correction of the discriminator.

2677 6.2.2 Readout architecture

2678 Fig. 6.3 shows the conceptual design of the entire HGTD ASIC with 225 channels. The
2679 channel matrix is represented on the top part by 15×15 small squares. The schema of a
2680 single channel, as presented in Fig. 6.1, is repeated for each small square. The readout of the
2681 channels is done by column, through an EOC cell, drawn at the bottom of the matrix. The
2682 information is passed on to the trigger and luminosity processing units. A diagram of the
2683 main ASIC common digital part is presented at the bottom.

2684 A fast command unit receives the fast commands from the central Trigger Data Acquisition
2685 system (TDAQ), which consist of 8 bits on every bunch crossing. The 320 MHz clock is
2686 extracted from the fast commands and clock received from the lpGBT, and from the 40 MHz
2687 clock is generated. Based on this, a phase-locked loop (PLL) generates all the different clocks
2688 needed to operate the ASIC, namely 80 MHz, 640 MHz, and 1.28 GHz. These clocks will be
2689 centred with an accuracy of ~ 100 ps using a phase shifter, further described in Sec. 6.4.3.

2690 A fast command unit receives the fast commands from the central Trigger Data Acquisition
2691 system (TDAQ) through an lpGBT chip. These commands are 8-bit long and are received
2692 in series at 320 Mbit/s, so a command per bunch crossing. The communication between
2693 the fast command unit and the lpGBT chip is done through two lines. One is for serial data,
2694 and the other to transmit a clock of 320 MHz which will be used, not only to establish the
2695 communication between the lpGBT and the ASIC, but also as a source clock from which all
2696 the internal clocks needed to operate ALTIROC2 will be generated. The 320 MHz clock from
2697 the lpGBT is divided by 8 and the passed to a Phase-Locked Loop (PLL) produces clocks
2698 of 40 MHz, 80 MHz, and 640 MHz. These clocks will be centred with an accuracy of 100 ps
2699 using a phase shifter. Further details about the clock generation and distribution are given
2700 in Sec. 6.4.3 and Sec. 6.4.4.

2701 The fast commands are processed by the Trigger Data Processing Unit (TDPU) which is
2702 responsible to read the timing information from the pixel matrix, pack these data into
2703 frames and serialize them. It is composed of a 12-bits bunch crossing counter to generate a
2704 bunch crossing identifier (BCID), a trigger table to store temporally trigger events for later
2705 processing, a data formatting unit that packs data into frames, and a serializer. More details
2706 are given in Sec. 6.4.1.

2707 The TDPU performs two tasks in parallel, one is to process incoming triggers and the other
2708 to readout data associated to a trigger event from the pixel matrix. In the incoming trigger
2709 processing task, the TDPU generates an internal trigger signal and a trigger identifier (TrigID)
2710 when an L0/L1 accept command is received. These are transmitted immediately to all the
2711 pixels. Then each pixel checks if it has data associated to that trigger event. If they have,

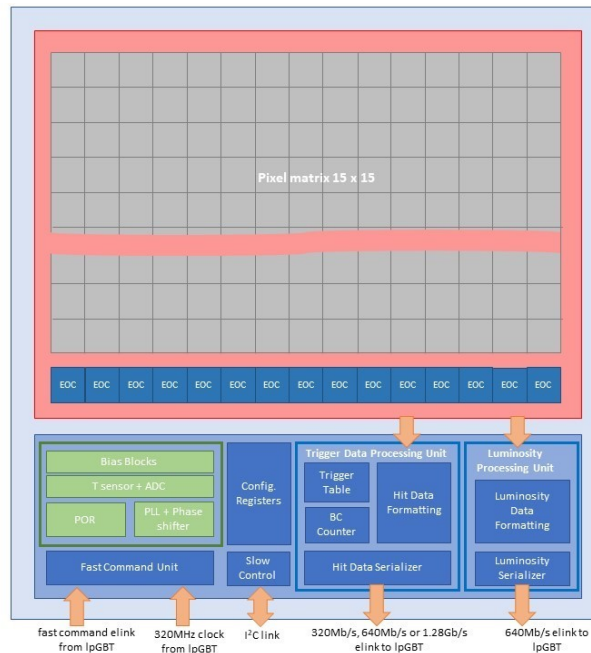


Figure 6.3: Schematic of the full HGTD ASIC. The top part represents the 15×15 channel matrix, while the bottom part shows the ASIC common digital part.

2712 these transferred, together with the corresponding TrigID to a secondary in-pixel buffer.
 2713 They remain there until are retrieved by the TDPU. The TrigID is used to tag a BCID with a
 2714 trigger event with only 5-bits, so it is not necessary to send the 12-bits of the BCID to the pixel
 2715 matrix. The trigger table is a FIFO that stores the correspondence between each BCID and its
 2716 associated TrigID. In the readout task, the TDPU is looking for a new entry in the trigger
 2717 table. When a new one is found, it requests to the EOCs to retrieve and store the data from
 2718 the pixels related to the TrigID fetched from the table. Then, the data are moved into the Hit
 2719 Data Formatting unit, where they are packed into frames, serialised and transmitted to the
 2720 peripheral on-detector electronics through e-links. The transmission speed of the e-link will
 2721 depend on the radial position of the ASIC, and will be set via and Inter-Integrated Circuit
 2722 bus I²C to one of three values: 320 Mbit s^{-1} , 640 Mbit s^{-1} , and 1.28 Gbit s^{-1} . It is connected to
 2723 an equal speed port in the lpGBT, described in Sec. Sec. 9.2.1.

2724 As mentioned previously, the luminosity measurement is carried out in three steps, each
 2725 one in a different region of the ASIC. The first step consists in determining whether the
 2726 hit occurred within one or both of the time windows. This windowing process is done
 2727 at the single-channel level and was described in the previous section. The windows are
 2728 generated in the logic at the end of each readout column, instead of at each channel, in order
 2729 to reduce power consumption. By distributing them to the channels as a clock tree, one can
 2730 compensate for the delays introduced by the long metal lines needed to reach each channel
 2731 and to minimize the skew between the channels in a column. In the second step, the result

2732 is collected at the EOC logic, where the number of hits in the column for each window is
 2733 computed. This information is passed on to the Luminosity Processing Unit (LPU), that
 2734 calculates the total number of hits in the ASIC within S1 and S2 windows. Then it performs
 2735 the subtraction of the hits within the larger and the smaller window (S1-S2). The 8 bits of S1
 2736 and the 8 bits of S2-S1 are truncated to respectively 7 and 5 bits to reduce the total bandwidth.
 2737 In the third step, each 12 bits package is transferred to the luminosity serializer where
 2738 data is encoded 6bto8b, leading to frames of 16-bits long. These are serialized at a rate of
 2739 40 MHz and sent to the lpGBT through a 640 Mbit s⁻¹ s-1 e-link. The measurement and data
 2740 transmission can be enabled/disabled by accessing one of the configuration registers. As
 2741 explained previously in Sec. 6.1, not all ASICs will be performing luminosity measurements.
 2742 Disabling the data transmission on those not performing the measurement will allow to save
 2743 power.

2744 The common digital part also includes several programmable digital to analog converters
 2745 (DACs) to generate different bias currents for all analog blocks of the ASIC, a band-gap, a
 2746 temperature sensor (under consideration) and some configuration registers. The latter are
 2747 used to set different features of the ASIC, such as the values of the DACs, the transmission
 2748 rate of the hit data and the PLL bias currents or frequencies. As mentioned previously, 4
 2749 configuration registers are also present for each channel. The I²C link mentioned previously
 2750 is also used to readout all configuration registers in order to check if SEU events have
 2751 corrupted their content, and to retrieve information from the control unit about the status of
 2752 the ASIC; the information related to data corruption is then passed on to the hit serializer.

2753 6.3 Single-channel readout electronics

2754 This section describes in detail the design of the single-channel readout electronics. As
 2755 introduced previously, it will receive the pulse signal from the LGAD sensor, and transmit
 2756 the TOA, TOT and luminosity information to the EOC logic. The preamplifier design is
 2757 first described in Sec. 6.3.1, while the discriminator and time-walk correction are presented
 2758 in Sec. 6.3.2. Concerning the digital blocks, the working principle of the time-to-digital
 2759 converters is presented in Sec. 6.3.3, while the designs of the local memory and the luminosity
 2760 processing unit are presented in Sec. 6.3.5 and Sec. 6.3.6 respectively.

2761 6.3.1 Preamplifier

The jitter due to electronics noise is often modelled as

$$\sigma_{\text{jitter}} = \frac{N}{dV/dt} \sim \frac{t_{\text{rise}}}{S/N} \quad (6.1)$$

2762 where N is the noise and dV/dt the slope of the signal pulse, of which S is the amplitude
 2763 and t_{rise} the rise time. Due to the fact that the noise scales with the bandwidth (BW) as
 2764 \sqrt{BW} , while the rise-time grows with the amplitude as S/BW , the most common timing
 2765 optimisations rely on using the fastest preamplifier.

2766 Most timing measurements in test beam have been carried out with broadband amplifiers,
 2767 which are voltage sensitive amplifiers with 50Ω input impedance. Some prefer using
 2768 a trans-impedance configuration and timing optimisation has been published for such
 2769 configuration [37, 42] However, in silicon sensors such as LGADs, the preamplifier speed is
 2770 not so crucial, both due to the fact that the current duration is not negligible with respect to
 2771 the preamplifier rise time and to the capacitive impedance of the sensor.

2772 The jitter with a voltage sensitive amplifier configuration can be easily calculated under some
 2773 simplifications and assuming that the detector current is a short pulse with a characteristic
 2774 time t_d . The corresponding input charge Q_{inj} is the integral of this current over t_d . The jitter
 2775 of a preamplifier can then be estimated through the following formula:

$$\sigma_{\text{jitter}} = \frac{e_n C_d}{Q_{\text{in}}} \sqrt{\frac{t_{r,pa}^2 + t_d^2}{2t_{r,pa}}} \quad (6.2)$$

where e_n is the noise spectral density and C_d the detector capacitance. The sensor drift time
 t_d and the preamplifier rise time $t_{r,pa}$ are combined in quadrature as an estimation of the
 total speed. It can be seen that the jitter is minimized when the preamplifier rise time is
 equal to the sensor drift time: $t_{r,pa} = t_d$. In that case, the jitter can be written as:

$$\sigma_{\text{jitter}} = \frac{e_n C_d \sqrt{t_d}}{Q_{\text{in}}} = \frac{ENC t_d}{Q_{\text{in}}} \quad (6.3)$$

2776 However this dependence is small: for instance for $t_d \sim 600$ ps, reducing or increasing by
 2777 a factor of two $t_{r,pa}$ with respect to the optimal matching value will deteriorate the jitter
 2778 by approximately 12%. Therefore to minimize the jitter, the sensor should have a small
 2779 capacitance, a small t_d and provide a large charge. For a $50\mu\text{m}$ thick LGAD in HGTD, a
 2780 $C_d = 4$ pF has been estimated when fully depleted (see Fig. 5.4(b)); typical $t_d \sim 0.6$ ns, and
 2781 for a gain of 20 it would give a $Q_{\text{inj}} \sim 10$ fC.

2782 The design of the ALTIROC uses a voltage sensitive preamplifier, presented in Fig. 6.4. This
 2783 is a broadband preamplifier with a cascoded Common Source configuration, consisting
 2784 of an input transistor ($M1$) and a follower transistor ($M2$). Both the gain and the noise
 2785 depend on the current that flows into the input transistor, which is why the drain current
 2786 I_d is tunable through configuration parameters. To this purpose two current sources are
 2787 combined: I_{d1} is a fixed current source of $150\mu\text{A}$, while I_{d2} can be varied from 0 to $850\mu\text{A}$.
 2788 Simulation studies have shown that the gain is small when increasing this current beyond
 2789 $600\mu\text{A}$. The rise time of the preamplifier can be modified, and so evaluate its impact on the

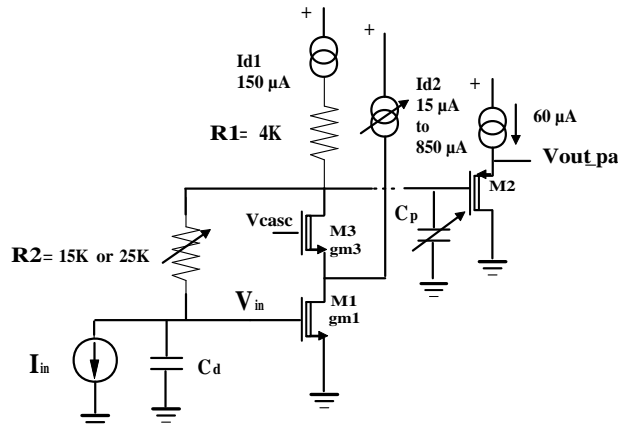


Figure 6.4: Schematic for the preamplifier implemented in the latest ASIC design, ALTIROC1.

2790 jitter. This is done through the pole capacitance, C_p , that is tunable by slow control (from
 2791 0 to 175 fF) allowing to set the preamplifier rise time between 300 ps and 1 ns. As for the
 2792 fall time of the preamplifier output, it depends on the input impedance of the preamp (R_{in})
 2793 that is given by the resistance R_2 divided by the open loop gain of the preamplifier. The
 2794 value of the input impedance depends therefore also on the drain current I_d . For example,
 2795 for an $I_d = 300 \mu\text{A}$ and $R_2 = 25 \text{k}\Omega$, the input impedance is around $1.6 \text{k}\Omega$. The value of
 2796 the resistor R_2 can be either $15 \text{k}\Omega$ or $25 \text{k}\Omega$. It can also absorb the sensor leakage current,
 2797 estimated to be below $5 \mu\text{A}$ after irradiation. The leakage current would cause the output
 2798 of the preamplifier to drift by an amount of the order of $R_2 \times I_{leak}$. The threshold of the
 2799 discriminator that follows the preamplifier must then be changed accordingly. This can be
 2800 done using the 7-bit DAC threshold correction that is integrated for each channel allowing a
 2801 correction within $\pm 50 \text{ mV}$.

2802 The preamplifier architecture, followed by a fast discriminator, has been simulated with
 2803 various detector capacitances and considering that 1 MIP would deposit a 10 fC charge,
 2804 which corresponds to an amplification gain of 20 in the LGAD. A calibration signal was used
 2805 in the simulation, and the result was convoluted with different input LGAD signals. The
 2806 LGAD pulses for different levels of irradiation obtained using the Weightfield2 software [50]
 2807 and presented in Fig. 4.4(b) were used as input, and the obtained preamplifier pulses are
 2808 presented in Fig. 6.5.

2809 6.3.2 Discriminator and time walk correction

2810 The measurement of the TOA of the particles is performed by a discriminator that follows
 2811 the preamplifier. The measurement of the time of the rising edge of the discriminator pulse
 2812 provides the TOA, while that of the falling edge, combined with the TOA, provides the
 2813 TOT. To ensure a jitter smaller than 10 ps, the discriminator is built around a high speed

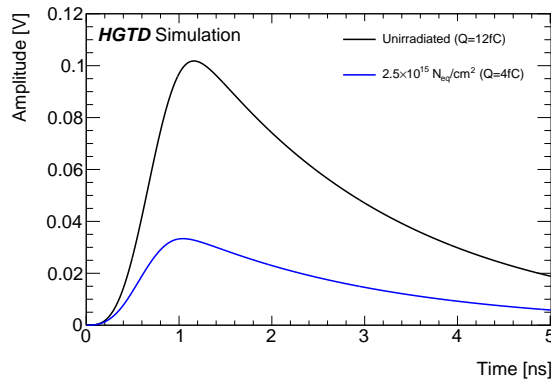


Figure 6.5: Simulation of the preamplifier output using as input the simulated LGAD signals presented in Fig. 4.4(b) for a non-irradiated sensor and after irradiation.

2814 leading edge architecture with hysteresis to avoid re-triggering effects. Two differential
 2815 stages with small input transistors are used to ensure a large gain and a large bandwidth
 2816 (aprox 0.7 GHz). The threshold of the discriminator (V_{th}) is set by a 10-bit DAC common to
 2817 all channels. An additional 7-bit DAC allows to make small V_{th} corrections individually for
 2818 each channel in order to compensate for differences amongst them or for different values of
 2819 leakage current.

2820 6.3.3 TDC

2821 The target timing resolution (quantisation step) of the TDC of the TOA is 20 ps, and is below
 2822 the gate-propagation delay in 130 nm technology, thus the Vernier delay line configuration
 2823 is employed. This configuration consists of two lines, each composed of a series of delay
 2824 cells implemented as differential shunt-capacitors, controlled by a voltage signal (V_{ctrl}) that
 2825 determines their delay. The timing resolution is determined by the difference in the delays
 2826 of the cells in each line. The TOA will be measured within a 2.5 ns window centred at the
 2827 bunch-crossing. As already mentioned before, the hits have a time dispersion with an RMS
 2828 of around 300 ps, so that such a window aligned with a precision of 100 ps contains all the
 2829 hits. The maximum conversion time for a 2.5 ns range must be below 25 ns so that hits
 2830 happening in the following bunch crossing can be converted.

2831 A graphic representation of the working principle of the TDC can be found in Fig. 6.6. In
 2832 the 'slow' line, the control voltage fixes the delay of each cell to 140 ps, while on the 'fast'
 2833 line it fixes it to 120 ps. The START signal (output of the discriminator) enters the "slow"
 2834 delay line while the STOP signal (end of measurement window) enters the 'fast' delay line.
 2835 Although initially the START signal is ahead of the STOP one, each delay-cell stage brings
 2836 them closer by an amount equal to the difference between the slow and fast cell delays, i.e.
 2837 20 ps. The number of cell stages necessary for the STOP signal to surpass the START signal

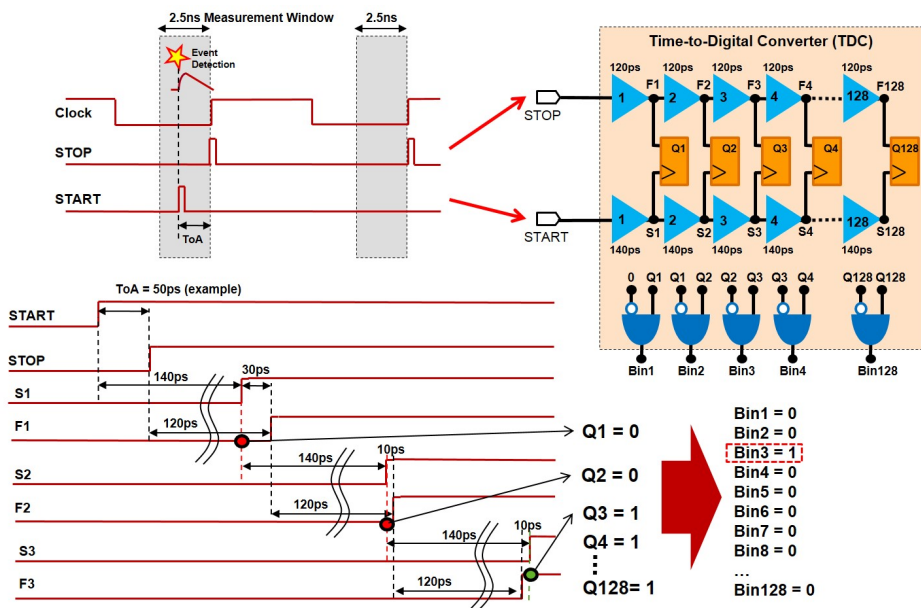


Figure 6.6: Graphic representation of the working principle of the TDC. The drawing on the top left shows how the START and STOP signals are generated, the first with the discriminator output upon event detection, the second corresponding to the next clock edge. The gray area indicates the 2.5 ns detection window. On the top right, the schema represents the TDC, with the 'slow' delay line (140 ps cells) that propagates the START signal, and the 'fast' delay line (120 ps cells) in which the STOP signal is propagated. The difference between delays defines the bin. After each cell the signals are compared (QX), and the bin number provides the converted measurement.

2838 represents the result of the time measurement with a quantisation step of 20 ps. A cyclic
 2839 structure is employed to reduce the number of cells per line and results in a smaller occupied
 2840 area. Since the time measurement is initiated only upon signal detection (instead of at each
 2841 time-measurement window), the reverse START-STOP scheme is used as a power-saving
 2842 strategy. The conversion time of a 2.5 ns input time interval is about 21 ns, finishing just in
 2843 time to be able to accept and process the signal in the next bunch crossing.

2844 The TOT TDC employs an additional coarse delay line for extending the range to 20 ns, while
 2845 the Vernier delay line (identical to the one used in TOA TDC) provides the high resolution
 2846 of 40 ps.

2847 The TDC power consumption is dependent on the time-interval being measured. For the
 2848 TOA TDC 2.5 ns (full dynamic range), the average power consumption over the 25 ns mea-
 2849 surement period is about 5.2 mW. It will become 3.5 mW for the time-interval equal to half
 2850 dynamic range. Thanks to the reverse START-STOP operation, the power consumption of the
 2851 TDC is much lower in the absence of a hit over threshold. This results in an average power
 2852 consumption per channel of 1.1 mW for both TDCs, assuming a time interval uniformly
 2853 distributed (1.25 ns average) and a maximal channel occupancy of 10%.

2854 6.3.4 Command pulser for calibration

2855 An internal pulser, common to all channels, is integrated to mimic input charges. This
 2856 pulser can be used to intercalibrate the gain of each channel by performing trigger efficiency
 2857 measurements for various input charges. This intercalibration allows to align the threshold
 2858 of each channel using a correction threshold DAC integrated for each channel. The pulser
 2859 consists of a programmable DC current (tunable with an internal 6-bit DAC) that flows
 2860 continuously in a 50 k Ω resistor until it is interrupted by a command pulse that shorts the
 2861 resistor to ground (see Fig. 6.7). A voltage step (V_{step}) equal to $-R \times I_{DAC}$, is then generated
 2862 and sent through the selected pixel internal 200 fF test capacitors (C_{test}). The simulated input
 2863 charge (Q_{inj}) is equal to $C_{test} \times V_{step}$. The dynamic range goes from 0 to 250 mV or 0 fC up
 2864 to ~ 50 fC (LSB = 0.76 fC). This pulser will also be used to calibrate the absolute value of the
 2865 phase. The command pulse that is encoded in the fast command elink is therefore distributed
 2866 as a clock tree inside the ASIC.

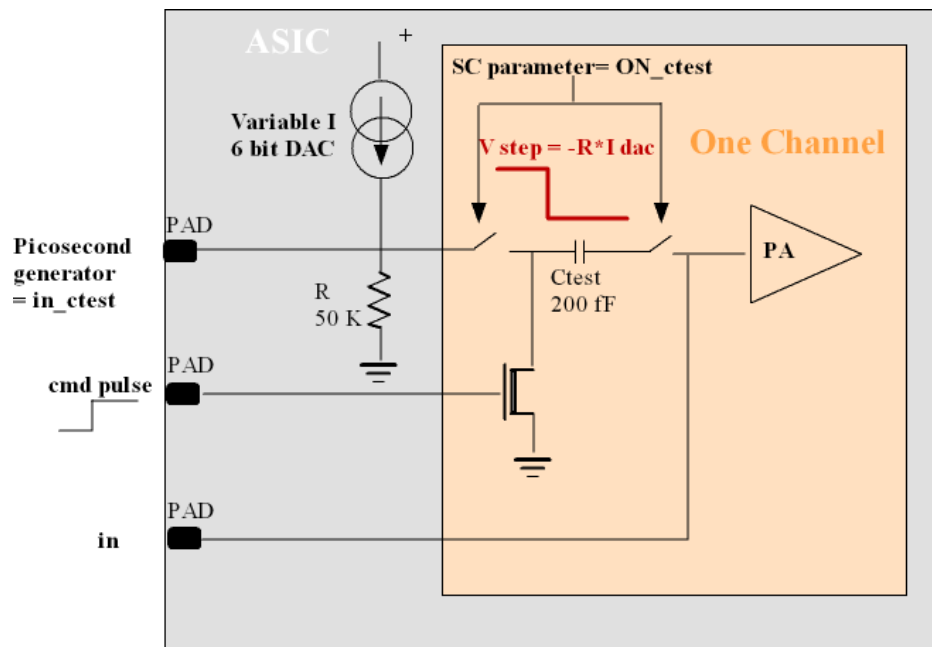


Figure 6.7: Pulser principle that shows the common 6-bit current DAC used to set the input charge as well as the pixel Ctest capacitor.

2867 6.3.5 Hit processor

2868 Each channel electronics is composed of an analog part, already described, and a digital part.
 2869 The latter is composed of three main blocks, as can be seen in the schematics of Fig. 6.1. The
 2870 hit processing unit, or hit processor, temporarily store the data related to a hit and select hits

2871 of events that have been triggered. The main circuit is the hit buffer which is composed of a
2872 memory of 1400 positions. Such size will allow to cope with trigger latencies of $35\ \mu\text{s}$, using
2873 one position per bunch crossing.

2874 The size of each buffer position is 19 bits: 7 for the TOA, 9 bits for the TOT, 1 bit for the hit
2875 flag, 1 bit for detection error (CRC) and 1 bit for the TOA overflow. This hit flag bit indicates
2876 if a hit has been detected in the bunch crossing. The buffer is implemented as a circular
2877 memory in order to store data in a continuous way. It has two pointers for memory reading
2878 and writing. A control unit in the hit buffer increments the write pointer in one unit each
2879 bunch crossing. The pointer goes from 0 to 1399 during an L0/L1 scenario (latency of $35\ \mu\text{s}$)
2880 and then goes back to position 0. For the L0 scenario, in which the trigger latency is $10\ \mu\text{s}$,
2881 the write pointer is incremented from 0 to 399. The latency is set through a configuration
2882 register at the periphery. In each bunch crossing the control unit checks if a hit occurs. In
2883 case of hit, the TOA and TOT measured by the TDCs in the analog front-end electronics
2884 stage are stored into the buffer and the hit flag of that position is set to 1. If not, the hit flag is
2885 set to 0 and no values are written in the TOT and TOT fields in order to save power.

2886 The hit buffer architecture is built around a two-port SRAM design. This configuration
2887 allows simultaneous Read/Write operations within the same clock period. 6 partitions of 256
2888 words are used in order to limit the lines capacitance and to optimize the power consumption.
2889 The power consumption, simulated taking into account parasitic elements and assuming a
2890 10% occupancy with a L0 trigger signal at 4MHz, is evaluated to 1.55 mW (at 25°C , 1.2V).
2891 This power dissipation decreases down to 1.2 mW with a 2.5% occupancy obtained with a
2892 L1 trigger signal at 1MHz. Concerning radiation tolerance, this SRAM architecture is less
2893 sensitive to SEU than DRAM as nodes levels are regenerated by the back to back inverters:
2894 ionizing radiations will significantly change the amount of charge on nodes but, assuming
2895 they don't completely flip the bits, the node levels will be restored to their normal value
2896 quite quickly, either by the feed-forward or by the feedback inverter. However, in order to
2897 improve the radiation tolerance, the memory cells are designed with large HVT transistors
2898 and with strong substrate/well contacts, sacrificing density for more robust and radiation
2899 tolerant design. The full active area of the hit buffer is $720\ \mu\text{m} \times 1080\ \mu\text{m}$.

2900 The reading pointer is handled by the next stage in the hit processing unit, the *trigger hit*
2901 *selector*. It transfers the TOA and TOT information to the *matched hit buffer* if it finds a hit
2902 flag equal to one when it receives a trigger. It also tags each trigger with an identifier TrigID
2903 provided by the TDPU of the periphery, which is stored together with the TOA and TOT
2904 in the matched hit buffer. This allows to know to which triggered event the data stored in
2905 the buffer is associated. This buffer operates as an average rate memory, storing the hits
2906 of triggered events until ready to be transferred. It will allow to cope with event-to-event
2907 fluctuations in the number of matched hits and to keep the bandwidth of the ASIC lower
2908 than $1.28\ \text{Gbit s}^{-1}$. It is implemented with a FIFO (*first in first out*), in which each position
2909 contains 21 bits: 16 for the TOA and TOT information, and 5 bits for the TrigID. The current
2910 design has a depth of 32 that could eventually be reduced in case simulations prove it

2911 possible. The writing of the data into the FIFO is done by the trigger hit selector block, while
 2912 the readout is performed by the EOC logic by placing a requested trigger ID (RqtTrigID).

2913 6.3.6 Luminosity processing unit

2914 As already described before, the windowing process of the luminosity measurement is
 2915 carried out on-channel, which is needed because of the large area of the chip. Transmitting
 2916 the output of the discriminator to the luminosity block at the periphery would imply the use
 2917 of a metal line of several millimetres. Such a long metal line would have large equivalent
 2918 RC that would delay the signal by several nanoseconds. The length of each channel-to-
 2919 luminosity block connection would vary from channel to channel and so would the delay.
 2920 As a result, these delays might cause that some hits inside one of the windows would be
 2921 outside, corrupting the measurement of the luminosity. The compensation of the delay for
 2922 each channel would be difficult. A simpler solution is to perform the windowing process
 2923 on-channel. This avoids the need to transmit the output of the discriminator to the periphery.
 2924 However, the windows must be distributed through the whole channel matrix. Again, long
 2925 metal lines are needed but their delays can be compensated by distributing them as a clock
 2926 tree.

2927 A scheme of the first step in the luminosity measurement is presented in Fig. 6.8. At the
 2928 channel level, an AND gate evaluates if the output of the discriminator is inside the window.
 2929 It generates a pulse that triggers a positive edge detector made of a flip-flop D with its D
 2930 input connected to a logic '1'. When a positive edge is detected, the output of the flip-flop D
 2931 goes high. This signal is asynchronous, so a synchronizer retimes the signal with a 40MHz
 2932 clock. The output of the synchronizer read out at each clock cycle and processed in the
 2933 end-of-column logic.

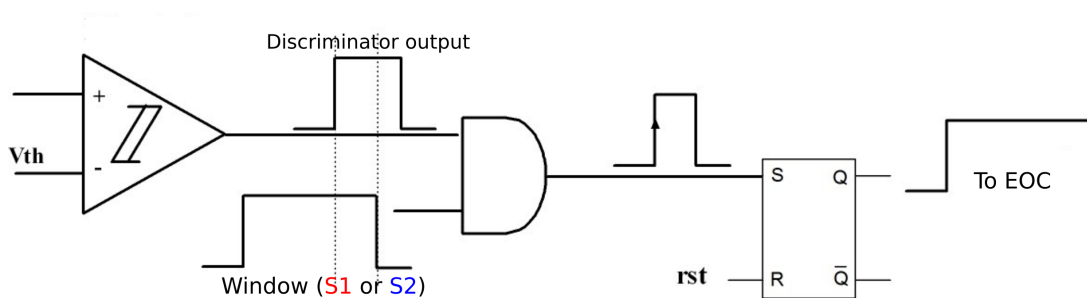


Figure 6.8: The signal of the discriminator is compared to the luminosity window (for each window) and a signal is transmitted to the end-of-column logic.

2934 6.4 ASIC End of Column logic and digital blocks

2935 This section describes the ASIC common digital part, including the readout process of
2936 the channels and the various blocks with specific functions. The design of many of these
2937 components is on-going.

2938 6.4.1 Matrix readout process

2939 Timing data readout

2940 As described previously, in order to read out the timing information it is necessary to create a
2941 table that matches the BCID provided by the TDAQ system and the internal trigger identifier.
2942 TDPU has a 5-bits counter to tag the trigger events that are being received. The counter
2943 can be initialized with the fast command used to reset the chip. When the trigger data
2944 processing unit receives a trigger command, it stores the content of the counter together
2945 with the corresponding BCID into the trigger table and increases the counter by one unit.
2946 When the counter reaches the largest value, it wraps up to 0. The trigger table is a FIFO
2947 with 32 positions of 17-bits each one: 12 bits for the BCID and 5 bits for the trigger identifier
2948 (TrigID). If the FIFO is full, an error message is generated and transmitted to the TDAQ
2949 through an e-link. The TDPU unit also generates an internal trigger signal (trig) with a
2950 duration of one clock cycle. This is immediately transmitted to all matrix channels as well
2951 as the trigger identifier trigID. Fig. 6.9 shows a block diagram of the main signals involved
2952 between the TDPU and the EOC. Both, the trigger signal and the identifier are processed by
2953 the hit processor as described in Sec. 6.3.5.

2954 The hit data formatting unit in the TDPU it is always checking if there is an entry in the
2955 trigger table. When it finds one, it fetches the entry and initiates the readout of the data
2956 stored in the matrix associated to that trigger event. The readout is carried out in two steps:
2957 first the retrieval of data associated to a given TrigID from the columns, and then the frame
2958 construction and data transmission. In the first step, the hit data formatting unit places the
2959 TrigID of the entry from the trigger table in the rqtTrigID bus and asserts the checkMatrix
2960 signal to indicate to all the EOC to retrieve data from the pixels. Then the EOC asks to all
2961 the pixels to check if they have data associated to that trigger identifier by asserting the
2962 checkTrigID signal. The hit data processor checks if there is a matched hit with the same
2963 trigger identifier as the requested trigger. If there is, a hit flag is asserted. Once all the pixels
2964 have checked if they have data, then the EOC starts reading all the pixels that have such flag
2965 asserted, one per clock cycle. The row address, TOT and TOA of each read pixel are stored
2966 in a FIFO placed at the EOC. When the data of a pixel have been read and stored, the flag
2967 of that pixel is set to low. Once all the pixels have been read, the EOC indicates to the hit
2968 data formatting block that is done by asserting the doneMatrix signal. In the second step,
2969 the hit data formatting block starts reading the FIFO of the end of the columns immediately

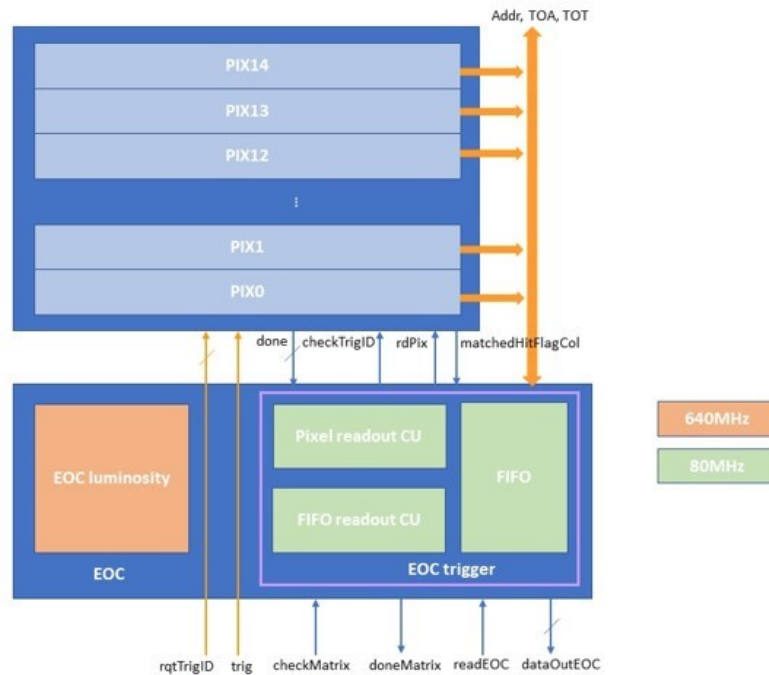


Figure 6.9: Block diagram of the main signals involved between the communication of the EOC with the pixels and the TDPU.

2970 at the moment that the FIFO has data available asserting the readEOC signal. The hit data
 2971 formatting does not wait that the EOCs have finished to retrieve from the pixels to start
 2972 reading the FIFOs. The column address is added to each FIFO entry and the data are placed
 2973 at the dataOutEOC bus. The TDPU packs the data in frames and serializes them. Once all
 2974 the buffers have been read and their data transmitted, the hit data formatting block waits for
 2975 a new entry in the trigger table and the loop is executed again.

2976 **Luminosity data readout**

2977 The instantaneous luminosity, that is, the number of detected hits in the pixel matrix per
 2978 bunch crossing, is measured every 25 ns. The process is carried out in three different regions
 2979 of the ASIC as already described in Sec. 6.2.2 The windowing is performed in pixel. The 2
 2980 windows are generated at the EOC and distributed to the whole column as a clock tree in
 2981 order minimize the skew from pixel to pixel. A trade off needs to be found between power
 2982 consumption and skew. The first trials of physical synthesis show a skew of 100ps. The
 2983 windows are generated with a programmable FSM running at 640 MHz clock. This FSM
 2984 divides the bunch crossing into 16 equal intervals of 1.5625 ns with a 4-bits internal counter
 2985 that continuously counts from 0 to 15. A control unit monitors asserts and deasserts the
 2986 two window signals window1 and window2 as a function of the value of the counter and

2987 of the 4-bits parameters minW1, minW2, maxW1, and maxW2 as shown in ???. The 4-bits
 2988 parameters indicate in which intervals window1 and window2 must be high or low. The
 2989 width of window 1 is fixed to 3.125 ns so the default values of minW1 and maxW1 are 1
 2990 and 14. However, both values can be modified in case it was necessary. The 1.5265 ns time
 2991 resolution of the window generator is not enough to center the position the windows respect
 2992 to the beginning of the bunch crossing. In order to provide the required resolution, the phase
 2993 of the 640MHz clock used by the EOC can be adjusted through the phase shifter. This 640
 2994 MHz clock is independent from the 640 MHz clock used in the serializers. More details are
 2995 given in Sec. 6.4.3 The windowing process at pixel level is described in Sec. 6.3.6. Every
 2996 pixel produces two measures per bunch crossing. The EOC sums the luminosity measures
 2997 of the whole column per bunch crossing. Those measures are passed to the luminosity
 2998 processing unit. This sums the measures of the columns. The luminosity of the window 1,
 2999 S1, is subtracted from the luminosity of window 2, S2. The result S2-S1 and S1 are trunked to
 3000 4 and 6 bits respectively. Both values encoded with 6b8b code, producing a 16-bit frame per
 3001 bunch crossing. Frames are serialized at 640 MHz. The whole bandwidth is occupied with
 3002 the luminosity data. In order to avoid desynchronization, a synchronization frame needs to
 3003 be sent periodically, which requires to lose the luminosity information of a bunch crossing.

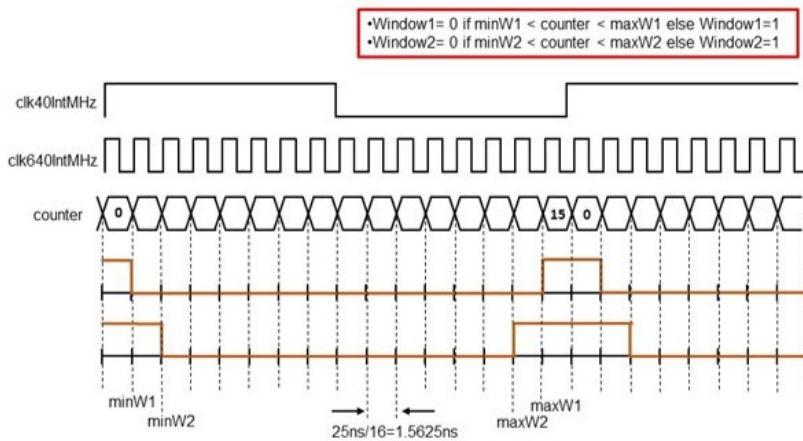


Figure 6.10: Block diagram of the eoc luminosity unit.

3004 6.4.2 Slow control

3005 The slow control is used to configure the ASIC as well as to retrieve information of its
 3006 internal status. For such a purpose, up to 1024 configuration registers of 8-bits each have
 3007 been implemented in ALTIROC. The memory map is not yet completely determined, but the
 3008 first 900 positions are dedicated to configure the channel registers. Each channel contains
 3009 4 configuration registers. The other 124 registers will be located at the periphery and will
 3010 be used to configure the hit data transmission rate, enable/disable the luminosity block, to

3011 program the length of the windows used for the luminosity, etc ... For the final ASIC, the
3012 configuration registers are read/write by using an I²C link while shift registered are used for
3013 the prtotype. The I²C link in the ASIC is slave to the master in the lpGBT in the peripheral
3014 electronics, described in Sec. 9.2.1.

3015 6.4.3 Clock generator and phase shifter

3016 The clock generator and phase shifter are a function block located in the ASIC common
3017 digital part used for the adjustment of the clock phase for the clock system in the ASIC. Some
3018 examples are the 40 MHz used by the TDCs and the 640 MHz used to generate the luminosity
3019 windows. A schematic of its design is presented in Fig. 6.11. The clock generator provides
3020 two clocks of 40 MHz and 640 MHz to the phase shifter from one of two clock sources.
3021 The first source, clk40MHz, is a 40 MHz clock coming from the fast command unit. A PLL
3022 multiplies by 8 this clock. The second source is an external clock of 640 MHz, clk640MHzExt,
3023 that is divided by 8. The clock source is selected through the clockSel signal.

3024 The phase shifter receives the clock CMOS signals of different frequencies (40MHz, the
3025 640MHz is present in the current prototype for debugging) generated by the PLL and
3026 outputs them with the same frequency but with an adjusted phase. This module is required
3027 to provide a shift step smaller than 100 ps, an additional jitter below 5 ps on the 40MHz
3028 clock, and a power consumption around 10mW.

3029 The design presented here is adapted from an lpGBT design in the 65 nm process. The core
3030 of the phase shifter is composed of two delay-locked loop (DLL) formed by delay line of
3031 16 delay cells (not including the dummy cell at the end of the delay line). The DLLs are
3032 used to delay the input clock signal of 640 MHz generated by the clock generator and to
3033 provide two 640 MHz clocks two the ASIC. One of these clocks will be used to serialize data,
3034 clk640MHzInt, and the other to generate the time windows, clk640MHLumInt. More details
3035 are given in ??.

3036 For the 80 MHz and 40 MHz clocks, coarse phase adjustment circuits are needed. Their
3037 output is re-sampled by the clk640MHzInt, with the result that these two clock signals
3038 acquire the same time resolution. Their time resolution is therefore 1/16 of the 640MHz
3039 clock period, equal to 97.6 ps.

3040 6.4.4 Clock distribution

3041 Fig. 6.12 shows a block diagram of the different clock domains inside ALTIROC2. All the
3042 internal clocks are derived from a source clock of 320MHz coming from the lpGBT chip.
3043 The fast command unit divides this clock by 8 producing a 40MHz clock, clk40MHz. This
3044 is passed to the clock generator, described in Sec. 6.4.3., that generates a 640MHz clock
3045 with a PLL. Both clocks are connected to the phase shifter that generates 1 clock of 40MHz,

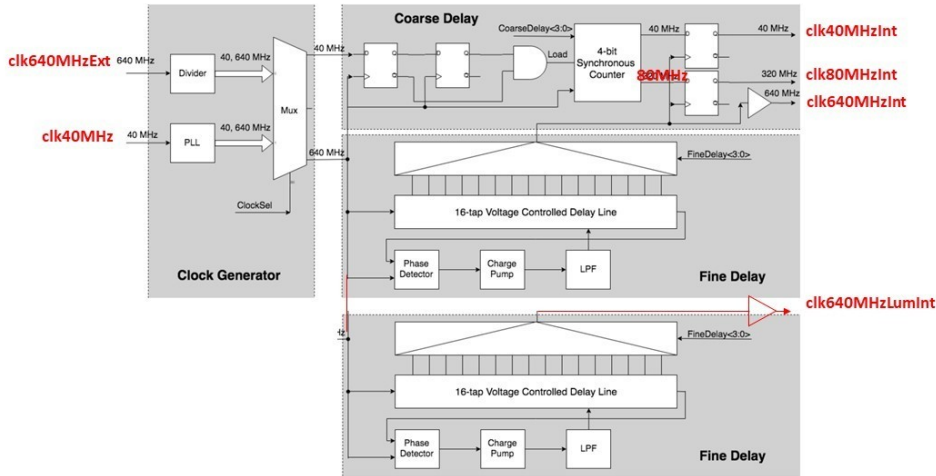


Figure 6.11: Schematic diagram of the clock generator and phase shifter.

3046 clk40MHzInt, 1 clock of 80MHz, clk80MHzInt, and 2 clocks of 640MHz named clk640MHzInt
 3047 and clk640MHzLumInt. Clk40MHzInt, clk80MHzInt, and clk640MHzInt are aligned in phase
 3048 but shifted from the clk40MHz as described in Sec. 6.4.3. Clk640MHzLumInt is used
 3049 to generate the time windows W1 and W2. Its phase can be adjusted with a resolution of 100ps
 3050 with the phase shifter in order to fine tuning the position of both windows from the bunch
 3051 crossing as described in Sec. 6.4.1.

3052 Most of the digital electronics run at 40MHz, these include the I²C and the configuration
 3053 registers. The 80MHz clock, clk80MHzInt, is used to readout the timing data from the pixel
 3054 matrix and to pack the data into frames in the trigger data processing unit. Finally, the
 3055 640MHz clock clk640MHzInt is used to serialize data.

3056 6.5 Radiation tolerance

3057 Two radiation effects must be taken into account: the Total Ionizing Dose (TID) that may
 3058 degrade the timing performance and the Single Event Effects which may corrupt the config-
 3059 uration registers and the time data. The worst expected TID and fluency are respectively
 3060 210 Mrad and $2.5 \times 10^{15} N_{eq}/cm^2$ taking into account the replacement of the inner modules
 3061 every 1000 fb⁻¹. The ASIC has been designed using TSMC130 nm technology which has
 3062 been tested up to 400 Mrad, which is two times above the requirement. Nevertheless known

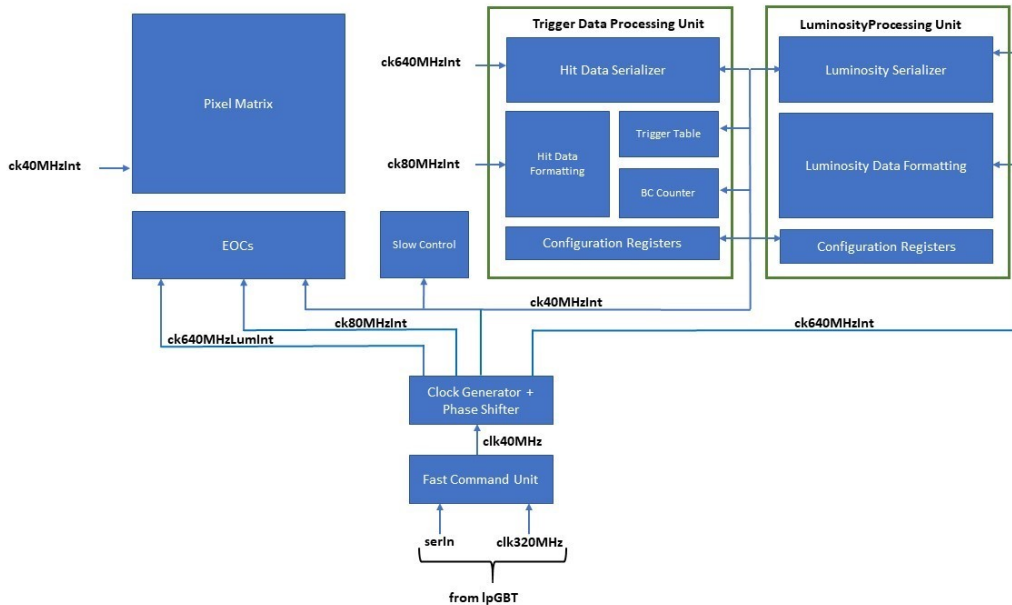


Figure 6.12: Schematic of the clock distribution.

3063 strategies have been used in the ASIC design to mitigate the radiation effects. TID radiations
 3064 degrade the performance of MOSFET by increasing their threshold and generating leakage
 3065 currents. To counter these effects, bias currents of analog blocks are set to quite large values
 3066 ($> 20 \mu A$) compared to the expected leakage currents and low VT transistors are avoided in
 3067 current sources. In addition, minimum size transistors are avoided, for PMOS transistors in
 3068 particular. At the layout level, substrate contacts are used to avoid latch-up. The DLLs of the
 3069 TDC part are designed to take care of radiation, temperature and voltage variations inside
 3070 the chip automatically. Besides, as the TDC bins are given by the difference of two delays,
 3071 it ensures compensation for variations under irradiations. As for the digital part and the
 3072 SEU tolerance, Triple Modular Redundancy (TMR) will be implemented on critical parts
 3073 of the 225 channel version (ALTIROC2). Simulations of upsets using CERN tools will be
 3074 performed to fully evaluate the effect of SEUs on the chip functioning.

3075 6.6 ASIC Power distribution and grounding

3076 To preserve the signal integrity and the jitter performance under an important digital activity,
 3077 the power distribution must be done carefully at the ASIC level. Each analog block (PA,
 3078 discri, TDC) is in a deep Nwell that is powered and grounded with its own power line and
 3079 ground line. All powers and grounds are therefore separated. Great care must be taken to

3080 reduce the resistance of the power lines, especially for the preamplifier power supply. The
3081 preamplifier Power Supply Rejection (PSR) has been simulated and found above 17 dB for
3082 frequencies up to 1 MHz and above 30 dB for high frequencies larger than 100 MHz, meaning
3083 that the noise from power supplies is attenuated by at least 17 dB. As for the digital blocks,
3084 they are in deep Nwell or directly on the substrate. The connection between all the digital
3085 grounds and vss will be done at the flex level. Tests at system level are necessary to decide
3086 whether the analog ground and digital ground (gnda and gnnd respectively) should be
3087 connected at the module level or at the PEB level. The same is done for the power supplies:
3088 all the analog power lines (vdda_block) are connected together at the flex level to a common
3089 Vdda and all the digital power lines of the digital blocks (vddd_block) are connected to a
3090 common vddd.

3091 6.7 ASIC prototype measurements

3092 The performance on the first prototype version ALTIROC0 containing only the analog part
3093 of the single-channel readout (the preamplifier and the discriminator) can be found in [45].

3094 In this section, the results concerning the second prototype ALTIROC1 are presented. This
3095 second version, ALTIROC1, consists of a 5×5 pad matrix instead of a 2×2 , in which
3096 the digital components have been added to the single-channel readout. Two iterations
3097 of ALTIROC1 have been produced. The second one ALTIROC1v2 corrects issues found in
3098 the TDC and only results from this iteration are presented here. Among the 25 channels,
3099 only 15 channels corresponding to 3 columns have the readout as described in Sec. 6.2.1 with
3100 voltage preamplifier. The 2 other columns are equipped with trans-impedance preamplifiers
3101 and their performance are not described in this document.

3102 Sec. 6.7.1 describes the testbench measurements which were performed with and without
3103 a sensor bump-bonded to it**ADD MORE INFORMATION OR CITE THE MODULE AS-**
3104 **SEMBLY CHAPTER.** In the case where no sensor is bump-bonded, on channel 4 of each
3105 column, a capacitor can be connected through a programmable switch to the preamplifier
3106 input, mimicking the LGAD sensor capacitance and thus allowing to study the performance
3107 as a function of the detector capacitance C_d . The capacitance is tunable from 0 to 7 pF with a
3108 step of 1 pF. The testbench measurements are performed thanks to a C_{test} capacitor of 200 fF
3109 integrated in ALTIROC1, and selectable by slow control and a calibration pulser described
3110 in Sec. 6.3.4 which generates a Dirac input charge with a relative precision between channels
3111 of $\sim 1\%$. All the measurements have been performed with only one channel activated at
3112 the same time. In order to understand the performance of the ASIC, an analog probe is
3113 integrated inside the prototype ASIC that allows to output the preamplifier signal to an
3114 oscilloscope. When this probe is enabled, the preamplifier output is not only sent to the
3115 discriminator but also to an amplifier with a gain of approximately 12. In a similar way, a
3116 digital probe allows to see the output of the discriminator, before going into the TDC.

3117 Two testbeam campaigns have been carried out during the year 2019 at DESY, in which data
 3118 was collected with ALTIROC1v2, bump-bonded to a non-irradiated LGAD sensor. The main
 3119 results are presented in Sec. 6.7.2.

3120 Irradiation tests were also performed at CERN using X rays up to 340Mrad. The results are
 3121 presented in Sec. 6.7.3.

3122 6.7.1 Test bench performance

3123 The first step towards the evaluation of the full single-channel readout is the calibration
 3124 of the TDC counts since the knowledge of the value of the LSB (Least Significant Bit) is
 3125 fundamental in order to obtain the real values of the TOA and TOT. This is achieved by
 3126 sending a delayed square pulse (called external trigger) directly to the TDC inputs bypassing
 3127 the preamplifier and the discriminator, and thus measuring the TOA as a function of the
 3128 delay as displayed in Fig. 6.13(a). The measured TOA TDC quantization step is found to
 3129 be around 22 ps, close to the nominal value of 20 ps. As a consequence, the maximum TOA
 3130 that can be converted is slightly larger than the nominal window of 2.5 ns. The uniformity
 3131 of the LSB for the TOA is shown on Fig. 6.13(b) and is better than 5%. The external trigger
 3132 has a variable width and can also be used to measure the LSB for the TOT. The averaged
 3133 measured LSB is around 170 ps close to the nominal value of 160 ps and the dispersions
 3134 are better than 5% as can be seen on Fig. 6.13(b). **NEED TO ADD A COMMENT ON THE**
 3135 **MISSING TOT point.**

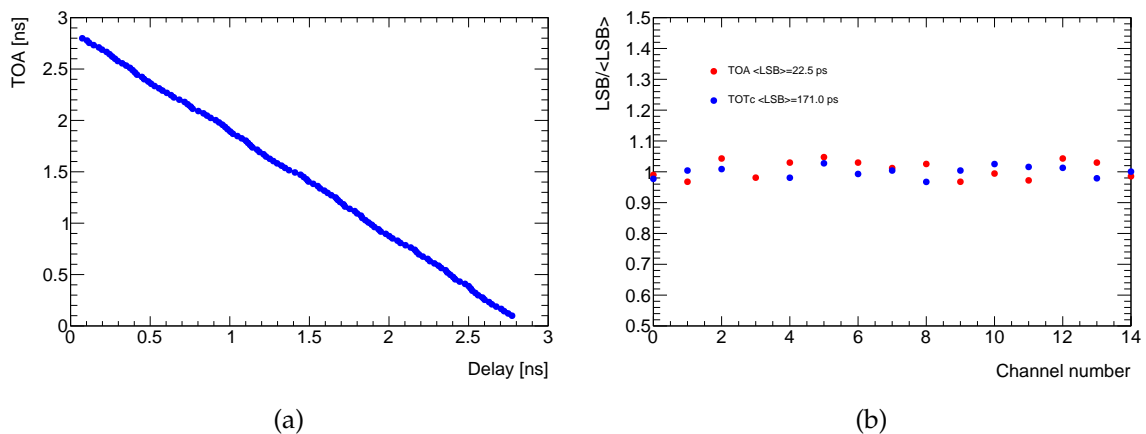


Figure 6.13: Average Time Of Arrival measurement with the TDC as a function of the programmable delay (a) and channel LSB divided by the averaged LSB as function of the channel number (b). All measurements are performed with an external trigger.

3136 The preamplifier jitter σ_{jitter} depends on the preamplifier rise time, which depends on the
 3137 drain current that goes into it. All the results below have been obtained with $I_d = 570 \mu\text{A}$

3138 **TO BE CHECKED** where the transistor enters in the strong inversion region and the gain
 3139 increases only with the square root of I_d and so the S/N doesn't increase significantly.
 3140 Fig. 6.14(a) shows the efficiency as a function of the input charge for an ASIC alone with
 3141 $C_d = 4$ pF in order to mimick the detector capacitance and with an ASIC bump-bonded to a
 3142 sensor. In both cases, full efficiency is achieved for charge greater than 3 fC.

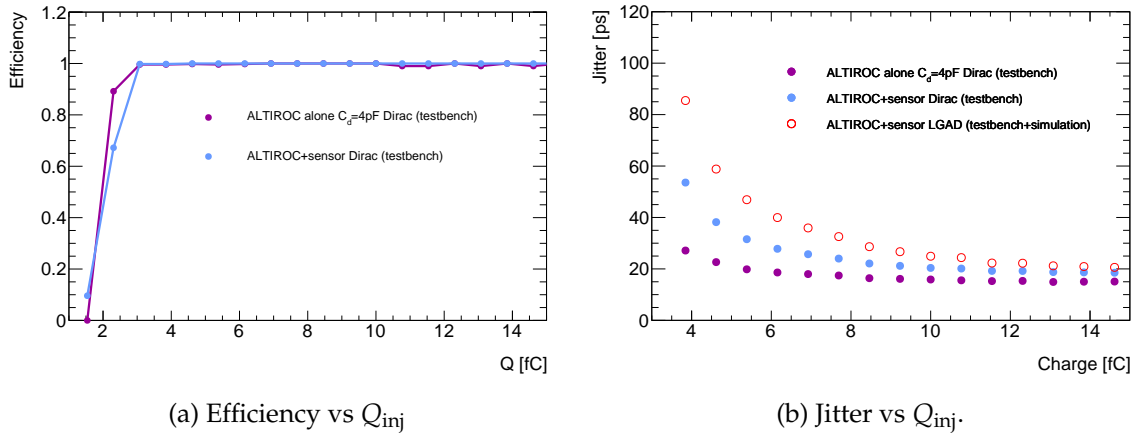


Figure 6.14: Efficiency (a) and jitter (b) measured as a function of the injected charge for an ASIC alone with $C_d = 4$ pF (purple) and with an ASIC bump-bonded to a sensor (blue) measured with the calibration setup. For (b), the open circle shows the jitter for a LGAD input signal estimated from the calibration data and the simulation.

3143 Fig. 6.14(b) shows the jitter variation as a function of the input charge for an ASIC alone
 3144 with $C_d = 4$ pF and with an ASIC bump-bonded to a sensor. For large charge a constant
 3145 jitter of about 15 ps is observed, which is attributed to the pulse command and clock jiter.
 3146 Even without subtracting this constant term, the jitter is smaller than 30 ps for $Q_{inj} > 6$ fC.
 3147 **NEED TO ADD AN EXPLANATION OF THE DIFF BETWEEN THE BLUE AND PURPLE**
 3148 **POINTS.** However, the performance obtained with the calibration signal can't be transposed
 3149 to a LGAD signal because the calibration signal is much faster. Based on the simulation, the
 3150 jitter obtained with the calibration needs to be multiplied by 1.65 to reproduce the results
 3151 obtained with a LGAD signal. Therefore, the jitter becomes smaller than 30 ps only for
 3152 $Q_{inj} > 8$ fC as shown on Fig. 6.14(b) and the jitter is 85 ps at 4 fC.

3153 **TO BE UPDATED:** The power consumption of the ASIC has been estimated through both
 3154 preliminary measurements and simulations for a 10% occupancy. Two operation modes can
 3155 be distinguished: physics runs and calibration runs. In the latter, up to a 10% occupancy
 3156 will be allowed. At the single channel level, the preamplifier and discriminator give a power
 3157 consumption of 1.57 mW, considering a drain current for the preamplifier of 1 mA. For each
 3158 time-to-digital converter, 0.55 mW has been estimated, while up to 2 mW have been allowed
 3159 for the digital part (hit processing unit, clock and luminosity unit). This yields a total of
 3160 4.67 mW per channel; considering that the current input to the preamplifier will be more
 3161 likely around 600 μ A, this gives a margin of 0.35 mW. In addition, an estimated allowance of

3162 250 mW for the common digital part seems reasonable, yielding a total power consumption
 3163 per ASIC of 1.2 W.

3164 **6.7.2 Test beam performance**

3165 An ALTIROC1v2 ASIC has been bump bonded to LGAD sensor arrays (HPK 3.2), with
 3166 1.3 mm × 1.3 mm pads and exposed in electron beam tests at DESY in the fall of 2019. The
 3167 LGADs have been operated with a bias voltage of 230 V, resulting in a MIP charge deposit
 3168 of about 18 fC. For an accurate timing reference, a fast Cherenkov-light emitting Quartz
 3169 bar of 6×6 mm² area transverse to the beam and 10 mm length along the beam, coupled
 3170 to a Silicon Photomultiplier (SiPM). The time resolution of this device was measured to be
 3171 37.6±0.7ps

3172 The Fig. 6.15(a) shows the TOA variation as a function of the TOT. The range of the TOT is
 3173 truncated since it was not possible to measure large values of TOT. This problem is attributed
 3174 to a coupling with the bias voltage distribution leading to sharp steps in the TOT versus
 3175 charge distributions². The Fig. 6.15(b) shows the time difference between LGAD+ALTIROC
 3176 and the reference time from the Quartz+SiPM system before and after time walk correction.
 3177 The distributions are Gaussian without any tails. After subtracting the contribution for the
 3178 Quartz+SiPM system, the time resolution decreases from 58.3±1.6 ps to 46.3±1.4 ps. The
 3179 time resolution is the quadratic sum of the intrinsic time resolution of the LGAD (about
 3180 25 ps) and the electronics jitter that is therefore deduced to be about 39 ps. **NEED TO ADD
 3181 AN EXPLANATION WHY IT IS WORTH THAN THE EXPECTATION!!!!**

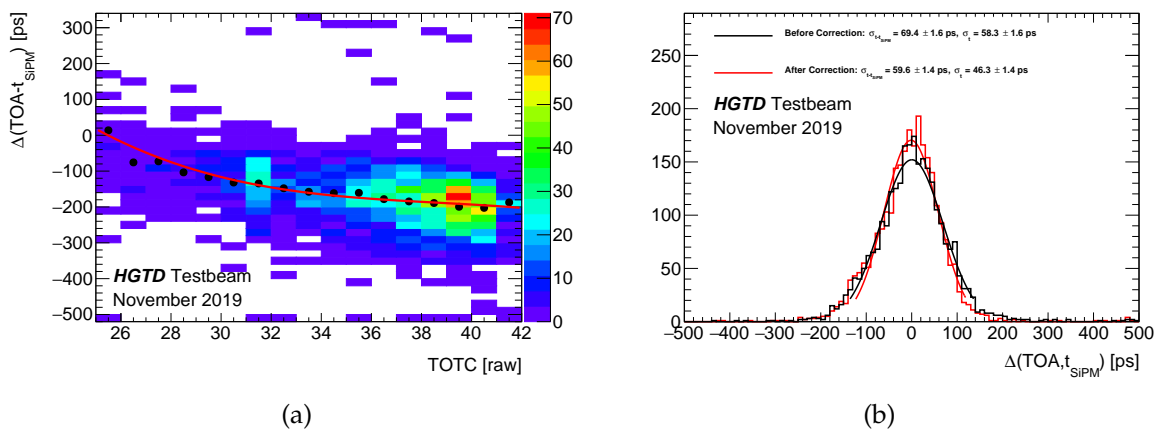


Figure 6.15: (a): TOA variation as a function of the TOT. (b): Time difference between LGAD+ALTIROC and the Quartz+SiPM system before and after time walk correction.

² When the ASIC is not bump bonded to a sensor, this effect is not observed.

3182 6.7.3 Irradiation tests

3183 Altiroc1 was irradiated at CERN with Xrays up to 340 Mrad. Tests were focused on the
 3184 preamplifier and the discriminator. A small decrease of the preamplifier amplitude was
 3185 observed (??) as well as a small increase of the discriminator jitter ((??)). DC voltages such as
 3186 the bandgap output, the 10-bit DAC used to set a common discriminator threshold as well
 3187 as the 7-bit DAC used for individual tuning of the threshold were also followed during the
 3188 irradiation and showed variations smaller than 20 mV (around 800 mV) between 20 Mrad
 3189 and 340 Mrad.

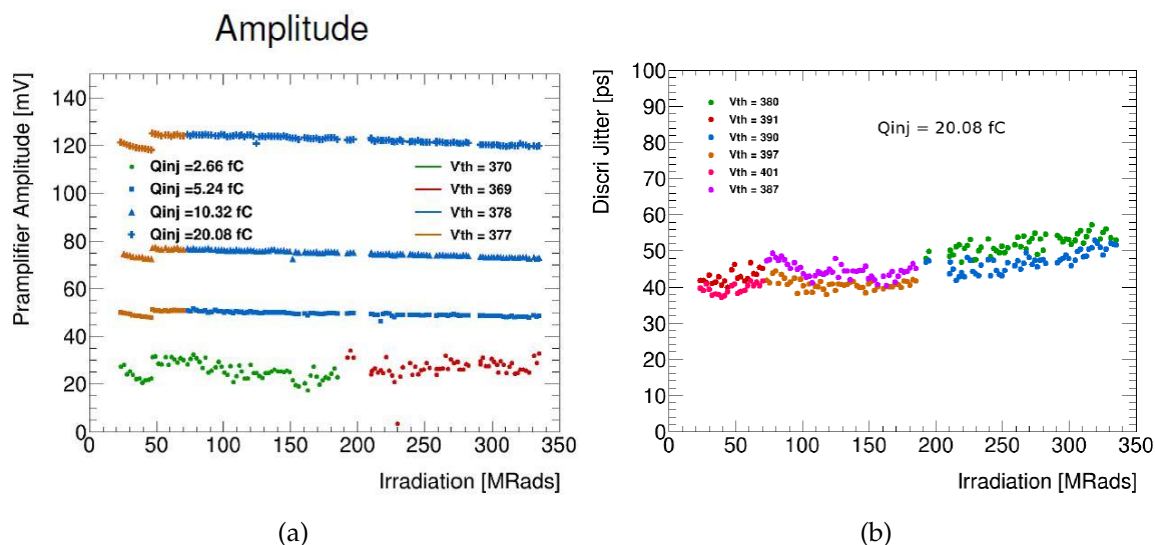


Figure 6.16: Preamplifier amplitude (Left) and discriminator jitter during irradiation tests.

3190 6.8 Monitoring

3191 6.8.1 Temperature monitoring

3192 An additional requirement of the ASIC is to be capable of monitoring two closely related
 3193 aspects of the LGAD: its operating temperature and its leakage current. While the electronics
 3194 themselves are not very sensitive to temperature changes, it is of utmost importance to
 3195 monitor the sensors in order to detect loss of cooling and thermal run-away, as explained
 3196 in Sec. 5.6. This information could also be used to estimate the particle fluence, since the
 3197 current increases linearly with it.

3198 A good estimate of the temperature dependence of the leakage current of a no-gain sensor is
 3199 a factor 2 increase for every 7°C. The temperature dependence of the gain is much lower,

3200 with an increase in gain of a factor 2 for a temperature decrease of 30–40 °C. Knowledge of
 3201 the sensor temperature with an accuracy of 0.5 °C would make it possible to determine the
 3202 leakage current to approximately 15% (while giving no relevant information on the gain).
 3203 The modules will be operating at room temperature (20-30 °C) during the R&D phase, and
 3204 during detector operation at about –30 °C as required by the sensor. Considering possible
 3205 temperature shifts within the chip plus some margins, two monitoring ranges have been
 3206 defined, [30- 40 °C] [-40 –10 °C], given a total temperature monitoring range of 80 °C. The
 3207 target resolution to determine temperature variations has been set to 0.2 °C (9-bit resolution);
 3208 the absolute value of the temperature is not relevant.

3209 The temperature sensor inside ALTIROC is based on a resistor which is sensitive to variation
 3210 of temperature, a constant current flowing through this device producing a voltage propor-
 3211 tional to the temperature. This current will be delivered by the current source present at
 3212 the ADC input of the lpGBT circuit. Four different types of resistor proposed by the TSMC
 3213 technology) have been evaluated for their ability to perform temperature measurement on
 3214 an irradiated environment. As they all present similar behaviour, only performance of the
 3215 N-diffusion resistor version is reported in Table 6.3. In important point is that the resolution
 3216 can be doubled using a current value of 200 μA , achieving then a resolution of 0.4 °C per
 ADC count.

Technology of the resistor	N+ diffusion resistor with salicide (rn1plus)
Value of the resistor	5 k Ω
Value of current flowing during test	100 μA
Sensitivity	+1.3 mV/°C
Variation of sensitivity with radiation	+15% at TID of 350 Mrad
Shift of temperature with radiation	–0.3 °C at TID of 200 Mrad –0.6 °C at TID of 350 Mrad
Resolution after conversion with ADC of LpGBT	0.8 °C per ADC count

Table 6.3: Evaluation of a N-diffusion resistor as temperature sensor under irradiation (TID).

3217

3218 6.8.2 Supply voltages monitoring

3219 The analog and digital supply voltages have also to be monitored. The first need is for
 3220 measuring and compensate the voltage drops in the power lines caused by the parasitic
 3221 resistances in the power wires of the flex cables (R_{FLEX} in Fig. 6.17). The VDDA and VDDD
 3222 voltages are sensed through dedicated wires on the flex and digitized by the ADC of the
 3223 lpGBT circuit on the peripheral board.

3224 The probing of the power voltages at the module level is also useful to detect latch-up events
 3225 on an ASIC. With the resolution of 1 mV of the lpGBT ADC and a parasitic resistance of
 3226 100 m Ω on the flex cable, minimal variation of 20 mA (considering an attenuation of 1/2 on

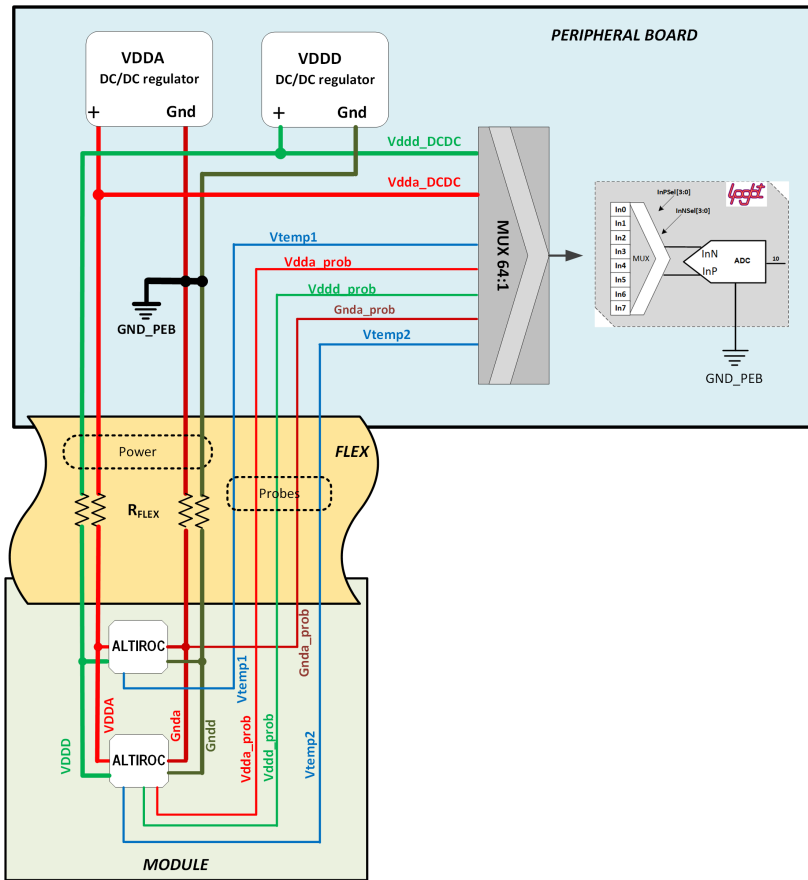


Figure 6.17: Complete schematic view of the voltage monitoring of a module using the ADC of the lpGBT circuit.

3227 the probing to respect the input dynamic range of the LpGBT ADC of 1V) can be detected,
 3228 much smaller than the expected current rise in a latch-up event.

3229 6.8.3 Complete monitoring system

3230 A complete schematic view of the proposed monitoring of ALTIROC using the ADC of the
 3231 lpGBT circuit is given in Fig. 6.17. Three signals (V_{dda_prob} , V_{ddd_prob} and $Gnda_prob$) for
 3232 the monitoring of the power supply voltages inside the two chips and two signals (V_{temp1} ,
 3233 V_{temp2}) for the measurement of the temperature inside the two ASICs are connected to the
 3234 ADC of the lpGBT circuit. The signal to be converted by the ADC is selected via multiplexers
 3235 controlled through the I²C interface of the lpGBT.

3236 A view of the complete interfacing of a peripheral board with the modules is represented
 3237 in Fig. 6.18. The analogue signals of monitoring coming from the modules are digitized by

3238 the converter implemented inside each lpGBT circuit of the peripheral board. The number
 3239 of channels of this ADC being limited to eight, a multiplexing is required at the input of
 3240 each channel. Multiplexers (MUX 64:1) are thus implemented to interface the signals coming
 3241 from the modules to the ADC on the peripheral board. With such multiplexer circuit, up to
 3242 8×64 signals can be interfaced to each lpGBT-ADC. With one multiplexer reserved for the
 3243 signals coming from the DC/DC regulators, 7 mux are available to interface the monitoring
 3244 signals coming from up to 84 the modules, which is larger than the maximum number of
 3245 modules expected per peripheral board. A full custom 64-to-1 multiplexing circuit is under
 3246 development with a radiation tolerance suitable with its implementation on the peripheral
 3247 board.

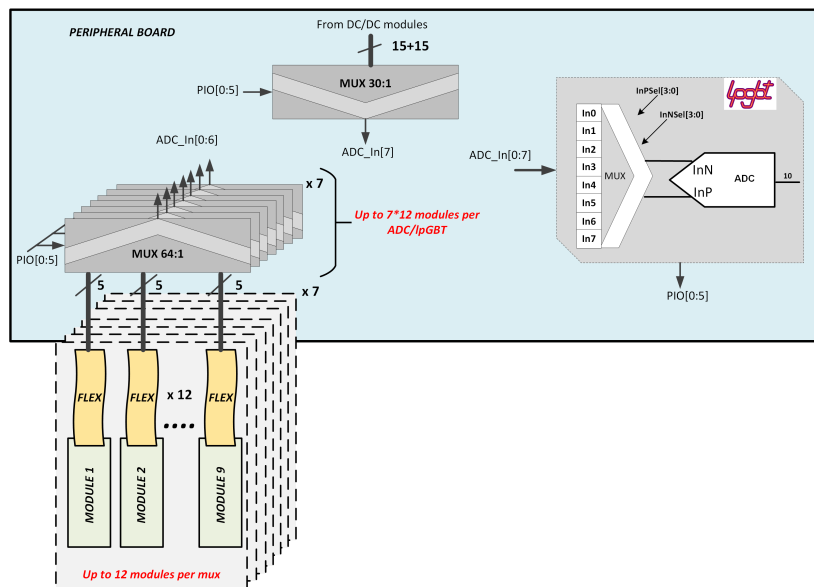


Figure 6.18: Interfacing of modules with a peripheral board for the monitoring.

3248 6.9 Roadmap towards production

3249 TO BE UPDATED

3250 As explained before, two iterations of the ALTIROC have already been submitted and
 3251 received. The strategy so far has been to validate with a first prototype the very front-end
 3252 readout electronics, and then secondly the full single-channel readout with ALTIROC1.
 3253 Concerning the latter, a new iteration has been submitted at the end of February 2019 in
 3254 which the implemented changes are expected to fix the issues found in the TDC. The main
 3255 modifications, besides other minor ones, have been:

- 3256 • Replacing all the 1.2 V MOS transistors used to filter the bias voltage of the ASIC by
3257 2.5 V thick oxide transistors which have a negligible gate leakage current
- 3258 • Increasing the size of the ASIC to make the debug pad accessible after bump bonding
3259 to the sensor
- 3260 • Modification of the TDC layout as described in Sec. 6.7.1.

3261 This prototype of ALTIROC has been designed so as to enable irradiation tests, which will
3262 be performed during 2019.

3263 The next prototype, ALTIROC2, will integrate the 225 channels and all the digital blocks
3264 currently existing in RTL code, having therefore all the functionalities of the final ASIC. The
3265 approach would be to use a *digital-on-top* design at the matrix level to avoid timing violations
3266 on all the digital signals, which are sent or received by the readout channel to or from the
3267 End Of Column logic. Digital-on-top design will also ensure an accurate timing distribution
3268 of the clocks and of the luminosity windows. On the other hand, because the power supplies
3269 distribution is critical for the analog performance, an *analog-on-top* design will be used for
3270 the peripheral electronics and for the floor plan.

3271 Triple voting logic against SEE will be implemented for all control signals and registers but
3272 not for read-out data. The submission of the ALTIROC2 is expected by the end 2019/early
3273 2020, followed by 6 months dedicated to its characterization and to perform tests concerning
3274 radiation hardness. A second iteration is scheduled early 2021 to be considered as the
3275 pre-production.

3276 Regarding the MUX 64:1, it has been designed and will be submitted in an MPW run in
3277 October 2019.

3278 **7 Module Assembly and Loading**

3279 **7.1 Introduction**

3280 The basic component of the HGTD is the module. A detector module consists of a sensor
3281 bump-bonded to two readout chips which are in turn connected to a flexible printed circuit
3282 (FPC, flex cable) for communication, power distribution and data output. The flex cable also
3283 provides high voltage for the silicon sensor. The HGTD is made up of 7984 modules mounted
3284 on intermediate plates. This chapter describes the module and its assembly process, together
3285 with the procedure of mounting them onto the intermediate plates. Quality assurance
3286 and control plans are presented. Results of the fabrication of various prototypes are also
3287 discussed.

3288 The module requirements depend on the final layout HGTD option that is chosen. Three
3289 possible scenarios are envisaged (see Sec. 15.4). One of these scenarios should be chosen
3290 once final estimates are available for radiation levels in the HGTD volume with realistic
3291 ITk services/supports and once the performance of real-size prototypes of the HGTD
3292 sensor+ASIC system are studied under irradiation.

3293 **7.2 The bare module**

3294 The bare module consists of an LGAD sensor interconnected through solder bumps to two
3295 ALTIROC front-end chips. The LGAD sensors and the ALTIROC chip have been described
3296 in Chap. 5 and Chap. 6. In this section the hybridization process, called bump-bonding, is
3297 discussed.

3298 Modules based on the 5×5 channel ALTIROC1 chip have already been fabricated and tested.
3299 A baseline hybridization process has been defined and the specifications agreed upon with
3300 two vendors. These vendors are currently being qualified (on ALTIROC1 devices). Full size
3301 prototypes will be produced as soon as the ALTIROC2 ASIC is available.

3302 7.2.1 Bare module assembly

3303 The LGAD sensor has a total size of 20.0 mm \times 39.5 mm, with an array matrix of 15 \times 30
3304 1.3 mm \times 1.3 mm pads and a dead region 0.25 mm wide around the active area. The readout
3305 ASIC has a total size of 21.7 mm \times 19.9 mm and a matrix of 15 \times 15 channels. The LGAD
3306 sensor has half the bump pads of each pad shifted from the central position by 250 μ m, while
3307 the other half of the pads are shifted by the same distance in the opposite direction. This
3308 allows to have a distance of 100 μ m between ASICs, with no gap in the sensor coverage or
3309 disruption from different pixel sizes (see Fig. 7.1).

3310 The LGAD sensors will be produced in 150 mm wafers of various thicknesses (depending
3311 on wafer and sensor providers), which will be thinned to the total sensor target thickness.
3312 Currently the baseline for the active thickness is 50 μ m and 300 μ m for the total thickness.
3313 The sensors will be probed at wafer level at the fabrication sites and this information will
3314 be made available to ATLAS. The under-bump metal will be deposited on the sensors at
3315 wafer level, a necessary step before bump-bonding with solder bumps. After under-bump
3316 metalization (UBM), the wafers will be diced and the selected sensors will be destined for
3317 hybridization.

3318 The ALTIROC ASIC will be produced in 200 mm wafers. The wafers will be thinned down
3319 to 300 μ m (current baseline). The front-end chips will then be probed to identify the good
3320 dies. This will be followed by UBM and solder bump deposition. The relatively large pad
3321 size of the HGTD sensors enables a less demanding bump-bonding technology process
3322 compared to the ITk Pixel detector. The low-cost electroless deposition of Ni/Au can be
3323 used to treat large pads (90 μ m diameter) of both sensor and ASIC wafers. Solder bumps
3324 (SnAg) with a baseline diameter of 80 μ m will then be deposited on the ALTIROC pads. A
3325 number of processes are available for the deposition of the bump balls, from solder laser
3326 jetting to electroplating. The most reliable, cost-effective technology will be selected.

3327 After UBM and bumping, the sensor and ASIC wafers have to be diced into single tiles.
3328 The next step of the hybridization process is flip-chipping. During flip-chipping, the sensor
3329 and ASIC tiles are aligned, heated and compressed so that each solder bump melts and
3330 connects the sensor and readout channels of the two substrates. It is foreseen that the
3331 bare assemblies will then be processed in a fluxless formic acid reflow oven in order to
3332 improve the connectivity of the solder bumps. The final step consists in the inspection of the
3333 bare assemblies with a high resolution (sub-micron) x-ray machine to discard devices with
3334 disconnected pad bumps. A small fraction of the mechanical bumps, that are located in the
3335 the periphery of each ASIC to provide rigidity to the assembly, can be faulty, as their role in
3336 not critical for the overall performance. Note tha electrical tests of the HGTD modules will
3337 be carried out after the bare assemblies are mounted (including noise and charge collection
3338 measurements that can reveal disconnected bumps not apparent with x-rays).

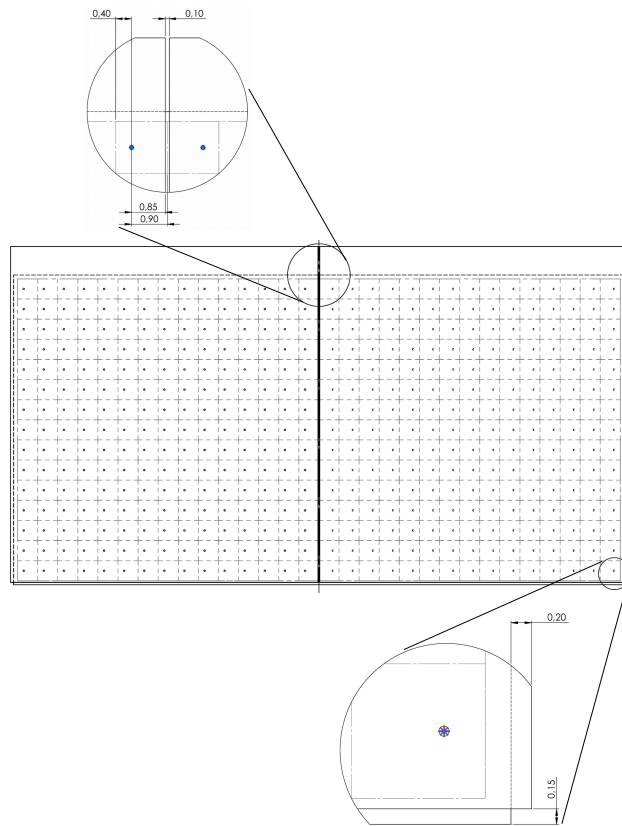


Figure 7.1: Sketch of the bare module (sensor and ASIC). Distances are in millimeters. The bump pads on the sensor are shifted by $250\ \mu\text{m}$ on each side of the sensor, to allow a $100\ \mu\text{m}$ separation between the ASICs (see text).

3339 7.2.2 First bare module prototypes: process and results

3340 The first ALTIROC1 devices have already been assembled. As described in Chapter 6,
 3341 the ALTIROC1 ASIC is a 5×5 channel prototype of the HGTD chip. The pad size is
 3342 $1.3\ \text{mm} \times 1.3\ \text{mm}$. The corresponding 5×5 pad sensors used in these first prototypes were
 3343 LGADs fabricated at CNM, in the context of an AIDA production (Run 11748), and at
 3344 Hamamatsu (Type 3-1, EXX28995). Both vendors deposited the UBM on the sensors (at
 3345 wafer level). In the case of CNM, a Ni/Au electroless process was used for UBM.

3346 The Ni/Au under bump metalization was also deposited on single ALTIROC1 tiles by CNM
 3347 through a chemical electroless process. SnAg solder bumps of $80\ \mu\text{m}$ diameter were then
 3348 placed on the chips using a laser jetting machine at IFAE. The bumps were prepared for
 3349 flip-chip with a formic acid reflow cycle. The bump strength was verified to be larger than
 3350 $60\ \text{gf}$ per bump through shear tests.

3351 The hybridization was performed by IFAE following the previous experience with the
3352 ALTIROC0 devices [45]. The same bonding cycle previously developed for the ALTIROC0
3353 devices was used for the hybridization of the first ALTIROC1 bare modules. The devices
3354 were reflowed with no weight and inspected with X-rays. Good alignment was observed as
3355 well as good connectivity in all the bumps (except those removed in one of the samples as
3356 part of the bump shear tests). CNM and HPK bare assemblies, along with the x-ray image of
3357 the bump connecting one of the readout channels, are shown in Fig. 7.2. The topology of the
3358 bumps was found to be mostly cylindrical, with a diameter of about $90\ \mu\text{m}$ and a height of
3359 approximately $50\ \mu\text{m}$. The hybridization specifications detailed below (see Sec. 7.2.3) follow
3360 the same process developed by IFAE, which is standard in the commercial sector and for
3361 which two companies have already been identified.

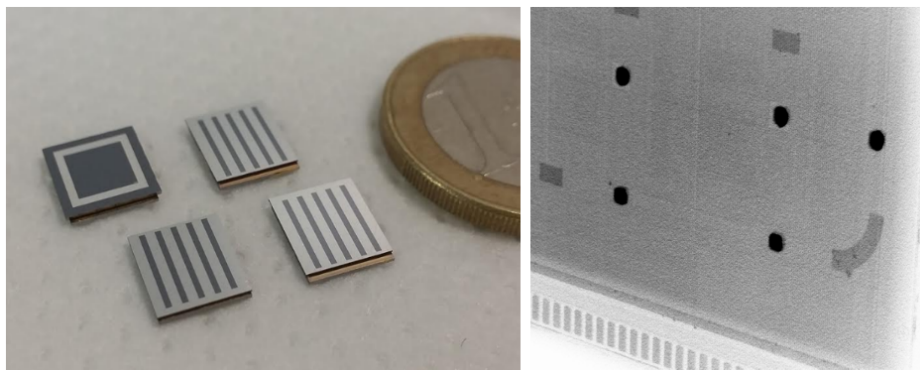


Figure 7.2: The first ALTIROC1 bare modules, with CNM and HPK sensors, and an x-ray image with a detail of a corner of one device are shown. In the x-ray image, the guard-ring solder bumps are in the periphery, while the bumps of two readout channels are visible in the center left of the image. The wire-bond pads of the ASIC are also apparent towards the lower part of the figure.

3362 The modules will experience thermal cycles during their lifetime, as the HGTD inner volume
3363 will be cooled with an input coolant temperature of $-35\ ^\circ\text{C}$. In order to verify the robustness
3364 of the bare assemblies, they were subjected to a long burn-in test (some glued to a to a PCB
3365 using Araldite 2011, see Sec. 7.4.3). During a total of two weeks the modules were thermally
3366 cycled between $-40\ ^\circ\text{C}$ and $130\ ^\circ\text{C}$. The solder connections were then verified with x-ray
3367 imaging and shear tests were carried out on the modules. The devices were able to sustain
3368 the maximum applied shear force of 1000 gf, between the ASIC and sensor and also between
3369 PCB and ASIC. One device was verified to sustain a perpendicular (with respect to the plane
3370 of the sensor) pull test of 100 gf before and after the two week thermal cycling. Fig. 7.3 shows
3371 the shear and pulling tests being carried out on an ALTIROC1 hybrid.

3372 The hybridization was also performed by the National Center for Advanced Packaging
3373 (NCAP China). NCAP is one of the leading companies in the integrated circuit packaging
3374 and testing industry in China. NCAP has more than $3200\ \text{m}^2$ cleanroom space and can
3375 provide bump-bonding service for 8 inch and 12-inch wafer. Its the production capacity for

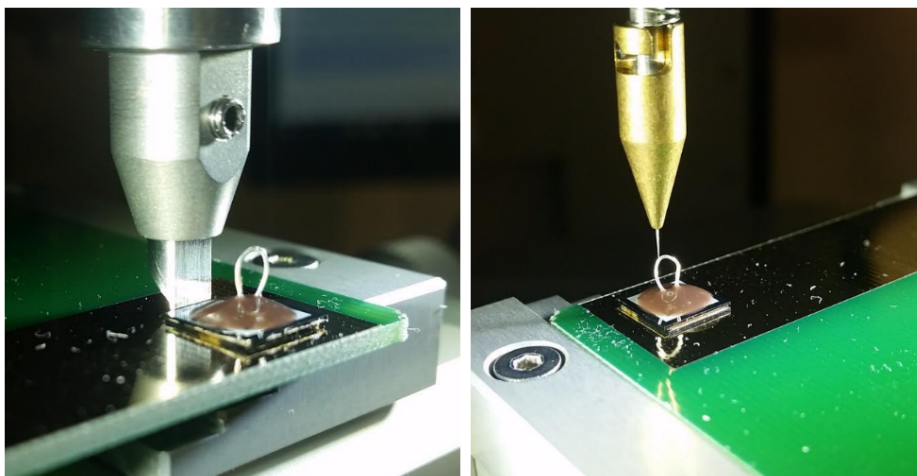


Figure 7.3: Shear and pull tests being carried out on an ALTIROC1 device. After thermal cycling during two weeks the device was able to sustain a maximum shear (pull) force of 1000 gf (100 gf).

3376 module hybridization can fully satisfy the requirement by the HGTD project. 15 ALTIROC1
 3377 bare module prototypes have been hybridized in NCAP. Part of them are shown in Fig. 7.4.
 3378 The 5×5 pad sensors used in these prototypes were LGADs fabricated at Hamamatsu (Type
 3379 3-1 and Type 3-2), and at NDL (Type 6 and Type 12). The solder connections were then
 3380 verified with x-ray imaging as shown in Fig. 7.5. These modules were sustain the maximum
 3381 applied shear force of 1000 gf during shear test. The bare modules hybridized in NCAP
 3382 have been tested in DESY test beam. Their performance in test beam will be described in
 3383 Sec.XXX.

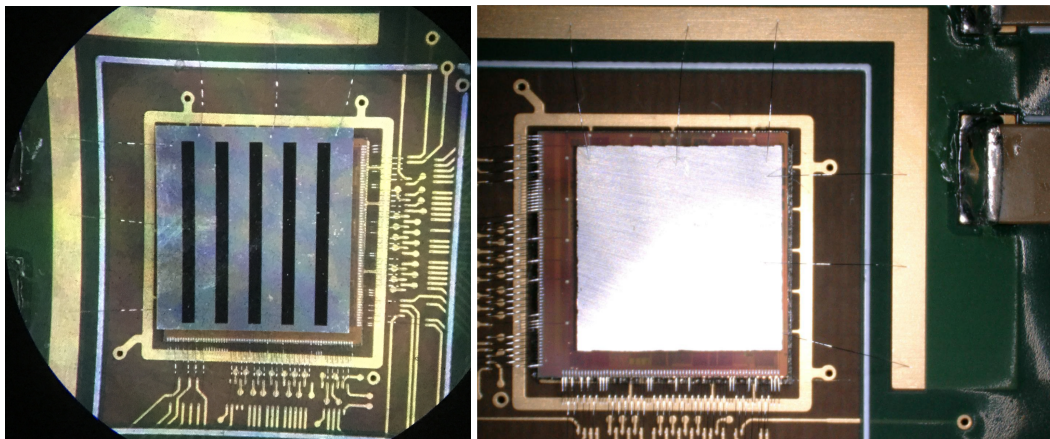


Figure 7.4: The bare modules hybridized in NCAP China. Left plot: The 5×5 pad sensors used in these prototypes were LGADs fabricated at Hamamatsu (Type 3-2). Right plot: The 5×5 pad sensors used in these prototypes were LGADs fabricated at NDL (Type 6).

3384 An alternative process explored during the initial R&D phase (but not intended for produc-

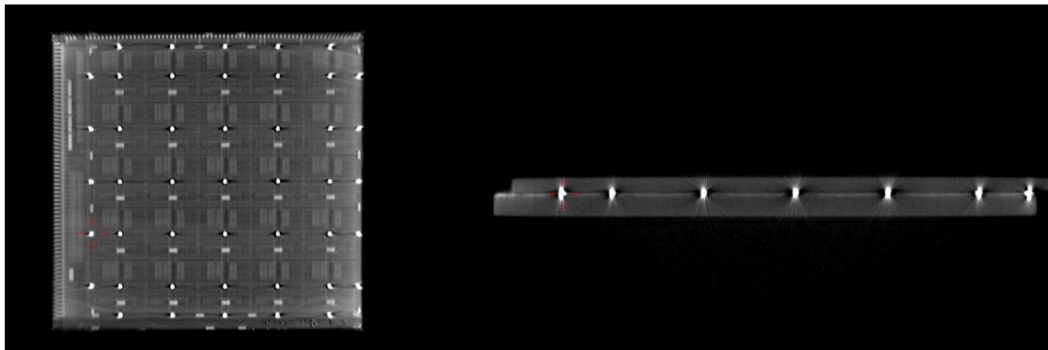


Figure 7.5: The x-ray image of the bare modules hybridized in NCAP China. In the x-ray image, the solder bumps are visible in top view and side view.

3385 tion) has also been developed to assemble ALTIROC0 devices. With Au bumps under bump
3386 metalization is not need since the ball bumps can be deposited directly on the aluminium of
3387 the front-end pads. An alignment and thermo-compression cycle is used to interconnect the
3388 channels of the sensor and ASIC. Studies determined that the bump topology resembled at
3389 conical frustum with a base of about 140 μm and a height of 15 μm .

3390 7.2.3 Hybridization specifications

3391 The baseline bump-bonding technology for HGTD relies on solder bumps. As described
3392 above, both solder bump and gold stud bump prototypes have been produced at different
3393 HGTD institutes. Gold stud bumping is a process that enables hybridization without the
3394 need of under-bump metalization, since the gold studs can be deposited directly on the
3395 aluminium. This makes it very useful for research and development activities. However,
3396 the process is not scalable for large productions. Thus, solder bumps are the baseline
3397 hybridization solution.

3398 The sensor fabrication sites will deliver silicon wafers. These wafers may, or may not, include
3399 under bump metalization, depending on the vendor fabrication capabilities and the overall
3400 HGTD hybridization strategy. In any case, it is expected that a fraction of the sensor wafers
3401 may have to be prepared for bump-bonding by a different vendor than the one producing the
3402 sensors. As explained above, given the large pitch and pad size of the sensors, the selected
3403 process for UBM is electroless deposition of Ni/ Au. Tab. 7.1 lists the relevant parameters
3404 related to the sensor wafer UBM.

3405 The HGTD ASICs will be produced in TSMC CMOS 130 nm technology. In order to perform
3406 the hybridization process, first UBM and then solder bumps have to be deposited on the
3407 ASIC wafers. As mentioned above electroless Ni/ Au deposition is selected as the baseline
3408 process for UBM, while solder bumps composed of SnAg (SAC305) would be deposited

Wafer material	Silicon
Wafer thickness	300 μm
Sensor size (R×C)	21 000 × 41 000 μm^2 (15×30)
Distance between pads	1.3 mm
Pad size (passivation opening)	90 μm
Pad metalization	Aluminum
Scribe line passivated	Yes
Baseline UBM process	Electroless Ni/Au

Table 7.1: Specifications of the HGTD sensor wafer under-bump metalization.

3409 through a laser solder jetting system. However, other procedures can be considered. Tab. 7.2
 3410 summarizes the requirements for the UBM and bumping of the HGTD ASIC wafers.

Wafer material	Silicon
Wafer thickness	300 μm
ESD sensitive	Yes
Passivation	8750A SiO ₂
ASIC size (rows×columns)	23 000 × 20 000 μm^2 (15×15)
Distance between pads	1.3 mm
Pad size (passivation opening)	90 μm
Pad metalization	Al
Baseline UBM process	Electroless Ni/Au
Solder bumps	SnAg (SAC305)
Baseline bumping process	Laser solder jetting
Bump shear strength	40 gf/bump

Table 7.2: Baseline specifications of the HGTD ASIC wafer UBM and solder bump deposition. Other UBM and bumping process will be studied.

3411 After UBM and bump deposition the sensor and ALTIROC wafers will be diced. The width
 3412 of the scribe line shall be 20 μm and the dicing precision $\pm 10 \mu\text{m}$. Break offs at the dicing
 3413 edge shall be limited to less than 75 μm .

3414 The flip-chip process is the final step in the hybridization procedure. The flip-chipping
 3415 will be done on single sensor tiles. Two ASICs have to be flip-chipped to a single sensor.
 3416 The cycle has to be consistent with the SnAg (Sn₃Ag_{0.5}Cu) solder bumps and result in a
 3417 high hybridization yield. Tab. 7.3 summarizes the flip-chip requirements for the HGTD
 3418 modules.

Alignment between ASIC and sensor	5 μm
Minimum distance between ASIC and sensor after flip-chip	20 μm
Maximum distance between ASIC and sensor after flip-chip	50 μm
Maximum failure rate per ASIC	0.044%
Shear strength after flip-chip	40 gf/bump

Table 7.3: Specifications of the flip-chip process for the HGTD modules.

3419 7.2.4 Quality assurance / quality control

3420 Each of the specifications listed above will be tested to show that they are met within the
3421 required acceptance criteria. Bare modules will be optically inspected and weighed. The
3422 distance between the substrates (bump height) will also be measured. Inspection with x-rays
3423 for disconnected channels before module assembly (dressing with the flex hybrid) will follow.
3424 If the yield of the bump-bonding process is found to be high after the initial production and
3425 the modules are found to be highly uniform, these time consuming steps (X-ray inspection
3426 and substrate separation) can be performed only on a small fraction of devices. Note that
3427 the channel connectivity will be anyhow tested during the module electrical tests. A small
3428 number of ASICs will be sacrificed to test the bump quality with shear tests before flip-
3429 chipping. Furthermore, a small number of devices will be tested destructively to verify the
3430 robustness of the hybridization process. Burn-in tests will be carried out on some devices to
3431 test specifically for the degradation of the bump-bonding.

3432 7.2.5 Production hybridization strategy

3433 The total surface covered by the HGTD (about 7 m²) requires a well planned approach to
3434 successfully carry out the hybridization of all the modules. The three step hybridization
3435 strategy consists of: process R&D and specification, search and qualification of bump-
3436 bonding vendors, and finally, module hybridization pre-production.

3437 As shown above, the baseline bump-bonding process has been developed and successfully
3438 tested. Initial specifications have been established. Full size tests will be carried out as
3439 soon as the final sized sensor and ASIC become available. However, an effort is being
3440 made to advance this critical step in module assembly to avoid possible bottlenecks later
3441 in the overall activities. The specifications have already been provided to two companies
3442 (one in Germany and one in China) and discussions of a possible early qualification of the
3443 bump-bonding process with the currently available devices (ALTIROC1) are on-going. Both
3444 companies have expressed their willingness to carry out the hybridization service for HGTD
3445 and can do the full process in-house (metalization, bump-deposition, dicing and flip-chip).
3446 The possibility of using dummy ASICs and sensors for the vendor qualification will also be

Component	Thickness [mm]
ASIC	0.30
Bump bonding	0.05
Sensor	0.30
Glue	0.10
Module flex	0.50
Connector	2.00
Total	3.25

Table 7.4: Contribution of each module component to its thickness.

3447 investigated. The target is to eventually carry out the final hybridization qualification on
 3448 two to four companies, though the impact on cost and schedule will have to be evaluated.

3449 7.3 Module design and assembly

3450 7.3.1 General description

3451 Baseline module design

3452 The bare module described above is glued with accurate positioning to the module flex, to
 3453 which a flex tail will be connected during detector assembly (see Sec. 13.1). ASIC signals and
 3454 low voltage, as well as bias voltage for the sensor (HV) will be connected by wire bonding.
 3455 Fig. 7.6 shows three modules with the different components stacked in the z direction of
 3456 the HGTD. The total thickness of a module, including ASIC, sensor and module flex with
 3457 all components and connectors, is 3.25 mm, with the contributions of each element listed in
 3458 Tab. 7.4

3459 Alternative module design

3460 Additionally to the development and test of the baseline design, alternative options are
 3461 being investigated, in particular with the aim of replacing wire bonding with mechanically
 3462 more robust solutions. In particular the usage of conductive glue for the connection of the
 3463 HV line to the sensor and of bump bonds to connect all signal and power lines between
 3464 the module flex and the ASICs is being studied and prototypes are in preparation . Fig. 7.7
 3465 shows the schematic drawing of the alternative module design.

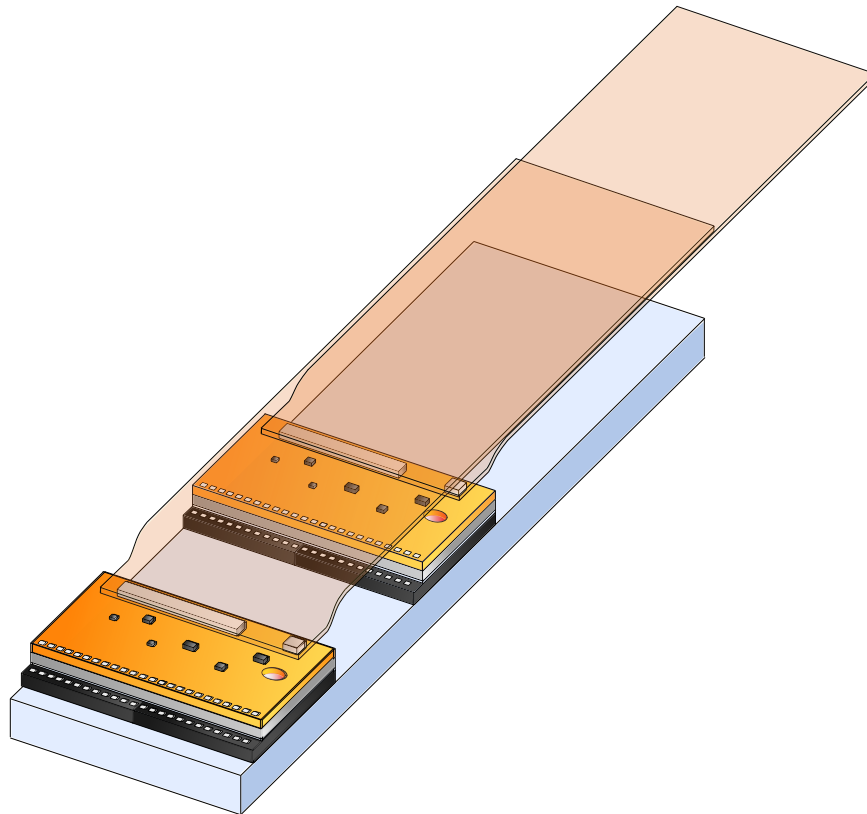


Figure 7.6: Schematic drawing of two adjacent modules on the top side and one on the bottom side of the cooling plates.

3466 7.3.2 Voltage distribution and signal readout: flex cables

3467 The module flex and the flex tail, based on the flexible electronics technology, connect
3468 the signals from the module to the peripheral on-detector electronics. The geometrical
3469 constraints on the flex tails are determined by the available space between two layers (see
3470 Tab. 11.2), the distance between the modules and the peripheral electronics and the maximum
3471 number of modules per readout row. Considering the harshest constraints, the flex tails
3472 must have a maximum length of about 600 mm, width of at most 36 mm, and thickness of
3473 less than 250 μm . The total length of flex cables in the HGTD is 2620 m.

3474 In terms of electrical requirements, one HV line has to be included in the design in order
3475 to bias the LGAD sensors (800 V maximum). The HV line must have a sufficient insulation
3476 resistance (IR) to not affect the other lines ($\text{IR} > 10 \text{ G}\Omega$). The types of signals to and from
3477 the ASICs in each flex cable include the transmission of high speed signals (1.28 Gbit s^{-1}) as
3478 well as clock and power signals. The total numbers of signals for each module are listed in

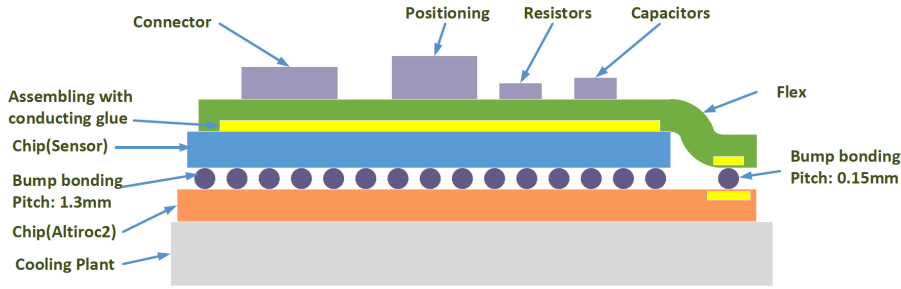


Figure 7.7: Schematic drawing of the alternative module design

Signal name	Signal type	No. of wires	Comments
HV	800 V max.	1	Clearance
POWER	$1 \times V_{dda}$, $1 \times V_{ddd}$, 1.2 V	2	2 planes, $R < 2.7 \text{ m}\Omega \text{ cm}^{-1}$
GROUND	Analog, Digital	1(2) plane(s)	Dedicated layer $R < 0.7 \text{ m}\Omega \text{ cm}^{-1}$
Slow control	Data, Ck (opt. + rst, error)	2 to 4	I ² C link
Input clocks	320 MHz, Fast command e-link (opt. 40 MHz (L1))	4 or 8	CLPS
Data out lines	Readout data (TOT, TOA, Lumi)	4 pairs	4 e-links differential CLPS
ASIC reset	ASIC_rst	1	Digital
Monitoring	Temperature, V_dda, V_ddd	4	DC voltage
Debugging	ASIC_debug	2	Analog

Table 7.5: Type and number of signal lines for two ASICs and a sensor included in the flex cable design

3479 Tab. 7.5. The impedance is required to be in the range of 85Ω – 115Ω for the differential pairs
 3480 and of 50Ω – 64Ω for the single lines, while the same radiation tolerance is required as for
 3481 sensors and ASICs, i.e. up to at least 2.0 MGy, as well as operation at a temperature of about
 3482 $-30 \text{ }^\circ\text{C}$ (see ??).

3483 As baseline design a module flex with a width of 39.5 mm and a length of 18.5 mm along the
 3484 readout row is planned to be produced as a 4-layer stack-up with a thickness of $500 \mu\text{m}$. The
 3485 flex tail is a 2-layer cable to be produced with different lengths, $220 \mu\text{m}$ thickness and a width
 3486 of 36 mm. A preliminary layout of the module flex and flex tail is shown in **Add Figure with**
 3487 **flex tail layout. Is it realistic to have one also for the module flex?** Two separate connectors,
 3488 one for HV and the other for all the other lines, will be used to connect the module flex to the
 3489 flex tail. A prototype of the flex tail has been ordered, while the design of the module flex is
 3490 being finalised. **This last sentence should be adapted to the actual status shortly before the**
 3491 **deadline**

3492 **Prototype characterisation**

3493 As part of the initial study phase, a prototype combining module flex and flex tail into one
 3494 L-shaped 4-layer design has been produced with the aim to understand the technology
 3495 requirements (materials, manufacturing capability, electrical and mechanical robustness)
 3496 and address any potential problems by representing a significant subset of the signals
 3497 (signal integrity, power distribution, HV insulation, interference and crosstalk). The direct
 3498 interaction with the CERN PCB Service allowed to optimise the manufacturing process
 3499 leading to the production of 4 prototypes of 750 mm length as depicted in Fig. 7.8. Apical,
 3500 Krempel and Kapton were the commercial materials chosen for this prototype. The length
 3501 was chosen based on a previous version of the detector layout and is significantly above the
 3502 size of the longest flex tail to be produced for the HGTD.

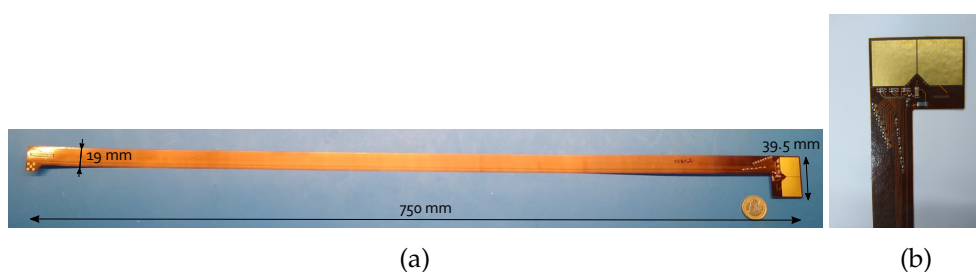


Figure 7.8: (a) Two flex cables prototypes. Top and bottom view. (b) Assembled extremities of the FLEX cable prototype from CERN PCB Service.

3503 The stack-up of the cable has layers numbered 1 to 4 from top to bottom. On the top layer
 3504 the single lines are routed following a micro-strip configuration. The differential pairs as
 3505 well as the HV line are placed in layer 3 in a stripline configuration in order to improve the
 3506 shielding of these lines. Layers 2 and 4 are supposed to be full planes dedicated to powering
 3507 and grounding. In order to perform the electrical tests the 4 flex cables have been assembled
 3508 with all the foreseen components.

3509 The qualification of the flex cables has been performed both at room temperature and in a
 3510 climate chamber reproducing the operating conditions of the HGTD in terms of temperature
 3511 (see Sec. 7.5), yielding very similar results.

3512 **Geometrical tests** The thickness and the width of the flex cable must be homogeneous
 3513 along its length. Several measurements of the thickness and width of the cables were
 3514 performed with a caliper every 5 cm. The mean values and standard deviations of the
 3515 measurements are shown in Tab. 7.6. The spread of the values is acceptable, as well as the
 3516 length and width average values.

	Length [cm]	Width [mm]	Thickness [μm]
Nominal	75	18	340
Measured	75.0 ± 0.2	17.99 ± 0.04	300 ± 10

Table 7.6: Mean values of the measured length, width, and thickness for 3 long flex cables.

3517 **Power integrity** A simulation of the voltage drop in each plane was performed with
 3518 the Cadence Allegro Sigrity PI software package [51] and the expectation for the longest
 3519 CERN prototype (750 mm) was estimated and compared with multimeter measurements
 3520 (see Tab. 7.7). While the measurements show an excellent agreement with the simulation,
 3521 the total resistance is above the specification for the power planes corresponding to about
 3522 200 m Ω for a flex cable of this length. **Add comments on ground planes.** Similar simulations
 3523 for the current baseline design of the flex tail (also shown in Tab. 7.7) are instead well within
 3524 specifications for all power and ground planes. The total power consumption is estimated
 3525 to be about 4.4 mW cm⁻¹, corresponding to a total of 1.8 kW over the whole detector (see
 3526 Tab. 11.1). **Update total power consumption consistently with updated total length**

plane type	CERN sim. [m Ω]	CERN meas. [m Ω]	tail baseline sim. [m Ω]
analog power	116 ± 7	150 ± 9	80
digital power	25 ± 1	30 ± 7	127
analog ground	111 ± 6	124 ± 2	32
digital ground	24 ± 1	26 ± 2	28

Table 7.7: Simulated and measured resistance of the analog and digital power and ground planes for the CERN prototype and simulation for the baseline design of the flex tail.

3527 **Insulation test** **to be updated**

3528 The insulation of the flex materials was checked for voltages up to 1 kV with the CAEN
 3529 DT5521HEN HV power supply [52] that can measure currents as small as 500 pA. Since no
 3530 current was observed over a long time, a lower limit was set on the insulation resistance at
 3531 2000 G Ω , well above the requirement.

3532 **Time Domain Reflectometry results** The Time Domain Reflectometry (TDR) test is per-
 3533 formed in order to check the impedance homogeneity of the tracks, which is crucial for
 3534 high-speed data transmission. Two assembled flex cables were used to measure two differ-
 3535 ential pairs and two single lines that are accessible from the adapter board. The TDR module
 3536 80E08 together with the DSA8200 oscilloscope by Tektronix [53] was connected through
 3537 SMA connectors to the adapter board. The impedance of the lines was measured and com-
 3538 pared with the impedance estimated from simulation, that was expected to be within the
 3539 requirements mentioned above. For all the measured lines the impedance is found to be
 3540 above the expectation by up to 20% (see Fig. 7.9). The same tests will be performed on the

3541 CERN prototypes and the discrepancies will be followed up with the manufacturer to adapt
3542 the design accordingly.

3543 **Integrated Bit Error Test (IBERT) and eye diagrams** Is the text up to date or only the
3544 figure?

3545 To emulate the signals from the ASIC an FPGA on the Kintex KC705 evaluation board
3546 [54] have been programmed and connected the flex cable via an adapter board to build an
3547 automatic test setup for all the types of cables that are under test.

3548 The FPGA injects test patterns at 1.25 Gbit s^{-1} and checks the response with the Integrated Bit
3549 Error Rate Test (IBERT). The SMA connectors placed on the adapter board route the signals
3550 to the oscilloscope for classical eye-diagram analysis. A wire bond between two differential
3551 pairs at the end of the flex cable creates a loopback path for the signals. Therefore, the
3552 transmission length of the signals is twice the FLEX length, 150 cm. The test configuration
3553 and the I/O drivers are compatible with the VC707 FPGA used by the LpGBT system. In
3554 this way we assure the same conditions as for on-field operation.

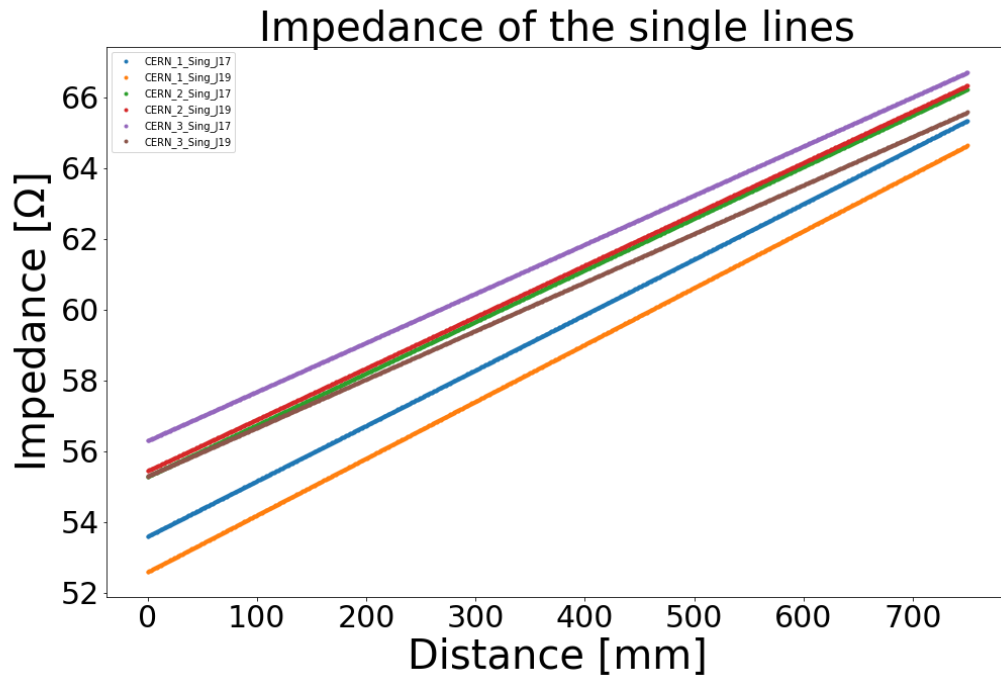
3555 The IBERT detected no errors over a few days, yielding a limit at 95% confidence level on
3556 the error rate for one of the long flex cables at 1.25 Gbit s^{-1} with BER no more than 6×10^{-15} .
3557 This value is well below the acceptable error rate of 1×10^{-12} .

3558 The same test was repeated with the HV up to 1000 V and showed no error for 8 days.
3559 The BER result obtained during this test was no more than 1×10^{-15} . The Kintex KC705
3560 evaluation board encodes the signals at the receiver after an equalization stage. The signals
3561 were measured prior to the equalizer by an oscilloscope. The signals amplitude ranges from
3562 $\pm 100 \text{ mV}$ to $\pm 200 \text{ mV}$. The eye diagrams in Fig. 7.10 measured without HV (a) and with HV
3563 (b) show a similar shape and opening area. The opening areas for both eye diagrams are
3564 larger than the no error accepted area limited by the mask. Tests over a longer time and a
3565 higher voltages are ongoing.

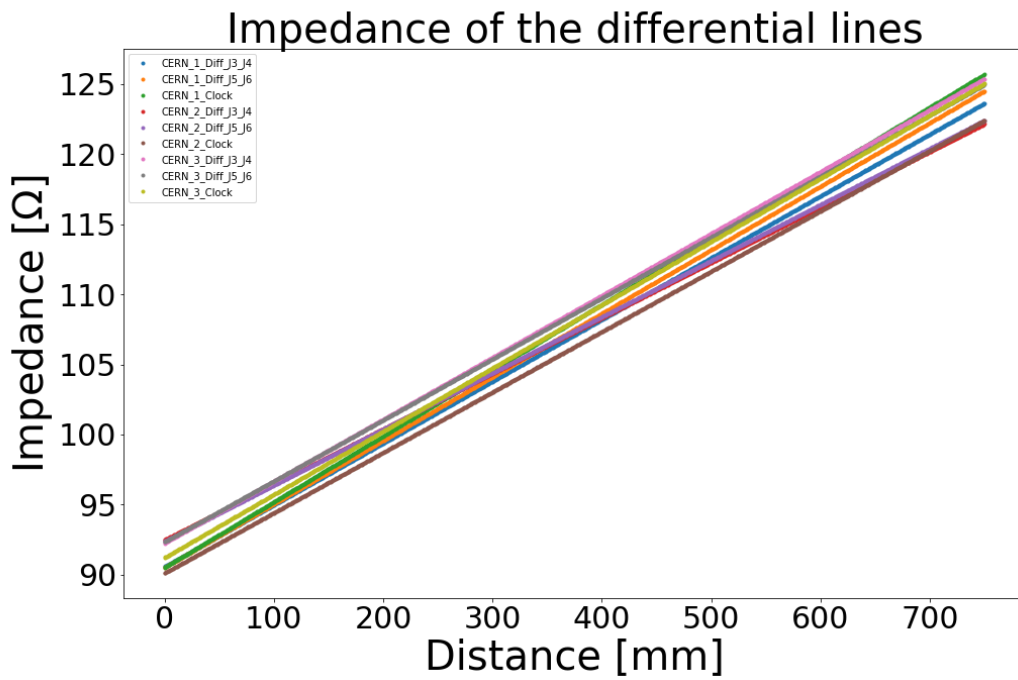
3566 7.3.3 Gluing and wire-bonding

3567 The assembly and interconnection of the bare module with the flex cable results in the HGTD
3568 module. The steps involved in the assembly process are the following:

- 3569 • Cleaning and preparation of the flex and bare module
- 3570 • Gluing of the flex on the bare module
- 3571 • Wire-bonding
- 3572 • Inspection, quality control and documentation



(a)



(b)

Figure 7.9: Impedance results for the single lines (a) and the differential pairs line (b) for the CERN prototype...

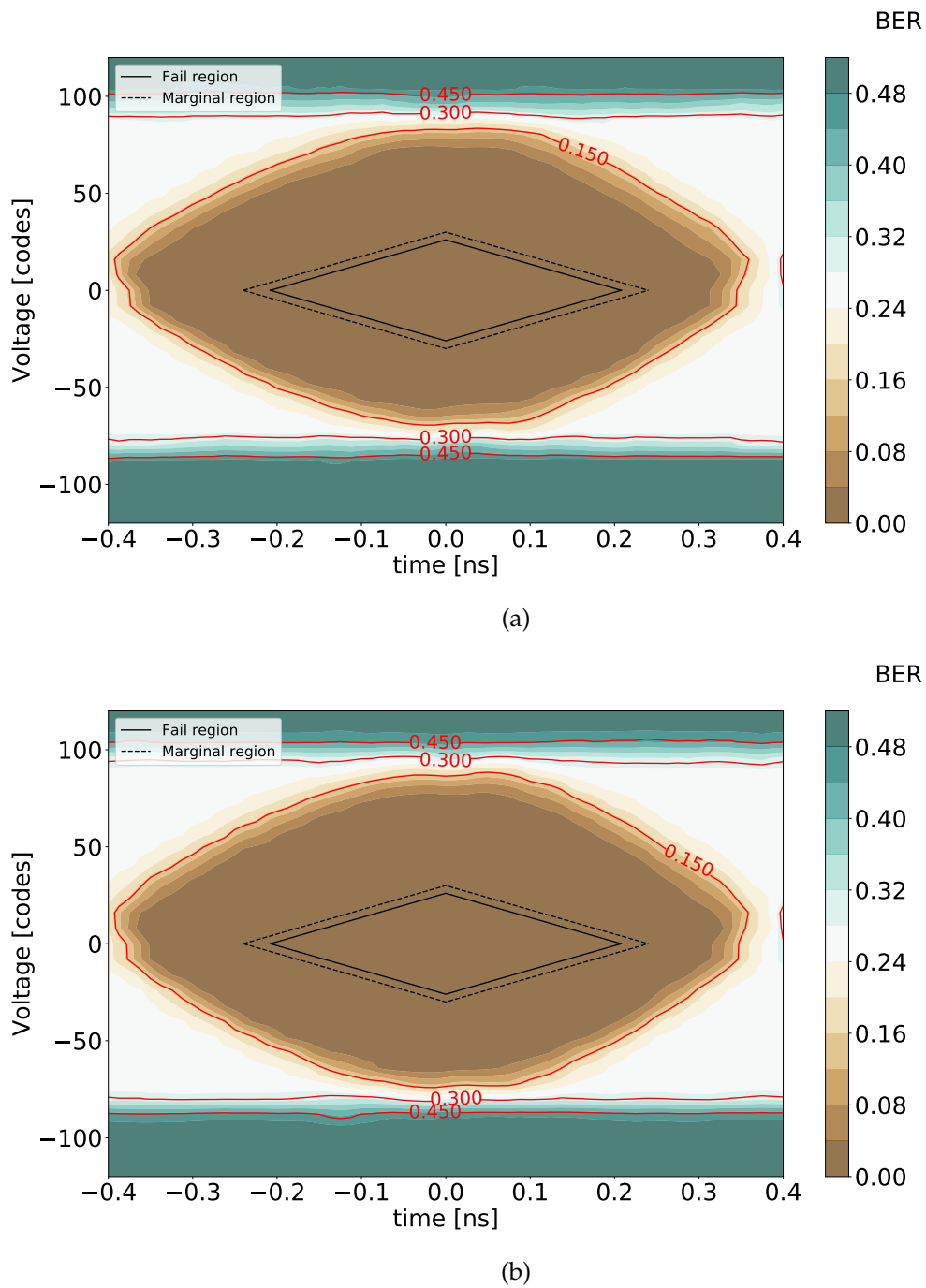


Figure 7.10: Eye diagrams for the flex cable prototype from the CERN PCB service. (a) $HV = 0V$ (b) $HV = 1kV$. The solid line indicated the mask in which no errors are acceptable, the dashed line is the marginal region in which few errors can be tolerated. **Figures updated!**

3573 These steps are discussed in more detail below.

3574 The assembly procedure for all HGTD modules will be identical and use the same tools. This
3575 facilitates the definition of the assembly procedure and increases yield. However, the details
3576 of the assembly procedure might differ between assembly sites, mostly in the gluing step,
3577 due to the availability of specialized equipment in the different institutes (dispensing and
3578 pick-and-place machines, for example). All module assembly and testing will take place in a
3579 clean environment equipped with temperature and humidity control and ESD protection.
3580 Specification for this environment will be developed and critical steps shall take place inside
3581 clean rooms. A database will be used to record the status of each module at every step
3582 of assembly. Electrical test results will also be added to the database. Given the number
3583 of modules needed for the HGTD, a few sites are foreseen to be qualified for the module
3584 assembly activities. To ensure uniform high quality in the module assembly process the sites
3585 will be asked to pass a site qualification stage.

3586 Initially the flex cables and the bare modules will be optically inspected for damage and
3587 anomalies. Components will be weighed and the surfaces where the adhesive will be
3588 deposited will be cleaned if needed. Bare modules and flex circuits will be mechanically
3589 joined using a dedicated adhesive. Several adhesives are currently being studied, for
3590 robustness, radiation hardness and other practical advantages (curing time, viscosity, etc).
3591 The baseline solution would be to use the same adhesive used in the ITk Pixel detector.
3592 Different options are available to carry out the gluing process. However, all assembly
3593 methods will be ensured to produce modules to the same specifications.

3594 One method to mechanically join the flex cable to the bare module relies on a pick-and-place
3595 machine, which typically achieves positioning accuracy of $\sim 10 \mu\text{m}$, and exists in a variety
3596 of automation options (from mostly manual to fully automated). Pre-tested components
3597 (flex cable and bare module) are loaded by vacuum tools of the machine. The operator then
3598 aligns the components through fiducials in the module (on the ohmic side of the sensor) and
3599 flex, visualized simultaneously in the machine monitor screen, and applies manually, or
3600 through a dispensing arm, the adhesive to the bare module and/or flex cable. The flex is
3601 then placed on top of the bare module and held in position until the adhesive is sufficiently
3602 cured.

3603 An alternative process relies in custom made jig gluing tools instead of the pick-and-place
3604 machine. A bare module placement jig is used to position the module using precision pins
3605 and hold it in position with vacuum. A similar jig is used to hold the flex and then the
3606 adhesive is applied. Bolts or guides are used to align the flex and bare module jigs and slide
3607 them together. Positioning accuracy of $\sim 100 \mu\text{m}$ is achievable with this method.

3608 Following mechanical assembly the front-end chips and the sensor bias voltage are electric-
3609 ally connected to the flex circuit through $25 \mu\text{m}$ diameter aluminium wire bonds using an
3610 automated ultrasonic wedge bonder. Wire-bond quality will be checked routinely through
3611 pull tests of sample wire bonds using a pull tester machine. Visual inspection of the wire

3612 bonds will also be performed. Fig. 7.11 shows the assembled ALTIROC1 device and the pull
 3613 testing procedure.

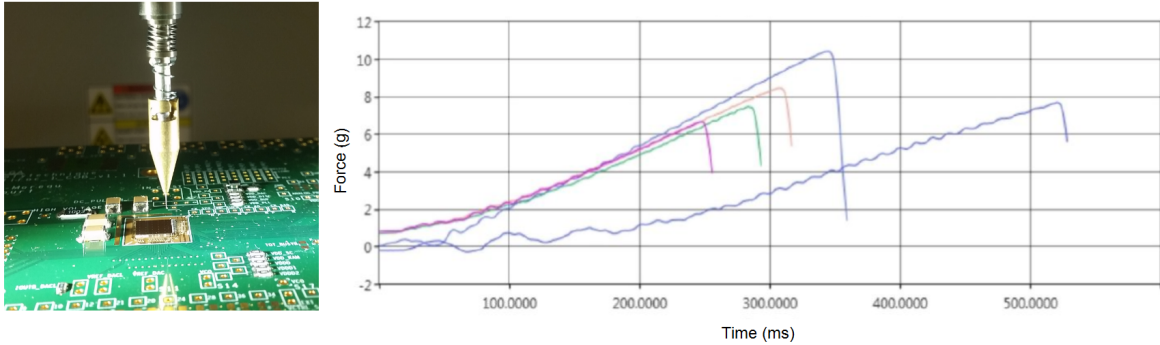


Figure 7.11: Photo of a mounted ALTIROC1 device being tested (left) and the measured wire strengths (right).

3614 **7.3.4 Assembly specifications**

3615 The flex cable specifications are described in Sec. 7.3 and summarized in Tab. 7.8. The flex
 3616 cables will be produced with different lengths and two orientations of the L-shape, so that
 3617 the general specifications can only be expressed in terms of tolerance of the actual size with
 3618 respect to the design.

Tolerance in length	1 mm
Tolerance in width	100 μm
Maximum thickness	350 μm
Insulation resistance of HV line	10 $\text{G}\Omega$
Maximum resistance of power planes	200 $\text{m}\Omega$
Impedance of single lines	45 Ω –55 Ω
Impedance of differential lines	80 Ω –100 Ω
Maximum allowed BER	10^{-12}
Radiation tolerance	5 MGy

Table 7.8: Specifications of the flex cable.

3619 The bare modules and flex cables that fulfil all the requirements will be used for the next
 3620 steps in the module assembly, i.e. gluing and wire bonding. The specifications for this stage
 3621 are aimed at ensuring the mechanical stability of the assembled module, see Tab. 7.9. These
 3622 need to be combined with the requirements in terms of efficiency, response and number of
 3623 working channels defined for the sensors and that are valid also for the assembled module.

Radiation tolerance	5 MGy
Lap shear force	5 MPa
Push-off strength	10 MPa
Wire bond pull force	6 gf
Positioning accuracy	100 μ m

Table 7.9: Specifications of the gluing and wire bonding processes.

3624 7.3.5 Quality assurance / quality control

3625 As described above, an exhaustive set of tests has been defined to qualify the flex cables
 3626 before connecting them to the hybrid. This set of tests constitutes also the baseline procedure
 3627 for quality control of the flex cables during production. However, since some of the meas-
 3628 urements (e.g. IBERT) are time consuming, the option of performing them only on a limited
 3629 fraction of flex cables will be considered once a low failure rate has been established. All
 3630 electric tests will be performed in a climate chamber reproducing the operating temperature
 3631 of the HGTD (about -30°C , see ??) and under controlled humidity conditions. Radiation
 3632 tolerance will only be tested on a small fraction of flex cables, which will not be usable for
 3633 assembly afterwards.

3634 After assembly all modules will be optically inspected and weighed, and their metrology
 3635 recorded in the database. As mentioned above, wire-bond pull tests will be carried out
 3636 periodically on a fraction of modules to ensure robust connectivity. All modules will be
 3637 tested for ASIC communication, current-voltage behaviour and response to a radioactive
 3638 source using a lightweight table top DAQ system. Short burn in tests, where the modules
 3639 are operated continuously for a day is foreseen. Furthermore a small fraction of the modules
 3640 will be subjected to long-term burn in tests where the devices will be subjected to thermal
 3641 cycles while being operated.

3642 7.3.6 Production strategy for module assembly

3643 The flex cable design will be finalized after testing it connected to the ALTIROC2 in the
 3644 demonstrator described in Chap. 14. Companies that are expected to be able to produce
 3645 long FCPs within specifications are being contacted and the production should be ideally
 3646 shared among a few of them that can provide the same quality of cables. The plan is to
 3647 involve them early on in the prototyping phase so that they can contribute to the design
 3648 optimizations specific to their manufacturing process.

3649 The standalone flex cable measurements, as well as gluing and wire bonding will be shared
 3650 among a few HGTD Institutes, so that the production rate does not exceed 4 modules per
 3651 day per Institute. An institute module assembly qualification procedure will be enforced to

3652 ensure that all sites uniformly produce modules according to specifications. A minimum set
3653 of equipment will be required (for example, wire-bonding and pull and shear machines) as
3654 well as a clean environment and minimum throughput capacities.

3655 7.4 Module loading

3656 7.4.1 General description

3657 The assembled modules have to be mounted on the cooling plates in readout rows, aligned
3658 along the x or y direction. Fig. 7.12 shows the positions of the modules on the front side (left
3659 plot in red) and back side (right plot in blue) allowing an overlapp of 20% for the inner part,
3660 55% for the middle part and 70% for the outer part. The longest rows contain 19 modules.
3661 For mechanical stability the modules will be glued to a thin support plate which is then
3662 screwed to the cooling plate. As described previously, the active area is divided in three rings
3663 (inner, middle and outer ring). Then, three types of support unit per side corresponding to
3664 the three rings are studied. The inner support unit consist of half disks of $120 < r < 230$ mm,
3665 the middle support unit consist of quarter disks of $230 < r < 470$ mm and the outer support
3666 unit consist of quarter disks of $470 < r < 660$ mm. The inner disks and middle disks will be
3667 replaced after every 1000 fb^{-1} and 2000 fb^{-1} respectively. The total number of support units
3668 for the eight sides of the detector is 80 (16 half inner supports, 32 quarter middle supports,
3669 32 quarter outer supports). Moreover, because of the positions of the modules are different
3670 for the two sides of the cooling plates, there are six different types of support units. The
3671 readout rows will therefore contain modules mounted on the support structures and they
3672 will be assembled independently. Fig. 7.13 shows a drawing of the detector units with the
3673 loaded modules.

3674 7.4.2 Support units and detector units

3675 Modules are installed and glued on plates (the support units) to be screwed on each side of
3676 one of the four cooling plates. Currently the design of the support units consists of a pattern
3677 plate, with modules inserted into (between the plate and the cooling plate). The full size
3678 plate is divided into three parts as described in Fig. 7.13. The current baseline is to use half
3679 disks for the inner part and quarter disks for the middle and outer parts. Depending on the
3680 feasibility (fragility, flatness, glue deposition), smaller supports could be considered. The
3681 maximum thickness of the support plate is typically 4 mm. The target material is currently
3682 carbon fibre. For example, Fig. 7.14 shows the current design of the half disk of the inner
3683 support unit.

3684 Windows are machined in order to encapsulate the modules which are glued on rectangular
3685 strips (see Fig. 7.14). The positioning of modules is given by the windows of the plate (see

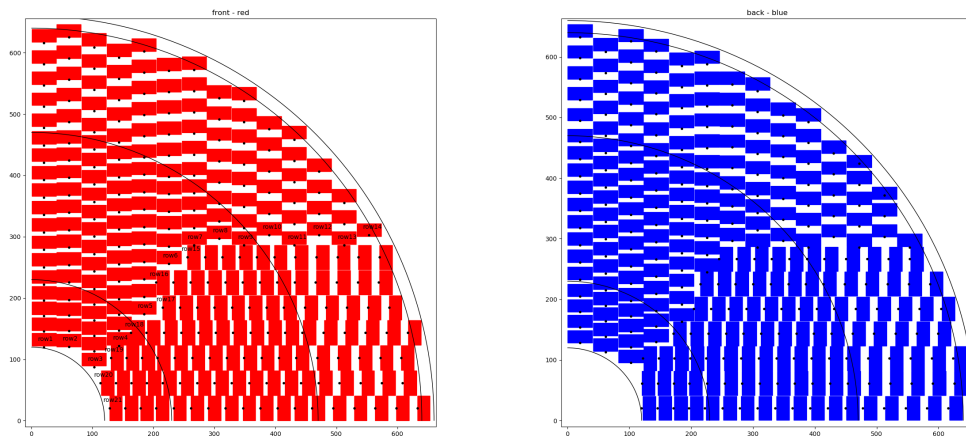


Figure 7.12: Position of modules and readout rows numbering for the front side of one disk (left plot in red) and for the back side of the same disk (right plot in blue). Smallest radius at 120 mm and maximal radius at 660 mm are shown. The 640 mm radius is the minimal target for the external instrumented area. 230 mm and 470 mm radius are shown as typical limits of the different parts of the support units.

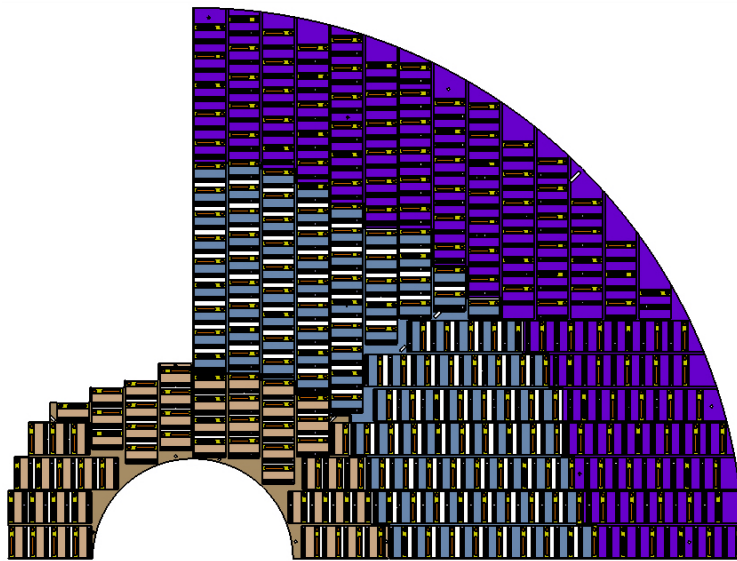


Figure 7.13: Detector units with modules assembled on the inner, middle and outer support plates

3686 Fig. 7.14). Once this detector unit is screwed to the cooling plate, the modules are in direct
 3687 contact with it, so that the thermal properties of the plate material and of the glue are not
 3688 critical. Moreover, thermal grease will be used to improve the contact.

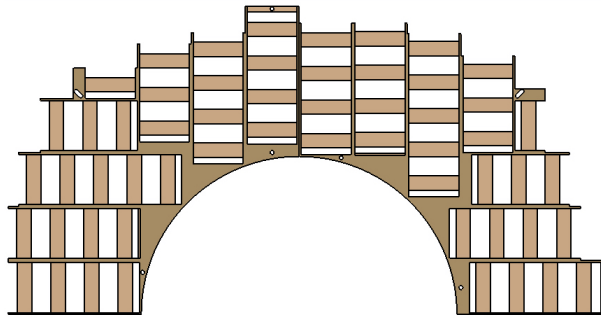


Figure 7.14: Drawing of inner support unit with holes for fixation on the cooling plate

3689 For better mechanical strength and rigidity, some reinforcement are added (Fig. 7.14). First
 3690 tests show that a single half disk for the inner part and a single quarter for the middle and
 3691 outer parts would guarantee the stability of the global structure. This type of support plate
 3692 is more complex than a simple plate, since the windows need to be defined precisely for
 3693 each module, but then the positioning of the module itself is straightforward. The structure
 3694 provides mechanical protection to the modules and the thicker plate has better rigidity. On
 3695 the downside, this design only allows for a small surface when gluing the module to the
 3696 plate and the mechanical strength and long term stability has to be studied. Some tests have
 3697 to be performed and several prototypes and the demonstrator will be useful to conclude
 3698 (see Chap. 14). In terms of thermal conductivity, this design has the advantage of placing
 3699 the modules in direct contact with the cooling plate, avoiding thermal constraints in the
 3700 material of the plate and the glue. Should a module be found to be faulty after gluing to
 3701 the support, rework should be relatively easy. Conclusive tests have already been carried out
 3702 and others will be done with the demonstrator.

3703 7.4.3 Gluing studies

3704 The modules are fixed to the support unit thanks to four glue dots 2 mm diameter each (see
 3705 Fig. 7.15). The glue dots are deposited onto the edges of the flex connector. The glue for
 3706 module loading into the intermediate plate should meet a priori, the parameters listed in
 3707 Tab. 7.10.

Radiation tolerance	> 5 MGy
Viscosity	< 100 Pa s
Lap shear force	1 MPa
Push-off strength	1 MPa

Table 7.10: Specifications of the glues parameters.

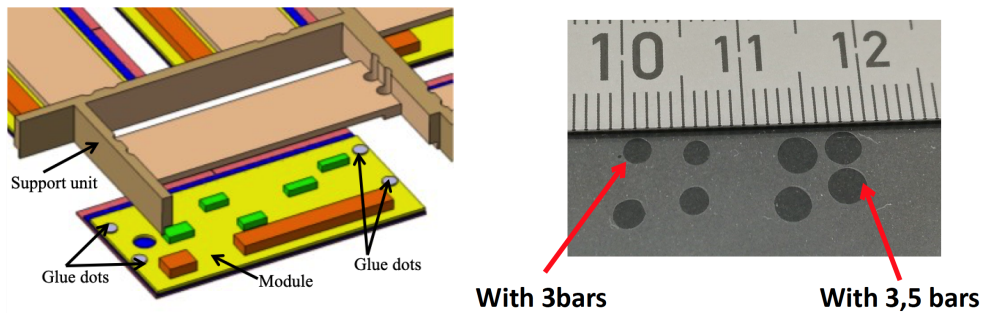


Figure 7.15: Schematic view of the module with the four glue dots allowing the fixation with the support unit (left) and test of glue deposition (right) - Pressure values are an example of tuned parameters, depending on the duration and temperature.

3708 With these constraints, six types of glues have been chosen to perform the tests:
 3709 Araldite2011; EG7655-LV; EG7655; EG7658; EG8050; Stycast 2850FT.
 3710 Their characteristics have been checked in the MaxRad (Materials and Adhesives for Ex-
 3711 treme Radiation Environments) CERN database. Ease of implementation (fluidity, life time,
 3712 duration and temperature polymerisation) has been evaluated and the ITk choice has been
 3713 considered for radiation hardness Moreover, the push-off strength and lap shear test have
 3714 been performed in several configurations. These tests have been done using some dummy
 3715 modules, with a piece of flex cable glued onto to mimic the flex connector. Other tests
 3716 with glass plate have been performed to determine the volume of glue to obtain the correct
 3717 thickness and surface (see Fig. 7.15). Finally, taking into account all the tests already done
 3718 and the results and recommendations from ITk project, ARALDITE 2011 is chosen as the
 3719 baseline for the loading of the modules onto the support unit.

3720 7.4.4 Procedure for assembly and qualification

3721 The procedure for stave loading will be tested when assembling the demonstrator (see
 3722 Chap. 14), which will be also used to improve the definition of the assembly procedure
 3723 and the qualification steps. Tools are being developed and tested for this design. Tests
 3724 are performed following a procedure first using glass plates then silicon glued to a small
 3725 flex prototype, all without any electrical functionalities, but with the correct geometrical
 3726 dimensions, instead of actual modules. Araldite 2011 is used as glue. Module loading on
 3727 support unit should follow this procedure :

- 3728 1. The modules are placed on a temporary plate with the pattern of the module positions.
 3729 They are maintained thanks to a suction system included in the plate.
- 3730 2. Four glue dots are dispensed on the left and right edge of the flex connector (see
 3731 Fig. 7.15); the thickness of the glue is insured by the automatic dispenser.

- 3732 3. The support unit is put in place and compressed on all modules at a nominal com-
3733 pression strength. A adjusting shim is used to ensure the correct thickness of the
3734 glue.
- 3735 4. The polymerisation is carried out (temperature and duration to be defined after final
3736 glue tests).
- 3737 5. The detector unit is removed and fixed on a plate for packaging and shipping.
- 3738 6. The system is turned over upside down and a second transport plate is fixed on the
3739 top.
- 3740 7. Electrical tests can be performed at this stage.

3741 During production, a visual inspection will be performed after module loading, looking
3742 for possible mechanical damages to the module, in particular to the edges of the hybrid,
3743 the flex components and the wire bonds. Signals will be injected into the sensors and the
3744 response will be tested with the same DAQ system used for the test of the single modules.
3745 Additionally it will be checked that there is no interference between the modules due to
3746 the stacked flex cables. For all detector units passing the qualification tests, the information
3747 on the nominal and measured position of the modules on the support unit, as well as any
3748 relevant performance results will be saved to a database. Once the initial characterisation is
3749 completed in the R&D phase, thermal tests are not foreseen during production. Mechanical
3750 stress tests could be performed on a small fraction of support units if it is deemed necessary.

3751 **7.4.5 Detector unit assembly strategy**

3752 Once the choice of the design is finalised, the production of the support units will be carried
3753 out by a company and the monitoring and control by an institute. Then, the plates will
3754 be shipped to the module loading sites that have been qualified. To minimize the amount
3755 of modules to be shipped and to avoid long distance transport, sites able to perform both
3756 module assembly and loading or geographically close to the module assembly sites will be
3757 preferred. Since the setup for mechanical and electrical qualification of the detector units
3758 is similar to the one needed for module assembly, the site qualification procedure will be
3759 mostly common to both activities (excluding the wire bonding capability in this case). As for
3760 module assembly, the exact procedure used for module loading might be slightly different
3761 among the institutes, but the same quality of assembled staves has to be delivered. 80
3762 support units (16 inner, 32 middle, 32 outer) will be produced. Since the module positions
3763 are different on the front side and on the back side of a cooling plate, the designs of the
3764 support units are different for the two sides. In total, there are 6 types of support units. The
3765 glue and the expendable supplies will be purchased from one or more companies. Most of
3766 the components of the electrical test benches are standard ones, available in the institutes.
3767 Some dedicated electronic boards will be developed to test the modules at many steps

3768 of the construction of the detector, included the loading step. The gluing and positioning
 3769 system will be developed in the institutes, using existing elements, complemented by specific
 3770 mechanical parts. Because of the non-standard shape of the detector units and the fragility,
 3771 different types of dedicated packaging will be necessary for transportation from loading
 3772 sites to CERN.

3773 7.5 Thermal calculation

3774 The power dissipation of the sensor depends strongly on the temperature of the sensor.
 3775 The irradiation of the sensors will increase the leakage current thus increasing the power
 3776 dissipation at a given temperature. Therefore the thermal properties of the system have been
 3777 studied following the strategy outlined in [55].

	Material	Thickness [mm] [mm]	Thermal Resistivity [W/ (m°C)]
Sensor	Si	0.25	124
Bumps	SnAg	0.05	79
ASIC	Si	0.25	124
Foil	Polymer	0.10	3.5
Structure	CarbonFiber	0.50	1
Cooling	Graphite foam	2.00	30
Tube	Al	0.30	135

Table 7.11: Material type and thickness used in the thermal simulation.

3778 In a first step the geometry of a stack with a single ASIC and (half an LGAD) sensor is built.
 3779 The material used in the thermal simulation of the module are shown in Tab. 7.11 along with
 3780 their thickness and thermal conductivity. The sensor, the ASIC, the foil, the structure and the
 3781 cooling are implemented each as a cuboid built of a square $2\text{ cm} \times 2\text{ cm}$ and the height given
 3782 in Tab. 7.11. The bumps connecting the sensor to the ASIC are implemented individually
 3783 as 225 cylinders with a radius of 0.045 mm and height of 0.05 mm. The cooling pipes are
 3784 half-cylinders embedded in the cooling material. The inner radius of the pipes is 1.5 mm
 3785 and the outer radius is 1.8 mm.

3786 The cooling is simulated as convection which is applied on the surface of the cooling pipes.
 3787 Their nominal temperature is -35°C . As baseline a power consumption of the ASIC of
 3788 1.2 W ($0.3\text{ W}/\text{cm}^2$) is used. For the sensor a power consumption of 0.4 W ($0.1\text{ W}/\text{cm}^2$) is
 3789 assumed.

3790 While the contact between the sensor and the ASIC via the SnAg bumps is assumed to be
 3791 perfect a thermal contact resistance of $0.01\text{ Wmm}^{-2}\text{C}^{-1}$ is applied to the contact between
 3792 ASIC and foil as well as foil and the carbon fiber structure. The contact resistance leads to a

3793 temperature step increasing the thermal resistance of the system. For a power dissipation of
 3794 1.6 W the temperature step is 0.4 °C at each material transition.

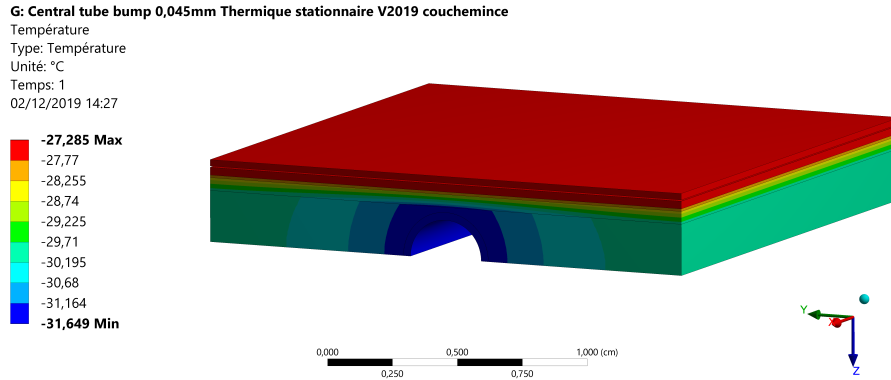


Figure 7.16: The temperature distribution is shown for the baseline power consumption with an ASIC and a sensor half.

3795 In Fig. 7.16 the result of the thermal simulation by ANSYS is shown. The maximum tem-
 3796 perature difference is 7.7 °C. If the ASIC is powered alone, the temperature difference is
 3797 5.6 °C, for the sensor alone, the temperature difference is determined to be 2.2 °C. The
 3798 thermal resistance for the sensor is therefore 5.4 °C/W and for the ASIC it is 4.7 °C/W. As
 3799 the difference between these two resistances of 0.7 °C/W is due to the soldering bumps, the
 3800 thermal resistance of the bumps was calculated analytically using a continuous equivalent
 3801 volume of SnAg instead of the discrete bumps. The approximation leads to a resistance of
 3802 0.5 °C/W, the larger value for the individual bumps can be understood as the heat transfer
 3803 will see also the resistance in the sensor plane before reaching the bumps in order to flow to
 3804 the cold reservoir.

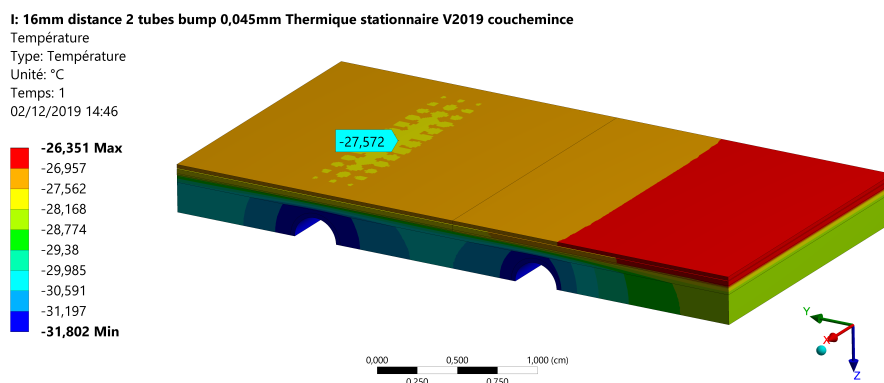


Figure 7.17: The temperature distribution is shown for the baseline power consumption with two ASICs and a sensor.

3805 As a second step the second half the sensor was added as well as the corresponding ASIC. The
 3806 current design of the cooling pipes calls for pipes every 16 mm, therefore a second cooling
 3807 pipe was added at the nominal distance leading to an asymmetric configuration shown in
 3808 Fig. 7.17. Compared to the previous simulation the temperature difference increases to 8.7 °C
 3809 peak to peak. However the temperature distribution of the sensor now shows variations
 3810 with a symmetry axis corresponding to the axis of the cooling pipe. Restricting the study
 3811 to a single cooling pipe \pm half the cooling pipe to cooling pipe distance, the temperature
 3812 increase is reduced to 7.4 °C which is close to the result of the previous simulation within
 3813 5%. For the simulations with only the sensor or ASIC dissipating power the temperature
 3814 increase is globally larger, however as the increase is less than a factor 2, but the power ins
 3815 doubled, the resulting thermal resistance is smaller. Therefore the single-ASIC simulation is
 3816 a good approximation of the system. Additionally the geometry is conservative as the next
 3817 cooling pipe is close to the second ASIC, but has not been simulated. This would further
 3818 reduce the thermal resistance.

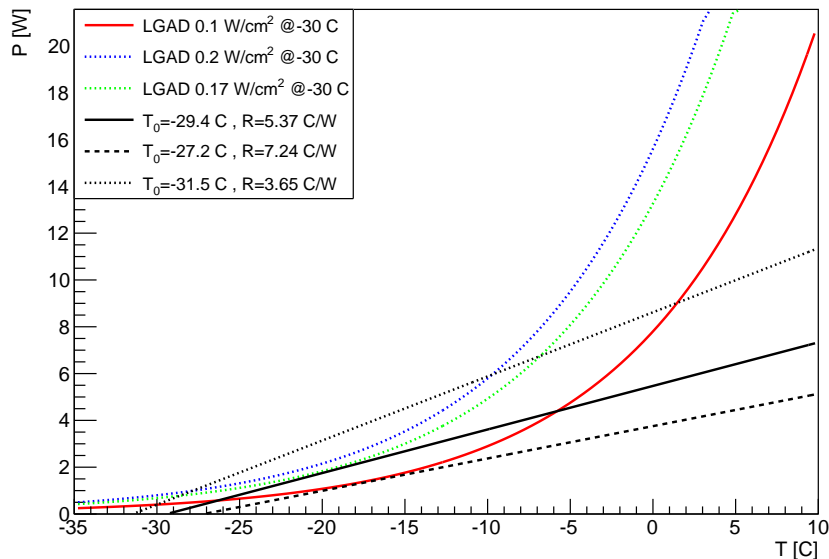


Figure 7.18: The power dissipation of the sensors as function of the temperature is shown as well as the thermal properties of the system.

3819 As the power dissipated by the sensor increases as function of the temperature, if the system
 3820 cannot evacuate the heat effectively, the temperature will increase, increasing the leaking
 3821 current, so that a thermal runaway condition is created as explained in [55].

3822 The power dissipation of the sensor is shown as a function of the temperature in Fig. 7.18.
 3823 The strong temperature dependence is clearly visible, e.g. in the red curve for the baseline.
 3824 The power dissipation of the ASIC increases the effective temperature delivered by the

3825 cooling system to $-29.4\text{ }^{\circ}\text{C}$. The black line has a slope which is the inverse of the thermal
 3826 resistance for the sensor. Once the power dissipation of the sensor crosses this line, thermal
 3827 runaway is excluded as the heat can be evacuated efficiently. At $-5\text{ }^{\circ}\text{C}$ stable operation
 3828 cannot be achieved anymore.

3829 The green dotted curve in Fig. 7.18 shows the maximal power dissipation the system can
 3830 handle. For about 0.17 W/cm^2 the power dissipation of the sensor is tangent to the black line.
 3831 Thus compared to the baseline a margin of 70% is included in the system. The temperature
 3832 dependence is modeled according to Sec. 5.5.8. The model was compared to the one used
 3833 in [55] by normalizing the models to the same power dissipation at a temperature of $-30\text{ }^{\circ}\text{C}$.
 3834 In a window of half-width $5\text{ }^{\circ}\text{C}$ around the normalization point, the two models agree within
 3835 15%.

3836 A different way of analyzing the properties of the system is to determine the resistance
 3837 for which the baseline sensor power dissipation is tangent to the line. The black dotted
 3838 line shows the result of increasing the thermal resistance starting at the ASIC by 40%.
 3839 Additionally the temperature increase of $5.6\text{ }^{\circ}\text{C}$ for the nominal system includes the contact
 3840 resistance degradation of $0.8\text{ }^{\circ}\text{C}$ proving a further margin of 14%. As the effective contact
 3841 area between materials is difficult to estimate, it is essential to have this margin built into the
 3842 system.

3843 If both the Carbon Fiber and the Graphite Foam were to be replaced by a system of Alu-
 3844 minium, the thermal resistances of the system would be improved further. The effective
 3845 operating temperature of system would decrease to $-31.5\text{ }^{\circ}\text{C}$ and the thermal resistance
 3846 would decrease to $3.7\text{ }^{\circ}\text{C/W}$ as shown in Fig. 7.18 leading to further margin in the operation
 3847 of the system.

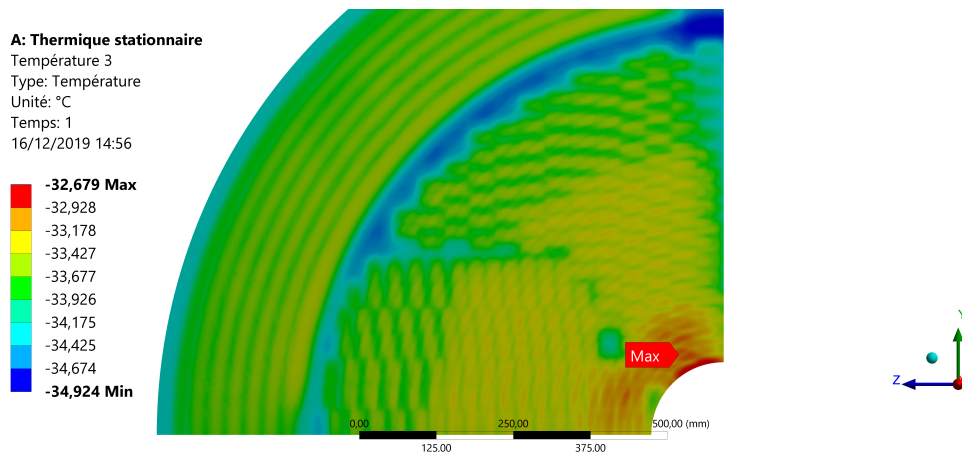


Figure 7.19: The temperature distribution on the surface of the cooling system is shown for a quarter disk when applying 0.4 W/cm^2 at the location of each ASIC.

3848 The detailed simulation of the stack for the full HGTD is not possible for the ASIC stack.
3849 Therefore a different approach is used. The cooling system is simulated fully for a quarter
3850 disk using only Aluminum as explained above. At the position of each module on the
3851 disk a power dissipation of 0.4 W/cm^2 corresponding to 1.6 W is applied. The resulting
3852 temperature variation is shown in Fig. 7.19. The maximal temperature is -32.7°C . The
3853 modules and the space between the modules explains the temperature variation. At the
3854 position of a module, in the inner ring of the HGTD, the resulting temperature is -33.3°C .
3855 In the space between the modules the temperature decreases only by about 0.1°C .

3856 Taking the maximal temperature of -32.7°C from the global model, the surface temperature
3857 of the cooling system of the detailed model with a single cooling pipe is fixed to this
3858 temperature. The temperature at the LGAD is determined to be -31.5°C , using the baseline
3859 power dissipation.

3860 The curvature of the cooling pipes could lead, for some modules, to a loss of cooling surface
3861 of the pipes, increasing the thermal resistance. However the temperature difference for
3862 the hottest module position with respect to the nominal temperature of the cooling system
3863 is 3.5°C . In the single cooling pipe model the temperature difference for the same power
3864 dissipation is larger with 5°C . Therefore there is no indication of such an effect. Additionally,
3865 the distance between two cooling tubes is smaller than the size of the module. Therefore
3866 the temperature difference in the simple model is increased artificially for the same power
3867 dissipation, increasing also the thermal resistance.

3868 The studies indicate that the simple model with a single straight cooling pipe is a conservative
3869 representation of the system.

3870 8 Power Distribution, Control, and Safety 3871 Systems

3872 This section covers the powering of the detector, including the layout of the High Voltage
3873 (HV) and Low Voltage (LV), from the supplies located in the USA15/UX15 services cavern,
3874 the DC-DC converters placed at the PP-EC patch panel area, up to the on detector peripheral
3875 electronics and modules sitting inside the vessel. The grounding and shielding schemes
3876 are also described. The details of the services needed to power the detector and respective
3877 connectivity are described in Chap. 12.

3878 8.1 High voltage

3879 Each of the 8032 LGAD sensor modules of the detector require individual bias voltage
3880 in a range up to 800 V. Such a high voltage is needed to power the sensors after being
3881 exposed to the high radiation conditions of the HL-LHC (detailed in Chap. 5). To allow
3882 an average leakage current up to 5 μ A per pad for radiated sensors, a 3 mA supply current
3883 will give sufficient margin. The bias voltage of the sensors has to be adjusted due to the
3884 gain degradation with the received fluence. Fig. 5.17 shows the required bias voltage as a
3885 function of the radial position for different fluence levels. In combination with the non-radial
3886 geometry, this results in a limited possibility to connect several modules to the same bias
3887 supply. The baseline choice is to use individual adjustable voltages to allow for optimal
3888 operation of the sensor modules. This requires 4016 HV supply channels per end-cap. The
3889 supplies will be based on commercial multi-channel rack mounted units located in the
3890 service cavern. A schematic layout of the high voltage system is shown in Fig. 8.1.

3891 The return line for HV channels will be done through a common ground, as for the liquid
3892 Argon calorimeter. This ground is available from the peripheral electronics boards via the
3893 modules analogue ground. The shielding will follow up to the filter units at the PP-EC patch
3894 panel boxes. A further low pass filter is placed on the flex cables near the sensor modules.

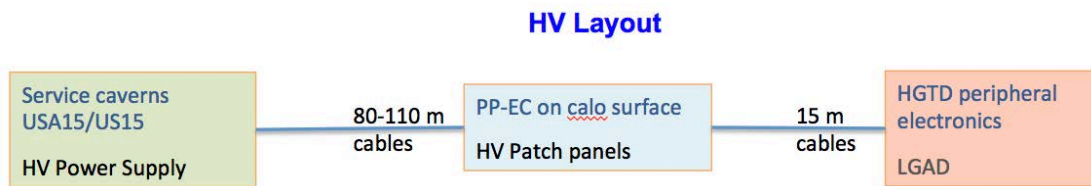


Figure 8.1: HGTD High Voltage Layout

8.2 Low voltage

For supplying the low voltages needed by the front-end end peripheral electronics a three stage system is used, as shown in Fig. 8.2. The system will have to be able to deliver almost 20 kW at 1.2 V. Bulk power supplies located in USA15 provide 300 V DC current to DC-DC converters placed in the PP-EC areas (described in Sec. 12.1.3). These second-stage multi-channel DC-DC units convert the 300 V to 10 V which is distributed to radiation hard DC-DC converters located on the peripheral electronics boards inside the vessel (details in Chap. 9). The last stage converts power to the front end ASICs on the detector chips and the peripheral electronics providing mainly 1.2 V DC power but also 2.5 V for optical links. The converters of the peripheral boards are based on the bpol12V ASIC developed by CERN for the HL-LHC upgrade. Due to space limitations on the peripheral boards, the 10 V to 1.2 V conversion will be done in a single stage (see Sec. 9.3). The exact output voltage for each converter on the peripheral boards is selected by a resistor chain to take the voltage drop of the flex cables into account.

Each ALTIROC ASIC requires 0.5 W analog power and 0.7 W digital power at 1.2 V. Separate DC-DC converters will be used for the analog and digital voltages. With two ASICs per module, one bpol12V based DC-DC converter can supply analog power to 4 modules or digital power to 3 modules. With 2008 modules per disk (or double-sided layer), 1196 DC-DC converters on the peripheral electronics per disk are needed to power the front end electronics, including power losses on the flex cables. A further 120 DC-DC converters per disk are required for powering the peripheral boards themselves.

These DC-DC converters on the peripheral electronics will need to provide almost 5.0 kW of power per disk. With an efficiency of 65%, each disk has to receive 800 A at 10 V which will be supplied by 52 channels providing 16 A each.

The 300 V will be provided by 14 rack-mounted units in the service cavern, each delivering 3 kW. Details on the low voltage units are given in Tab. 8.1.

With an 80% efficiency of the 300 V to 10 V DC-DC power converters located in the PP-EC area, a total cooling power of 4 kW is required at these locations. A water leak-less cooling system, providing water at $\approx 18^\circ\text{C}$, and corresponding pipes/manifolds on the calorimeter

Voltage	Location	Current/channel	Nb of channels/units
300 V	USA15	10 A	14
300 V → 10 V	PP_EC	16 A	208
10 V → 1.2 V (or 2.5 V)	On peripheral board	4 A	5360

Table 8.1: Type of LV units, location, delivered current per unit and number of units/channels.

3924 surface will be needed. Details on the services, patch panels area and cabling are given in
 3925 Sec. 12.1.6. A schematic layout of the low voltage system is shown in Fig. 8.2.

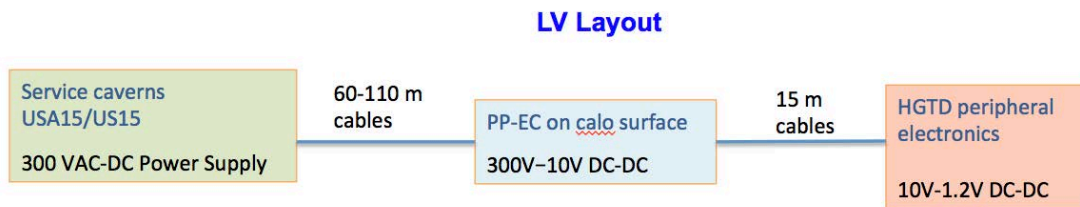


Figure 8.2: HGTD Low Voltage/power Layout

3926 8.3 Grounding/shielding

3927 The grounding and shielding of HGTD follows similar requirements as defined for ITK.
 3928 The ground reference point for the HGTD itself will be the inside of the detector vessel.
 3929 The inside of the vessel is covered with a thin high conductive foil to ensure the function
 3930 as a Faraday cage. Both end caps will be independent Faraday cages. The cage will be
 3931 extended up to the patch panels at PP-EC through the shields of the LV, HV and control
 3932 cables. The patch panels as well as the vessels are electrically insulated from the detector
 3933 walls and from the mechanical structures on which they are mounted. The Faraday cage
 3934 will be connected through a single ground line to the ATLAS common ground. This will
 3935 constitute the reference potential.

3936 This requires the CO₂ transfer line to be electrically insulated at the cooling junction box
 3937 located on the end cap calorimeter surface. Shielding for cables will be discontinued appro-
 3938 priately to avoid ground loops.

9 Peripheral Electronics

The on-detector peripheral electronics transfers data between the detector modules and the TDAQ, DCS, and luminosity systems as well as having a central role in the monitoring of sensor temperatures and supplied low voltage. The system is based on the CERN-developed lpGBT ASICs [56]. The modules are connected via flex cables, while signals to and from the TDAQ and the luminosity systems are transferred on optical fibres where the DCS data and commands are embedded in the data streams via the TDAQ optical fibres.

Each flex cable serves a module consisting of two ALTIROC ASICs and contains two differential electrical CERN Low Power Signalling (CLPS) e-links transmitting timing data at different rates (320 Mbit s⁻¹, 640 Mbit s⁻¹, or 1.28 Gbit s⁻¹) depending on the ASIC position. Flex cables for modules placed at a radius above 320 mm carry a further two differential e-links at 640 Mbit s⁻¹ with luminosity data. Each cable also contains four e-links with clock and fast commands to the ALTIROC ASICs at 320 Mbit s⁻¹ as well as the lines for the ASIC low voltage power supplies, control signals and the bias voltage of the sensor. The digital output data from several ASICs are merged in lpGBTs on peripheral boards (PEB) and transmitted on optical fibres to the off detector DAQ system. Control signals to and from the ALTIROC ASICs are transmitted via I²C bus where the commands and data are embedded in the data streams transmitted to and from the detector DAQ system. An overview of the HGTD readout chain is presented in Fig. 9.1.

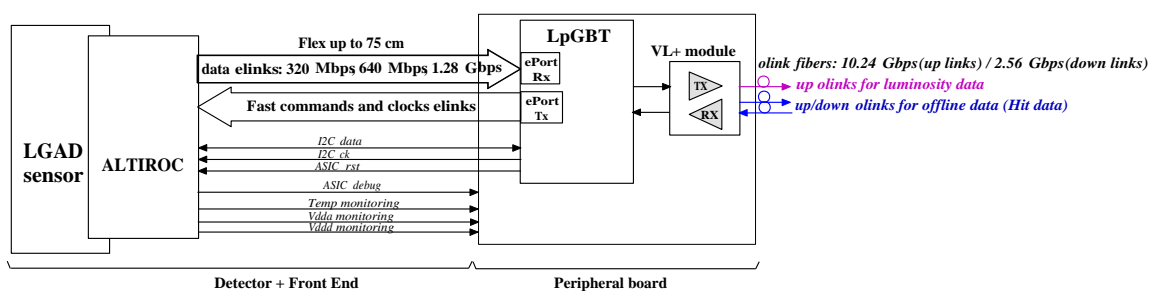


Figure 9.1: Upstream and downstream data flow. The e-links transmit data, fast commands and clocks between the ALTIROC ASIC and the lpGBT. VL+ is the Versatile Link+ VTRX+ module. I²C-bus, ASIC control and monitoring lines from the ALTIROC are also shown.

The peripheral electronics also includes the 10 V to 1.2 V DC-DC converters for the digital and analogue voltage supplies to the ALTIROC ASICs. The supply voltages are monitored

3960 using the internal multiplexed ADC on the lpGBTs. The ADCs are also used to measure
 3961 actual voltages received by the ALTIROC, as well as the sensor temperatures, where voltages
 3962 to be measured are selected by analog 64-to-1 multiplexers mounted on the peripheral
 3963 boards as described in Sec. 9.4.

3964 The lpGBTs that are used for transmitting luminosity data do not a priori need to receive
 3965 downlink data via optical fibres and will thus only send data off the detector. These lpGBTs
 3966 will receive the required 40 MHz clock from lpGBTs connected to the off-detector DAQ
 3967 system.

3968 A schematic block diagram of the PEB electronics for one module connected to off detector
 3969 electronics is shown in Fig. 9.2.

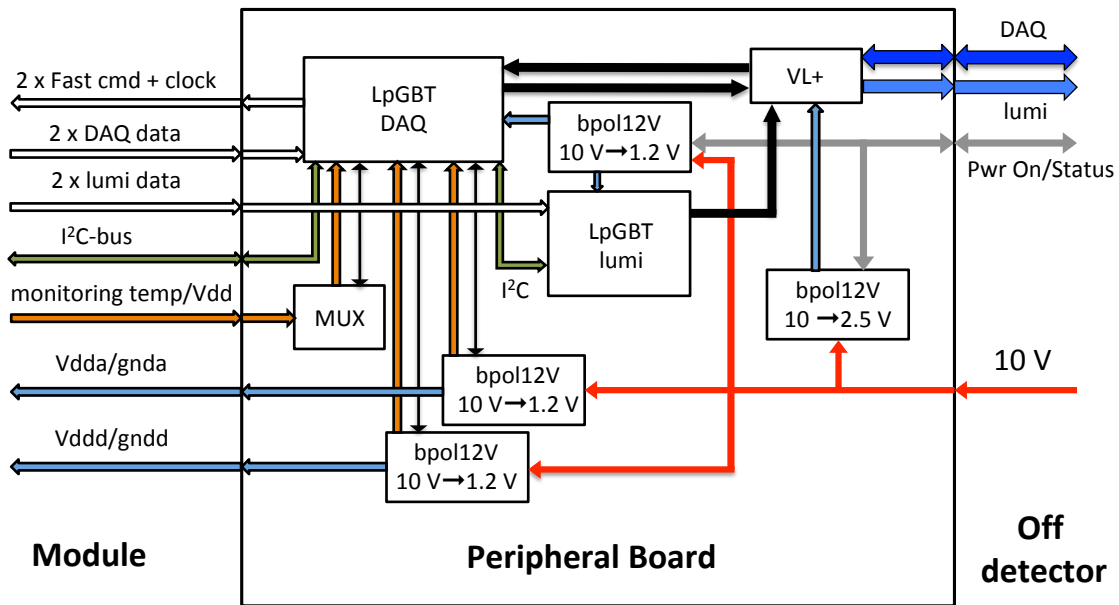


Figure 9.2: Block diagram of the peripheral electronics for powering and read out of a module. Multiple modules are connected to the same DC-DC converters and lpGBTs. The brown lines indicate voltages measured by the multiplexed ADC on the lpGBT. Light blue lines are low voltage power supplied from bpol12V DC-DC converters. The thin black lines are control signals via the general purpose I/O lines of the lpGBTs. The thick black lines are high speed electrical links to and from the VL+ optical module. Other lines are explained in the figure.

3970 As introduced in Sec. 2.3, each HGTD vessel contains two cooling disks (shown in Fig. 2.4),
 3971 with detector modules mounted on both sides, thus having two instrumented layers per

3972 disk. The baseline design for the peripheral electronics is to have five PEBs per quadrant per
3973 side of each cooling disk. Such a layout yields 80 boards per HGTD vessel, and 160 boards
3974 in total. Each board covers three or more readout rows in order to have similar number of
3975 ALTIROC ASICs connected per board to allow to optimize the use of the lpGBTs by sharing
3976 across readout rows.

3977 9.1 Requirements

3978 9.1.1 Data transfer

3979 The bandwidth required for the digital data output from the ALTIROC ASICs is given by
3980 the number of pads hit in an event. The expected average number of hits depends on the
3981 radius of the module position. The hit rate has been studied using simulation and results are
3982 presented per ASIC for an $\langle\mu\rangle = 200$ in Fig. 9.3. The radial dependency is clearly seen, with
3983 a maximum below 20 at the innermost radius. Such a rate can be accomodated within the
3984 maximum available rate for the data from the ALTIROC, which is 1.28 Gbit s^{-1} . For larger
3985 radius the bandwidth per e-link can be reduced, using 640 Mbit s^{-1} at radii above 200 mm
3986 and 320 Mbit s^{-1} at radii above 300 mm. These rates are chosen in order to minimise the
3987 numbers of lpGBTs and optical links while allowing a 50% increase of the expected number
3988 of average hits per ASIC at a readout rate of 1 MHz.

3989 In addition each ALTIROC ASIC requires a 320 Mbit s^{-1} fast command link to supply both
3990 the bunch crossing information and the TDAQ commands. A 320 MHz clock extracted inside
3991 the lpGBT from the command data packages is also sent to each ASIC.

3992 9.1.2 Physical limitations

3993 The available physical space for the peripheral electronics is very limited. It is constrained in
3994 the radial direction by the end of the instrumented area and the limit of the HGTD vessel,
3995 therefore ranging from 660 to 920 mm. Because the allowed width of the HGTD is only
3996 75 mm, the space available for the electronics in the z-dimension is also very small: 9 mm
3997 with a 1 mm margin. The layout of the PEBs will keep the same quadrant symmetry as has
3998 been described for the readout rows, but taking into account that a 10 cm path has to be left
3999 empty in the area close to the shortest readout row. This is to allow enough space for the
4000 connection of the cooling services, and in principle is only needed in two out of the four
4001 quadrants per disk.

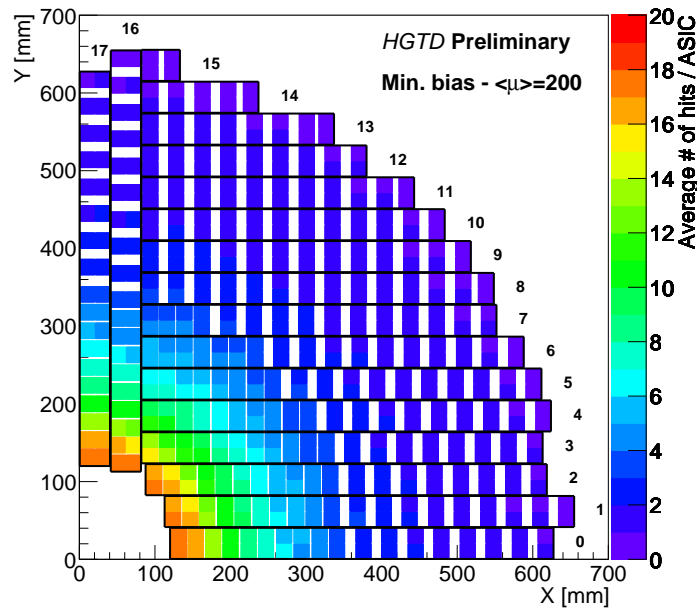


Figure 9.3: Average number of hits in each ASIC in one quadrant of the second layer in a simulated sample with $\langle \mu \rangle = 200$. The rectangles indicated by the black lines correspond to the readout rows. The numbering of the readout rows is also shown.

4002 9.1.3 Radiation tolerance

4003 All the active electronic components will be located at radii above 720 mm. Extrapolating
 4004 from Fig. 2.13, the maximum expected fluence which these components have to withstand
 4005 will be below $1 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ and the TID below 0.2 MGy.

4006 9.2 Data transfer

4007 The data transfer is, as mentioned above, based on the lpGBT ASIC. The data transfer
 4008 between the peripheral electronics and the off detector systems uses optical fibres based on
 4009 the VTRX+ optical transceiver developed within the Versatile Link Plus project.

4010 9.2.1 LpGBT

4011 A block diagram of the lpGBT is shown in Fig. 9.4 and more details concerning its specifica-
 4012 tions can be found in [56].

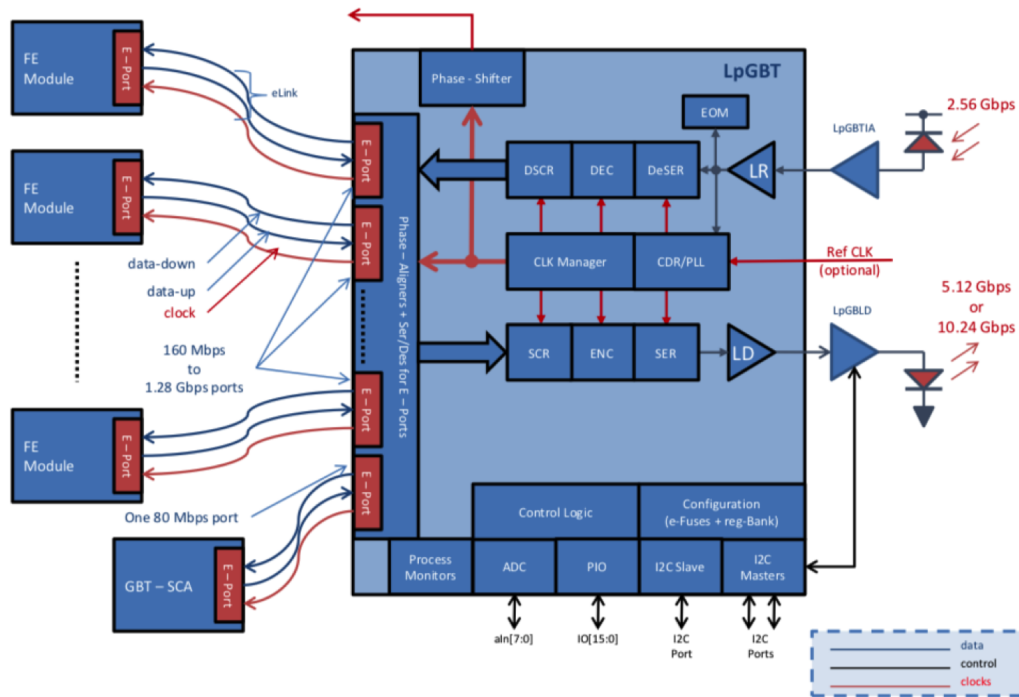


Figure 9.4: Block diagram of the lpGBT ASIC.

4013 The lpGBT ASIC is able to transmit data to an optical link at $10.24 \text{ Gbit s}^{-1}$. When using
 4014 FEC5 error correction code the bandwidth can be shared by 7 groups of 32 bit data received
 4015 on differential (CLPS) e-links. The 32 bits can come from one 1.28 Gbit s^{-1} , two 640 Mbit s^{-1} ,
 4016 or four 320 Mbit s^{-1} e-links. The phase aligner circuit for each input e-link of the lpGBT will
 4017 be used to ensure that the received data is sampled by the lpGBT at the optimal phase. This
 4018 allows data from flex cables with different lengths to be connected to the same lpGBT. The
 4019 total package length of the transmitted data, including headers, error correction codes, and
 4020 2 bits of internal and 2 bits external DCS data, is 256 bits at a rate of 40 MHz.

4021 Each lpGBT is able to receive four independent 320 Mbit s^{-1} bit streams encoded in the
 4022 2.56 Gbit s^{-1} , 64 bit frame, data from an optical link. Each package includes headers, FEC12
 4023 error correction bits as well as 2 bits internal and 2 bits external DCS data.

4024 The lpGBTs require configuration commands for setting up registers controlling their beha-
 4025 viour, e.g. bit rates and phase shift adjustment. This is normally done though their I²C bus
 4026 slave port, however to avoid external I²C bus cables, the lpGBTs receiving data via optical
 4027 links on each peripheral electronics board will be programmed by e-fuses to receive their
 4028 configuration via the 2.56 Gbit s^{-1} downlink bit stream. The same lpGBTs will in turn be
 4029 used via one of their I²C bus master ports to configure the lpGBTs for the luminosity system
 4030 of the same board.

- 4031 • **Fast commands and clock distribution.** Each data package received by the lpGBT via
4032 the optical links contains up to four independent 320 Mbit s^{-1} data streams. These can
4033 be mirrored to four different outputs of the lpGBT, allowing one lpGBT to control 16
4034 ALTIROC ASICs using 8 bit words. The 320 MHz clock required by the ALTIROC
4035 is extracted from the data streams by the lpGBT where also four independent clock
4036 streams are mirrored four times. Preliminary measurements done by the CERN
4037 lpGBT group show an excellent random component of the jitter (2.1 ps) but a sizeable
4038 deterministic part. The minimum number of lpGBTs required for the peripheral
4039 electronics is defined by the above limitation of not more than 16 ALTIROC ASICs
4040 connected to the same lpGBT.

- 4041 • **DAQ data.** The different e-link bit rates 1.28 Gbit s^{-1} , 640 Mbit s^{-1} , and 320 Mbit s^{-1}
4042 allow for an average number of hits bunch crossing per ALTIROC at $\langle \mu \rangle = 200$ of up
4043 to about 40, 20 and 10, respectively, at 1 MHz of event readout. Based on the expected
4044 average hit rate per ASIC (Fig. 9.3), 1.28 Gbit s^{-1} will only be used for ASICs placed at
4045 radii of less than 190 mm and 320 Mbit s^{-1} will be used for radii greater than 275 mm
4046 in order to keep the number of lpGBTs to a minimum, in view of the limited space
4047 available for the peripheral electronics.

- 4048 • **Luminosity data.** Each ALTIROC ASIC at radii larger than 320 mm provides 16-bit
4049 luminosity data for each bunch crossing, transmitted to the lpGBTs via the flex cables.
4050 Two 640 Mbit s^{-1} e-links are merged into a 32 bit lpGBT group, allowing 14 luminosity
4051 e-links to be connected to a single lpGBT for transmission to the off detector electronics
4052 via an optical link. In the baseline design, no downlink data are foreseen for these
4053 lpGBTs, which will be operated in simplex transmitter mode. The clock signal will
4054 instead be obtained as a 40 MHz clock from DAQ lpGBTs. Operation parameters and
4055 controls for the luminosity lpGBTs, e.g. phase adjustment delays, are set up via the
4056 I²C bus also from the DAQ lpGBTs.

- 4057 • **I²C bus.** Each lpGBT has three I²C bus masters and one slave. Only the master ports
4058 on the DAQ lpGBTs can be used since the luminosity lpGBTs do not receive optical
4059 downlink data. One I²C bus master will be connected to up to eight ALTIROC ASICs
4060 on four modules for DCS control. I²C-bus master ports are furthermore used, as
4061 previously mentioned, to configure all lpGBTs of the luminosity readout. Depending
4062 on the final optical link units, these may require I²C-bus connection for configuration
4063 of the laser drivers. Since the I²C-buses will only be used for configuration, traffic
4064 will be minimal during data taking limiting the risk of generating noise inside the
4065 ALTIROC ASICs.

4066 9.2.2 Optical links

4067 Each lpGBT connected to the DAQ system will need one up and one down optical link, while
4068 the lpGBT connected to the luminosity readout will only need an uplink to the off detector
4069 system. The VTRX+ optical transceivers under development within the Versatile Link plus
4070 project at CERN are specified to handle four fibres for transmission and one for receiving.
4071 The dimensions are specified as $20 \times 10 \text{ mm}^2$ footprint and 2.5 mm in height, which fits
4072 within the available 9 mm. The specified radiation hardness, 1 Mrad and $1 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$,
4073 exceed the required levels at radii greater than 85 cm, where they will be located. The VTRX+
4074 modules are pluggable via electrical connectors and are delivered with a pigtail ending in a
4075 12 fibre MT type optical connector.

4076 9.3 DC-DC converters

4077 The peripheral electronics will contain DC-DC converters supplying the required 1.2 V
4078 required by the ALTIROC ASICs and the lpGBT ASICs. The Versatile Link plus require 2.5 V
4079 for the laser driver and limited current at 1.2 V for the receiver. The DC-DC converters use
4080 the bpol12V ASIC developed at CERN. The bpol12V will be used as a single stage converter
4081 from the 10 V input to the 1.2 V output (or 2.5 V for the laser driver). The motivation for this
4082 choice is the limited available surface available for peripheral electronics. The footprint of
4083 one DC-DC converter circuit is assumed to be $12 \text{ mm} \times 30 \text{ mm}$ in the final layout, with a
4084 height of 5 mm including the necessary Faraday cage surrounding the inductance.

4085 The maximum output current of the bpol12V is 4 A, with an estimated efficiency of around
4086 65% at currents between 3 A and 4 A. When operating near the maximum current the input
4087 voltage should not exceed 10 V to reduce switching transients. The ASIC is designed for
4088 radiation tolerance up to 150 Mrad and $2 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$. The converters need a 460 μH in-
4089 ductance as well as further filtering components. The printed circuit board layout optimized
4090 by the bpol12V developer team will be used.

4091 The analogue and digital voltages are supplied separately to the ALTIROC ASICs. Each
4092 ALTIROC requires at most 0.5 W analogue and 0.7 W digital power. The two ASICs on the
4093 same module share supplies. The current consumption is dependent on the average number
4094 of hits within an ASIC and thus has a radial dependence. The power consumption of an
4095 lpGBT will not exceed 0.75 W. One DC-DC converter is able to supply voltage to the analog
4096 part of ALTIROCs for 4 modules, while a single converter can only supply the digital part of
4097 3 modules. Converters supplying the ALTIROC ASICs are switched on via general purpose
4098 I/O-lines from lpGBTs. A second I/O-line is used to report the converter status. The DC-DC
4099 converters supplying the PEB itself, i.e. the lpGBTs and the optical links are switched on
4100 by an external 1 V signal. The status of these converters is read out via external electric
4101 cables (open drain) and on I/O-lines on lpGBTs other than those they supply to allow to

4102 differentiate between possible power failures and lpGBT failures. Care has to be taken that
4103 the external electrical cables do not violate the grounding rules.

4104 The DC-DC converters supplying voltages to the ALTIROC ASICs are switched on by
4105 applying a voltage (at least 850 mV) which is generated via general purpose I/O-lines from
4106 DAQ lpGBTs. The status of the DC-DC converter is reported via an open drain Power Good
4107 output and monitored via lpGBTs.

4108 **9.4 Control and monitoring**

4109 The DCS control and monitoring of the front-end electronics, the monitoring of the sensor
4110 temperature and the delivered and received low voltage of the electronics is handled through
4111 the lpGBTs. The DCS information is embedded in the up and down bit streams of the optical
4112 connections at a rate of 80 Mbit s⁻¹. Two bits per data package at 40 MHz, in both directions,
4113 can be used for the general purpose I/O-port, ADC or I²C bus masters of the lpGBT.
4114 Depending on if used in transceiver or simplex mode, a further two bits may be available
4115 for I²C buses. Since, in the baseline option, the lpGBTs of the luminosity system will not
4116 have optical downlinks, only the lpGBTs connected to the DAQ system will be used for DCS
4117 handling.

4118 Each flex cable will, as described in Chap. 7, carry 5 voltages: 2 temperatures from sensor
4119 of each of the two ALTIROC; received analogue and digital supply voltages and analog
4120 current return voltage. Due to the resistance of the conductors on the flex cable, the latter
4121 three voltages serve to measure the current consumption and detect latch-up. Each lpGBT
4122 has an 8 input 10-bit multiplexed ADC allowing 1 mV resolution for a 1 V range. To handle
4123 all voltages to be measured, a 64:1 multiplexer Sec. 6.8 is used. Each multiplexer, which can
4124 switch the received voltages from up to 12 modules, is controlled by 6 I/O lines from an
4125 lpGBT.

4126 The peripheral electronics boards will each as mentioned above receive an external control
4127 signal to switch on the DC-DC converters supplying the lpGBTs and the optical links. The
4128 status of these converters is read out on I/O-lines on lpGBTs other than those they supply to
4129 allow to differentiate between possible power failures and lpGBT failures. Further I/O-lines
4130 on the DAQ lpGBTs are used for switching on and monitoring the status of the DC-DC
4131 converters supplying voltages to the ALTIROC ASICs.

4132 The I²C bus will be used to control and configure the ALTIROC ASICs as well as to configure
4133 the luminosity system lpGBTs and the DAQ lpGBTs that are not pre-fused to receive
4134 configuration via optical links.

4135 9.5 Connectors

4136 The limited space available for the peripheral electronics puts severe constraints on connect-
4137 ors. The PEB ground will be connected to the reference ground of the detector vessel.

- 4138 • The flex cables from a readout row will enter the peripheral electronics in two bundles
4139 of up to 10 cables. Each flex cable is 18 mm wide. Development work is ongoing, in
4140 collaboration with commercial companies, aiming at connectors with a height less
4141 than 3 mm, a width for two parallel cables less than 40 mm, and with a length of
4142 closely stacked connectors in the direction of the cable less than 6 mm per connector.
4143 The total occupied surface of these connectors on the peripheral boards will thus be
4144 40 mm × 60 mm or less per readout row, dependent on the number of modules per
4145 row.
- 4146 • Although D-sub type connectors (37 and/or 25 pins) are good candidates for connect-
4147 ing the high voltage, custom made connectors are being investigated. All modules
4148 have individual high voltage. The ground plane of each peripheral board defines the
4149 ground of the modules to which they are connected via the analog ground plane on
4150 the flex cables.
- 4151 • The peripheral boards will each require up to 2 cables with 10 V for the on-board
4152 DC-DC converters. Suitable connectors are under study.
- 4153 • The optical fibre pigtailed of the VTRx end in a 12 fibre MT-type connector to which the
4154 patch cables of the fibre feed-throughs at the detector vessel have to be connected.

4155 9.6 Peripheral boards

4156 9.6.1 Layout considerations

4157 The peripheral electronics will be split up in five peripheral boards (PEB) per quadrant with
4158 a similar number of sensor modules connected per board. This is achieved by letting the
4159 readout rows 0–2, 3–5, 6–9, 10–14, and 15–17 be combined into one board each. (The read
4160 out row numbering is shown in Fig. 9.3). The combination allows to reduce the number of
4161 lpGBT ASICs, multiplexers and VTRx used and to better use the available surface area at the
4162 outer radius of the disks.

4163 The number of modules, e-links for DAQ at different transfer rates and luminosity e-links per
4164 peripheral boards are shown in Tab. 9.1. The bit rates of the DAQ e-links will be re-optimised
4165 for the final layout based on further simulations, making full use of the available lpGBT
4166 capacity.

4167 A number of considerations have to be taken into account for the actual PEB design.

Peripheral board	Read out rows	Nb of modules	1.28 Gb/s	640 Mb/s	320 Mb/s	Luminosity
Front						
0	0-2	56	24	24	64	56
1	3-5	56	8	32	72	66
2	6-9	59	0	2	116	110
3	10-14	45	0	0	90	90
4	15-17	39	16	16	46	42
Back						
0	0-2	55	20	26	64	56
1	3-5	54	6	34	68	62
2	6-9	56	0	4	108	104
3	10-14	42	0	0	84	84
4	15-17	37	12	18	44	38

Table 9.1: Number of module readout e-links at different rates for the different peripheral boards of the front and back layers. The tables also show which readout rows are connected to which board and the number of modules per board.

- 4168 • To limit the implications of possible failing components, care must be taken such that
4169 as few modules as possible and preferably not in a contiguous area will be affected.
4170 This requires that modules sharing the same DC-DC converters are not contiguous.
4171 They have to share the same lpGBT for readout which will also transmit their clock
4172 and fast commands, control their DC-DC supply as well as DCS control via I²C bus
4173 and handle the module voltage monitoring. Optimised schemes for this exist and will
4174 be implemented.
- 4175 • For the same reason, modules sharing the same readout lpGBTs should as far as
4176 possible share luminosity lpGBTs who will receive their configuration control and
4177 clock from the readout lpGBT.
- 4178 • The power dissipation of the peripheral electronics is used to preheat the CO₂ cooling
4179 requiring a suitable radial arrangement of the DC-DC converters and lpGBTs.
- 4180 • The distance between lpGBTs and VTRx must be kept small to minimize the risk of bit
4181 error at the high data rate.

4182 9.6.2 Layout

4183 Combining multiple readout rows on the same PEB, as described above, the required number
4184 of different components is worked out per PEB as shown in Tab. 9.2. The number of DC-

4185 DC converters is relying on the real power consumption of the final ALTIROC still under
4186 design.

4187 With the number of components from the table, assuming 20 mm × 10 mm for the VTRx,
4188 20 mm × 20 mm for the lpGBT including the necessary decoupling capacitors, 15 mm × 15 mm
4189 for the multiplexers, and 12 mm × 30 mm for the DC-DC converters, the total surface for the
4190 components of the front layer of a quadrant is 1030 cm². The available board surface per
4191 quadrant, removing the 15° for the cooling manifold and 450 cm² for flex cable, HV, and LV
4192 connectors, is 2270 cm² giving sufficient surface for the active components.

4193 The full electric layout of the PEBs has yet to be performed. Fig. 9.5 shows only a conceptual
4194 design demonstrating the complexity of these boards due to the component density and
4195 limited area and thickness available. This conceptual design does not take into account any
4196 PCB design rules yet and can only be considered as an illustration of the final PEB design.
4197 A functional prototype of the PEBs is scheduled in the demonstrator program in 2020 in
4198 parallel with a complete PCB design of a real size PEB.

Peripheral board	lpGBTs DAQ	lpGBTs Luminosity	DC-DC converters	MUX	VTRx
Front					
0	8	4	36	5	8
1	8	5	37	5	8
2	8	8	39	5	8
3	7	7	31	4	7
4	6	3	26	4	6
Back					
0	8	4	36	5	8
1	8	5	36	5	8
2	8	8	37	5	8
3	6	6	31	4	6
4	6	3	26	3	6

Table 9.2: Numbers of lpGBTs, analog multiplexers, VTRx, and DC-DC converters for the different Peripheral Electronics Boards.

4199 9.7 Power dissipation

4200 The dominant source of the power dissipation on the PEBs is the power loss in the DC-DC
4201 converters. With an average power consumption of 1.12 W per ALTIROC ASIC (Fig. 11.2),
4202 the total required power delivered to the ASICs from the DC-DC converters will be 4.5 kW
4203 per double sided layer. At 65% efficiency, the power loss due to the ASIC supplies will be

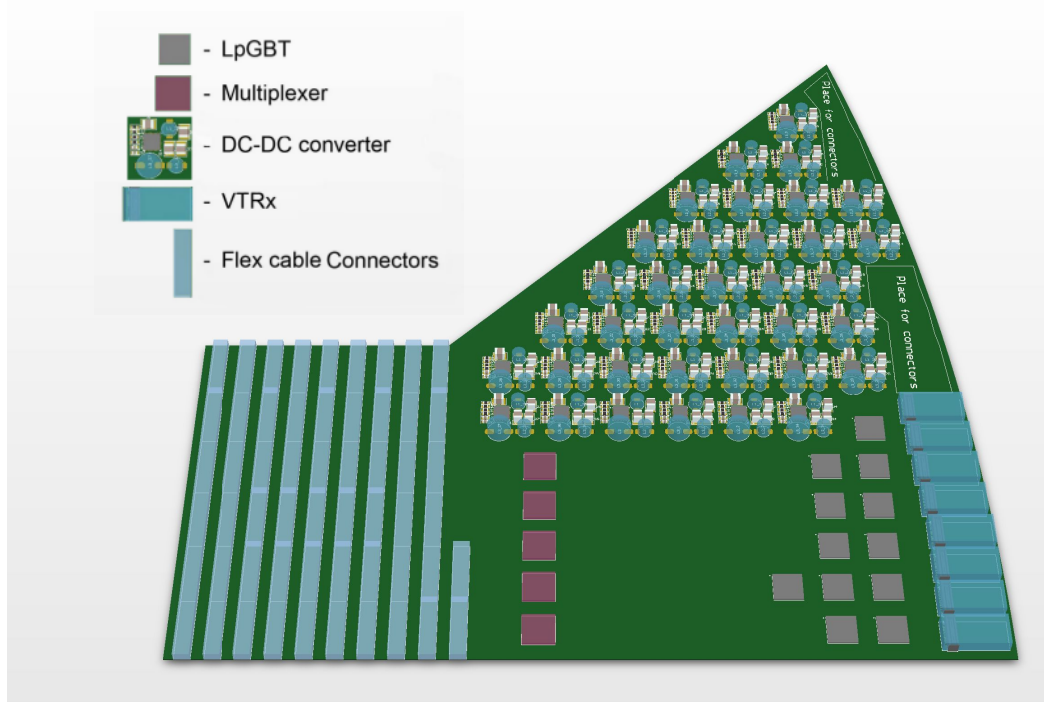


Figure 9.5: Conceptual design of a PEB 0 for the front layer. The left blue rectangle are the flex connectors (56). The pink squares are the analog multiplexer ASICs (5). The grey squares are the lpGBT used for the luminosity and data read-out. The right blue rectangles are the VL+ components (8). Finally at the top are located the 36 DC-DC converters.

4204 1.6 kW per double sided layer. Including an estimated power consumption of 600 W per
 4205 double sided layer for the lpGBTs and VTRx, the total power dissipation of the peripheral
 4206 electronics will be 2.2 kW per double sided layer. The total power dissipation will be 4.4 kW
 4207 per endcap.

4208 10 DAQ and Luminosity Measurement

4209 10.1 DAQ interface

4210 The HGTD data acquisition system will be embedded in the ATLAS TDAQ common read-
4211 out [31]. The proposed HGTD architecture is shown in Fig. 10.1 and can be divided in
4212 two main blocks: on-detector electronics located in the experimental hall and off-detector
4213 electronics located in the USA15 counting room. The on-detector electronics consist of
4214 ALTIROC modules connected via flex cable to the Peripheral Electronics Board, as described
4215 in Chap. 9.

4216 The interface between on-detector and off-detector electronics is performed via optical link
4217 using lpGBT chip set and VTRx+ optoelectronics, which provides different data paths for
4218 Timing, Trigger and Control (TTC), DAQ and DCS. Two optical links with different purpose
4219 data streams are proposed: the main data stream that provides ToT and ToA information per
4220 triggered event and the luminosity stream that contains bunch-by-bunch hit information for
4221 luminosity measurements. The main data stream is used for the propagation of clock, fast
4222 commands and configuration to the modules, as well as the data information for the ATLAS
4223 event processor. The luminosity stream only send hit information through the uplink and
4224 will be described in Sec. 10.3.7.

4225 10.1.1 Off-detector electronics

4226 The off-detector electronics is based on the general-purpose FELIX system [57], which is
4227 the main interface between the off-detector back-end and the on-detector electronics. The
4228 proposed back-end architecture is shown in Fig. 10.2. FELIX receives event data from the
4229 on-detector electronics and transmit them to the Data Handler via multi gigabit network. In
4230 addition, FELIX interfaces to the TTC system via Local Trigger Interface (LTI) and to DCS
4231 for control, configuration and monitoring.

4232 The FELIX downlink will follow lpGBT encoding, which is composed of 64-bits frames that
4233 are transmitted at every LHC bunch crossing period with a data rate of 2.56 Gbit s^{-1} . The
4234 clock is propagated to the lpGBT and thus to the modules by sampling the data stream. The
4235 downlink frame has different fields for data (fast commands), internal and external config-
4236 uration meant for lpGBT, module and DCS handling. The uplink will also follow lpGBT

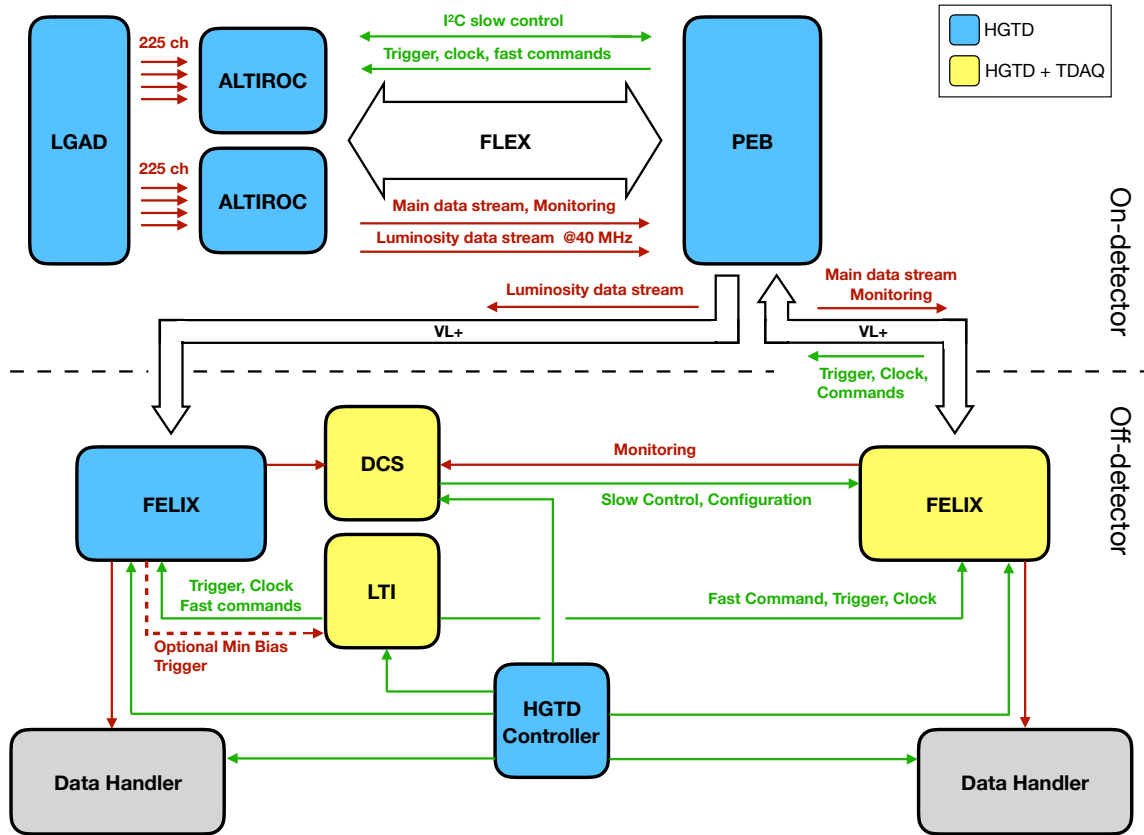


Figure 10.1: Data transmission paths among the ALTIROC, Peripheral Electronics Board (PEB), and DAQ components for hit data, luminosity data, clock, fast commands, and DCS/slow controls.

4237 encoding with a data rate of $10.24 \text{ Gbit s}^{-1}$, the different frame fields for data, configuration
 4238 and DCS will be decoded in the FELIX board. Upstream the data will be forwarded to the
 4239 Data Handler using multi-gigabit network. In addition, monitoring information, like errors
 4240 and timing will be computed in FELIX and a prescaled sample of the events will be send to
 4241 the monitoring unit.

4242 The Data Handler will receive data from FELIX via multi-gigabit network. It will decode
 4243 HGTD specific information providing event building and monitoring within a common
 4244 TDAQ infrastructure. Afterwards, the data will then be send to the Dataflow system for
 4245 further processing by the Event Filter. In addition, the Data Handler will also receive trigger
 4246 information via FELIX for monitoring and automatic recoveries. On the other hand, a
 4247 software application called HGTD Controller, running in a dedicated computer will be used
 4248 to manage the module and IpGBT configuration. The Controller will be also used to manage
 4249 HGTD calibrations via dedicated software that will be described in the following section.

4250 Requirements on the number of FELIX boards are set by the number of optical links and it is

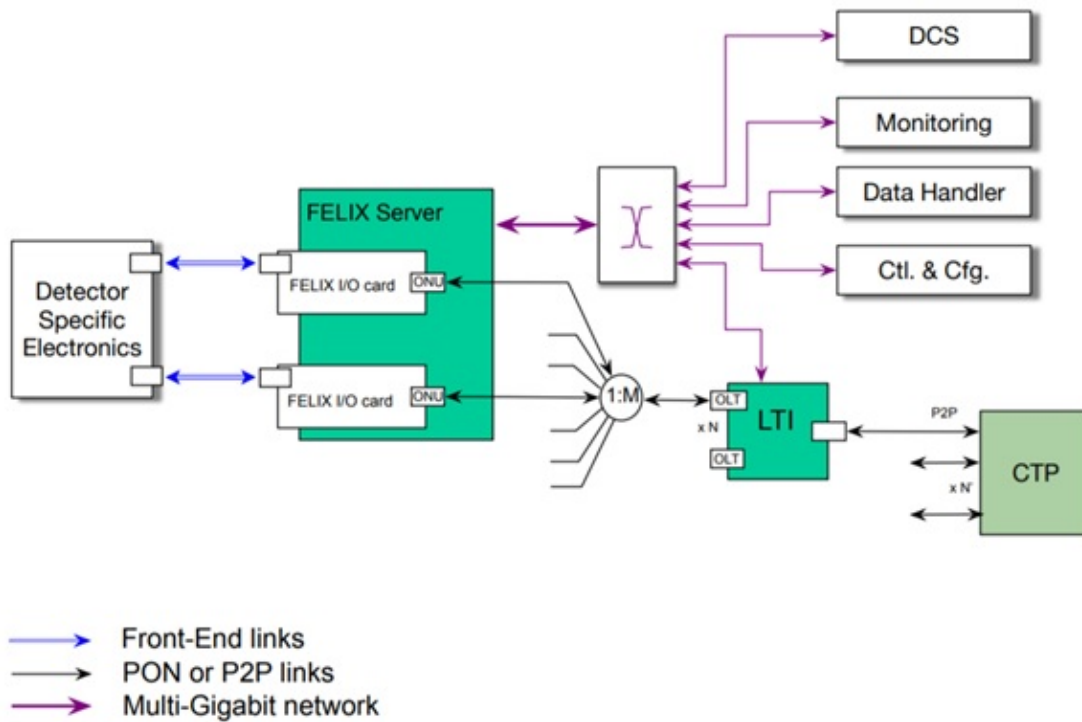


Figure 10.2: Proposed off-detector back-end architecture for Phase II. Plot taken from [31].

4251 driven by the HGTD layout. Current estimates call for a total of 48 FELIX I/O cards and 48
 4252 Data Handlers for the main data stream. The luminosity back-end electronics will require 32
 4253 FELIX I/O cards.

4254 10.1.2 Calibration and timing

4255 Regular calibrations will be performed in HGTD in which different parameters like ToA
 4256 and ToT will be monitored and tuned. A dedicated HGTD software running on the HGTD
 4257 Controller will be used for this purpose. It will consist on different nested loop with a specific
 4258 module configuration followed by a charge injection and a trigger command with a proper
 4259 timing. The HGTD Controller will interface with FELIX for the handling of the module
 4260 configuration and generation of a particular bitstream for the fast commands. Downstream
 4261 the event data will be processed and stored in histograms for a further analysis and may
 4262 be used as input inside the nested loop for tuning purposes. The implementation of an
 4263 histogramming unit inside FELIX in order to speed up the calibration procedure will be
 4264 studied.

4265 On the other hand, a precise timing of the modules is critical for operation, for this purpose
4266 a dedicated timing calibration will be performed. In a first stage the detector will be timed
4267 using standard calibration procedure, it will consist on different charge injections while
4268 looping over coarse and fine delays DAC values of the TDC. Afterwards, the delay values
4269 which set the 7-bit ToA in the middle range will be selected. In a second stage, the detector
4270 will be timed during stable beams by using dedicated LHC fills with single colliding bunches
4271 (similar to 3b fills during Run2 used for Pixel and SCT timing scans), alternatively the same
4272 operation could be done during standard fills with a dedicated trigger stream with single
4273 colliding bunches. During these fills different delays values of the TDC will be scanned. The
4274 data will be analyzed offline and the delays that ensure the proper timing will be selected.
4275 Further timing corrections taking into account clock jitter variations will be described in the
4276 following section.

4277 **10.2 Timing correction**

4278 Despite regular calibrations and timing of the detector, dynamic and static contributions to
4279 the clock has to be taken into account and will be studied in this section. The master clock
4280 will be distributed to the lpGBT downlinks and then to the individual ALTIROC readout
4281 chips, in which a clock tree will be used to distribute the clock as uniformly as possible.
4282 Any temporal or spatial variation in the time discriminator may compromise the ultimate
4283 resolution of the detector unless it is understood and controlled.

4284 The sensors themselves will have a resolution as good as 30 ps per hit, as described in
4285 Chap. 5. The contributions to the time resolution from the on-detector electronics (UX15)
4286 and from the clock distribution (USA15) has to be smaller than this. For instance, the clock
4287 dispersion for HGTD should be less than 10 ps across a wide range of frequencies and
4288 over the detector acceptance. Static contributions to the timing resolution, i.e. those fixed
4289 by geometry or varying on time scales longer than a run, include the time-of-flight and
4290 detector alignment; the propagation times to distribute the clock to each ASIC as a whole;
4291 and uneven clock propagation paths within an ASIC to each TDC. Dynamic contributions,
4292 like the variation of the clock with time, can occur through a variety of mechanisms across a
4293 wide range of frequencies, including high-frequency jitter, noise in the flex cables, and low-
4294 frequency day/night temperature changes. These effects must be monitored and calibrated
4295 to minimise static and dynamic contributions to the timing measurements. In the case of
4296 dynamic contributions, sufficient data may not be recorded to calibrate away fast effects,
4297 and therefore in this section we study how to determine the timing correction in real-time
4298 using all of the data flowing through the FELIX for each trigger accept.

4299 For relativistic particles produced in an LHC collision, the time-of-arrival distribution of
4300 each measurement channel will consist of a Gaussian core derived from the time dispersion
4301 of the LHC collisions convolved with the combined hit time resolution of the sensor and

4302 electronics, as shown in Fig. 10.4. The mean of the distribution encodes information on the
 4303 relationships between the global LHC clock on arrival to ATLAS, the mean LHC collision
 4304 time for a given bunch, and the reference clock phase at a given TDC. This mean shifts from
 4305 zero through the cumulative effects of time-of-flight, clock propagation delays, and dynamic
 4306 shifts of the clock phase during data-taking. Assuming that the relationship between the
 4307 clock at the TDC and the LHC clock is stable within a given time interval, data collected
 4308 during the interval can be used to sample the t_{hit} distribution and estimate its mean, t_0 . This
 4309 mean can then be used to correct the cumulative time offset of each channel individually.

4310 Assuming a trigger rate of 1 MHz and after 250 ms of data collection, the t_0 can be measured
 4311 with a precision of 5 ps for a single channel at 150 mm radius. If t_0 is calculated on a per-
 4312 ALTIROC level, combining the hits of up to 225 channels, the same precision can be reached
 4313 in 2 ms. Integration times are shown in Tab. 10.1.

Radius [mm]	150	250	350	450	550
$\sigma(t_0)$ after $T_{\text{int}} = 100$ ms for 1 channel	8 ps	12 ps	20 ps	29 ps	44 ps
$\sigma(t_0)$ after $T_{\text{int}} = 100$ ms for 15×15 channels	0.6 ps	1.0 ps	1.7 ps	2.6 ps	4.2 ps
T_{int} required for $\sigma(t_0) < 5$ ps for 15×15 channels	2 ms	5 ms	13 ms	38 ms	92 ms

Table 10.1: Precision of the t_0 determination, $\sigma(t_0)$, vs integration time T_{int}

4314 10.2.1 Sources of clock jitter

4315 The data path from the ALTIROC up to the DAQ is shown in Fig. 10.1 and described
 4316 in Sec. 10.1. Different contributions to the clock jitter are expected in the readout system:

- 4317 1. Front-end electronics: the clock distribution within the ALTIROC to each TDC will
 4318 be shifted due to path-length differences and possible internal jitter. A conservative
 4319 random Gaussian-distributed 5 ps jitter is included to account for jitter in the ALTIROC.
- 4320 2. FLEX cable: it is made of Kapton and copper could pick up noise from the environment
 4321 and might have some inherent time jitter performance. A random Gaussian-distributed
 4322 5 ps jitter is included to account for jitter in the FLEX.
- 4323 3. lpGBT: a preliminary measurement of the lpGBT clock performance in [58] indicated
 4324 that a large-non Gaussian deterministic time jitter might be expected for the lpGBT.
 4325 However, any front-end chip with a phase-locked loop can filter this effect to a small
 4326 2.2 ps jitter. Both of these scenarios are included in the t_0 calibrations study, and are
 4327 shown in Fig. 10.3.
- 4328 4. FELIX: the clock jitter from the FELIX system will depend on the final chips used and
 4329 bandwidth filtering applied, as studied in [59]. A conservative 5.2 ps jitter is added to
 4330 represent the worst jitter expected for the FELIX.

4331 Additional sources of timing jitter and t_0 variation are expected to affect the t_{hit} measurement
 4332 and are included in this study. The LHC radio frequency systems which compensate the
 4333 beam loading and maintain bucket stability result in a periodic collision point time shift
 4334 in the ATLAS Detector. The variation in the average time of collision with bunch number
 4335 was studied in [60], and the expected bunch crossing time offset for the ATLAS detector is
 4336 included as a bunch-dependent variation. The collision time is expected to shift by a few
 4337 ps per bunch, but can be corrected to a jitter in the order of 5 ps. Finally, a time-of-flight
 4338 variation is added as a static radially-dependent offset from 0 to 70 ps as a function of sensor
 4339 radius.

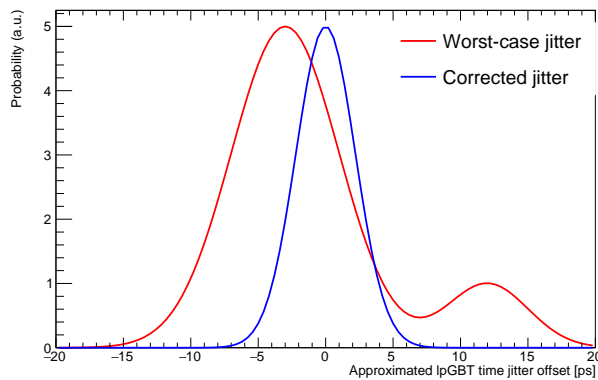


Figure 10.3: Timing jitter distribution assumed for the lpGBT in the corrected (blue) and uncorrected (red) scenarios. These distributions are approximations of the timing jitter expected in the lpGBT.

4340 Random event-by-event fluctuations cannot be calibrated away, although they are included
 4341 as part of the hit time resolution. Instead, the performance of the timing correction procedure
 4342 will depend on how many longer-term variations (heat cycles or other effects) affect the time
 4343 measurement, which are largely unknown. For the purpose of this study, these unknown
 4344 longer-term variations are parameterised as a sinusoidally varying 100 ps time offset with
 4345 variable period.

4346 10.2.2 Timing correction procedure

4347 The hit time offset t_0 is calculated at regular intervals as the arithmetic mean of the t_{hit}
 4348 distribution. The length of the time interval strongly affects the performance of the timing
 4349 correction. The t_0 can be calculated to better precision with averaging over a longer time, but
 4350 shorter times can correct for faster variations. The average timing can be computed online in
 4351 the FELIX in order to collect as much data as possible, and then applied offline.

4352 The hit time distribution before and after the timing correction is shown in Fig. 10.4. Fig. 10.5
 4353 shows the timing performance as a function of the integration time and the variation period

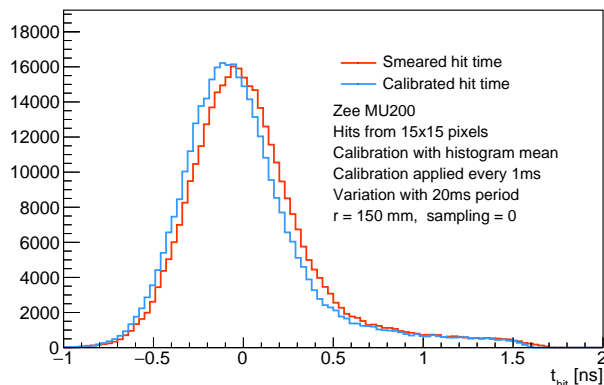


Figure 10.4: Hit time distribution before (red) and after (blue) the timing correction procedure. Non-Gaussian tails arise from late particles, backscatter, and other effects. The hit time distribution is obtained from the HGTD simulation described in Sec. 3.1.

4354 for channels at three different radii, calculating the t_0 correction from a 15×15 grid of
 4355 channels, and including all of the sources of time variation discussed above, including the
 4356 sinusoidally varying 100 ps offset with period plotted along the x -axis. Smaller calibration
 4357 window sizes can reduce t_0 jitter when shorter-term variations affect the hit time. However,
 4358 longer calibration windows, which can collect more statistics and therefore more precisely
 4359 determine t_0 , result in a better hit time correction. Variations with period smaller than 1 ms
 4360 cannot be corrected with this procedure because of insufficient statistics, and variations with
 4361 period greater than 20 ms can be corrected in all regions of the detector. The timing correction
 4362 procedure should also work well for longer-term variations on the scale of 1×10^5 s (1 day).

4363 The procedure outlined above and the corresponding results are a preliminary plan for the
 4364 timing correction scheme using conservative values of clock jitter contributions. Conservat-
 4365 ive estimates for the expected ALTIROC and FLEX timing jitter were used, and the study
 4366 will be updated when final numbers are available. When accounting for the expected jitter
 4367 from components of the readout system and LHC bunch crossing time drift, the clock jitter
 4368 of approximately 10 ps can be reached, in accord with the specifications outlined in Sec. 4.2.2.
 4369 If additional unknown sources of jitter are included, the timing correction procedure can
 4370 reduce the total jitter to 20 ps for the time variations studied. In general, more accurate
 4371 corrections can be calculated to correct for longer-term variations, and should result in
 4372 smaller total clock jitter.

4373 The timing correction procedure assumes that time offsets across different channels are not
 4374 correlated. However, the time offsets in each channel are expected to be somewhat correlated
 4375 from both global (i.e. offsets in the LHC collision time and the ATLAS clock) and local effects
 4376 (i.e. tree structure of the clock distribution creates correlations between modules of the same
 4377 branch), the timing correction procedure assumes the worst-case scenario of no correlation

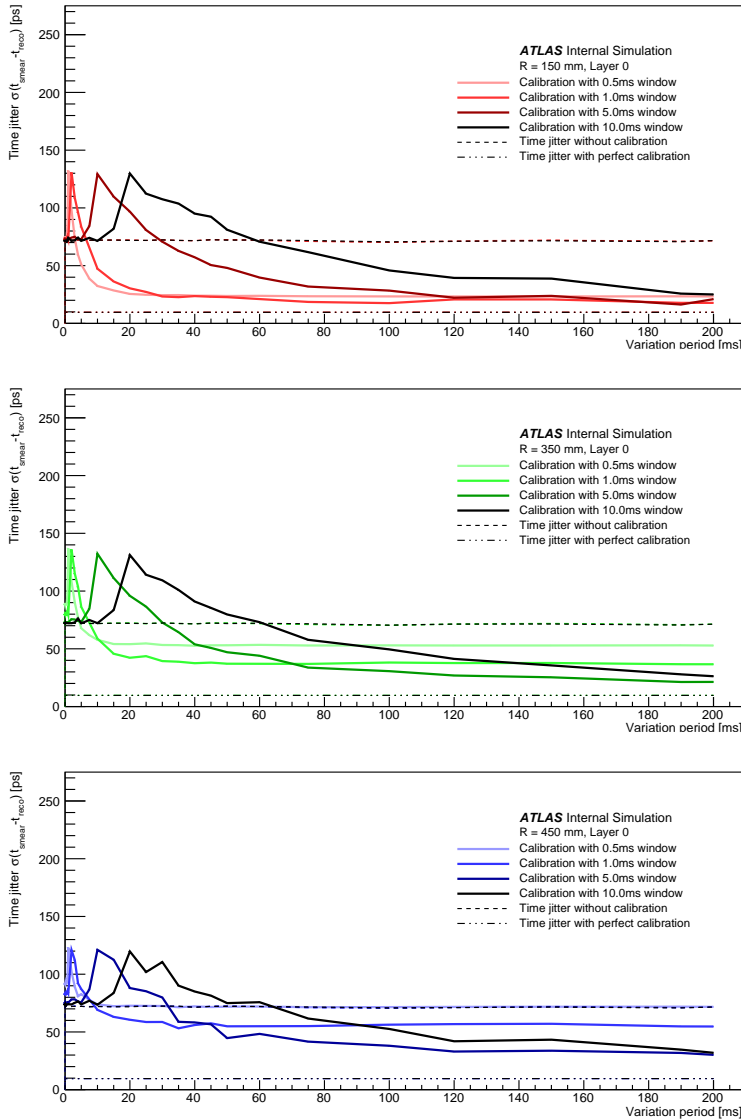


Figure 10.5: Hit time resolution $t_{\text{smear}} - t_{\text{reco}}$ after the timing correction procedure as a function of the variation period, and for several different choices of calibration window time, shown for $r = 150$ mm, $R = 350$ mm, and $r = 450$ mm. t_{reco} is the hit time taken from simulation and includes inherent hit time resolution effects from the sensor and electronics and the collision time spread. The t_{smear} term adds additional sources of time jitter from the ASIC, FELIX, flex cable, lpGBT, and ATLAS collision time drift, with an additional sinusoidally varying 100 ps offset of variable period. The time jitter without any correction applied is shown as the dashed line, and the time jitter without any long-term timing variation effects is shown as the dotted-dashed line. For a variation period of greater than 20 ms, and with the right choice of calibration window size, the calibration procedure will always improve the t_0 precision.

4378 and applies corrections per-ASIC level. Timing corrections targeting global or more broadly
4379 correlated effects can combine hits from more channels, achieving more statistical precision
4380 and a better correction across even shorter timescales. Furthermore, the t_0 jitter at the ASIC
4381 level can be corrected on a per-channel basis by using the hit times of single pixels, although
4382 a factor of 225 would be lost in statistics.

4383 10.3 Luminosity

4384 The measurement of the integrated luminosity delivered by the LHC is critical for almost all
4385 physics analyses, as discussed in Sec. 3.3.2.

4386 Any luminosity detector (luminometer) attempts to measure some observable which is
4387 assumed to be proportional to the instantaneous luminosity, or equivalently, to the average
4388 number of inelastic interactions per bunch crossing $\langle\mu\rangle$. Conceptually simple examples
4389 are the average number of charged-particle tracks reconstructed in the inner tracker [17]
4390 or the noise-corrected number of clusters in the pixel detector [61]. In the early years of
4391 LHC operation, many luminometers used the so-called *event-counting* method [62], also
4392 known as *zero counting*, which exploits Poisson statistics to infer the pile-up parameter μ
4393 from the fraction of bunch crossings in which no interaction was detected. As the mean
4394 μ of the Poisson distribution increases, the fraction of bunch crossings with no detected
4395 interaction decreases, and eventually reaches zero. The μ value at which this saturation, or
4396 “zero starvation”, occurs depends on the geometrical acceptance and the efficiency of the
4397 luminometer considered. Already in LHC Run 2, the baseline ATLAS luminometer [18] was
4398 forced to exploit its 16-channel granularity to switch from event counting to hit counting.
4399 This latter method [62] applies a Poisson formalism very similar to that of event counting, to
4400 extract μ from the average number of detector hits recorded per bunch crossing; the finer the
4401 granularity of the luminometer, and the smaller the acceptance of its individual channels, the
4402 higher the pile-up value at which the method eventually saturates. In the limit of a very large
4403 number of channels, as is the case in a pixelated detector such as the HGTD, the per-channel
4404 occupancy becomes small enough for the Poisson non-linearity to become almost negligible.
4405 The average number of hits in randomly selected colliding-bunch crossings then depends
4406 linearly on the luminosity (except perhaps at the highest μ values expected at the HL-LHC,
4407 where the hit-counting Poisson formalism may need to be invoked again).

4408 The primary calibration technique to determine the absolute luminosity scale of a bunch-by-
4409 bunch luminometer employs dedicated van der Meer (vdM) scans [17] to infer the delivered
4410 luminosity at one point in time from the measured parameters (primarily the intensity and
4411 the transverse area) of the colliding bunches. The conversion factor from luminometer
4412 counting rate to measured luminosity is then determined by comparing the luminosity
4413 computed from the above-mentioned accelerator parameters to the visible, uncalibrated
4414 interaction rate reported by the luminometer at the peak of the beam-separation scans. The

4415 beam conditions during vdM scans are different from those in normal physics operation,
4416 with lower bunch intensities and only a few tens of widely spaced bunches circulating.
4417 These conditions, which are optimized to reduce various systematic uncertainties in the
4418 calibration procedure [63], typically result in a pile-up parameter μ of about 0.5 at the peak
4419 of the scans, and as low as $\mu \sim 2 \times 10^{-5}$ in the tails of the scans, where the beams are barely
4420 overlapping. Since the same luminosity-calibration procedure is foreseen at the HL-LHC,
4421 the luminometer response will have to remain linear over more than six orders of magnitude
4422 in μ , from vdM conditions ($\mu \sim 2 \times 10^{-5}$ to $\mu \sim 0.5$) up to high-luminosity physics data
4423 taking at an $\langle \mu \rangle$ of around 200.

4424 The online and offline environments impose different, and sometimes conflicting, constraints
4425 on the luminometers and the associated luminosity-determination methods, with processing
4426 speed being of the essence during data taking (possibly at the expense of absolute accuracy),
4427 and offline luminosity requiring the best possible precision on much longer time scales. For
4428 instance, track counting [17], which proved essential to control the dominant luminosity
4429 uncertainties in both LHC Runs 1 and 2, can only be used offline as it requires a dedicated,
4430 randomly triggered event stream that must be subjected to extensive offline analysis before
4431 usable luminosity values can be provided.

4432 Bunch-by-bunch luminosity estimates are required not only for offline physics analysis, but
4433 also in the online environment, for instance to apply bunch- and μ -dependent corrections to
4434 calorimeter data in the high-level trigger algorithms; to optimize the trigger menus on the
4435 fly; and to monitor, analyze and improve the accelerator performance over the long term. An
4436 additional requirement is the availability of a bunch-integrated, fast and reasonably accurate
4437 luminosity measurement, provided at ~ 1 Hz as input to the collision-optimization and
4438 luminosity-leveling accelerator protocols.

4439 As discussed further in Sec. 10.3.5, the precision of the offline determination of the integrated
4440 luminosity has so far been limited not by statistics, but by systematic uncertainties. An
4441 essential lesson from LHC Runs 1 and 2 is that the dominant systematic uncertainties can
4442 only be determined, or at least constrained, by confronting the response of a redundant set
4443 of luminometers, each based on a different technology, with complementary capabilities and
4444 independent instrumental biases.

4445 10.3.1 HGTD as a luminometer

4446 As a fast high-granularity detector in the forward region, the HGTD provides unique
4447 capabilities for measuring the luminosity at the HL-LHC. The idea for using HGTD as a
4448 luminometer is straightforward: the occupancy will be linearly correlated with the number
4449 of interactions (i.e. the luminosity). The high granularity gives a low occupancy, and
4450 therefore excellent linearity between the average number of hits and the average number
4451 of simultaneous pp interactions over the full range of luminosity expected at the HL-LHC,

4452 as discussed in Sec. 10.3.2. With detector signal durations in the few-ns range, the charged-
4453 particle multiplicities within the acceptance can be determined accurately for each individual
4454 bunch crossing separately. With the occupancy information sent at 40 MHz, i.e. for every
4455 bunch crossing independent of the ATLAS trigger (further discussed in Sec. 10.3.6), the
4456 HGTD will provide both online and offline unbiased per-BCID luminosity measurements.
4457 The measurement is made in a reduced $|\eta|$ range, and in this proposal the plan is to read
4458 out the ASICs for sensors at $470 \text{ mm} < r < 640 \text{ mm}$ (equivalent to $2.4 < |\eta| < 3.5$) for
4459 the luminosity determination. The HGTD is designed to have capabilities to constrain
4460 many systematic uncertainties by itself, with the goal of reducing the total uncertainty on the
4461 integrated luminosity in HL-LHC compared to Run 2 despite the much harsher experimental
4462 conditions, as is discussed in Sec. 10.3.3 and Sec. 10.3.5.

4463 10.3.2 Linearity of the luminosity determination

4464 For the $|\eta|$ range proposed above, the average number of hits per double-sided layer and
4465 per inelastic pp collision is 44.6, and approximately 7% of these collisions result in 0 hits.
4466 Fig. 10.6(a) shows the average number of hits per bunch crossing registered in the first
4467 double-sided HGTD layer (both sides of the innermost cooling plate) as a function of the
4468 number of simultaneous inelastic pp interactions. The black points at μ of 1 and around
4469 175–225 are determined from fully simulated minimum-bias events with $\mu = 1$ and $\langle\mu\rangle$
4470 in the range 190–210, respectively. The green stars represent samples where several $\mu = 1$
4471 minimum-bias events have been overlaid to produce samples with intermediate numbers of
4472 interactions, while making sure not to double-count multiple hits in the same channel. A
4473 linear fit to the points in the hatched region at low and intermediate μ values is extrapolated
4474 to the $\mu \sim 200$ region where its prediction can be compared to the hit multiplicities extracted
4475 from fully simulated high-pile-up samples. The small discrepancy between the extrapolated
4476 linear fit and the simulated points in the bottom left frame around $\mu \approx 200$ is attributed to
4477 multiple particles hitting the same pad; it could easily be corrected by applying the Poisson
4478 hit-counting formalism. Also discuss doing a correction for double-hits, and show result for
4479 that. Describe plot 10.5b, and what it implies for the precision that is achieved for the vDM
4480 scan.

4481 10.3.3 Noise and afterglow subtraction

4482 The HGTD is affected by three distinct background contributions to the luminosity sig-
4483 nal: single-beam backgrounds, instrumental noise, and afterglow, in order of increasing
4484 importance.

4485 Single-beam background arises from activity correlated with the passage of a single beam
4486 through the detector. This activity is caused by shower debris from beam-halo particles,
4487 that impinge on the luminosity detectors in time with the circulating bunch. Although

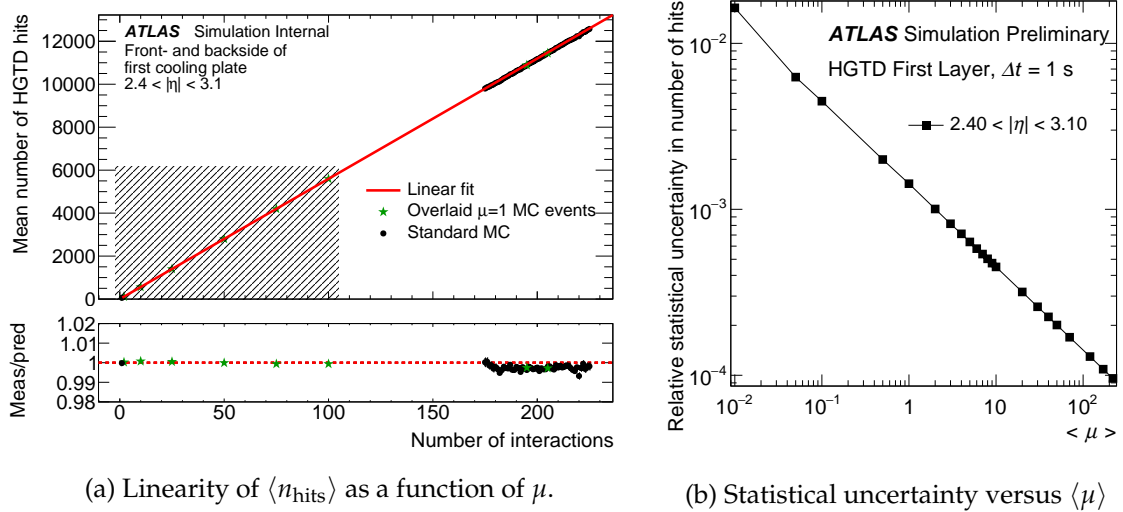


Figure 10.6: Left: mean number of HGTD hits per bunch crossing as a function of the pile-up parameter μ . **Be more rigorous about calling this "number of interactions"?** The straight line is a linear fit to the low pile-up points, and is compared to the simulated data in the bottom panel (see text). Right: pile-up dependence of the statistical uncertainty affecting the simulated hit counts shown in the left panel. **Both of these plots need to be updated using the new samples. Add that the right plot is per BCID.**

4488 its impact remains to be simulated, single-beam background is expected to be close to
 4489 negligible (on the scale of the luminosity signal), based not only on experience with phase-1
 4490 luminometers, but also on HGTD-specific features: on the incoming-beam side, not only
 4491 should the shielding provided by the endcap calorimeter absorb all of the high-radius
 4492 backgrounds (except for a few muons), but the surviving background particles will be
 4493 out-of-time by several nanoseconds wrt the collision products traveling back from the IP.
 4494 Residual HGTD backgrounds on the outgoing-beam side, if any, can be roughly estimated
 4495 from a few non-colliding bunches injected in each ring for this specific purpose, as was
 4496 frequently done during LHC Runs 1 and 2.

4497 Instrumental noise can arise from thermal noise in detector electronics, or from high-rate
 4498 contributions from "noisy pixels" (such as caused by radiation-induced "single-event upsets").
 4499 Thermal-noise (and, up to a point, noisy-pixel) contributions can be subtracted by the same
 4500 method as that used for afterglow, which is discussed in the next paragraph. Alternatively,
 4501 noisy pixels can be masked, if only to prevent excessive dataflow rates (in which case their
 4502 unavailability will have to be accounted for when normalizing the measured hit counts).
 4503 **Discuss here (or elsewhere) the possibility to flag according to measurements of the sideband**
 4504 **rates.**

4505 As detailed in Ref. [62], all Run-2 bunch-by-bunch luminometers (with the exception of

4506 track counting) observe some activity in the BCIDs immediately following a collision,
4507 which in later BCIDs decays to a baseline value with several different time constants. This
4508 afterglow is attributed to slow particles (such as neutrons) and to delayed decays (e.g.
4509 from stopped muons), that originate from the hadronic cascades initiated by pp collision
4510 products. **Add that for CMS also electronic signal residuals?** For a given bunch pattern,
4511 the afterglow level is observed to be proportional to the luminosity in the colliding-bunch
4512 slots. Its magnitude relative to the luminosity signal, and its time structure, both depend
4513 on the spectral sensitivity of the luminometer considered (and therefore on the detector
4514 technology it uses), as well as on the location and the physical environment (geometry,
4515 chemical composition of neighbouring equipment) in which this luminometer operates. The
4516 magnitude of the afterglow contamination observed in Runs 1 and 2 varies widely, from
4517 10^{-4} for LUCID in vdM scans, to 0.2-0.4% for BCM in high- μ bunch trains; it can be as high
4518 as 10% in pixel detectors during routine physics running, therefore requiring a delicate
4519 correction that contributes sizeably to the total luminosity uncertainty.

4520 The time resolution of the HGTD is a unique capability that is essential to mitigate the large
4521 impact of instrumental noise and afterglow that is intrinsic to the pixel-cluster counting
4522 technique. As described in Sec. 6.1, and illustrated in Fig. 6.2, the ASIC will send occupancy
4523 information in two different time windows:

- 4524 • a *central time window*, 3.125 ns wide, centred on the nominal bunch crossing time;
- 4525 • a *sideband window*, nominally covering 3.125 ns before the central time window and
4526 3.125 ns after the central time window.

4527 This double-sideband window will be programmable. Here it has been chosen symmetric,
4528 such that its occupancy provides, after appropriate scaling, an estimate of the noise and
4529 afterglow contributions as interpolated under the luminosity signal in the central time
4530 window, separately for each BCID. This ability to perform an in-situ measurement of the
4531 noise and afterglow level for each bunch crossing, using data from empty RF buckets just
4532 before and after the filled bucket within the same nominally filled 25-ns bunch slot, is a
4533 unique capability of the HGTD compared to other luminometers.

4534 10.3.4 Statistical precision of the luminosity determination

4535 **Should come before the previous section. Add the vdM discussion.** To confirm that statistical
4536 uncertainties are small for the online luminosity measurements, the size of the uncertainty
4537 has been studied as a function of the duration of the averaging period and $\langle\mu\rangle$. The average
4538 number of hits per bunch crossing is simulated using a toy Monte-Carlo method with inputs
4539 extracted from fully simulated samples. For each value of $\langle\mu\rangle$, a random number of pp
4540 interactions is drawn from a Poisson distribution with a mean equal to $\langle\mu\rangle$. For each pp
4541 interaction, a number of HGTD hits is then generated randomly based on the distribution
4542 of hits per pp interaction extracted from full-simulation samples. By repeating this process

4543 11 000 times (for the number of turns the LHC beams will make) and averaging the number
 4544 of hits, the statistical precision achieved in each individual BCID during 1 s of LHC running
 4545 is emulated. Fig. 10.6(b) shows the relative uncertainty expected from statistical fluctuations
 4546 as a function of $\langle\mu\rangle$ using this method. The coverage of $|\eta| < 3.5$ presented here gives a
 4547 statistical uncertainty of 0.14% at $\langle\mu\rangle = 1$ and 1.6% at $\langle\mu\rangle = 0.01$. For measurements in the
 4548 low- μ regime (e.g. during van der Meer scans) better precision can be achieved through a
 4549 longer averaging time. **Have the plot extend to even lower mu if possible.**

4550 10.3.5 Systematic uncertainties affecting the luminosity determination

4551 A detailed discussion of the systematic uncertainties affecting the 2012 luminosity determin-
 4552 ation at $\sqrt{s} = 8$ TeV is presented in Ref. [17]; the sources and the magnitude of the luminosity
 4553 uncertainties in LHC Run 2 at $\sqrt{s} = 13$ TeV are comparable. Of the dominant uncertainties,
 4554 two are luminometer-specific (rather than related to, for instance, beam conditions or accel-
 4555 erator instrumentation): the time stability of the luminometer response, and the calibration
 4556 transfer.

4557 The time stability of relative-luminosity measurements is potentially affected by different
 4558 sources, depending on the time scale considered.

- 4559 • *Long-term stability* refers to potential drifts of the luminometer response on the time
 4560 scale of days to months, compared to its response at the time of the vdM-calibration
 4561 session. Such drifts have been seen to arise, for instance, from gain fluctuations in, or
 4562 flux-induced ageing of, LUCID photomultipliers (PMTs); darkening of TILE scintil-
 4563 lators; cumulative radiation damage to inner-tracker silicon-strip or pixel modules;
 4564 or unaccounted-for dead or inefficient channels. In LHC Runs 1 and 2, this class of
 4565 effects contributed from 0.5% to 1.3% to the systematic luminosity uncertainty, a large
 4566 number compared to the luminosity-precision goal of 1% at the HL-LHC.
- 4567 • *In-run stability* refers to variations in luminometer response on the time scale of one
 4568 ATLAS run (a few hours). The reference ATLAS luminometers (BCM in most of Run 1,
 4569 LUCID in Run 2) proved mostly immune to such drifts. In contrast, inner-tracker based
 4570 luminosity measurements, such as track- or pixel-cluster-counting, were significantly
 4571 more sensitive, typically because of unaccounted-for changes in either tracker condi-
 4572 tions (noisy or misbehaving modules), or in effective coverage (disabled modules).
 4573 Because the luminosity, and therefore the pile-up parameter μ , typically decays during
 4574 an LHC fill, such drifts are difficult to disentangle from a genuine μ -dependence of
 4575 the detector response. It is therefore essential, especially for pixel-counting methods,
 4576 to keep track of variations in both the number and the radial location of misbehaving
 4577 channels on the time scale of a few minutes: for instance, a few noisy pixels that
 4578 suddenly start firing at a high rate may bias the luminosity measurement and prove

4579 hard to correct for after the fact.

4580

4581 The *calibration-transfer uncertainty*, which in LHC Run 2 typically amounted to a 1.0–1.5%
4582 uncertainty on the absolute luminosity scale, refers to how precisely one controls potential
4583 shifts in detector response, that occur between the beam conditions of vdM scans ($\langle\mu\rangle \sim 0.5$,
4584 a few ten low-intensity isolated bunches, no bunch trains) and those of physics data-taking
4585 ($\langle\mu\rangle \sim 200$, hundreds to thousands of high-intensity bunches grouped in trains with diverse
4586 patterns). Such shifts can arise, for instance, from rate-dependent effects in inner-tracker
4587 solid-state sensors or LUCID photomultipliers; from bunch–pattern-dependent “out-of-time
4588 electronic pileup” (in which the electrical signal from a given 25 ns bunch slot leaks into the
4589 following bunch slot); in the case of track counting, from a residual pile-up dependence of
4590 the tracking efficiency; or, in randomly triggered readouts of inner-tracker luminosity data,
4591 from subtle deadtime effects through which a higher-luminosity bunch can shadow a small
4592 fraction of the triggers in the immediately following bunch slot. All of these effects have
4593 been observed at some level in Run 2, and required μ - and time-dependent corrections to
4594 the luminosity scale that could exceed 10% during high-luminosity operation.

4595 The HGTD has several characteristics that will aid in constraining, and hopefully reducing,
4596 such systematic uncertainties. To better monitor the time stability, the region instrumented
4597 with the luminosity readout will be segmented into 16 sub-regions, with 4 divisions in η
4598 and 4 divisions in ϕ , as shown in Fig. 10.7. Each region has sufficient statistical sensitivity
4599 to determine the luminosity independently of the other regions. **Check specifically if we
4600 can calibrate a single region in the vdM scan. Also maybe revise the radial divisions, at
4601 least look at how it can be made with the new layout.** Regions at different η will accrue
4602 radiation damage at a different rate, therefore comparing their response can help determine
4603 the degradation due to radiation. The partitioning of the regions can be controlled in the
4604 luminosity back-end electronics firmware, so that a different optimisation than the one
4605 described here can be accommodated.

4606 While such internal consistency checks will undoubtedly prove valuable, they are unlikely
4607 to be sufficient, if only because any bias or drift that is correlated across all 16 regions
4608 remains undetectable by the HGTD alone. Experience at LHC has repeatedly shown that
4609 independent checks based on several luminometers using different technologies are essential
4610 to controlling the systematic uncertainties to the level suggested by the physics program.

4611

4612 Built into the HGTD design are several features that are expected to reduce the magnitude
4613 of the calibration-transfer correction (if any), as well as help constrain the associated uncer-
4614 tainties:

- 4615 • the pixel-cluster counting technique is intrinsically linear, and no μ -dependent cor-
4616 rections are expected to be necessary (except possibly at the very highest bunch

- 4617 luminosities expected at HL-LHC, where the well-established Poisson hit-counting
4618 formalism may have to be invoked);
- 4619 • for a given bunch pattern, the most likely reasons for the hit count to deviate from
4620 strict proportionality to the true luminosity are afterglow and instrumental noise. The
4621 exquisite time resolution of the HGTD, combined with the methodology outlined in
4622 Sec. 10.3.3, provides a unique strategy to control these effects to the level needed;
 - 4623 • the most likely reason for a bunch-pattern dependence of the HGTD hit count is
4624 again the afterglow, the magnitude of which is sensitive to the length of, and the
4625 separation between, bunch trains. The above-mentioned afterglow subtraction at the
4626 bunch-by-bunch level should eliminate this potential bias;
 - 4627 • electronic out-of-time pileup from one BCID to the next is presumably eliminated by
4628 the extremely short pulse duration of HGTD pixels;
 - 4629 • eliminating deadtime effects associated with large μ variations from one BCID to the
4630 next, is one of the motivations for the trigger-less, 40 MHz readout of the luminosity
4631 information discussed in Sec. 10.3.6).

4632 10.3.6 Occupancy readout at 40 MHz

4633 Experience with luminosity determination at the LHC shows that the capability to read out
4634 a luminometer at 40 MHz, i.e. on every single bunch crossing, is critical to its function as
4635 an independent device that must provide bunch-by-bunch (bbb) luminosity measurements,
4636 with the best possible precision both online and offline. In LHC Runs 1 and 2, this require-
4637 ment was satisfied only by LUCID and BCM; the fact that it was out of reach for track and
4638 pixel-cluster counting methods proved a significant limitation to the final precision of the
4639 integrated luminosity in both ATLAS and CMS in Run 2.

4640 In view of the more exacting luminosity-precision requirements of the HL-LHC physics
4641 program, the 40 MHz readout of the occupancy is key to a full exploitation of the HGTD
4642 potential as a stand-alone, high-precision luminometer for both online and offline use. This
4643 becomes apparent when one considers

- 4644 • the TDAQ implications of a readout triggered by sampling randomly selected colliding-
4645 bunch pairs,
- 4646 • some of the requirements associated with the van der Meer calibration,
- 4647 • use cases of bunch-by-bunch luminosity measurements in both the online and the
4648 offline environment, and
- 4649 • some features specific to the HGTD-based luminosity determination.

4650 Reading out every single bunch crossing avoids any potential trigger-induced bias, which can
4651 occur even with so-called “random” triggers; it also ensures that the maximum possible per-
4652 bunch statistics are available for luminosity determination. If the luminosity measurement
4653 were to be carried out using a detector which is not read out on every bunch crossing, the
4654 following considerations would have to be addressed.

- 4655 • The luminosity must be determined from an unbiased sampling of collisions, therefore
4656 data passing physics triggers cannot be used. Such triggers normally require a lot
4657 of activity in the detector, and the presence of e.g. high momentum leptons or jets.
4658 They are typically sensitive to pile-up effects, and therefore not representative of the
4659 luminosity; they also are severely statistics-limited.
- 4660 • The traditional method for overcoming the trigger bias is to use a dedicated random
4661 trigger, sampling each bunch crossing evenly. The bandwidth for such a trigger comes
4662 at the expense of that available for physics, thus effectively representing a loss in
4663 data-taking efficiency.
- 4664 • A random trigger does not result in a completely unbiased dataset for the luminosity
4665 determination. There is a shadowing effect from the standard trigger deadtime, in
4666 which more luminous bunches shadow collisions in subsequent, less luminous bunch
4667 slots. The associated corrections are unlikely to be negligible (they proved noticeable
4668 in the 2017 track-counting based luminosity measurements), and presumably too
4669 complex to be carried out in the luminosity back-end electronics boards.
- 4670 • Even if the luminosity extraction could be performed online using the luminosity
4671 back-end electronics to analyze Level-0 triggered data, it would reduce the available
4672 statistics by several orders of magnitude: this would make the HGTD inadequate as
4673 an online luminometer, as further argued below.
- 4674 • If the luminosity-extraction analysis can only be carried out offline, the data has
4675 to be saved to disk, further degrading the statistics usable for luminosity-related
4676 applications.

4677 The vdM calibration technique requires evaluating, as a function of the transverse beam
4678 separation, the four-dimensional integral (over x , y , z and time) of the proton-density
4679 distributions in each colliding-bunch pair. Since the proton population and the transverse-
4680 density distributions vary significantly from one bunch to the next, fitting a vdM scan curve
4681 obtained by summing the interaction rate over all colliding-bunch pairs (rather than fitting
4682 a separate scan curve for each pair) would result in unpredictable and non-reproducible
4683 biases to the absolute luminosity scale. This fundamental requirement, on its own, implies
4684 that the HGTD must provide statistically precise bunch-by-bunch luminosity measurements
4685 over the full μ range covered during a vdM scan (2×10^{-5} to 0.5).

4686 The above span in interaction rate, combined with the LHC bunch-revolution frequency of
4687 approximately 11 kHz and with a typical integration time of 60–100 s during individual vdM-

4688 scan steps, implies that a readout based on randomly triggered colliding-bunch crossings
4689 would be unable to even approach the necessary rate capability. Triggering the HGTD
4690 readout during the vdM scan using some kind of independent track- or hit-multiplicity
4691 trigger is not an option, since the *absolute* efficiency of such a trigger cannot be determined
4692 to the necessary precision of a few per mille (except by comparison with another, already
4693 absolutely-calibrated luminosity monitor). This chain of arguments explains why, during
4694 LHC Run 2, no direct vdM calibration of track- nor pixel-cluster-counting algorithms was
4695 ever attempted by ATLAS. These inner-tracker-based luminosity algorithms, which relied
4696 on a random trigger, suffered from such low statistics during vdM scans that they required
4697 a couple of hours of data-taking at $\mu \sim 0.5$ during the vdM session, in order to be cross-
4698 calibrated to LUCID instead, thereby making their absolute calibration fully correlated with
4699 that of LUCID.

4700 During routine physics running, the need – both online and offline – for a 40 MHz readout
4701 of the luminosity data is fundamentally related to the intrinsic variation of the instantaneous
4702 luminosity across the colliding-bunch pairs. These bunch-to-bunch variations are driven by
4703 fluctuations in both bunch intensity and emittance; during Run 2, they sometimes exceeded
4704 20–30%.

4705 In the present ATLAS online-luminosity architecture, bunch-by-bunch luminosity measure-
4706 ments provide the basic input to the computation of the bunch-integrated luminosity value,
4707 that is used, for instance, to select the most appropriate ATLAS trigger settings; inform
4708 the online monitoring tools of various ATLAS subdetectors; optimize collisions; control
4709 the luminosity-leveling protocols; monitor accelerator performance, etc. Depending on the
4710 application considered, the required refresh times vary from one to a few ten seconds.

4711 In addition to a precise bunch-integrated measurement, bunch-by-bunch measurements that
4712 are statistically stable ($\ll 0.5\%$) and reasonably accurate on an absolute scale, are required
4713 online for several purposes, such as:

- 4714 • providing μ -dependent corrections to liquid-argon-based triggers, with a refresh rate
4715 of a few minutes;
- 4716 • supplying the accelerator with diagnostics such as bunch-by-bunch specific-luminosity
4717 values, which offer a better estimate of the beam-averaged emittance than state-of-the-
4718 art accelerator instrumentation. Such diagnostics have proven essential to the steady
4719 improvement of LHC performance. They are needed not only during physics running
4720 for detailed analysis of accelerator operation, but also in real time with refresh rates of a
4721 few seconds for periodic emittance scans, as well as for some accelerator-development
4722 sessions, during which the beam parameters are tailored on a bunch-by-bunch basis
4723 and the required refresh rates are at the few-seconds level.

4724 Use cases for bunch-by-bunch measurements in the offline environment include, for in-
4725 stance:

- 4726 • computing the bunch-integrated luminosity eventually used in physics analyses from
4727 the sum of per-bunch luminosity values, after recalibration and application of bunch-
4728 dependent corrections, such as residual μ -dependence or afterglow subtraction;
- 4729 • refined μ - (and therefore bunch-) dependent corrections to the cell-by-cell energy
4730 measurements in the liquid argon calorimeter;
- 4731 • bunch-by-bunch comparisons of the relative consistency of the luminosity values across
4732 multiple luminometers. Such studies have revealed significant μ - and bunch-position
4733 dependent biases in all the bunch-by-bunch luminometers available in Run 2, and
4734 again demonstrated that confronting independent luminometers is a key ingredient to
4735 precision luminosity measurements.

4736 Finally, the afterglow-subtraction capability detailed in Sec. 10.3.3 and the potential use of the
4737 occupancy information in the Level-0 trigger outlined in Sec. 10.3.10 are entirely dependent
4738 on the availability of dedicated occupancy data at 40 MHz.

4739 10.3.7 Luminosity back-end electronics

4740 **Add something about the occupancy data flow?** For every bunch crossing of the LHC, each
4741 ASIC in the region $2.4 < |\eta| < 3.5$ will send occupancy counts in the central time window
4742 and in the sideband time window. These counts are encoded into 7 and 5 bits, respectively.
4743 In addition 4 bits are used for encoding, using the 6b8b encoding scheme, resulting in 16 bit
4744 sent per ASIC for every bunch crossing. Thus there is a steady data rate of 40 MHz times
4745 16 bits, or 640 Mbit s^{-1} , from each ASIC. The luminosity data is sent via lpGBTs dedicated to
4746 the luminosity readout to the back-end electronics, requiring 212 lpGBTs for each of the four
4747 disks of the HGTD, i.e. 848 links for the whole detector. The data sent by the lpGBTs are
4748 collected by the luminosity back-end electronics, consisting of dedicated FELIX units. These
4749 units are separate from the FELIX units handling the timing data. Each FELIX card can take
4750 up to 24 input fibres at 10 Gbit s^{-1} . With an integer number of FELIX units per HGTD disk,
4751 this results in 20 FELIX units needed for the luminosity data (5 per double-sided layer).

4752 **Update this, after investigating the feasibility of storing per-ASIC information** The lumin-
4753 osity back-end electronics aggregates the central time window data and the sideband data
4754 separately, for each of the 16 regions described in Fig. 10.7. This reduces the massive data rate
4755 from the ASICs to only two integers for each of the 16 regions, on each of the HGTD double-
4756 sided layers (128 integers in total), separately for each of the 3600 BCIDs. **Use exact number.**
4757 The FELIX will store these sums in registers in the FPGAs, and update them continuously
4758 with the new data for every bunch crossing. **Likely true, but investigate possibility to transfer**
4759 **the data to the server processor.** These sums are the raw data needed for determining the
4760 luminosity, which is only needed with a frequency of about once per second. Assuming
4761 that the luminosity data gets pushed out of the FELIX at a rate of 10 times per second, and
4762 using 64 bits to encode each of the integers, the total data rate out of the luminosity back-end

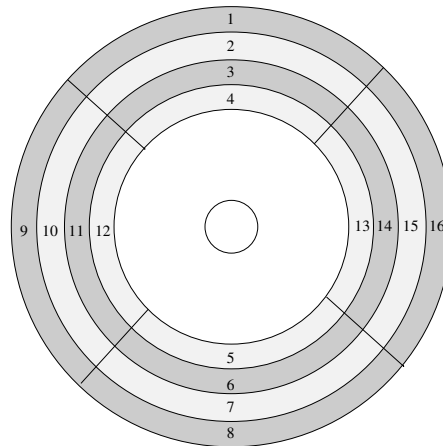


Figure 10.7: Sketch of the partitioning of the sensors into 16 regions for the luminosity determination. Each of the regions can be used to determine the luminosity independently of the others. Regions at different radius will be subject to different levels of radiation over time.

4763 electronics is only $128 \times 64 \times 3600 = 29.49 \text{ Mbit s}^{-1}$, or 3.69 MB s^{-1} . Thus, the luminosity
 4764 data represents a negligible strain on the network downstream of the back-end electronics,
 4765 and the data flow is independent of the trigger. The conversion from the occupancy sums
 4766 to a calibrated luminosity will happen in dedicated software algorithms. These software
 4767 algorithms can run on any downstream computer, most likely in the Data Handler. **Discuss**
 4768 **if something is done specifically for the offline luminosity.**

4769 10.3.8 Per-event luminosity information stored in the ATLAS raw data

4770 In the processing of the luminosity data by the back-end electronics, the per-event informa-
 4771 tion is lost when the data is aggregated. To allow for per-event occupancy data to be stored
 4772 in the raw data for events passing all the stages of the trigger, the luminosity back-end
 4773 electronics have to implement a buffer to store the data for each of the 16 regions per disk
 4774 for each event separately, until a L0 trigger accept is received and the corresponding occu-
 4775 pancy information can be sent. Whether this capability will be implemented, and per-event
 4776 occupancy information will be recorded in the ATLAS raw data, has not yet been decided.
 4777 The per-event occupancy in the central time window provides no unique information over
 4778 what can be calculated from the HGTD precision timing data, it would merely serve as
 4779 validation of the luminosity and precision timing data. The information about the occupancy
 4780 in the sideband time window however does provide unique information compared to the
 4781 HGTD timing data, and could have use cases in e.g. searches for new, slow-moving particles.
 4782 Provided that the capability to buffer per-event luminosity data is implemented in the
 4783 luminosity back-end electronics, the payload to be stored in the ATLAS raw data would be

4784 occupancy counts for each of the 16 regions per disk encoded as 4-byte integers, in total 512 B
4785 per event. Calculate the implication of the amount of memory needed for the pipeline.

4786 10.3.9 Operation in non-Stable Beams conditions

4787 As with other silicon sensor-based detectors close to the LHC beamline, the HGTD will only
4788 ramp up the full High Voltage on the sensors once Stable Beams have been declared, in order
4789 to avoid destroying the detector in case of catastrophic beam losses. At the same time, there
4790 is a need from the accelerator operations perspective to have an estimate of the luminosity at
4791 the ATLAS interaction point in conditions where Stable Beams have not been declared. This
4792 situation occurs at the start of every physics run, and can also be necessary during periods
4793 of machine commissioning.

4794 Providing an online luminosity estimate in non-Stable Beams operation reinforces the need
4795 for ATLAS to have several different luminometers at the HL-LHC, employing different
4796 detector technologies. Less precise, but more radiation tolerant, detectors could then be the
4797 primary sources of luminosity measurements by ATLAS when Stable Beams have not been
4798 declared. Whether a safe operation mode can be found for the HGTD during non-Stable
4799 Beams conditions is still to be investigated. A possibility of operating just the top and
4800 bottom luminosity regions (regions 1 and 8 in Fig. 10.7) at a reduced HV setting could be
4801 safe. The reduced HV setting would result in a lower hit efficiency, and thus a different
4802 relationship between the instantaneous luminosity and the average number of hits expected
4803 in the HGTD, compared to operating at nominal HV conditions. A separate calibration
4804 of the luminosity determination for such a operating mode can be accommodated in the
4805 luminosity back-end electronics. Whether a safe operating mode of the detector in non-Stable
4806 Beams conditions can be found will require extensive tests of the sensors and possibly also
4807 operating experience with the full detector.

4808 10.3.10 Minimum-bias trigger at Level-0

4809 The data made available at 40 MHz for the luminosity measurements can also be used by
4810 the L0 trigger to record minimum-bias events under low- μ data-taking conditions. Such
4811 data-taking conditions are expected during e.g. heavy-ion runs, van der Meer scans or
4812 for runs dedicated to soft-QCD measurements. The HGTD will be installed where the
4813 current MBTS detector is located. The MBTS detector has been used extensively for these
4814 purposes during Run-1 and Run-2, e.g. during the heavy-ion runs where it played a crucial
4815 role. However, the MBTS will not survive at the HL-LHC. With improvements of several
4816 orders of magnitude in both granularity and time resolution, the HGTD can provide all
4817 the functionality of the MBTS. The number of hits in the time window centred around the
4818 nominal collision time provides good separation between empty bunch crossings and those
4819 with pp collisions. A simple threshold for the minimum number of hits using the occupancy

4820 information is straightforward to implement in the luminosity back-end electronics. Such a
4821 binary trigger decision can then be communicated directly to the central trigger. The latency
4822 for reaching the Level-0 global trigger processors in time for a decision is not expected to be
4823 a problem.

4824 **10.4 Detector Control System**

4825 This section covers the Detector Control System (DCS).

4826 In order to ensure the coherent and safe operation of the HGTD, a Detector Control System
4827 (DCS) will be put in place. The main tasks of the DCS are to bring the detector in any
4828 desired operational state, to monitor its operational parameters and to signal any abnormal
4829 behaviour, thus allowing manual or automatic corrective actions. The DCS provides a
4830 homogeneous interface between the operator and the detector and its infrastructure, enabling
4831 tasks such as detector calibration, commissioning and operation.

4832 The DCS elements are distributed over various detector components: front-end electronics,
4833 services, back-end electronics and DCS servers. A Finite State Machine (FSM) structure
4834 will be implemented and integrated in the ATLAS FSM tree during data taking, and will
4835 allow to operate in stand alone mode during commissioning and maintenance. Real-time
4836 monitoring of critical parameters will be implemented, and alerts will be raised as soon as
4837 critical conditions are reached or connection to one or more hardware devices is lost. All
4838 relevant DCS parameters will be archived for debugging, performance tuning and offline
4839 studies.

4840 The DCS will control and monitor the following parameters: the power, both high- and
4841 low voltages, supplied to the detector; temperatures of the detector modules, peripheral
4842 electronics and cooling; humidity and overpressure inside the vessel.

4843 **10.4.1 High Voltage**

4844 The high voltage supply system, that will be purchased, must include the hardware and
4845 software components for being connected to the DCS for control and monitoring of both
4846 voltage and current. For the HV supplies the behaviour at the selected current limit is
4847 preferably programmable to not only be in trip mode but also for current limiting operation.
4848 The supplies will be based on commercial multi-channel rack mounted units located in the
4849 service cavern.

4850 As detailed in Sec. 8.1, in order to compensate for the damage due to radiation, the bias
4851 voltage will be raised. The read-out thresholds of the discriminators in the front-end
4852 electronics must also be adjusted accordingly. This operation will be performed by scripts

4853 implemented in the DCS. Monitoring the leakage current and the TOT as an indicator of the
4854 collected charge will give a good estimate of the sensor gain evolution during data taking,
4855 allowing to perform the necessary HV adjustments. The scripts will take the estimate into
4856 account and calculate the optimal bias voltage and the read-out thresholds according to an
4857 optimized algorithm.

4858 **10.4.2 Low Voltage**

4859 The bulk 300 V supplies as well as the 300 V to 10 V DC-DC converters are assumed to be
4860 commercial products. Both of them must provide provisions for communication with DCS
4861 allowing for control and monitoring of voltage and current. The voltages from the DC-DC
4862 converters on the peripheral boards and the voltages received at the front-end ASICs are
4863 monitored via multiplexers and ADC channels on the lpGBT ASICs of the peripheral boards,
4864 as described in Sec. 9.4. From the lpGBTs the information is sent via optical fibres to FELIX
4865 boards of the DAQ system for transmission to the DCS system. The optical links to the
4866 lpGBTs from the DAQ FELIX boards will exchange data bits, embedded in the data streams,
4867 for switching on and monitoring the status of the DC-DC converters powering the front-end
4868 ASICs. However, several DC-DC converters per peripheral boards must be controlled by
4869 DCS over wires, as they will power the lpGBTs, which will control the rest of the DC-DC
4870 converters on the board.

4871 **10.4.3 Temperatures**

4872 The temperatures of the sensor modules are monitored as voltages from temperature sensors,
4873 embedded in each ALTIROC front-end ASIC, via the same multiplexers and ADCs that
4874 used for the modules voltage monitoring. The temperature at the peripheral boards will be
4875 monitored through temperature sensors inside the lpGBTs. Information on temperatures
4876 inside the detector vessel, when the peripheral electronics is not powered, the DCS will
4877 be obtained from two sources: by means of temperature sensors located on the cooling
4878 plates, directly connected to off-detector ELMB++ units installed in the PP; and from the
4879 Interlock system, which will monitor the NTC sensors installed on the detector modules, as
4880 it is described further below in Interlock section.

4881 **10.4.4 Pressure and humidity**

4882 To keep a dry atmosphere inside the detector volume, an overpressure of the flushing N₂ gas
4883 must be maintained at all times. It is important to monitor the humidity inside the vessel and
4884 the pressure difference between the vessel volume and the UX15 cavern atmosphere. The
4885 overpressure monitoring can be implemented using pressure difference sensors, which can

4886 be located in the USA15 cavern and connected to the detector volume and the environment
4887 via two rigid pipes keeping the sensors away from high radiation areas.

4888 Radiation hardness is an issue for humidity sensors. Studies are needed to select appropriate
4889 radiation tolerant sensors. One option would be sensors based on optical fibres, that are
4890 being developed in ATLAS for ITk.

4891 **10.4.5 Configuring**

4892 Configuration of the front-end electronics and the lpGBT ASICs are in a similar manner
4893 controlled via commands, either I²C-bus or direct, embedded in the DAQ data stream
4894 between FELIX and lpGBT boards. Sec. 6.8 and Sec. 9.4 give more details on the control and
4895 monitoring of the ASICs and peripheral electronics respectively.

4896 **10.4.6 DCS software**

4897 The HGTD DCS structure is shown in Fig. 10.8. The DCS software will run on a local
4898 control station (LCS) in the ATLAS service cavern USA15. All DCS operations will be
4899 performed from this server. The DCS project will be integrated in the global ATLAS DCS. At
4900 a higher level, the ATLAS Global Control Station (GCS) controls all sub-detectors, collects
4901 data from external systems interfaced to the ATLAS DCS, such as the LHC collider status
4902 information or the Detector Safety System, and sends the data to sub-detectors via dedicated
4903 DCS Information Servers (IS).

4904 A Finite State Machine (FSM) structure will be implemented with rules for performing actions
4905 on the detector modules, the front-end and the back-end electronics and the infrastructure,
4906 while states will be propagated to the appropriate upper nodes. The DCS software consists
4907 of three layers. The lower layer establishes communication with different hardware (device)
4908 units. An intermediate layer is responsible for overall data processing, storing data to
4909 databases, mapping and calculations. The upper layer is responsible for overall detector
4910 operation and visualisation. The JCOP Finite State Machine FSM toolkit will be used to
4911 build a representation of the detector as a hierarchical, tree-like structure of well-defined
4912 subsystems, called FSM units. The HGTD FSM tree is shown in Fig. 10.9. The tree consists
4913 of two main nodes: the infrastructure and the detector. The infrastructure node includes
4914 all common devices, while detector nodes are split first on a functionally level into high
4915 voltage, low voltage and temperature, and then in a geographically level into the two vessels
4916 and down to the individual modules.

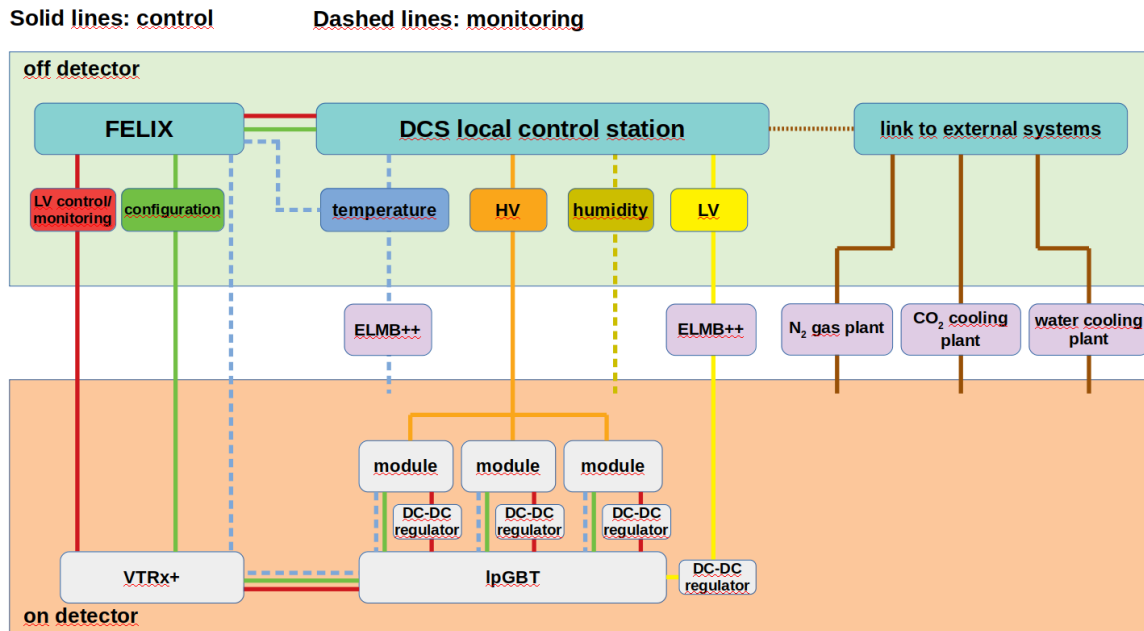


Figure 10.8: HGTD DCS layout

4917 **10.4.7 Interlock system**

4918 As detailed in ??, the HGTD Interlock system (HIS) is a standalone safety system that protects
 4919 the detector against a variety of risks. The Interlock system must always be running and its
 4920 components must never be disconnected.

4921 HIS hardware will be implemented in an Interlock Matrix Crate (IMC) located in USA15.
 4922 The Interlock crates are monitored by DCS...

4923 As one of the main dangers for silicon detectors is overheating, several hundred temperature
 4924 sensors will be installed on detector modules to monitor their temperature. Temperature
 4925 information from NTC sensors will be provided to DCS using ELMB, also located in the
 4926 IMC crate.

4927 **10.4.8 External systems**

4928 **Infrastructure** Beside the control and monitoring of the detector parameters, the DCS will
 4929 help to protect the detector from various risks raised from infrastructure failures. The
 4930 information on CO₂ cooling and N₂ gas plants state, as well as on other infrastructure
 4931 systems, including beam status, will be shared by their control systems. This will allow DCS
 4932 to monitor the infrastructure and bring the detector to a safe state if abnormal conditions
 4933 occur. In the event of any severe infrastructure failure or safety alarm the ATLAS Detector

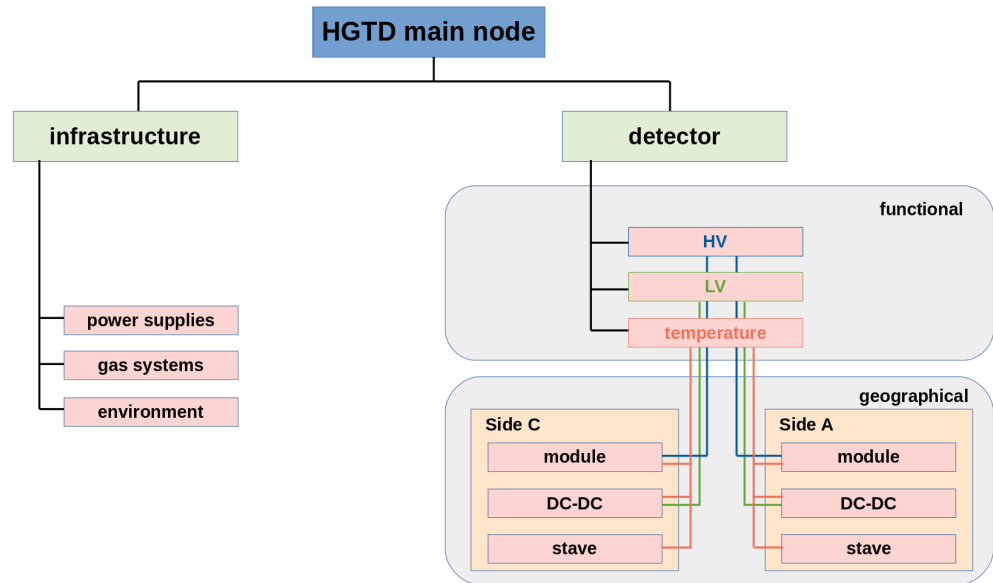


Figure 10.9: HGTD FSM layout

4934 Safety System (DSS) will act on the detector equipment via HGTD Interlock System. Such
 4935 actions are imminent and may have a coarse impact on the detector. To deal with this the
 4936 DSS actions can be delayed, allowing the DCS to implement more sophisticated control
 4937 sequences on the equipment before the actions triggered by DSS are executed.

4938 Signals related to risks due to common infrastructure failures or safety issues.

4939 Water leak-less cooling system for PP.

4940 DSS: Various fault signals from CO₂ cooling plant are processed by the DSS safety or
 4941 environmental alarms, such as smoke or flammable gas detection, magnet vacuum or
 4942 cryogenics failures (risk from water due to condensation melting), ATLAS emergency stop,
 4943 flooding signals corresponding to failures of common infrastructure, such as UPS power,
 4944 rack cooling, N₂ gas system failures, will be available from DSS

4945 (Un)stable beam conditions signal from the Beam Interface (for HV ramping).

4946 11 Detector Mechanics

4947 TODO: UPDATE TEXT , FIGURES FOR: 1) BASELINE WITHOUT ALUMINIUM INTERME-
4948 DIATE PLATE BETWEEN MODULES AND COOLING PLATE. 2) COOLING PIPES MADE
4949 OF TITANIUM 3) COOLING PLATE MADE OF ALUMINIUM 4) FEEDTROUGH OUTER
4950 VESSEL PART WITH NEW LAYOUT , RECALCULATE POWER AND THERMAL STUDIES
4951 WITH 400 MW/CM² INSTEAD OF 350 MW/CM², MAKE FIGURES WITH 3 RINGS BLUE
4952 COLOURS AND READOUT RAWs,.....

4953 11.1 Engineering design overview

4954 This chapter describes the global detector structure, the main mechanical sub-assemblies,
4955 as the hermetic vessel, front and back covers, inner and outer rings, the moderator, the
4956 support and cooling disks, the bolting and alignment device to LArg cryostat wall. The
4957 cooling system, common project with ITK and CMS, is also presented including cooling
4958 requirements and main components from the chiller up to detector cooling channels.

4959 As presented in previous sections, the space allocated to the HGTD equipped vessel is limited
4960 in (r,z) . In addition, the routing of the services should fit inside a gap of 17 mm in z against
4961 the end-cap calorimeter wall. These requirements are challenging many of the engineering
4962 parameters, like the stiffness and thermal insulation of the hermetic vessel, the thickness
4963 of the flex and connectors, the size of the support and cooling plates with imbedded CO₂
4964 channels and manifolds, the peripheral electronics boards and feed-throughs.

4965 In addition, the detector must be designed for easy and fast integration into the ATLAS
4966 detector, and it should be constructed to permit quick removal and re-installation of the active
4967 layers in the high-radiation environment while maintaining the beam pipe in position.

4968 The HGTD system includes two identical detectors fixed at both calorimeter end-caps. The
4969 various components of a single detector are shown in Fig. 2.4. They consist of a cylindrical
4970 hermetic cold vessel (front cover with heaters and back cover bolted to the inner and outer
4971 rings) that encapsulates two instrumented disks and an inner part of the neutrons moderator.
4972 Each instrumented disk (Fig. 11.1) represents a cooling support plate composed of two
4973 separate half disks with silicon modules installed on both sides, as shown in Fig. 2.9.

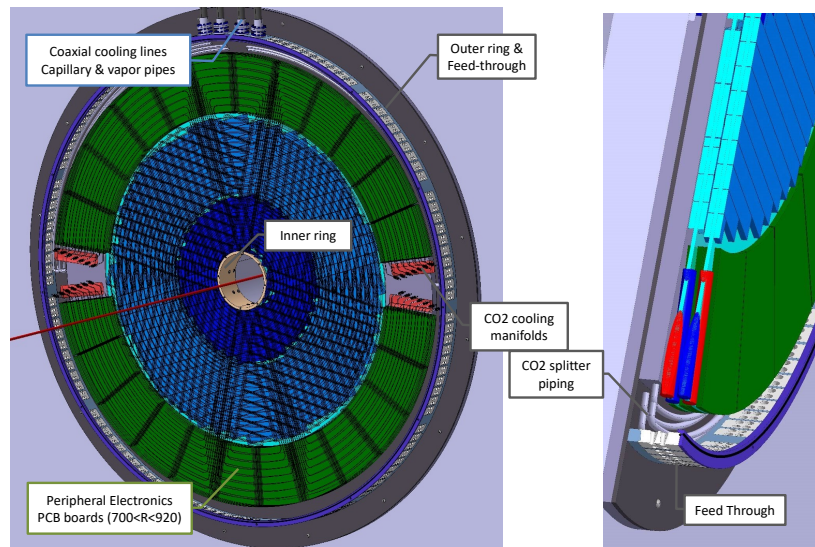


Figure 11.1: General view of the HGTD detector showing the silicon sensors inside the hermetic vessel. The green outer crown is the peripheral electronics limited by the outer ring which is holding tight electrical feed-through and cooling transfer lines. **UPDATE FIGURE WITH 3 RINGS BLUE COLOURS AND THE PEB REGION WITH ONLY 1 GREEN REGION AND NEW OUTER VESSEL FEEDTHROUGH.**

4974 The radial extent of the active area is 120 mm to 640 mm, which yields an acceptance from
 4975 pseudo-rapidity of 2.4 to 4.0. To protect the ITk and the HGTD from back-scattered neutrons
 4976 produced in the end-cap and forward calorimeters, 50 mm of moderator is included, as in
 4977 the current ATLAS detector. The envelope in z for the full detector, including the moderator,
 4978 supports, front and back covers, and the free gap with calorimeter front wall is 125 mm (or
 4979 75 mm excluding the moderator). The moderator is made up of two disks of different radii
 4980 to provide more peripheral space inside the vessel. This space allows electrical services,
 4981 feed-through, connectors and CO₂ distribution lines to fit inside the restricted envelope.

4982 The detector will partially occupy the ATLAS end-cap regions that presently house the
 4983 Minimum-Bias Trigger Scintillators (MBTS) and moderator. The cold vessel will be located
 4984 at z positions of $3420 \text{ mm} < z < 3545 \text{ mm}$ from the interaction point. The mid-plan of first
 4985 and last active layers will be located at $z = 3446 \text{ mm}$ and $z = 3472 \text{ mm}$. The position of the
 4986 two HGTD end-caps within the ATLAS detector is shown in Fig. 2.3. The overall dimensions
 4987 are summarised in Tab. 2.1. The total weight per end-cap is estimated to be 350 kg including
 4988 the moderator disks and to be 275 kg without the external moderator disk. The heaviest
 4989 components are the internal and external disks of the moderator, amounting to 75 kg each,
 4990 followed by the half-circular instrumented disks, weighing 30 kg each.

4991 11.2 CO₂ cooling system

4992 The cooling system is based on the evaporating CO₂ 2-Phase Accumulator Controlled Loop
4993 (2PACL) concept. It will be integrated with the general cooling system developed for the
4994 ATLAS ITk [64]. CO₂ cooling is chosen because it makes significant mass savings inside the
4995 detector possible due to the use of tubes of smaller diameter than in systems which are based
4996 on conventional cooling liquids. CO₂ evaporates at much higher pressures than common
4997 refrigerants, keeping the vapour compressed and therefore the volume low. The boiling
4998 temperature depends on the pressure and, as this pressure is relatively high, a pressure drop
4999 in the lines due to small-diameter piping does not cause much change in the evaporative
5000 temperature. In addition to the benefit of high pressure, CO₂ also has a low viscosity and
5001 high latent heat, so that less flow is needed than with other refrigerants. The narrower pipes
5002 can accommodate much higher flow speeds, which is a benefit for the overall boiling heat
5003 transfer coefficient. Taking into account the radiation environment in which the HGTD will
5004 operate, CO₂ is one of the most appropriate refrigerants because of its radiation hardness
5005 and low activation.

5006 The CO₂ will be pumped in liquid state from an external primary chilling source and will
5007 partially evaporate as it absorbs the heat dissipated by the HGTD components. Within each
5008 pipe, a small amount of CO₂ flows at high pressure in the form of small drops, and enough
5009 space is left for the vapour to circulate. A highly-efficient heat extraction is achieved by
5010 making use of the large latent heat for a liquid to vaporise, meaning that not only less fluid is
5011 needed to extract a certain amount of heat but also that the temperature of the liquid phase
5012 remains constant, while that of the vapour increases only slightly. The cooling power is then
5013 determined by how much CO₂ is left in a liquid state. Because it is used in mixed states
5014 (liquid and vapour), a significant mass reduction is introduced when comparing with other
5015 liquid mono-phase refrigerants.

5016 11.2.1 Requirements

5017 An operation temperature of -35°C must be maintained inside the HGTD vessel, in the
5018 level of cooling channels close to the modules, with a stability of a few degrees Celsius. As
5019 discussed in Chap. 5, the operating temperature must be kept as low as possible because,
5020 after irradiation, the leakage current of the sensors increases with temperature. In addition,
5021 the ASIC performance (S/N and jitter) will benefit from low temperatures, with observed
5022 improvements of up to 10% at -30°C compared to room temperature. These conditions
5023 will limit the heat dissipation and ensure good performance of the sensors and ASICs.
5024 The operating temperature of the peripheral on-detector electronics is flexible. It can be
5025 in the range of -35°C up to 20°C , making the cooling and stability requirements of these
5026 components less stringent. Taking into account that these electronics are located within the
5027 cold vessel, they will need to be maintained at a temperature close to the sensor operation

5028 point to avoid excess heat flowing towards the sensors. They will be used as pre-heaters to
 5029 stabilise the cooling parameters before coolant reaches the modules.

5030 Tab. 11.1 summarises the power consumption estimated for the various components of the
 5031 detector. This defines a need for maximum cooling power of 38 kW in total (19 kW per
 5032 end-cap).

HGTD Component	Power consumption	Total [kW]
Sensor	30 to 100 mW cm ⁻²	2.0–6.4
ASIC	< 300 mW cm ⁻²	17.6–19.2(*)
Flex cable	4 mW cm ⁻¹	1.8
HGTD vessel heaters	100 W m ⁻²	0.6
Total in active region		22–28.0
Pre-heaters (Perip. electr.)		8.8
Ambient pick-up		2.5
Total power dissipation		33.3–39.3

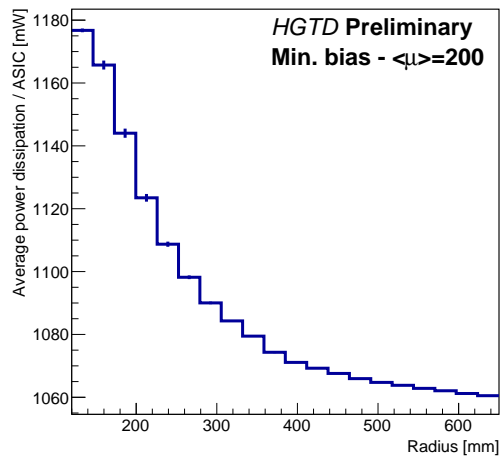
Table 11.1: Total power consumption estimates for the HGTD and breakdown for the various components, for a total number of 8032 sensors of $2 \times 4 \text{ cm}^2$ each, 16064 ASICs of $2 \times 2 \text{ cm}^2$, and 8032 flex cables of different lengths. (*) The 19.2 kW corresponds to 1.2 W (or 300 mW cm⁻²) consumed by each ASIC when calibration is taking place and is equivalent to 10% occupancy of all channels of an ASIC. During normal data taking, the total power consumed by the ASIC is smaller since it decreases with increasing radius.

5033 The ASICs, followed by the sensors, consume the most power, with up to 300 mW cm⁻²
 5034 by the ASIC and up to 100 mW cm⁻² by the sensors at the innermost radius. The power
 5035 dissipation of the ASICs decreases as a function of their radial position because the hit rate
 5036 decreases at larger radius, as shown in Fig. 11.2. Taking this radial dependence into account,
 5037 the total power consumed by the ASIC increases to 17.6 kW during data taking. The total
 5038 power consumed by the ASIC increases to 19.8 kW when calibrations are taking place and is
 5039 equivalent to 10% occupancy across all channels in the ASIC.

5040 The power dissipated in the flex cables is expected to be 4 mW cm⁻¹, leading to less than
 5041 300 mW per flex cable for the longest flex cables of 75 cm and 1.8 kW in total for all the flex
 5042 cables.

5043 The peripheral electronics boards will act as pre-heaters for the cooling system. On these
 5044 boards, the DC-DC converters will be the component with the highest power dissipation. As-
 5045 suming a 65% efficiency for the DC/DC converters, the peripheral electronics will dissipate
 5046 an estimated 8.8 kW.

5047 The total cooling power needed for the cooling station is 39.3 kW in total (19.7 kW per
 5048 end-cap). Given the uncertainties on current estimates of the power dissipation of some
 5049 components, a cooling unit dedicated to HGTD of 50 kW will be constructed. A spare cooling
 5050 station shared with ITk is also foreseen.



(a)

Figure 11.2: (a) Average power consumed per ASIC (in mW) as a function of the radius of the ASIC. REDO THIS FIGS WITH NEW READOUT ROW ORIENTATION BY CHRISTINA IF NOT POSSIBLE LEAVE ONLY RIGHT FIGURE THAT SHOULD NOT CHANGE.....

5051 11.2.2 Cooling design

5052 The cooling design is based on the technology implemented for the ATLAS Insertable B-
 5053 Layer detector and on industrial standards. Tri-axial vacuum-insulated transfer lines will be
 5054 used to connect the CO₂ cooling station located in USA15 and a junction and distribution
 5055 box to be located on the outer radius of the end-cap calorimeter on the HO side, close to the
 5056 HGTD patch panel area, detailed in Sec. 12.1.3. One such box per end-cap will be used to
 5057 disconnect/re-connect the rigid transfer lines for opening or closing ATLAS and to distribute
 5058 the CO₂ flow from one big transfer line to four smaller proximity lines.

5059 A permanent extension of transfer lines will be installed to allow the connection of the cooling
 5060 station to the HGTD cooling box when the end-cap calorimeter is in the open position. This
 5061 set-up will provide cooling during the yearly shutdowns and maintenance periods. While
 5062 the opening and closing of the end-cap calorimeter is taking place, the cooling will be
 5063 disconnected and will be reconnected after the movement is finished. During this period, the
 5064 temperature inside the vessel could increase up to room temperature, since the N₂, blowing
 5065 at 20 °C, will continue to flow at up to 750 l h⁻¹, improving the convection heat transfer
 5066 of inner parts of the detector. The anti-condensation heaters on the front cover and feed-
 5067 through crown should also be switched on during calorimeter end-cap movement phase.
 5068 The time estimated to reach 20 °C from -35 °C of the HGTD cold mass (200 kg, covering
 5069 mostly the on-detector system and moderator inner part (see Tab. 11.2), is determined by the
 5070 equivalent specific heat capacity (c in J kg⁻¹ K⁻¹) of the cold mass. Considering the thermal
 5071 power input as 200 W, mainly from the heaters, and the equivalent specific heat in the range

5072 of $c = 750 \text{ J kg}^{-1} \text{ K}^{-1}$, the unit time is about 12 min per degree increase, for a total of 10–12
5073 hours to reach room temperature.

5074 Rigid proximity transfer lines are under development for Phase-II upgrade applications
5075 for ATLAS and CMS. These aim for a transfer capacity of about 5 kW per unit. The HGTD
5076 design places an inner hose, with inner diameter of 5 mm for the CO_2 liquid, inside a 16 mm
5077 mid-hose for the vapour return. This hose, made of a multi-layer insulated (MLI) pipe,
5078 is enclosed within a vacuum hose of outer diameter less than 50 mm. The vacuum level
5079 inside the transfer lines must be less than 1×10^{-4} mbar in order to avoid convection and
5080 condensation on the outer wall. The relatively small outer diameter of such lines, less than
5081 50 mm, will facilitate their routing in the gap between the barrel and end-cap calorimeters,
5082 through a dedicated slot in ϕ allocated inside the original ITk envelope, as agreed with the
5083 ITk and Technical Coordination groups.

5084 The on-detector cooling layout is illustrated in Fig. 11.1. The four tri-axial rigid lines, one for
5085 each half-disk cooling plate, enter the HGTD vessel at the top position. They are holding
5086 capillary lines of 0.75 mm in diameter and up to 5 m long, ended inside the hermetic vessel at
5087 the manifold r-phi location. They supply CO_2 liquid to the 8 cooling loops that are embedded
5088 in each half-disk cooling plate on a semi-circular concentric pattern, as shown in Fig. 11.3.
5089 The radial distance between the concentric pipes in the loops at $120 \text{ mm} < r < 640 \text{ mm}$
5090 is 16 mm. This is the region covered by active modules placed on either side of cooling
5091 disk with overlap from 20% up to 70%. In the peripheral electronics area, at $r > 680 \text{ mm}$,
5092 where the dissipated power is used as pre-heaters, the distance between pipes is increased to
5093 30 mm to take into account the lower heat dissipation, thus keeping a uniform temperature
5094 distribution on the total area of the cooling disk.

5095 The on-detector cooling channels are designed to be made out of titanium pipes T40 grade 2
5096 or equivalent, baseline titanium allow similar to ITK production program. The non-magnetic
5097 stainless steel 304L is an alternative material, due to its wide usage in this technical field
5098 of particle physics equipment and manufacturing expertise, such as bending and welding
5099 processes. All pipes and fittings outside the on-detector area should be standards stainless
5100 steel products. The cooling plant is protected against over-pressure with safety valves set
5101 to 130 bar. This value is used as maximum design pressure on the cooling loops. To ensure
5102 that the pipes can sustain such levels of CO_2 pressure, the thickness of the pipes must be
5103 at least 0.3 mm. The outer diameter of the pipes is 4.0 mm. Their length varies from 4 to
5104 6 m for different loops. The maximal transfer capacity of the cooling loops corresponds to
5105 100 W m^{-1} . The characteristics of the loops are defined in close collaboration with the CERN
5106 Cooling group.

5107 The first prototype of the cooling loops has been manufactured from stainless steel 304L at
5108 the CERN workshop. This prototype corresponds to the inner zone of the cooling half-disk
5109 with the radial spacing of 16 mm identical to the final version). It has been successfully
5110 tested up to 165 bar at CERN proof pressure facility. The thermal tests will be undertaken

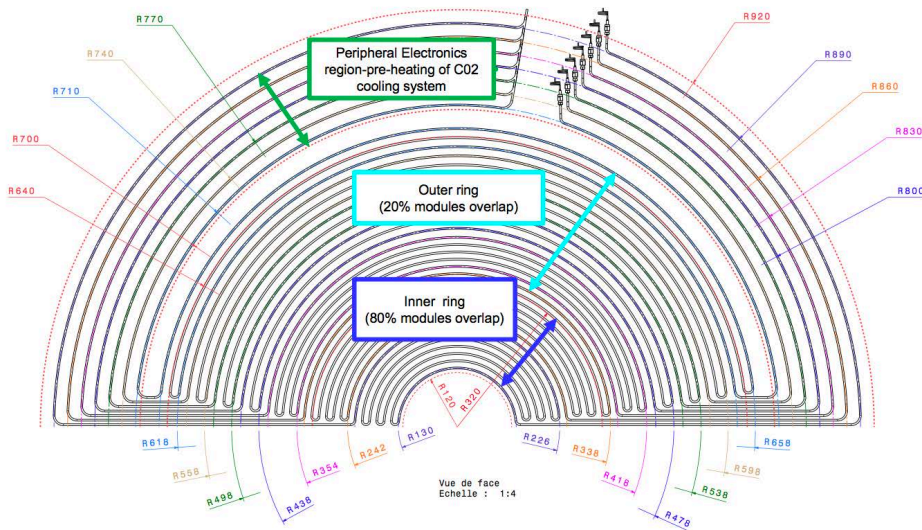


Figure 11.3: Layout of the cooling loops on a half disk support. The maximum power dissipation at the modules is 2.8 kW per half-disk, which loads the cooling pipes by about 90 W m^{-1} . The pitch of the inner, middle, and outer loops is 16, 20, and 30 mm, respectively. CHANGE FIGURE WITH LAST THAT HAS 1 MORE LOOP

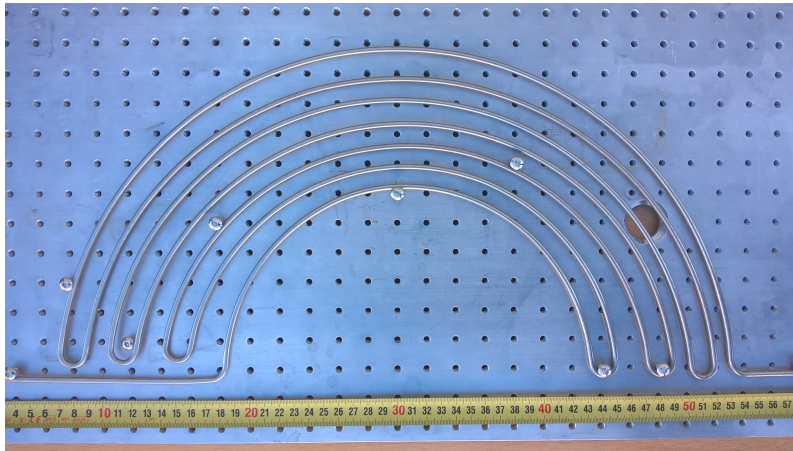


Figure 11.4: Cooling loop prototype corresponding to the inner part of the half-disk support.

5111 with Baby-Demo setup before being integrated into the sandwich structure of the cooling
 5112 support (see Fig. 11.4).

5113 Cooling pipes made of Aluminum may be considered for the final detector to reduce the
 5114 radiation length between the active layers and thereby improve the ability to associate ITk
 5115 tracks with HGTD hits. In addition, Aluminum is less activated by radiation and therefore
 5116 may allow greater access to the detector. This is important for replacing the inner ring

5117 midway through the HL-LHC and for maintenance during long shutdowns.

5118 The half disks with embedded cooling loops are the main support structure for the instru-
5119 mented active layers, as described in Sec. 11.5

5120 Given the challenging performance of the on-detector cooling system, one full scale pro-
5121 totype of cooling half-disk support will be produced, including aluminum panels and
5122 embedded cooling loops, equipped with appropriate heaters to simulate the silicon modules
5123 power dissipation. This prototype will be submitted to several thermal cycles to study
5124 thermo-mechanical behaviour, temperature distribution, CO₂ cooling parameters, and the
5125 performance of conductive media needed in between the modules, the support plates and
5126 the cooling channels.

5127 11.2.3 Cooling plant demonstrator

5128 One important milestone for the cooling development is the proof that the CO₂ evaporation
5129 temperature of -35°C can be achieved at the local HGTD support disks with realistic transfer
5130 lines and coolant distribution. Because of the crucial importance of this technology in the ITk
5131 and HGTD systems, a CO₂ cooling test facility called "Baby demonstrator" was set up by the
5132 CERN cooling team in collaboration with ATLAS and CMS. This is being tested (Fig. 11.5).
5133 This facility is installed in Building 180, next to the mock-up of ATLAS calorimeter, and
5134 will be used for tests of prototypes of ITk and HGTD cooling components with a real-scale
5135 geometry.



Figure 11.5: CO₂ cooling plant demonstrator located in Building 180 at CERN.

5136 This demonstrator will operate at low temperature with a limited cooling power of 5 kW.
 5137 The fluid transfer is subject to losses, which, in a two phase system, appears as a drop of
 5138 saturation temperature on the return line due to the frictional pressure drop of the flowing
 5139 media and static height differences. The main results were already presented in [64]. As
 5140 an example, Fig. 11.6 shows a typical temperature distribution in the cooling system from
 5141 the CO₂ plant to ITk on-detector loops and back, reaching the temperature of -40°C , the
 5142 target temperature for the ITk modules. To provide this temperature in the pixels staves, the
 5143 cooling plant temperature needs to deliver -45°C to account for the estimated 5°C lost in
 5144 the distribution and transfer lines.

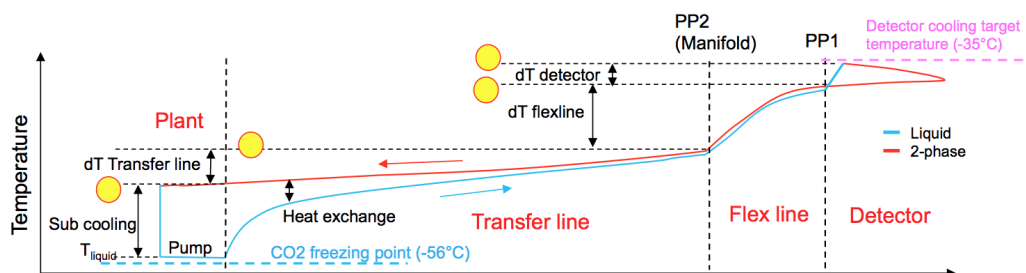


Figure 11.6: Typical temperature distribution between the CO₂ cooling plant and ITk loop [64].

5145 In order to optimise the performance of HGTD local supports at -35°C , specific prototypes
 5146 as well as the half disk cooling supports will be submitted to real scale CO₂ tests at the
 5147 Baby-Demo facility at CERN.

5148 11.3 Moderator

5149 The moderator, to be placed between the end-cap calorimeters and the active layers of the
 5150 detector, will protect both the ITk and HGTD against the back-scattered neutrons that are
 5151 produced by the end-cap calorimeters.

5152 The moderator disks will be made of borated polyethylene with a density of 0.95 kg L^{-1} ,
 5153 similar to the one used in the present ATLAS detector. As seen in Fig. 2.4, the new moderator
 5154 will be divided into two disks per end-cap, one inside and one outside the HGTD vessel.

5155 The moderator on the outside is mechanically separated from the HGTD hermetic volume. It
 5156 will be directly screwed to the LAr cryostat wall and will provide the necessary flat surface
 5157 on which to install the HGTD and accessible bolting brackets. It has a thickness that varies
 5158 along the radius, 10 mm in the region $180\text{ mm} < r < 342\text{ mm}$ (covering the LAr calorimeter
 5159 cryostat central flange and the bolting spots) and 20 mm elsewhere ($140\text{ mm} < r < 180\text{ mm}$
 5160 and $342\text{ mm} < r < 1100\text{ mm}$). The weight of this disk is 75 kg.

5161 The part of the moderator to be placed inside the vessel has a thickness of 30 mm, a radial
5162 coverage of $120 \text{ mm} < r < 900 \text{ mm}$, and a weight of 75 kg. It provides appropriate support
5163 for the instrumented layers and, because it does not extend to radii higher than $r =$
5164 900 mm , it leaves enough free space for the cooling services as shown in Fig. 11.1 right.

5165 In each end-cap, the total moderator thickness in z , summing the two disks, will then be
5166 50 mm , except at the inner and outermost radii. There, it is 40 mm in the region between
5167 110 mm – 342 mm and 20 mm for $r > 900 \text{ mm}$. During the maintenance, and when the
5168 replacement of the inner modules takes place at the surface, the two moderator disks may
5169 stay bolted in the LAr cryostat, together with the rear vessel cover.

5170 11.4 Hermetic vessel

5171 The hermetic vessel is the primary integration structure of the HGTD detector. It is construc-
5172 ted of four main components made of composite structures in carbon fiber (Fig. 2.4): the
5173 front and back covers, the inner ring and the outer ring (which will hold all the services
5174 feedthroughs), and includes the internal moderator. The vessel measures 1100 mm at the
5175 outer radius and 110 mm at the inner radius. The thicknesses of front and rear covers are
5176 15 mm and 7 mm , respectively, and weigh 25 kg and 15 kg , respectively.

5177 11.4.1 Requirements

5178 The hermetic vessel provides a robust support structure to the detector disks in a cold and
5179 dry volume. All materials chosen must satisfy safety requirements related to the expected
5180 radiation levels, described in Sec. 2.4, and the temperature range. Including safety factors
5181 and assuming no replacement of components during the HL-LHC, the materials used should
5182 survive $8.3 \times 10^{15} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ and 7.5 MGy . Components that will be replaced midway through
5183 the HL-LHC will see these criteria divided by two.

5184 The safe temperature range is defined by the minimum coolant temperature, -35°C , and
5185 the expected module interlock temperature, 40°C , with a margin of 20°C on both sides.
5186 This results in a safe range from -55 to 60°C . In addition, all mechanical components inside
5187 the vessel, including adhesives and bolting design, should withstand CTE (Coefficient of
5188 Thermal Expansion) mismatches over the temperature range specified above.

5189 This lower temperature limit of -55°C is also critical because it approaches the freezing
5190 point of the CO_2 coolant, which is -56.6°C for the given operating pressure.

5191 One of the requirements is to ensure the detector volume dry, keeping the dew point at about
5192 -60°C or below, to avoid condensation on the detector components. This can be achieved by
5193 permanent flushing with dry N_2 with 1% over pressure above atmospheric reference. The

5194 N₂ flow will renew gas in the vessel volume 3–5 times per hour (360 to 600 l min⁻¹). For this
 5195 purpose, the HGTD vessel was designed to be as hermetic as possible.

5196 Another requirement is to keep the temperature of the outer surface of the HGTD vessel
 5197 safely above the cavern dew point ($\sim 17^{\circ}\text{C}$). This will be done by placing flat heaters on the
 5198 external face of HGTD front cover and near the service feedthroughs, as described further
 5199 below.

5200 11.4.2 Front cover and heaters

5201 The front cover is designed as a sandwich structure, consisting of a honeycomb core placed
 5202 between two thin Carbon Fibre Reinforced Panels (CFRP). As a means to reduce the front
 5203 cover deflection from over pressure and CTE mismatch, radial stiffeners are integrated into
 5204 the structure during the curing process of the composite, as shown in Fig. 11.7. Considering
 5205 the tightness and stiffness requirements of the front cover, the front cover was designed as a
 5206 single piece. When replacing the inner part of the detector half way through the HL-LHC
 5207 life time, the front cover can be easily taken away because the beam pipes will already be
 5208 removed. In the opening scenario, when the beam pipe is in place, the front cover must be
 5209 slid over the beam pipe. Such a solution requires the use of dedicated tooling to properly
 5210 control the position of the cover with respect to the beam pipe. This complicates the opening
 5211 procedure and the conditions of access to the internal components of HGTD. We are studying
 5212 design options to split the front cover in two parts to allow removal or easier displacement
 5213 of the cover during YETS maintenance.

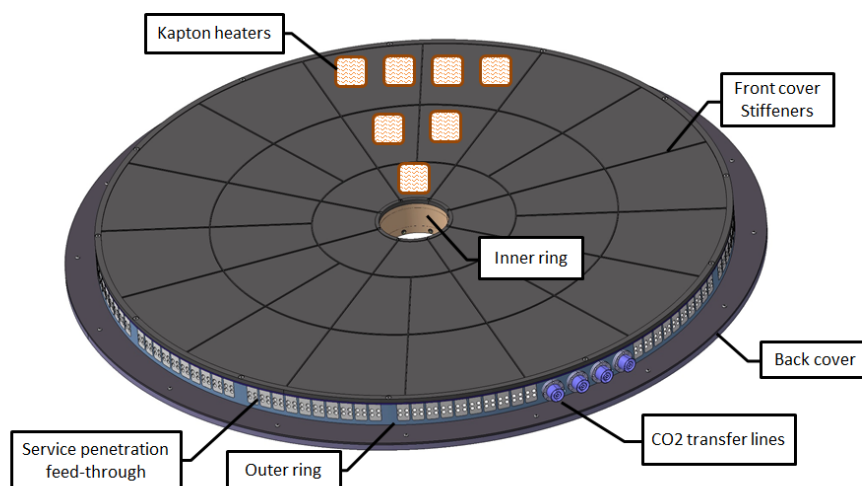


Figure 11.7: 3D view of the hermetic vessel with its main components. In particular, the front cover, equipped with kapton heaters uniformly distributed on the full surface, is visible. Not all heaters are shown in the picture. The heaters do not extend beyond the top of the three-millimetre stiffeners. CHANGE FEEDTROUGH WITH THE NEW LAYOUT AND FRONT COVER SPLITTED IN 2 PIECES

5214 The HGTD inner volume will be cooled down to as low as -35°C , therefore heaters will be
 5215 required on the external face of the front cover to prevent condensation on the vessel outer
 5216 surface. In a way similar to what is done on the LAr end cap cryostat front face, heaters
 5217 will be placed on the external face of the front cover in between the radial stiffeners. Their
 5218 purpose is to ensure a minimal temperature of 20°C outside the HGTD vessel, safely above
 5219 the cavern condensation temperature of about 17°C . The expected power of the heaters on
 5220 the vessel front cover is 100 W m^{-2} . This leads to a total contribution of approximately 300 W
 5221 per end-cap expected from the heaters, which is included in the CO_2 cooling plant budget
 5222 summarized in Tab. 11.1.

5223 The temperature distribution expected on the HGTD front cover and on the LAr cryostat wall
 5224 is shown in Fig. 11.8. In the temperature calculations, which were performed using Finite
 5225 Element Analysis (FEA), the ambient temperature of 23°C and heat exchange coefficient of
 5226 $10\text{ W m}^{-2}\text{ K}^{-1}$ were taken as input parameters. A temperature distribution in the range of 19
 5227 to 21°C has been estimated outside the vessel and on the LAr cryostat wall.

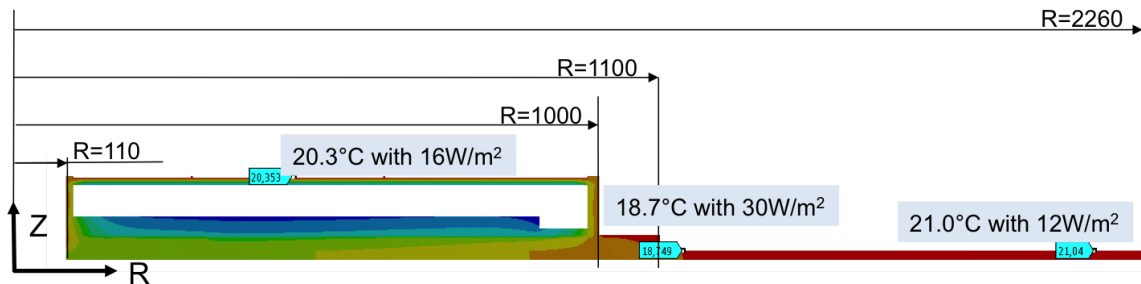


Figure 11.8: The calculated temperature distribution on the HGTD front cover and LAr front wall with heaters powered on. Transverse view of half the hermetic vessel and the cryostat wall is shown ($0\text{ mm} < r < 2260\text{ mm}$). The total power consumption of heaters is 50 W on front cover, 150 W near the feedthroughs, and 100 W on the LAr cryostat wall. REPLACE FIGURE WITH THE NEW ONE WITH LAST CALCULATIONS

5228 11.4.3 Back cover and interface with LAr cryostat

5229 In order to minimize the mechanical impact on the LAr end-cap cryostat, the vessel interface
 5230 with the cryostat wall will be made using the same threaded holes that are at present used
 5231 to mount the MBTS. The alignment of the hermetic vessel on the calorimeter end-cap will
 5232 be done with respect to the axis of the LAr warm tube, taking into account the existing
 5233 moderator, as shown in Fig. 11.9. Potential conflicts with the cooling pipe, currently installed
 5234 on the moderator and used for cooling of beam pipe during the bake-out procedure, require
 5235 verification during the LS2 and may require some optimisation. To optimise the vessel
 5236 installation procedure, the bolting/unbolting the back cover to the cryostat wall should be
 5237 possible without opening the HGTD hermetic vessel.

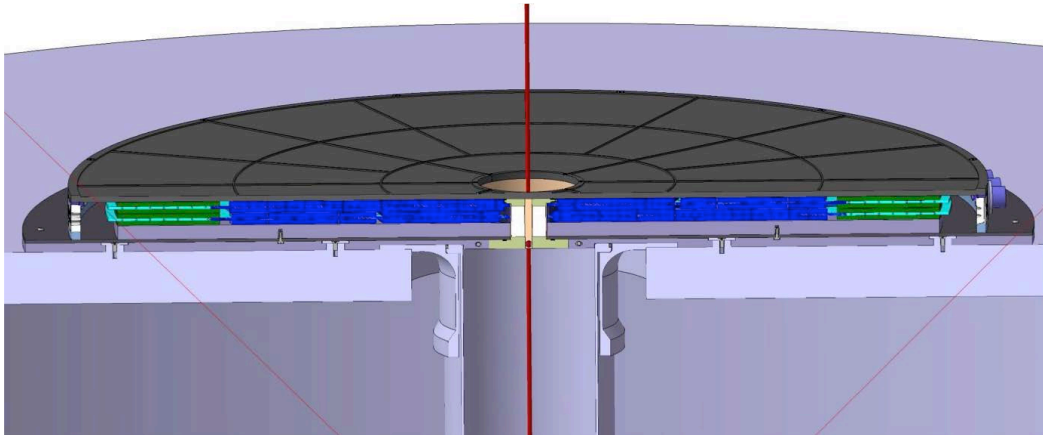


Figure 11.9: Transverse view (r,z) of the HGTD components and moderator bolted into the end-cap LAr calorimeter cryostat.

5238 The stiffness of the vessel when mounted on the cryostat wall was studied using FEA. In
 5239 this calculation, 10 mbar over-pressure have been applied, corresponding to dry nitrogen
 5240 blowing inside the vessel to prevent any ambient humidity leak from outside. The results
 5241 are presented in Fig. 11.10, showing a maximum deflection of 0.7 mm on the front cover.
 5242 This is equivalent to a maximum stress (Von Mises) of 70 MPa, giving a comfortable safety
 5243 margin compared to the over-pressure setting of 1020 mbar on the safety valves.

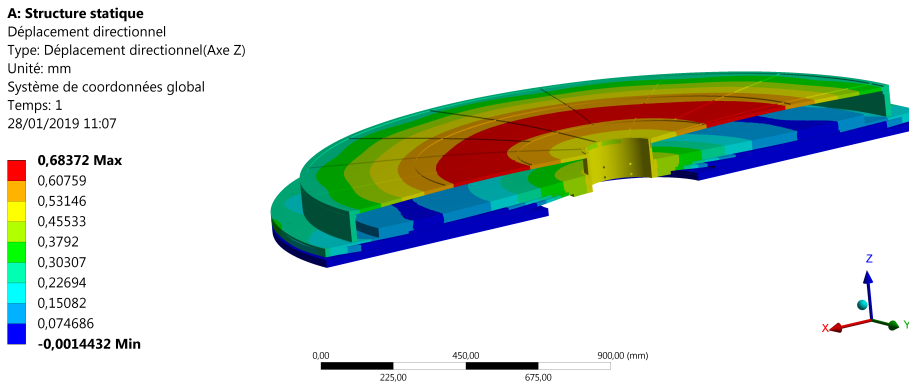


Figure 11.10: Finite Element Analysis (FEA) of the hermetic vessel with an over-pressure of 10 mbar. The red area corresponds to a maximum deflection of 0.7 mm in the vessel front cover. REPLACE WITH NEW FIGURE WITH FRONT COVER SPLITTED IN 2 PIECES

5244 11.4.4 Inner ring design

5245 The inner ring of the hermetic vessel borders on the beam pipe, resulting in a high level of
 5246 radiation and heat exposure. Design efforts are ongoing to select the best material with high

5247 radiation resistance and low thermal conductivity to provide a shielding barrier during the
 5248 beam pipe bake-out. Earlier projects with a similar environment, such as the ATLAS IBL
 5249 and the LHC beam-pipe, have demonstrated good performance from carbon fibre structures
 5250 and the aerogel insulating layers. With the actual design, shown in Fig. 11.11, the inner ring
 5251 is composed of a sandwich structure consisting of six millimetres of aerogel core enclosed
 5252 between two thin sleeves made of carbon fibre reinforced panels. Further research on high
 5253 performance materials, such as Kevlar panels and honeycombs, is being undertaken to
 5254 address the specified stiffness, thermal protection, and radiation resistance, taking into
 5255 account the low space allocated to the inner ring.

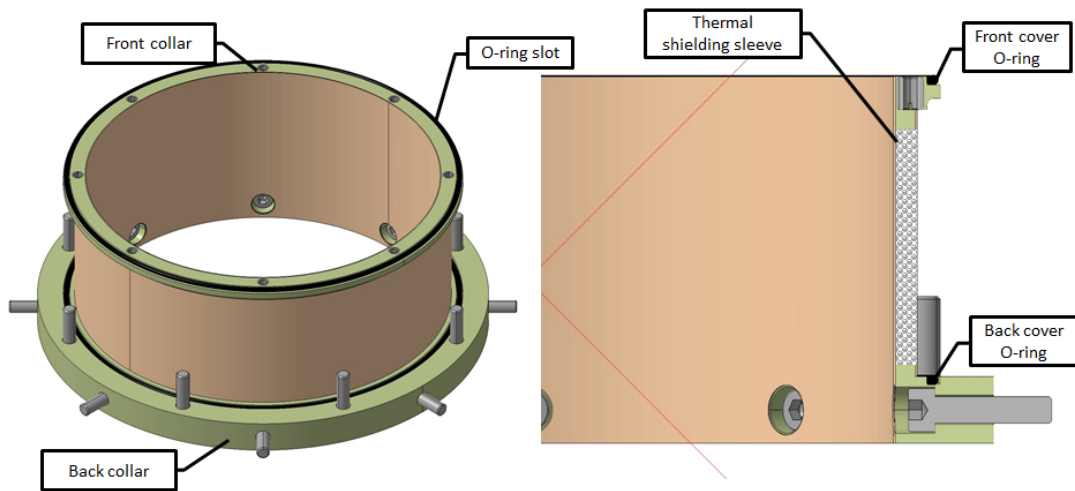


Figure 11.11: Central inner ring with its front and back collars. It is the central structure of the hermetic vessel, which ensures stiffness and tightness, thermal shielding, and HGTD positioning on the LAr cryostat.

5256 To provide tightness as well as the alignment of the vessel with respect to ATLAS coordinate
 5257 system, precisely-machined collars made of low thermal conductivity material, such as
 5258 Glass Fibre Resin Epoxy (GFRE) or high performance PEEK polymer, will be installed on
 5259 both extremities of the inner ring. Appropriate threaded inserts will be incorporated into the
 5260 front collar to allow bolting of the front cover. The circular slot will hold the sealing O-ring
 5261 made of PUR or EPDM material. The back collar will be bolted to the central flange of the
 5262 moderator, providing the hermetic vessel alignment with respect to the central tube of the
 5263 LAr cryostat.

5264 11.4.5 Outer ring design

5265 All available passages between detector volume (dry and cold) and the outside world run
 5266 through the outer ring, which holds conductor cables, optical fibres, CO₂ cooling lines, and
 5267 nitrogen blowing tubes. The outer ring structure, which is an assembly of several parts,

5268 must be made of stiff material with low thermal conductivity. As for the inner ring collars,
 5269 the main candidate materials are Glass Fibre Resin Epoxy (GFRE) and high performance
 5270 polymer PEEK. Taking into account the large diameter of this part (up to 2000 mm), the
 5271 manufacturing process is still under study to meet our specifications with a reasonable
 5272 cost.

5273 The maximum amount of service feed-through is equivalent to 480 cables with 12 mm in
 5274 diameter. In addition, enough space has been also allocated to hold four CO₂ transfer lines
 5275 with an outer envelope of 50 mm in diameter each, and few dry nitrogen holes (12 mm
 5276 in diameter) as shown in Fig. 11.12. The feedthrough concept is based on a resin potting
 5277 section on each cable and a detailed routing map. Given the limited space for services
 5278 inside the vessel, and to guarantee complete sealing of the detector hermetic vessel, the
 5279 feedthrough layout will be adopted to cable diameters and positions in ϕ , clamping the
 5280 cables at their correct locations. The final potting will provide a tight seal to ensure dry
 5281 detector volume and to minimize heat leaks. The potting mixture under study is PUR
 5282 (Poly-Urethane Rubber already demonstrated in IBL program), which is easy to handle and
 5283 is radiation resistant. From a maintenance standpoint, the cable clamps are removable and
 5284 replaceable individually.

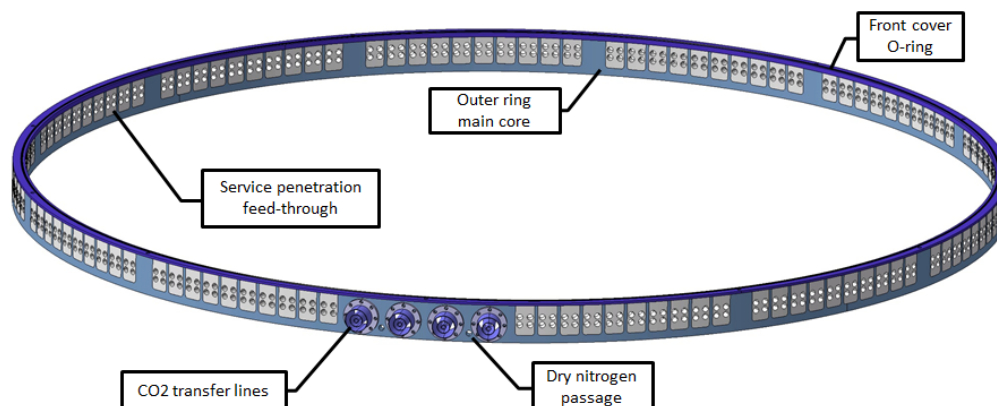


Figure 11.12: The outer ring assembly. The largest part of the hermetic vessel, with 2 m in diameter, it contains the service feedthroughs for cables, CO₂ transfer lines, and dry N₂ pipes. REPLACE WITH NEW DRAWING FROM CHINA

5285 The CO₂ transfer lines will pass through the cold-warm interface of the outer ring using
 5286 standard conical sealing made of PUR or EPDM (Ethylene-Propylene-Diene Monomer),
 5287 currently used in vacuum technology. The design of these cooling lines will be developed in
 5288 common with CMS Phase-II HGICAL, which will transport a similar amount of heat (4.7 kW
 5289 for CMS and 4.0 kW for HGTD per line) under similar cooling specifications. In general, it
 5290 is planned to work closely with the present program for both ATLAS and CMS trackers to

5291 develop and implement common solutions, such as appropriate improvements which can
5292 be made to the feed-through design and potting techniques.

5293 **11.5 Local supports and cooling disks**

5294 The design of local supports features four half disks per end-cap to provide the cooling and
5295 support on both sides for the module staves and peripheral electronics boards. Cooling
5296 piping with a semi-circular concentric pattern is embedded into sandwich structure of local
5297 supports to extract heat dissipation produced in the modules and peripheral electronics, as
5298 described in Sec. 11.2.1 and Sec. 11.2.2.

5299 **11.5.1 Geometry and design**

5300 The cooling support plates are composed of a carbon fibre structure with two high stiffness
5301 panels and a foam core inside. A good candidate for the foam is a composite pyrolytic
5302 graphite foam, similar to that planned for ITk. It has good thermal conductive characteristics
5303 and absorbs the different thermal expansion of the embedded stainless steel cooling pipes
5304 and carbon-fibre panels. As an alternative, a thermally conductive epoxy is also under study,
5305 due to its bonding, thermal performance, and reasonable cost. The high thermal conductivity
5306 of carbon fibre panels gives uniform temperature distribution over the large cooling disks.

5307 All the support disk edges will be sealed by pultruded carbon fibre U-shaped crowns, which
5308 will be the direct interface with the HGTD global support. As for the hermetic vessel inner
5309 ring, glass fibre epoxy resin is also considered as a good candidate to seal the panels edge.
5310 In order to perform the long term stability and alignment in the ATLAS coordinate system,
5311 these edges will be directly connected to the inner ring at small side radius and the outer
5312 ring vs moderator at the large side radius.

5313 If the intermediate plates/staves, shown in ??, will be chosen to load the modules in the
5314 final detector, these will be made from a material with a high thermal conductivity (ASIC,
5315 PEEK graphite reinforced, carbon fibre low epoxy) and will be bolted to the cooling support
5316 disks with thermal grease. If instead the “pattern thicker plate” option is chosen, seen in
5317 ??), the modules will be in direct contact with the cooling and support plate, insuring better
5318 thermal conductivity.

5319 Given the challenging performance of the support plates, one full scale prototype will be
5320 produced with a half-disk composite plate, a few embedded cooling loops, and one stave
5321 of dummy modules as heaters. This prototype will be connected to the Baby-Demo facility
5322 at CERN and submitted to several thermal cycles to study thermo-mechanical behaviour,
5323 temperature distribution, CO₂ cooling parameters and glue layers integrity between modules
5324 and carbon fibre skins.

5325 11.5.2 Thermal performance

5326 Thermal Figure of Merit (TFM) is used as a baseline parameter of the thermal performance
5327 of the design. TFM is defined as the ratio of the temperature difference built up from the
5328 hottest point of the coolant to the module power per unit of area. It has units of $^{\circ}\text{C cm}^2 \text{W}^{-1}$.
5329 It can be thought as the thermal impedance of the on-detector support assembly. For a
5330 given heat flux, the goal is to have the TFM as low as possible to minimize the temperature
5331 gradient between the sensor and the cooling channel. By selecting stainless steel material,
5332 the heat convection between the coolant and the inner channel wall is optimal. On the other
5333 hand, by reducing the thermal bridges between the modules and the channel, the thermal
5334 conductivity is improved and the TFM is reduced in the same level.

5335 A finite element analysis of the current CO_2 cooling design has been performed for one
5336 half disk face. The model considers a uniform power dissipation for all the modules of
5337 350 mW cm^{-2} but takes into account the higher density of modules in the inner ring com-
5338 pared to the outer ring, as described in Fig. 11.3. It considers the input cooling temperature of
5339 -35°C at the centre of each cooling loop with equivalent heat convection of $0.5 \text{ W cm}^{-2} \text{K}^{-1}$.
5340 The calculation has been made in a static regime, with no external heat exchange.

5341 The results are presented in Fig. 11.13. The temperature is uniform over the full surface
5342 of the cooling disk and close to -32°C . The TFM is near $17^{\circ}\text{C cm}^2 \text{W}^{-1}$ ($4^{\circ}\text{C cm}^2 \text{W}^{-1}$ as
5343 convective and $13^{\circ}\text{C cm}^2 \text{W}^{-1}$ as conductive).

5344 The model used in these calculations will be improved by implementing a more realistic
5345 representation of the modules, the reduction of the power dissipation with radius, and
5346 the thermal contact between components. The results presented assumed a pessimistic
5347 estimation of 350 mW cm^{-2} for all the modules, the expected maximum power dissipation
5348 in the modules located at the inner radius. In addition, a study of thermal runaway will be
5349 performed to cover for possible excesses in heat productions from the electronics or a lack of
5350 CO_2 liquid. Such studies will provide important input to the optimization of the cooling
5351 design.

5352 11.6 Detector overall layout

5353 An illustration of the HGTD detector components was shown in Fig. 2.4.

5354 The front view of the two double-sided layers that will be placed on each end-cap are shown
5355 in Fig. 11.14. They have a rotation of 15° with respect to each other to facilitate the entrance
5356 of the cooling pipes inside the cooling disks.

5357 A detail of the detector in the (r,z) direction, in the inner radius region close to the beam
5358 pipe, is shown in Fig. 11.15. It includes two cooling/disk supports where the double-sided

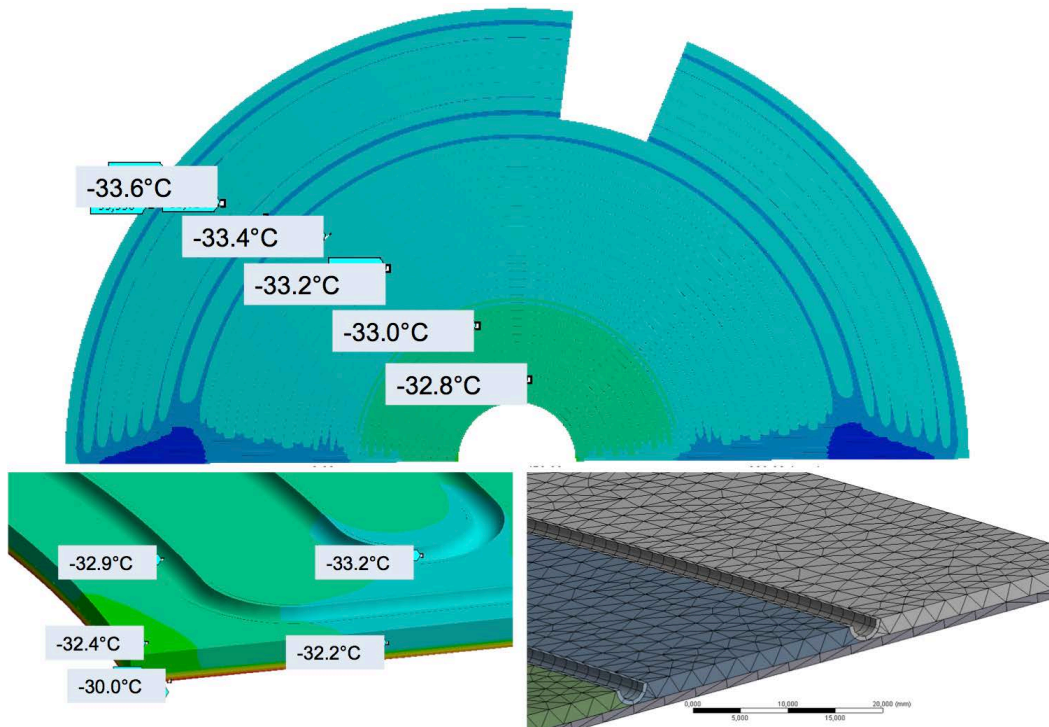


Figure 11.13: Expected temperature distribution over one half disk face. The bottom right figure shows FEM model used. REDO WITH 400 mW/cm²

5359 layers of the detector are mounted, the front and back covers of the vessel and the inner and
 5360 outer layers of moderator.

5361 The full assembly, including 50 mm of moderator, will match the envelope of 125 mm in the
 5362 z direction. A detailed breakdown of the (r,z) dimensions of the detector components is
 5363 presented in Tab. 11.2, and also the materials and estimated weight of various components.
 5364 The bottom of the table lists each component of a double-sided layer of detector modules
 5365 mounted on the cooling support.

5366 The measured thickness of the current prototype of the sensor-ALTIROC ASIC assembly
 5367 is about 1 mm thick. This gives a comfortable margin with respect to the final envelope
 5368 assembly protocol, with an expected thickness of module package (module and support)
 5369 of 3.5 mm. Since the longest readout row will include 19 modules, 10 will be the maximum
 5370 amount of stacked flex cables. Taking into account the estimated thickness of one flex cable of
 5371 0.3 mm, it gives the total thickness of flex cables stack of 3.0 mm per side. With an allocated
 5372 envelop for flex cables of 4 mm, considering it together with 1 mm integration gap, it should
 5373 be possible, though challenging, to fit all the components within the design envelope.

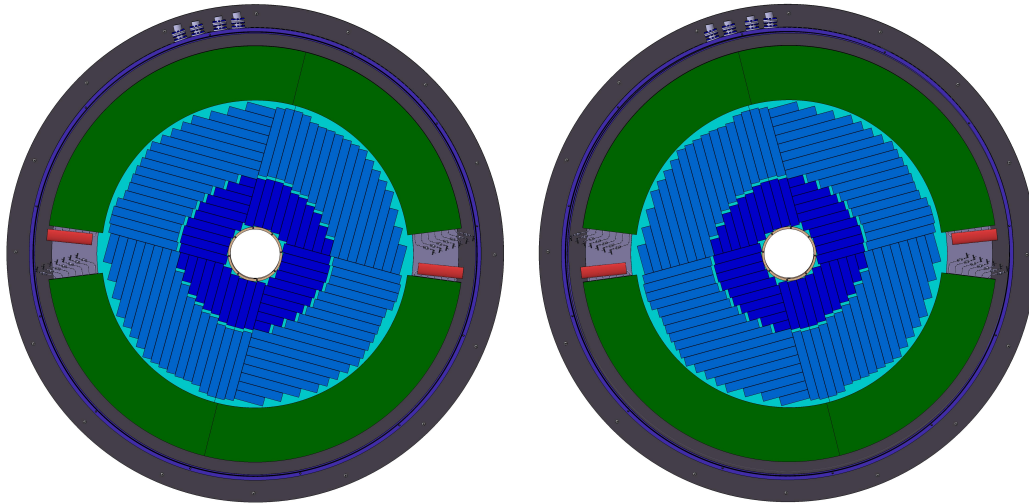


Figure 11.14: Front view of the two double sided layers that are placed on each HGTD end-cap. These two disks (right and left figures) have a rotation of 15° with respect to each other to facilitate cooling pipes interconnection with peripheral transfer lines. REPLACE WITH FIGS THAT HAVE THE READOUT RAWs

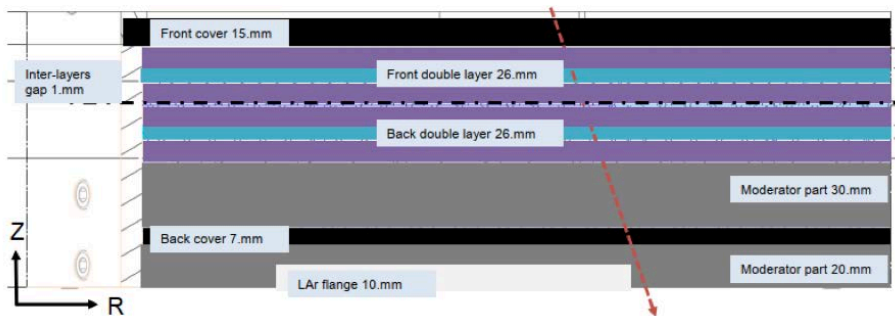


Figure 11.15: A detail of the detector in the (R,z) direction in the inner radius region close to the beam pipe, including two active double sided layers (installed on the cooling support plates), front and back covers and internal moderator. An extra 20 mm moderator is located outside the vessel in close contact with the end-cap cryostat.

HGTD components per end-cap	Thickness (mm)	z_{in}/z_{out} (mm)	R_{in}/R_{out} (mm)	Weight (kg)
Vessel Front cover	13.0	3420/3433	110/1000	25
Front double side layer (2 half disks)	26.0	3433/3459	120/980	60
Rear double side layer (2 half disks)	26.0	3459/3485	120/980	60
Internal Moderator	30.0	3485/3515	120/900	75
Vessel Back cover	7.0	3515/3522	110/1100	15
Vessel inner ring	10.0	-	110/120	5.0
Vessel outer ring	20.0	-	980/1000	35
External Moderator	20.0	3522/3542	110/1100	75
Air gap with LAr cryostat	3.0	3542/3545	110/1100	75
Total/end-cap (w/ mod.)	125.0	3420/3545	110/1100	350
Total/end-cap (w/o ext. mod.)	75.0			275
Double side layer breakdown	Thickness (mm)			
Air gap with vessel or with moderator	2			
Flex tail packing (0.22 mm per unit)	4.2			
Module package	4.2			
Cooling + support plate	6			
Module package	4.2			
Flex tail packing (0.22 mm per unit)	4.2			
Inter-layer gap	1.2			
Total per double sided layer	26.0			

Table 11.2: HGTD components per end cap. The top part of the table shows the components with their dimensions in z , r and their weights. Each double sided layer is divided in two half circular disks of 30 kg each. The total weight of the detector, including the moderator is 350 kg (275 kg without the external moderator). The bottom part of the table shows a breakdown of the front double sided layer. The breakdown of the back layer is identical.

5374 12 Detector Infrastructure

5375 12.1 Services

5376 12.1.1 Specifications

5377 The HGTD services (cables, fibres, pipes) can be grouped in several categories depending on
 5378 their role: optical fibres for data transmission; bias voltage for the sensors (high voltage-HV);
 5379 power for the electronics (low voltage-LV); DCS control, temperature sensors, heaters; dry
 5380 gas flushing; and CO₂ cooling.

HGTD Services	Number	Diameter (mm)	Routing
Optical bundles	40	9.5	HGTD - USA15
HV proximity cables	160	16	HGTD - (PP-EC)
DC-DC power control	40	14	HGTD - USA15
Interlock temp. sensors cables	32	16	HGTD - USA15
Sensors cables	10	12	HGTD - UX15
10 V power cables	72	15	HGTD - (PP-EC)
N ₂ gas pipes	2	15 and 18	HGTD - USA15
CO ₂ cooling lines	4	<50	HGTD - (PP-EC)
Total in barrel-endcap gap	356		
HV cables	170	15.3	(PP-EC) - USA15
300 V LV	10	14.4	(PP-EC) - USA15
300 V LV control	10	12	(PP-EC) - USA15
DCS cables	16	14	UX15 - USA15

Table 12.1: Summary of HGTD services required for each end-cap, including spares. In the upper part of the table are listed the cables, fiber bundles and pipes, which start on HGTD vessel. Some of them are routed directly to racks located in USA15 or UX15. Others go to PP-EC area on calorimeter end-caps. From the PP-EC the other group of cables are routed to service caverns, they are shown in bottom part of the table. The local cables routed inside service caverns are not included in the table.

5381 The services will include patch panels (PP-EC), which will be located on calorimeter ex-
 5382 tended barrels in several accessible places, close to the small wheel ($z \approx 6$ m). The main
 5383 purpose of the PP-EC is to provide a disconnection point for the services, which cannot be
 5384 accommodated by flexible chains due to lack of space and must be disconnected at ATLAS

5385 opening. The PP-EC will also allow to realise remapping between connectors on back end
5386 electronics and on the detector. More details on PP-EC are given below in Sec. 12.1.2 and in
5387 Sec. 12.1.3.

5388 An estimate of the required services per end-cap is summarised in Tab. 12.1 and is discussed
5389 in detail below. The table does not include the pigtailed, which serve for interconnection
5390 between cables and peripheral electronics boards inside vessel.

- 5391 • The number of optical links per end-cap is 1464?, including 520? up-links for data
5392 readout, 520? down-links for electronics configuration and fast signals (clocks, trigger,
5393 etc), and 424? up-links for luminosity readout. Multi-mode optical fibers will be used
5394 for data transmission, they will be grouped in bundles containing 48 fibers connected
5395 to 2 MTP connectors, 24 fibers per connector. The fibers will be encapsulated in a
5396 common sheath with reinforcement filler in order to be safely routed on cable trays and
5397 in the flexible chains. The number of fibers per bundle and per connector is optimised
5398 taking into account the routing of the fibers inside the HGTD vessel and the space
5399 available in flexible chains. Including spare fibers, a total of 40 bundles per end-cap
5400 are needed. Optical patch panes will be implemented in USA15 to organise the correct
5401 mapping for DAQ and luminosity readout.
- 5402 • THIS PARAGRAPH TO BE REVISED ONCE THE GROUNDING SCHEME HAS
5403 BEEN DECIDED (COMMON OR INDIVIDUAL RETURN LINES). The baseline for the
5404 bias-voltage distribution is to provide individually adjustable voltage per each HGTD
5405 module. Consequently 3992 HV lines are needed per end-cap. Return lines of the
5406 HV channels, belonging to the same power supply module, will be merged together,
5407 requiring fewer wires per HV module. Assuming 48 channels per HV module, 4200
5408 lines are needed. They will be grouped into 84 cables with an outer diameter of about
5409 16 mm. Adding 4 spare cables, it gives a total of 88 cables per end-cap, to be installed
5410 between the HV power supplies located in USA15 and the HGTD PP-EC. On PP-EC
5411 the HV lines will be re-mapped into 80 cables to match connectors on the peripheral
5412 electronic boards.
- 5413 • The powering is organised in three stages. The bulk power supplies located in service
5414 caverns provide 300 V DC current to the DC-DC converters that will be placed in the
5415 PP-EC area. These second-stage multi-channel DC-DC units convert 300 V to 10 V to
5416 supply the radiation hard DC-DC converters that will be located on the peripheral
5417 electronics boards inside the vessel. The last converters power the on-detector chips
5418 and peripheral electronics providing the 1.2 V DC power and also 2.5 V for the optical
5419 links. The 10 V voltage can be adjusted to take into account voltage drop on the cables.
5420 With such a layout the following cables are needed per end-cap: 4 cables to deliver
5421 300 V DC power, 4 cables for control and monitoring, 4 for interlock for interlock and
5422 4 cables for monitoring the DC-DC converters on PP-EC, all of them to be routed
5423 between service caverns and PP-EC area. In addition, 72 proximity cables are needed

5424 to connect the DC-DC (300 V to 10 V) units located on the PP-EC area to the peripheral
5425 electronics boards, inside the vessel.

5426

5427 • The DCS requirements and related components are described in Chap. 8 and Chap. 9.
5428 The DCS services include the following cables per end-cap:

5429 – Control and monitoring for peripheral electronics, 40 cables.

5430 – Readout of temperature sensors on cooling loops, pressure sensors, mechanical
5431 interlocks etc., 10 cables.

5432 – Interlock temperature sensors on detectors, 32 cables.

5433 The readout of sensors will be organized using ELMB II, part of which will be located
5434 in the experimental cavern, the rest, which provide the information from Interlock
5435 temperature sensors to DCS, will be placed in Interlock Matrix Crates in USA15 cavern.

5436 • The heaters, similarly to the ones currently installed on the LAr cryostat flange, will
5437 be installed on HGTD vessel front cover and in the proximity of the feed-throughs.
5438 Several power and temperature sensor cables will be needed for the HGTD heaters.

5439 • The HGTD hermetic vessel will be flushed with dry nitrogen to prevent condensation
5440 on the detector components. For N₂ gas circulation 1 inlet pipe and 1 outlet pipe, with
5441 an inside diameter of 16 mm and 13 mm respectively, will be installed to each vessel.

5442 • Four CO₂ cooling pipes <50 mm in diameter will be routed from the vessel feed-
5443 throughs to the cooling box located in the PP-EC area. The routing of the transfer lines
5444 between cooling box and CO₂ cooling plant located in USA15 is discussed in next
5445 section.

5446 12.1.2 Services layout

5447 The overall HGTD service layout is illustrated in Fig. 12.1.

5448 As it was described above, the detector vessel will be fixed on calorimeter end caps, which
5449 move at ATLAS opening. In the present ATLAS detector all end cap services are installed
5450 in flexible chains to avoid their disconnection before movement. Currently all the chains
5451 are fully occupied, but it is expected that they will be partly rearranged at Phase-II ATLAS
5452 upgrade and some space became available for a fraction of the HGTD services. Also a new
5453 small flexible chain per end cap is considered to be installed for HGTD. The priority for
5454 installation in flexible chains will be given to the most critical regarding disconnection cables
5455 and pipes. The other services should be disconnected before calorimeter end caps are moved.
5456 For that purpose the patch panels (PP-EC) will be organised on the calorimeter surface in
5457 accessible places. The 300 V to 10 V DC-DC converters as well will be installed in the PP-EC

5458 area in order to make LV cables as short as possible, that is necessary to minimise the power
 5459 losses (and voltage drop) on cables.

5460 The CO₂ transfer line will include rigid and flexible parts. The rigid line will be installed
 5461 between CO₂ cooling plant, located in USA15 and manifold box located on voussoir plat-
 5462 forms in ATLAS toroid area above calorimeter end caps. From the box two flexible lines, one
 5463 inlet, another outlet, will be routed to the splitter box on the top of calorimeter end cap on
 5464 IP side, close to HGTD. From the splitter box 4 smaller rigid lines will be installed on the
 5465 calorimeter front wall and connected to the HGTD vessel. The use of flexible lines avoids
 5466 the disconnection of CO₂ cooling lines at short openings during YETS. However, on the
 5467 platforms, there is not enough room to accommodate the flexible pipes long enough for long
 5468 openings in LS periods, when calorimeter end caps can be moved about 12 meters. For such
 5469 openings the flexible lines must be disconnected from the splitter box on the calorimeter and
 5470 can be extended with additional flexible lines to supply the HGTD with CO₂ cooling in open
 5471 position.

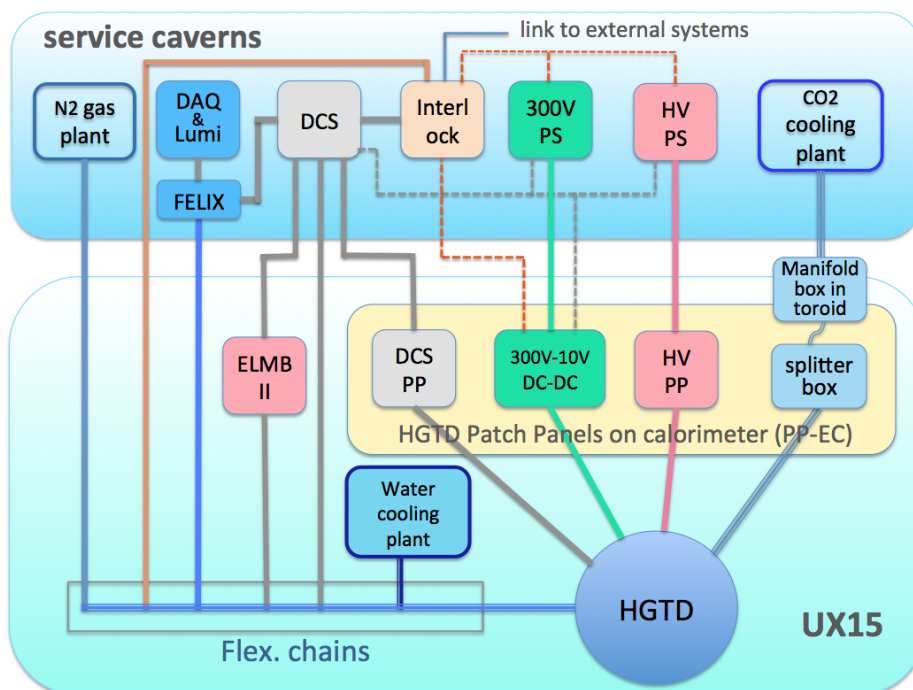


Figure 12.1: Overall HGTD services layout from the detector to USA15 or UX15. The optical fiber bundles, N₂ gas pipes, interlock and cooling temperature sensor cables, part of DCS cables and, still to be confirmed, the 300 V power supply cables are planned to be installed in flexible chains. The HV cables and rest of DCS cables will be routed through the patch panels, where they will have a disconnection point.

5472 To allow commissioning of the detector after installation in the pit and for maintenance
 5473 during shutdown periods, it should be possible to operate the HGTD when ATLAS is in

5474 open configuration, which requires reconnecting the services in the open position. For that
 5475 purpose, the extenders of cables and CO₂ cooling lines will be installed between respective
 5476 positions of the patch panels in closed and open configurations. Most of these extensions
 5477 must be permanently held in place, which will help minimise the time required to put the
 5478 HGTD in working order after each opening.

5479 12.1.3 Patch panels in PP-EC area

5480 The positions of the patch panels (PP-EC) and DC-DC units on the calorimeter end caps
 5481 will be selected by Technical Coordination, in four sectors in accessible areas to allow
 5482 disconnection of services. It will also be possible to replace any faulty DC-DC converter at
 5483 short access during the run. The probable patch panel locations are shown in Fig. 12.2.

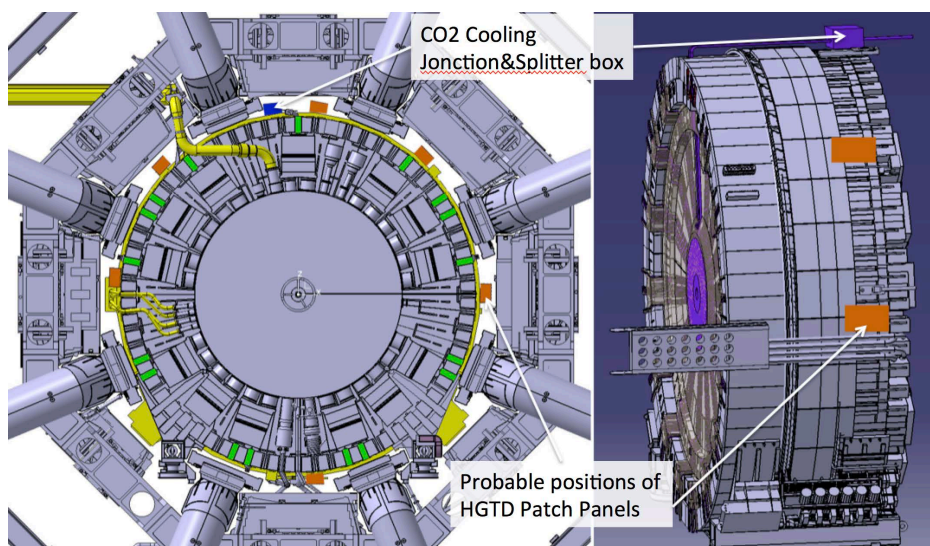


Figure 12.2: PICTURE AND CAPTION TO BE UPDATED. View of the probable locations of the HGTD patch panels (PP-EC) on the surface of the end-cap calorimeters. The DC-DC (300 V to 10 V) converters and cooling splitter box will also be located in this region. The exact position in z and ϕ is still to be allocated by Technical Coordination.

5484 The value of the magnetic field, along with radiation levels, are critical parameters for the
 5485 design of the DC-DC power converters. The magnetic field in the patch panel region is
 5486 shown in Fig. 12.3, varying from 0.05 T up to 0.5 T. The power supplies should be placed in
 5487 areas where the field is weaker, midway between two barrel toroids and as close as possible
 5488 to the surface of the calorimeter. Radiation levels in these areas was estimated using FLUKA
 5489 calculations, giving a maximum of 15 Gy and less than $1 \times 10^{12} \text{ n}_{\text{eq}} \text{ cm}^{-2}$ at the outer radius
 5490 of calorimeter end cap, where the patch panel boxes will be located.

5491 The DC-DC power converters located in the PP-EC area will require water cooling. Assuming
 5492 80% efficiency, about 4 kW of cooling power is needed in all PP-EC locations, combined per

5493 end cap. The existing in ATLAS leak-less water cooling systems have a sufficient capacity to
 5494 supply the HGTD detector on both end caps. The dedicated connecting pipes and manifolds
 5495 on the calorimeter will be required.

5496 12.1.4 Services routing on the calorimeter front wall

5497 As already mentioned, the space available to route the HGTD services in the gap between
 5498 the calorimeter barrel and end cap is very limited, making the design and installation of
 5499 the services a challenging task. This space is shared between ITK and HGTD services, and
 5500 also the scintillator counters, belonging to the Tile calorimeter system, are installed here. In
 5501 the present ATALS configuration, the counters are fixed on the Tilecal and LAr front face,
 5502 where the HGTD cables will be routed. In LS3 the scintillators must be replaced by new
 5503 ones. It was agreed with Tilecal system and Technical Coordination that the scintillator
 5504 counters will be installed on top of the HGTD services, while the last will be fixed on the
 5505 wall of the calorimeter. Such layout will allow access to the counters and their replacement
 5506 during HL-LHC lifetime. In order to provide more robust support and fixations for HGTD
 5507 cables and for scintillator counters and, at the same time, to protect the Tile calorimeter
 5508 scintillator tiles and fibers on front face, thin aluminium support plates will be fixed on the
 5509 Tile calorimeter modules.

5510 The envelop for HGTD services is shown in Fig. 12.4. All space in ϕ on the front wall of
 5511 the LAr end cap cryostat is available for HGTD services, while at bigger radius they have
 5512 to be grouped to fit in space between LAR barrel crates and further between Tilecal barrel
 5513 fingers, sharing the space with ITK services installed on the calorimeter barrel. However
 5514 room in two gaps between LAr barrel crates on top cannot be used to root HGTD cables.

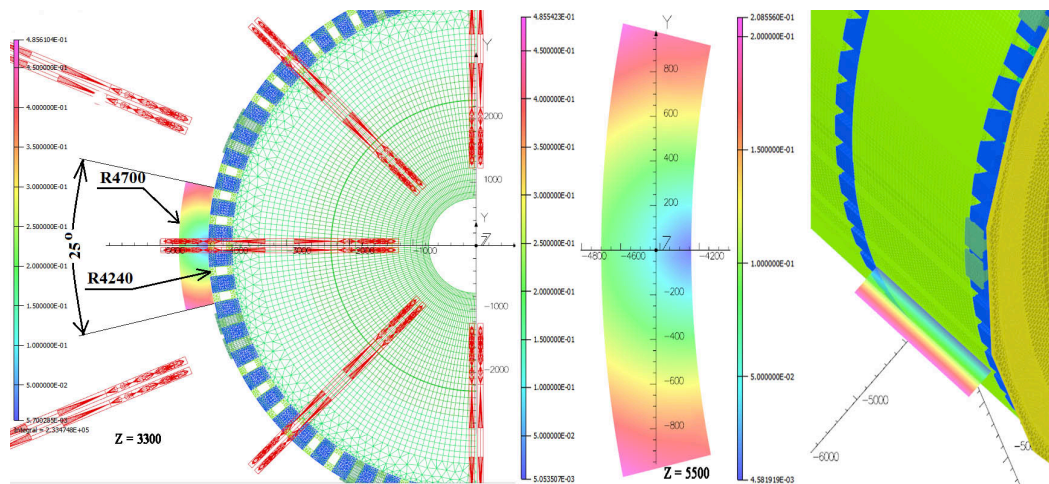


Figure 12.3: Magnetic field in the region of the HGTD PP-EC patch panels. FONT SIZE OF THE SCALE NUMBERING TO BE INCREASED.

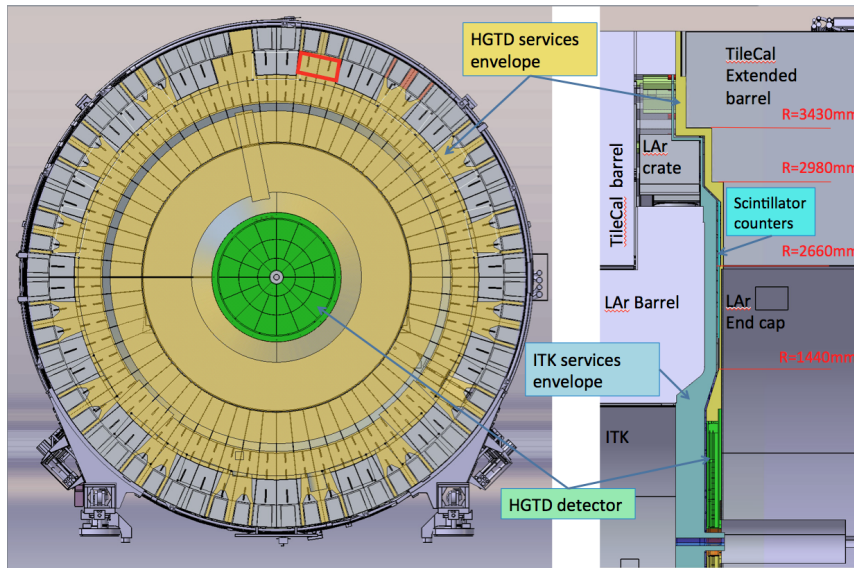


Figure 12.4: THE COLORS TO BE UPDATED. The envelope for HGTD services. On the left: front view of the calorimeter end cap on side A. The space available for HGTD services is shown with yellow color. With red rectangles is shown the area, where the room the HGTD services is very limited. On the right: the HGTD services envelope in the gap between calorimeter barrel and end cap. The envelopes for ITK services and the Tilecal scintillator counters are also shown.

5515 One constrain comes from the requirements to keep free access to the end plates of 3 Tilecal
 5516 modules, located at the top of the calorimeter, to allow remove these plates when accessing
 5517 the electronics of the modules. Space in another gap is blocked by LAr HV filter box. The
 5518 space in z available for HGTD services on LAr end cap cryostat wall at radius >1.4 m is only
 5519 17 mm. The exception will be a dedicated slot for four CO_2 cooling pipes, as described in
 5520 Sec. 11.2.2.

5521 The HGTD services routing on the calorimeter end cap front wall is shown in Fig. 12.5. The
 5522 cables, connected to the outer ring of HGTD vessel in four layers, will pass to one layer at
 5523 $r > 1.4$ m to fit within the envelope of 17 mm. Below the Tilecal barrel fingers, the cables
 5524 will be regrouped to a few layers to come out on the calorimeter surface through the gaps
 5525 between the fingers. As it was discussed above, the HGTD cables cannot be routed in two
 5526 top gaps between LAr barrel crates. Due to that the cables from top section of HGTD deviate
 5527 towards neighbour gaps.

5528 On the surface of the calorimeter end cap the cables will be routed towards PP-EC located in
 5529 4 places around the calorimeter.

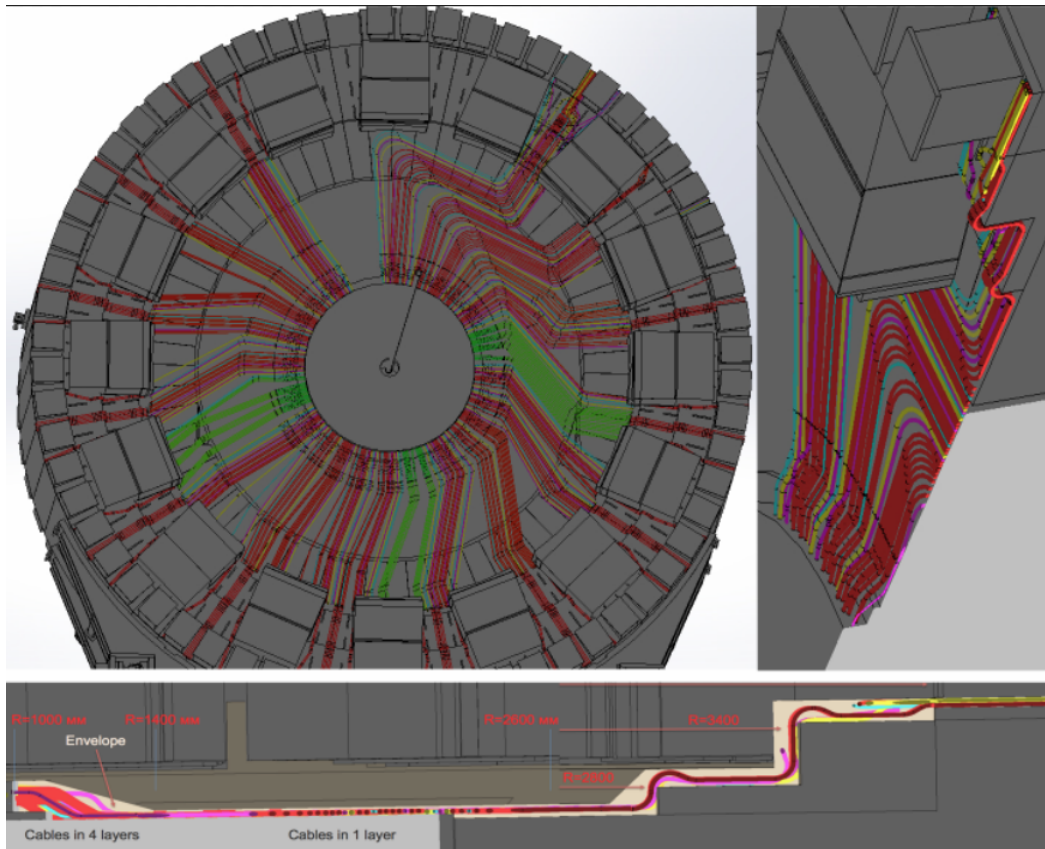


Figure 12.5: PICTURE AND CAPTION TO BE UPDATED. HGTD preliminary services routing on the calorimeter wall.

5530 12.1.5 Services connection to outer ring and inside the vessel

5531 The outer ring of HGTD vessel provides interface for all services of the detector. With
 5532 such approach, the HGTD detector can be completely assembled and tested at surface and
 5533 brought down to the pit for installation with closed vessel. Once the vessel fixed to the
 5534 front wall of LAr cryostat, the pipes, cables and optical bundles will be connected to the
 5535 detector. To realise such scenario, the cooling and gas pipe fittings, electrical and optical
 5536 connectors will be embedded in outer ring, as it was discussed in (REF. TO THE OUTER
 5537 RING SECTION OF CH 11 ???). The layout of outer ring is shown in Fig. 11.12.

5538 The organisation of services inside the HGTD vessel is schematically shown in Fig. 12.6. The
 5539 short pigtailed, one per cable, will interconnect the cables and peripheral electronics boards
 5540 (PEB). The optical bundles, connected to the outer ring, will be terminated with 24-fiber
 5541 MPO connectors. The optical pigtailed will be used to distribute these 24 fibers from each
 5542 bundle to several VTRx+ optical link modules installed on the PEB. One bundle is required
 5543 per PEB, including spare fibers. The optical pigtailed will also contain spare fibers terminated

5544 by connectors.

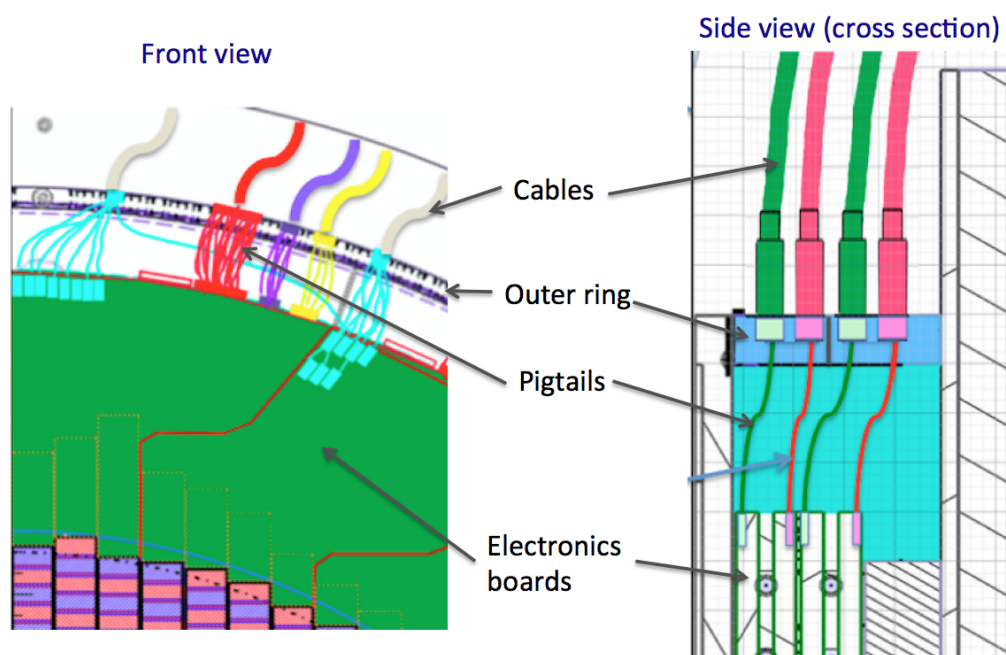


Figure 12.6: PICTURE TO BE UPDATED

5545 12.1.6 Services installation

5546 The installation of services and patch panels will be done in close collaboration with Technical
 5547 Coordination. The delivery of CO₂, under-pressure water cooling stations and the N₂ gas
 5548 plant is the responsibility of Technical Coordination and the CERN support cooling and gas
 5549 teams.

5550 The various components should be available at different times depending on the delivery
 5551 and final location in the ATLAS cavern. To decouple the installation of cables, patch panels
 5552 and the detector, the mock-ups of patch panels and vessel feed-throughs can be installed on
 5553 their final place with the aim to precisely indicate the cable connection points.

5554 In an environment as complex as ATLAS, cable routing requires numerous turns and trans-
 5555 itions between cable trays, which does not allow the installation of cables with an accuracy
 5556 of several centimetres at connection points. As a consequence, some extra length should be
 5557 allowed for each cable, which could then be accommodated on cable trays, however it is
 5558 usually not always possible due to the lack of space.

5559 Therefore, the common approach for installing long cables is to pull cables with the connect-
 5560 ors attached only on the detector side, to allow adjustments to the cable length on the other

5561 side. The connector at the second end of the cable should be attached in situ, though that
5562 is not always feasible due to connector complexity or lack of space or time for this work.
5563 Given all this, the different installation scenarios are foreseen for different HGTD services,
5564 as described below. All cables, except optical bundles, must be tested before installation in
5565 the cavern.

5566 Taking all that into account, the various installation scenarios are foreseen for different
5567 HGTD services, as it is described below.

5568 The **Optical bundles** will be delivered with connectors attached, tested and protected at
5569 both ends at the factory. Some space should be reserved to accommodate an extra length on
5570 the cable trays below the racks in USA15. The optical bundles will be routed through small
5571 plastic flexible chains available in sector 11, to avoid disconnecting them at ATLAS openings.
5572 An optical patch panel will be used in USA15 to remap the fibers between luminosity and
5573 data readout.

5574 For **HV cables** two installation scenarios are considered. If space on cable trays below the
5575 racks in USA15 is available to accommodate an extra length, the HV cables will be delivered
5576 with connectors fixed at both ends. Otherwise, the cables will be made in double length,
5577 folded in the middle, with connectors attached at both ends, to be routed to the PP-EC
5578 patch panel. Such a configuration makes it possible to test the cables and connectors before
5579 installation. After pulling such cable pair into the service cavern, the loop will be cut out
5580 to precise length and the missing connectors attached in-situ. One of these scenarios will
5581 be chosen when the layout of the racks and the services in the service caverns are available
5582 from Technical Coordination.

5583 The **LV and DCS cables** to be routed between the experimental and service caverns will be
5584 installed with one connector (detector side), the second connectors will be attached in-situ
5585 near racks. The same scenario will be applied for DCS cables between the HGTD or patch
5586 panels and racks in UX15.

5587 The **Proximity cables** listed in Tab. 12.1 must be delivered with connectors attached at both
5588 ends, because it would be extremely difficult to install them in-situ, near the calorimeter.
5589 Before installing the connectors, the length of these cables must be precisely measured
5590 in-situ, by pulling the pilot cables between the mock-up of outer ring and the patch panel
5591 box, placed in their final positions.

5592 The installation of the patch panels, services and respective connectivity will be done when
5593 access is permitted by Technical Coordination. These activities will start well before the
5594 HGTD installation and will be spread over time. In the current schedule, these activities are
5595 planned over approximately 16 months, from January 2024 to April 2025 (??).

5596 The installation of transfer lines for the CO₂ cooling system and pipes for the N₂ gas system
5597 is the responsibility of Technical Coordination, who will plan and organise it.

5598 **13 Detector Assembly, Installation, and** 5599 **Commissioning**

5600 **13.1 Assembly and commissioning on surface**

5601 In order to prevent any contamination of the active sensors (dust, metallic chips), all detector
5602 assemblies and testing must take place in a clean environment, equipped with temperature
5603 and humidity control gauges. The floor should be ESD protected (Electro Static Discharge)
5604 for personnel and components at all work-stations and setups. Specifications for this
5605 environment are under development considering that all critical assembly steps shall take
5606 place in a clean room class ISO-8 or better.

5607 **13.1.1 Half disks instrumentation**

5608 The detector units (see Sec. 7.4.2), corresponding to 16 inner half disks, 32 middle quarter
5609 disks and 32 outer quarter disks, will be assembled and qualified in different Institutes and
5610 shipped to CERN to be mounted on the cooling plates. Separately, the peripheral electronic
5611 boards and the flex tails will be qualified at collaborating Institutes and shipped as well to
5612 CERN. The QA procedure on the detector units, peripheral electronic boards and flex tails
5613 will be repeated at least on a sample of elements, to confirm that no damage has occurred
5614 during shipping.

5615 The half disks will be instrumented by mounting first the peripheral electronic boards,
5616 followed by the outer ring and the detector units. After mounting the outer ring and
5617 before adding the detector units, the pigtailed will be connected. The last step consists in the
5618 connection of the flex tails between the modules and the peripheral electronic boards.

5619 At this stage the full on-detector readout chain is connected for the first time and the first
5620 full calibration will be performed, as described below.

5621 Each instrumented half disk will be a single piece of 30 kg with 12 cm inner radius and 98 cm
5622 outer radius. **Dimensions to be updated.** Dedicated tools will be developed to allow the
5623 disks assembly in the optimal position (horizontal vs vertical) with appropriate rotation to
5624 fully instrument the two faces of the half disk.

5625 13.1.2 Detector assembly on hermetic vessel

5626 After completion of all tests and replacement of faulty components if necessary, the in-
5627 strumented half disks will be installed on the hermetic vessel that has been previously
5628 mechanically tested at the Institutes taking care of the machining. The mechanical tests will
5629 be repeated at CERN to exclude damage due to shipping and the air tightness of the vessel
5630 will be checked. **Add description of assembly procedure.** After this step the detector can
5631 be connected to the baby-demo CO₂ cooling system to perform pressure and cooling tests,
5632 additionally to further performance tests, as described below.

5633 13.1.3 Quality Assurance after assembly

5634 The first set of tests after instrumentation of the half disks will probe the electrical connectiv-
5635 ity, followed by a first full calibration of all the on-detector elements and a full-chain test with
5636 real particle signals. The test bench consists of a cold box with a radioactive source, as well
5637 as a portable system for powering, read-out and control to used at the surface for the tests
5638 described in this Section and in the cavern before the detector is connected to the ATLAS
5639 DCS and DAQ systems. **Add more details if necessary, keep compatible with BoE.**

5640 A database will be used to record the status of each component at all assembly steps, in
5641 particular electronic and thermal parameters of the instrumented half disks. The aim is to
5642 have a full history tracking from the production process up to the final assembly and testing.
5643 Already existing ATLAS databases will be adapted to avoid duplication of the software
5644 development effort. The database identification protocol of all mechanical components
5645 will be based on a serial number and/or QR code (bar-code if any). In addition to the files
5646 recording detailed technical parameters (raw material, chemical composition, manufacturing
5647 process, testing), the database will allow an easy monitoring of the construction progress.
5648 After completion of the detector installation in the experimental cavern, the database will
5649 evolve towards a collection of system configuration data, necessary to analyze the detector
5650 operation conditions and performance.

5651 The detector assembly and QA on surface is expected to take 9 months per end-cap, between
5652 October 2023 and June 2024 for HGTD-A and between August 2024 and April 2025 for
5653 HGTD-C, finishing 4–12 months before the final HGTD installation in the ATLAS cavern, as
5654 detailed below.

5655 13.2 Installation in the cavern and commissioning

5656 13.2.1 Access scenarios

5657 The access for installation and maintenance of the detector and the off-detector electronics
5658 located in UX15 can only occur in breaks of LHC operation, and the exact actions depend
5659 on the duration, induced radiation levels and ATLAS opening scenarios. The back-end
5660 electronics situated in USA15 will be accessible at any time, but other actions will be limited
5661 when taking data. Different kinds of stops are expected at HL-LHC, similar to the present
5662 LHC breaks:

- 5663 • **Short access** for a few hours only, primarily for LHC machine interventions and
5664 usually announced on short notice. In these periods, electronic components located
5665 in the HGTD patch panel (PP-EC) areas can be accessed for simple interventions, for
5666 example for the replacement of 300 V-10 V DC-DC converter modules. Access to DCS
5667 equipment in racks in UX15 will also be available.
- 5668 • **Technical Stop**, typically of one week duration, for maintenance of the LHC and
5669 experiments. The same areas as for the short access periods will be accessible, but it
5670 will be possible to perform more complex and long operations.
- 5671 • **Year-End Technical Stop (YETS)**, the yearly maintenance for about 12 weeks. In this
5672 period the ATLAS detector is opened, keeping the beam pipe in place as illustrated in
5673 Fig. 13.1. The access to the HGTD is possible, only if radiation levels will allow for it
5674 (see Sec. 13.2.2). The distance between the barrel face and the HGTD face is typically
5675 3.5 m. To get access to the detector the front vessel cover has to slide along the beam
5676 pipe, using dedicated tools, in order to not damage the beam pipe nor the HGTD
5677 vessel. A priori only the peripheral electronics boards of the front face of the front
5678 disks are easily accessible with minimal manipulations.
- 5679 • **Long shutdown (LS)** of typically 2 years, for large upgrade or consolidation programs
5680 for Experiments and LHC. The ATLAS detector will be in large opening position,
5681 with the beam pipe removed, as shown in Fig. 13.2. The distance between the barrel
5682 calorimeter face and the HGTD face is maximum 12 m. After the LS3, where the HGTD
5683 should be installed, the next Long shutdowns should be used for deep maintenance of
5684 the detector. It will be possible to bring on surface the half disks of HGTD to replace
5685 malfunctioning components and to replace the inner ring at the middle life time of the
5686 HL-LHC.

5687 The access and manipulation of the HGTD components sitting inside the vessel need to
5688 follow strict safety rules due to expected high radiation levels has described in the next
5689 section.

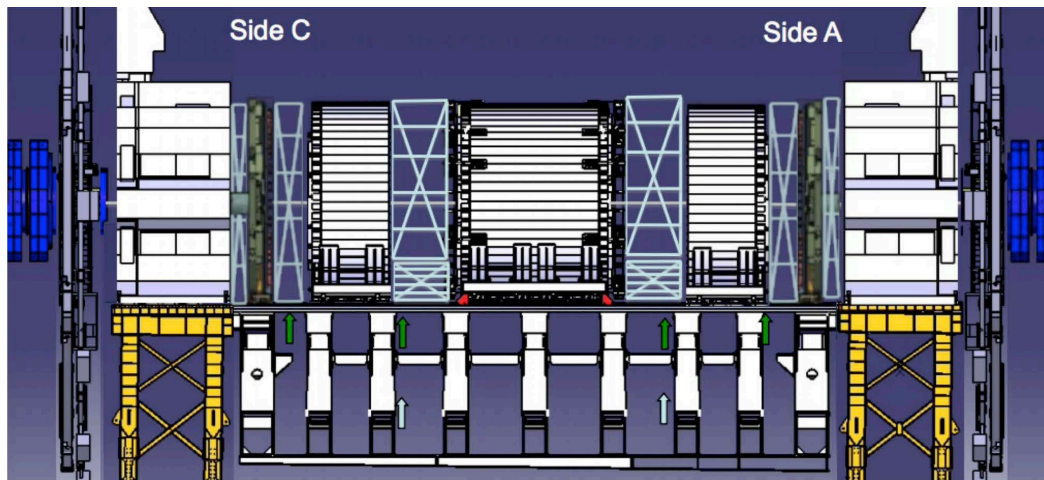


Figure 13.1: ATLAS in short opening configuration.

5690 13.2.2 Maintenance, radiation environment, and radio protection

5691 Dedicated simulation studies have not yet been performed to estimate the radiation levels
5692 expected in the region of the HGTD end-caps, during its installation planned for June–July
5693 2025, giving more than 500 days of cooling time after the LHC LS3 shutdown. During future
5694 YETS maintenance periods the access to the HGTD should occur typically after 28 days
5695 of cooling time. This is based on the assumption that each YETS is preceded by one week
5696 of technical stop and three weeks needed for the opening process of ATLAS. If the LHC
5697 run is terminated by 4 weeks of heavy-ion operation, the effective cooling time will be
5698 approximately 56 days.

5699 The radiation environment at installation can be evaluated using FLUKA simulations, which
5700 have been performed by ITk/RP in order to estimate the expected radiation levels during
5701 ID removal in LS3, assuming 297 fb^{-1} of accumulated data. The dose equivalent rate map
5702 after 28 days of cooling time is shown in Fig. 13.3, for the geometry corresponding to the
5703 completed large opening, with all beam pipes and inner detector removed. The calculations
5704 uncertainty have been estimated comparing the simulated and measured radiation levels
5705 during the Extended YETS in 2016–2017. From the results presented in Fig. 13.4, it can be
5706 seen that the simulations underestimate the actual radiation level, and this underestimation
5707 increases progressively with the distance from the beam pipe. The simulation results are
5708 about 50% lower than measured data at 1 m radius away from the beam pipe, close to the
5709 end-cap. This can be taken as an indication of the uncertainties of the simulations, which
5710 are used to predict dose levels in LS3, when the HGTD will be installed, and for future
5711 maintenance periods.

5712 The HGTD installation will take place after more than 500 days of cooling time; for that
5713 period a reduction of radiation levels by a factor of 2 to 3 from that shown in Fig. 13.3

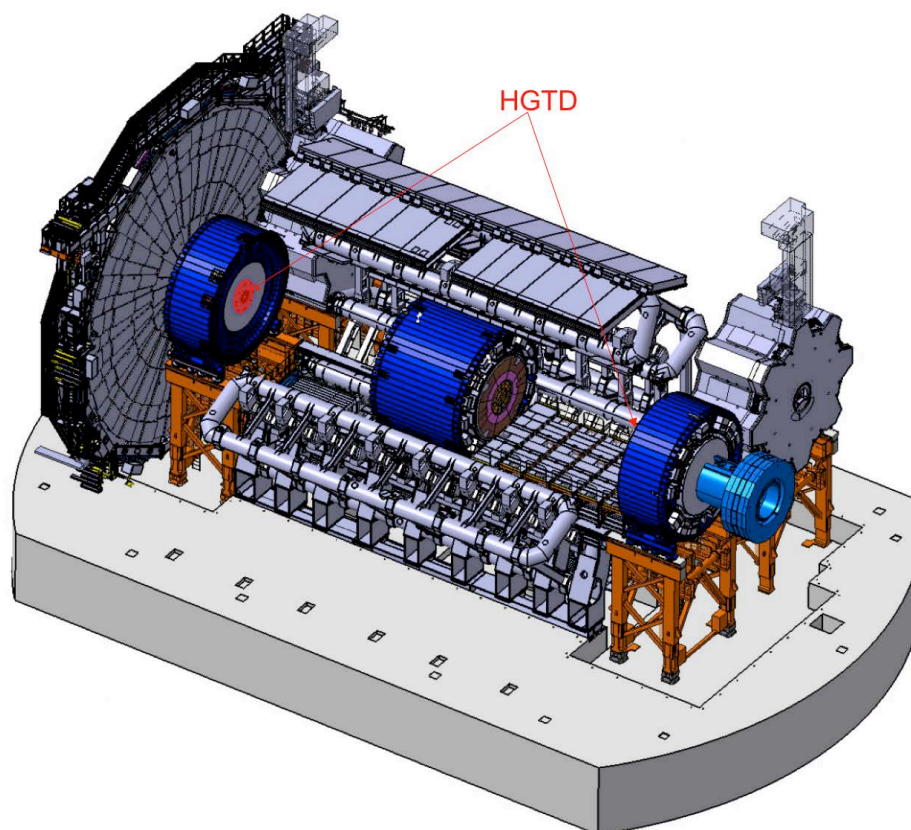


Figure 13.2: ATLAS in large opening configuration. The HGTD is superimposed to the MBTS scintillators that are presently installed on calorimeter end-cap cryostat.

5714 is expected. This reduction, obtained as a function of cooling time according to Sullivan-
5715 Overton formula [66] for a cooling time of 500 days, has yet to be confirmed by detailed
5716 simulations. Presently, considering the calculations uncertainties, we can take as reference
5717 the dose map shown in Fig. 13.3.

5718 The exact estimations of the dose levels at the annual shutdowns, to occur after 28–56 days
5719 of cool-down time, will determine which access and exact maintenance can be envisaged
5720 to the peripheral electronics siting inside the HGTD vessel. As an example, the presently
5721 available calculations give the estimate of the dose equivalent rate in front of HGTD in LS4
5722 after 56 days of cooling time in range from 35 to 150 $\mu\text{Sv h}^{-1}$, varying with distance from
5723 beam pipe. At such levels the access to any HGTD component inside the vessel may be
5724 compromised, subject to the detailed simulations still to be done in collaboration with CERN
5725 RP group.

5726 The replacement of the inner part of the detector half-way through the HL-LHC program
5727 should occur in LS5, after about 2000 fb^{-1} of accumulated data. The work will be done when
5728 ATLAS will be in long-open configuration with beam pipes removed, after relatively long

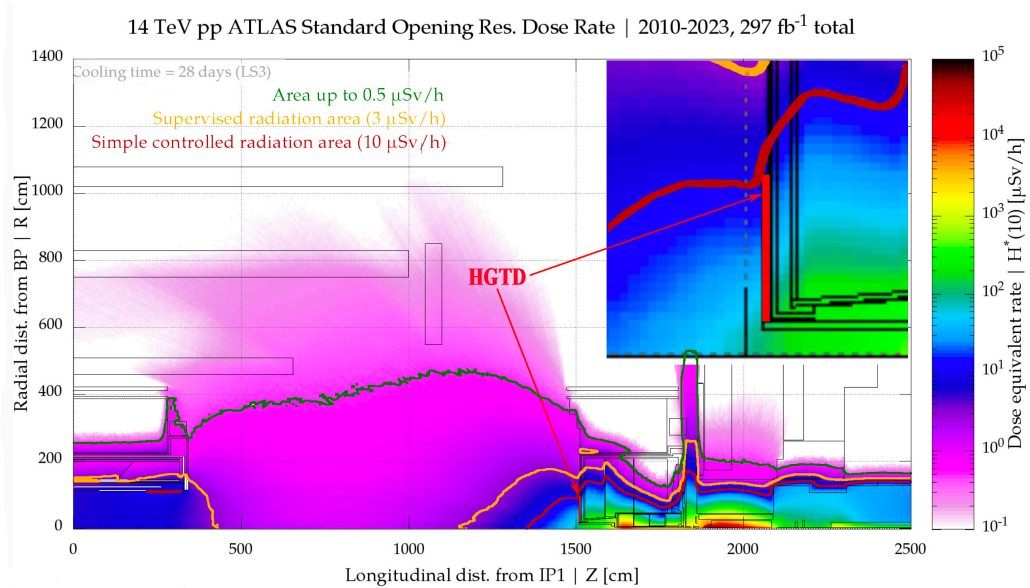


Figure 13.3: FLUKA simulations of ambient dose equivalent rate in LS3, after 297 fb^{-1} of accumulated data and 28 days of cool-down period. ATLAS is in the large opening configuration, all beam pipes and inner detector are removed. With coloured lines the boundary of various radiation areas are shown. This figure is the combination of plots given in Ref. [65].

5729 cooling time. Calculations dedicated to this configuration must be performed to estimate
 5730 the radiation environment during the works with reasonable accuracy. Available at the
 5731 present time FLUKA simulations of dose equivalent rate for LS5 period assuming 2177 fb^{-1}
 5732 of accumulated data have been done for short-opening geometry with beam pipes in place,
 5733 after 181 days of cool-down period, as shown in Fig. 13.5. In this configuration, the radiation
 5734 levels expected in the HGTD region should be in first approximation in range of 30 to
 5735 $50 \mu\text{Sv h}^{-1}$ (from outer to inner radius). When replacing the inner part of the detector, the
 5736 expected dose rates should be slightly lower due to longer cooling time and absence of beam
 5737 pipes. Nevertheless, it will be well above the threshold defining the simple controlled area
 5738 ($50 \mu\text{Sv h}^{-1}$), therefore the work duration will be severely limited.

5739 The replacement of detector modules on cooling supports will be done at the surface, and
 5740 will certainly require additional cooling time before accessing the components of the detector.
 5741 In order to minimise the radioactivity of the detector, less prone to activation materials must
 5742 be used in the construction, in particular by avoiding the use of stainless steel components
 5743 and replacing them as much as possible by those made of aluminium or plastic. First of
 5744 all, the possibility of manufacturing the aluminum pipes integrated in the cooling supports
 5745 should be considered.

5746 During all ATLAS upgrade and maintenance activities, as on the CERN site in general, the
 5747 radio-protection ALARA (As Low As Reasonably Achievable) principle should be strictly

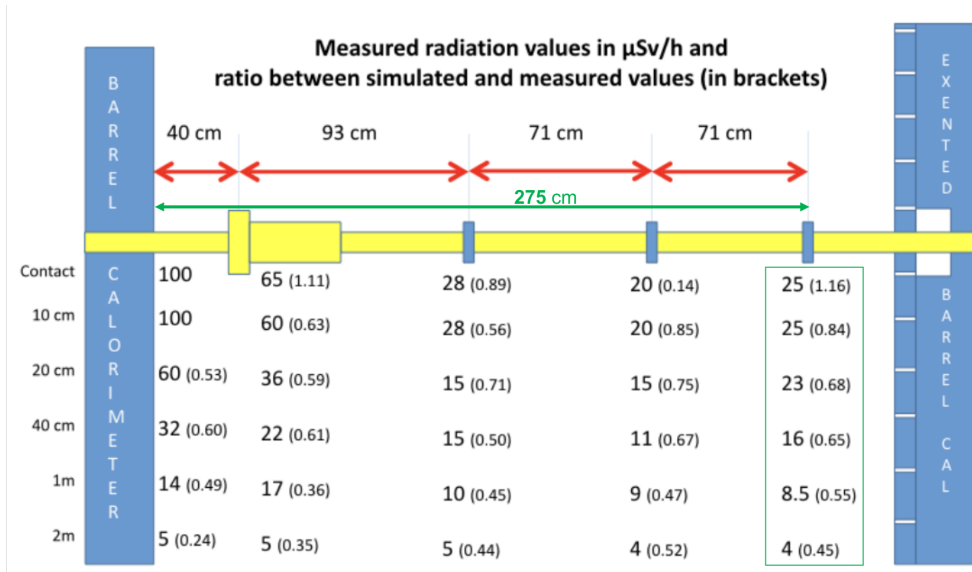


Figure 13.4: Comparison between simulated and measured radiation values in the region between the ID end-plate (in the left) and the End-cap Calorimeter (in the right, where the HGTD will be installed) during a standard opening scenario in the EYETS 2016–20117. The measured values, given in $\mu\text{Sv h}^{-1}$, have been taken on 15 December 2016 after 51 days of cool-down. The ratio of simulated over measured values is given in brackets [64].

5748 followed. It will certainly be implemented during the installation and maintenance activities
 5749 of the HGTD, in accordance with the rules and recommendations of the CERN Radiation
 5750 Protection service and in close collaboration with Technical Coordination.

Classification criteria	Level 1	level 2	level 3
Individual dose equivalent	<100 μSv	100 $\mu\text{Sv/h}$ - 1 mSv	>1 mSv
Collective dose equivalent	<500 μSv	500 $\mu\text{Sv/h}$ - 5 mSv	>5 mSv
Ambient dose equivalent rate	<50 $\mu\text{Sv h}^{-1}$	50 $\mu\text{Sv h}^{-1}$ - 2 mSv h^{-1}	>2 mSv h^{-1}
Airborne activity	<5 CA	5 CA - 200 CA	>200 CA
Surface contamination	<10 CS	10 CS - 100 CS	>100 CS

Table 13.1: ALARA classification criteria.

5751 It is expected that the HGTD installation zone will be classified at least as a “simple controlled
 5752 radiation area”, which is defined as the area whose ambient dose equivalent rate $H^*(10)$
 5753 does not exceed 10 $\mu\text{Sv h}^{-1}$ at workplaces or 50 $\mu\text{Sv h}^{-1}$ in low occupancy areas. All work
 5754 in controlled radiation areas must be planned and optimised including an estimate of the
 5755 collective dose and the individual effective doses to the personnel participating in the activity.
 5756 This should be described in the DIMR file (Dossier D’Intervention en Milieu Radioactif),
 5757 which must be prepared for each intervention. The Radiation Protection service will assign

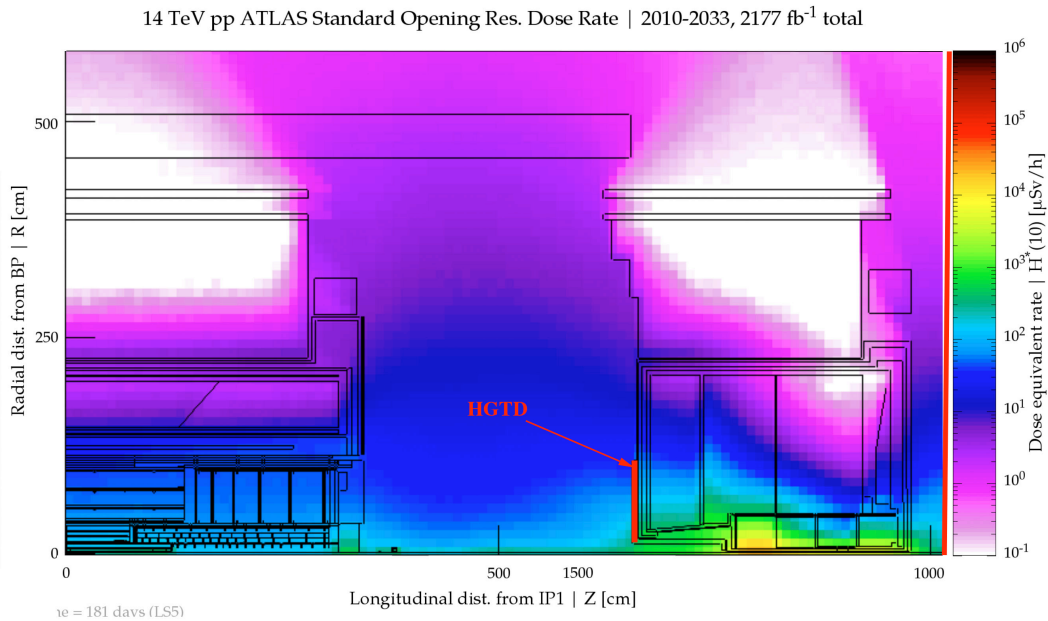


Figure 13.5: FLUKA simulations of ambient dose equivalent rate in LS5, after 2177 fb⁻¹ of accumulated data and 181 days of cool-down period. ATLAS is in the short opening configuration [67].

5758 an ALARA level to each type of activity, accordingly the CERN classification criteria, which
 5759 are shown in Tab. 13.1. Since the airborne radioactivity and contamination can be ruled out,
 5760 the ALARA level classification will be primarily determined by individual and collective
 5761 effective dose. As seen from the table, HGTD installation activity will be situated between
 5762 ALARA Level 1 and Level 2, considering the ambient equivalent dose. However, the
 5763 collective dose during replacement of the inner part of HGTD at half-life time of HL-LHC, on
 5764 both end caps, may approach the limit of 5 mSv, which corresponds to the Level 3 threshold.
 5765 In this case Level 3 scenario is applied, which involves additional optimisation efforts and
 5766 implies that dose planning and work organisation are reviewed by the ALARA committee.
 5767 DIMR level I and level II will be prepared and discussed between the intervening personnel
 5768 and ATLAS radiation safety officer (RSO) and GLIMOS prior to intervention, which can only
 5769 start when the DIMR is approved. All the activities will be followed by RSO and GLIMOS
 5770 on everyday basis, involving CERN Radio Protection experts when necessary.

5771 Beside the careful work optimisation, additional measures, which will help minimise the
 5772 exposition of personnel to radiation, should be considered. Among such measures are
 5773 provisions of shielding, which will reduce the dose rate to the human body; use of tools for
 5774 remote handling; organising working place in such a way, that people are placed in the outer
 5775 radius of HGTD avoiding exposure to the area near the beam line, where the dose rate is
 5776 much higher.

5777 13.2.3 Transport to the cavern and lowering

5778 After the pre-assembly on surface the internal moderator, the vessel and the half circular
5779 instrumented disks will be transported to the pit. Two scenarios for transport are being
5780 considered, either in separate pieces or transport the two fully assembled HGTD end-
5781 caps. In case of moving the fully assembled detector a total weight is 275 kg per end-cap,
5782 assuming that the external moderator part will always be transported separately. The overall
5783 dimensions are 1100 mm radius and 105 mm in thickness. These parameters should be taken
5784 into account for the transport truck and lowering, but they are well below the lifting capacity
5785 limit of the crane in ATLAS SX1 surface building and the dimensions of both shafts. Each
5786 end-cap, HGTD A and HGTD C, could be lowered on side A and side C, respectively, and
5787 lowered directly from the surface to the minivans that will be in place during the long
5788 shutdown LS2, needed also for the ITk installation. These minivans have already been used
5789 in LS1 for the replacement of the MBTS scintillators, as shown in Fig. 13.6 (where the HGTD
5790 disks are superimposed to the MBTS disks to give an idea of the overall dimensions).

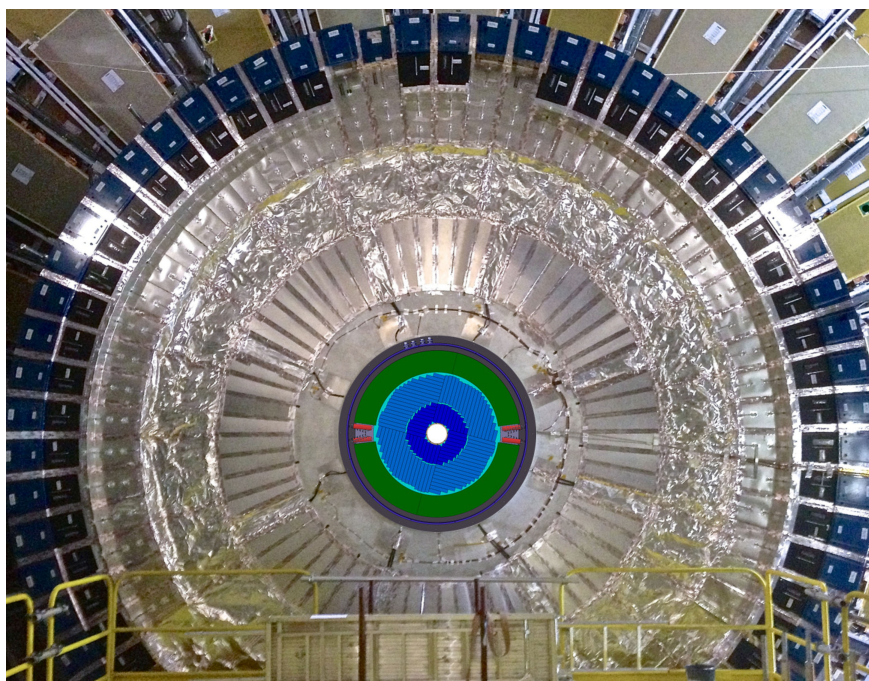


Figure 13.6: ATLAS in large opening configuration. HGTD detector superimposed on MBTS scintillators, that are currently installed on the LAr end-cap cryostat.

5791 Specific tools are needed to perform the transport, lowering and final installation of HGTD
5792 and to insure the positioning of the instrumented disks into the vessel. All these tools are
5793 still at a conceptual stage and will need to be carefully designed, and, where possible, use
5794 synergies with tools already developed for other sub-detectors.

5795 **13.2.4 Detector Installation and Commissioning**

5796 The installation of the detector will be done in the "Large Detector Opening" configuration as
5797 shown in Fig. 13.2. This operation can start only after the removal of the MBTS scintillators
5798 and the moderator that is presently installed in ATLAS.

5799 As already mentioned in the previous section, temporary access platforms will be in place,
5800 the same to be used for the installation of ITk. A local small lifting tool is needed to lift
5801 each component of the detector, or each fully assembled end-cap detector side that will
5802 weight 275 kg in total (excluding the external moderator piece of 75 kg that will be mounted
5803 separately). The accurate alignment of the detector with respect to ATLAS coordinates
5804 system will be based on the LAr cryostat central hole. The hermetic vessel inner ring, which
5805 is the direct interface to central hole, will be adjusted according to latest survey group
5806 measurements records. The hermetic vessel inner ring will be carefully assembled to avoid
5807 any conflict with the beam pipe ionic pump and its power connector as seen in Fig. B.6.

5808 The installation of each end-cap will take 1 month each and is planned for June and July
5809 2025 for the A side and C side, respectively.

5810 The overall commissioning will start immediately after the connectivity of the services to the
5811 detector, described in detail in Sec. 12.1.6. The access to the detector components during the
5812 commissioning should be possible until approximately March 2026, close to the expected
5813 end-cap calorimeters closure. This will leave at least 6 months of intense commissioning
5814 while access is still possible. Both the installation and commissioning of HGTD will be done
5815 with the participation of several collaborating Institutes.

5816 **14 Demonstrator**

5817 **14.1 Introduction**

5818 The R&D period will extend up to early 2022 to validate the choice of many components
5819 before the Final Design Review. In addition, it is essential to also validate some key aspects
5820 of the integration during this period, building a realistic demonstrator. The plan is to have a
5821 two step schedule decoupling the mechanics/cooling aspects from the full electronics/DAQ
5822 demonstrator activities. The heater demonstrator will be based on a silicon-based heater
5823 substrate to study the thermal performance of the system, instead of a real sensor and ASIC
5824 module. The full demonstrator will be similar to the heater demonstrator but equipped with
5825 some HGTD modules and read-out through a prototype of the peripheral electronics and
5826 back-end. A dedicated organisation is being set up to ensure coherence of the numerous
5827 parallel activities and monitor the schedule.

5828 **14.2 Heater demonstrator**

5829 The goals of this demonstrator are:

- 5830 • Use the simple cooling plate system to validate the CO₂ thermal calculation which will
5831 be used for the final design of the HGTD cooling loops.
- 5832 • Choose and validate the module loading procedure (intermediate plate, gluing, flex
5833 cable stacking...) by equipping the demonstrator with heaters in a geometry similar to
5834 the HGTD modules.

5835 The demonstrator will consist of a rectangular cooling plate covering about 6 cm × 800 cm
5836 as displayed in Fig. 14.1, corresponding to the longer stave in the HGTD. The cooling system
5837 will be made of a single loop (technical details given in Annex) embedded in a carbon fibre
5838 structure and will be used first to validate the thermal calculation of the CO₂ cooling on
5839 a simple design: CO₂ cooling parameters such as pressure and flow will be varied and
5840 the temperature on the plate will be measured with RTDs embedded into the heaters. A
5841 dedicated vessel should be also built, allowing dry nitrogen flushing and a feed-through for
5842 electrical connections. The convection conditions should be as close as possible to the final
5843 ones.

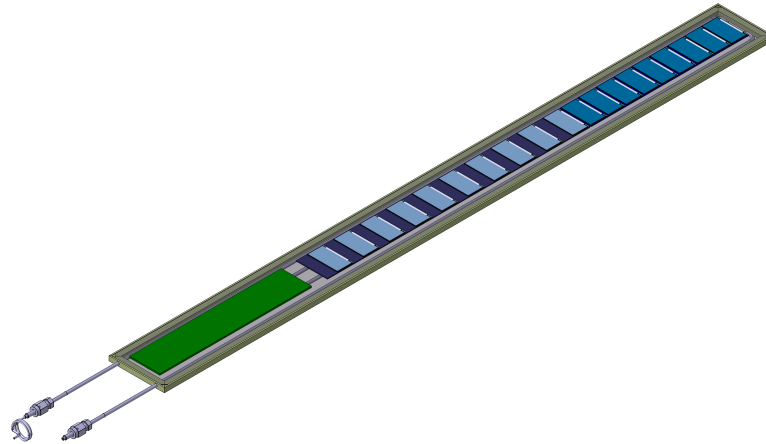


Figure 14.1: Schematic view of the cooling structure equipped with heater modules in blue. The green area corresponds to the peripheral electronics board.

5844 After this first set of measurements, a stave should be mounted on the top and bottom faces
 5845 of the cooling plates. Fig. 14.2 shows preliminary calculations of the temperature uniformity
 5846 for both options which will be compared to the measurements. As expected, the calculation
 5847 predicts a uniform temperature with the pattern intermediate plate, 0.4 K between inner and
 5848 outer module, while up to 1.8 K is observed with the full intermediate plate.

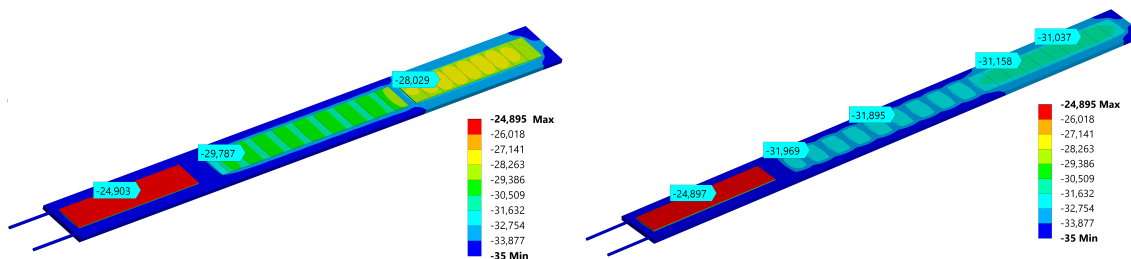


Figure 14.2: Expected temperature uniformity on the demonstrator equipped with the full intermediate plate (left) or the pattern intermediate plate (right)

5849 Real HGTD modules will not be available before 2020. Consequently to mimic the radial
 5850 heat dissipation expected in the HGTD, silicon heater devices similar to the ones used by the
 5851 pixel ITk demonstrator will be used for the module loading. Thus the heater demonstrator
 5852 program will address two important aspects of the HGTD system: module loading and
 5853 thermal performance. A schematic drawing of the silicon heater is shown in Fig. 14.3.

5854 The heaters consist of a silicon substrate with a similar geometry (area) of the modules



Figure 14.3: Heater transverse view

5855 and a thickness of 300 μm . A geometry slightly smaller than the final HGTD module was
 5856 chosen due to ease of production by the manufacturer. The heaters will have a size of
 5857 20.2 mm \times 38.4 mm. They will be made of a TiW continuous layer of size XXX produced
 5858 on a 300 μm thick wafer. The heaters dissipate power by applying a current through a
 5859 thin metal layer embedded in the silicon substrate. The amount of generated heat can be
 5860 controlled through the provided current. In order to monitor the temperature of the heater,
 5861 resistance temperature detectors (RTDs) are implanted on top of the thin metal separated by
 5862 an oxide layer. The RTDs will then be placed on top of an oxide layer separating the heater
 5863 from the RTDs, which will also be made from TiW. They are operated by applying a voltage
 5864 and reading the current which is previously calibrated to provide temperature information.
 5865 The RTDs are operated through a flexible cable that also provides the current to the heater
 5866 element. The flex is glued to the top of the heater and its pads are wire-bonded to the heater.
 5867 The heater flex PCB design can be found in Fig. 14.4.

5868 The heater flex will be designed to mimic the HGTD module flex cable in terms of geometry,
 5869 material and rigidity. It will contain a connector similar to the one being considered for
 5870 the final flex design, which can provide power to the heaters and individual readout lines
 5871 for the RTDs on each heater. The flex tails will be layered one on top of each other out to
 5872 the peripheral readout boards. Though the final specifications of the peripheral readout
 5873 boards will not be available, a compact connector scheme is foreseen. The system will be
 5874 controlled by external power supplies that will provide the desired operational thermal
 5875 range to fully study the system performance. The nominal power dissipation foreseen for
 5876 the innermost part of the heater stave is 400 mW cm^{-2} , but deviations from this value will be
 5877 explored. The entire heater demonstrator will be placed within an isolated container box to
 5878 maintain temperatures close to -30 degrees and allow for nitrogen or dry air to be flushed

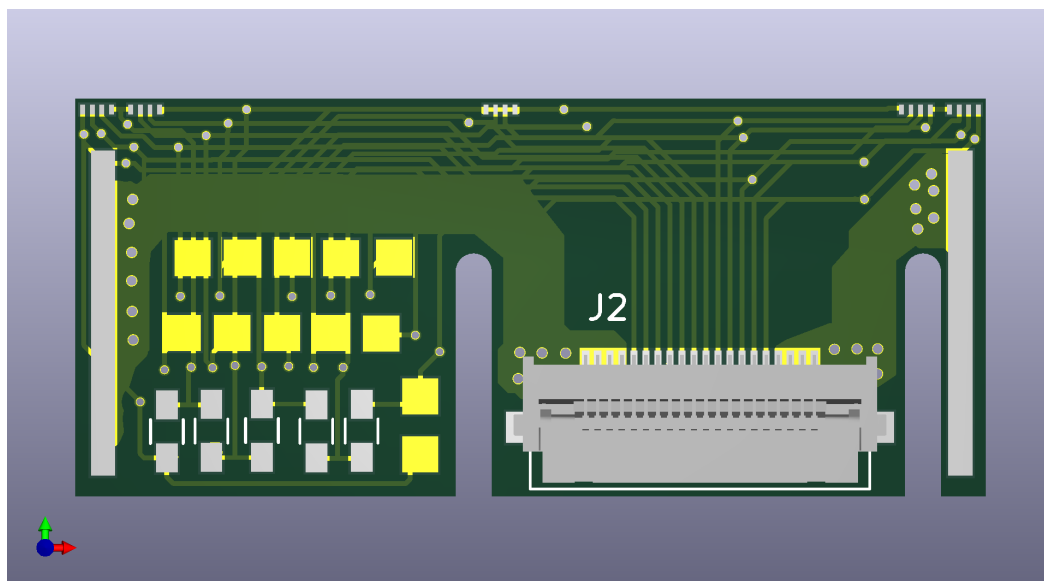


Figure 14.4: Heater flex PCB layout.

5879 into the apparatus to maintain a dry atmosphere. The CO₂ cooling will be provided by the
5880 CO₂ baby demo cooling plant, sitting nearby, as shown in Fig. 11.5. The design of the heater
5881 demonstrator apparatus can be found in Fig. 14.5.

5882 The institutes that plan to participate in the HGTD module assembly and loading effort will
5883 also participate in the heater (and/or full) demonstrator effort and will thus gain expertise
5884 on the module assembly process. The calibrations of the RTDs will also be carried out by
5885 the institutes, before and after module loading. The assembly of the intermediate plates
5886 around the cooling plane will be carried out at CERN, where the full cooling tests will be
5887 conducted.

5888 In summary, the heater demonstrator will allow to validate the thermal performance of the
5889 HGTD, by using heaters loaded into a long stave and combined with a CO₂ cooling system.
5890 Furthermore, the exercise of assembling the heater modules, populate the intermediate
5891 plates and mount the full heater demonstrator is expected to provide valuable experience
5892 towards the final HGTD stave assembly and loading effort.

5893 *Add information on production details - number of sensors ordered, vendor info? Date of arrival -*
5894 *date of end of studies*

5895 **14.2.1 Mechanical structure**

5896 *Describe the work on Didier Laporte - LPNHE. Different size than full size modules.*

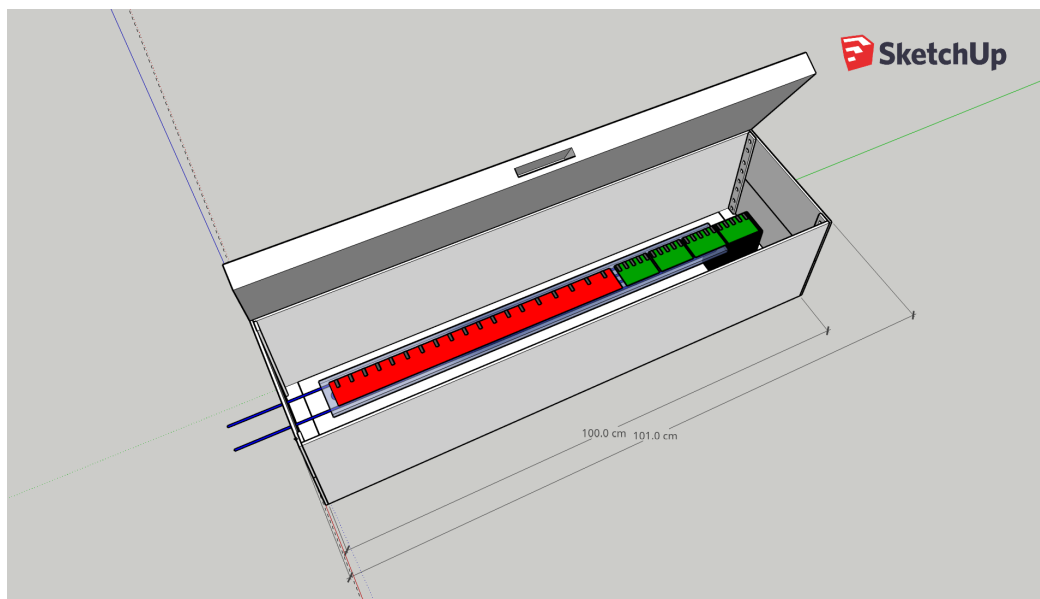


Figure 14.5: Heater apparatus. *To be updated by Afonso*

5897 14.3 Peripheral and back-end electronics, data acquisition

5898 This demonstrator will exercise the final HGTD read-out path and will be used to validate
5899 the PEB, the clock distribution and the FELIX board used for the data acquisition.

5900 14.3.1 Peripheral electronics demonstrator

5901 The peripheral electronics demonstrator will evaluate the different paths from the module
5902 flex to the PEB via flex cables like the data transmission, high voltages and the power
5903 distribution. In addition, it will exercise the assembling, connection and integration of the
5904 peripheral electronics. It consist of a PEB connected up to 55 HGTD modules via a stack of
5905 flex cables. In a first stage, an Spartan-7 FPGA will be used to emulate the ALTIROC2 ASIC
5906 and a Kintex-7 FPGA to emulate the lpGBT chipset, while the VTRx+ and the BPOL12V will
5907 be replaced by similar commercial components (SPF+ and TPS56428RHLR respectively),
5908 given the unavailability of the different items. An scheme of the peripheral electronics
5909 dremonstrator is show in Fig. 14.6. The design of the different items has already started and
5910 a peripheral electronics demonstrator will be ready by April 2020. On a second stage, the
5911 different components will be replaced by the ones of the final design and will be integrated
5912 in the full demonstrator set-up.

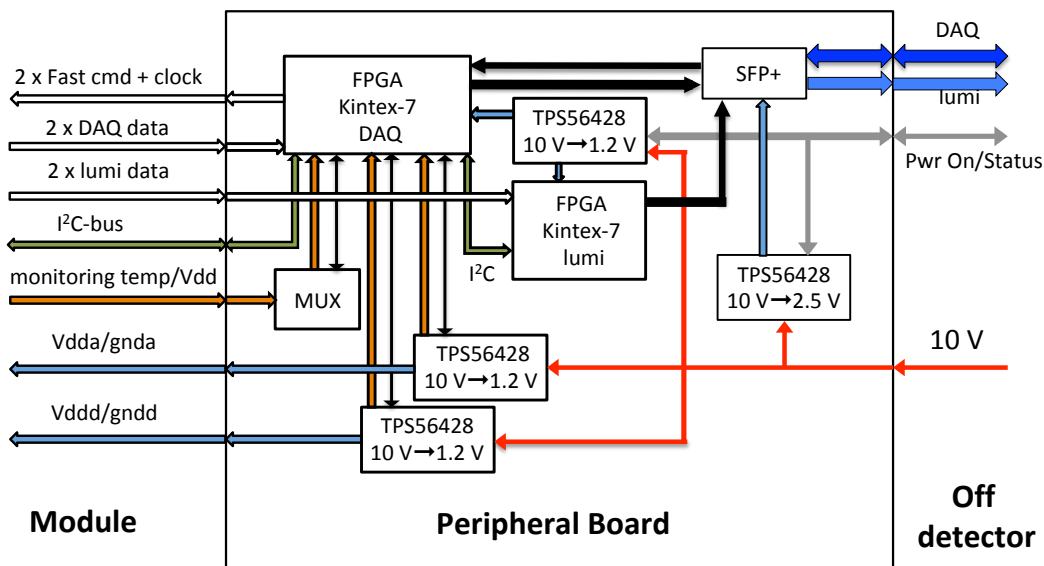


Figure 14.6: Block diagram of the peripheral electronics board demonstrator. A Kintex-7 FPGA will be used to emulate the lpGBT chipset, a SPF+ will replace the VTRx+ and a TPS56428RHLR will be used instead of BPOL12V DC-DC converter.

5913 14.3.2 DAQ demonstrator

5914 The DAQ demonstrator will exercise the entire read-out path up to the off-detector back-end.
 5915 Activities at CERN has already started and a Phase-I FELIX board and its DAQ PC have
 5916 been purchased. On a first stage, the HGTD e-link data will be emulated inside FELIX
 5917 in order to test the read-out chain. Afterwards, the FELIX board will be connected to an
 5918 FPGA emulator that will send HGTD data in FULL mode in order to validate the readout
 5919 chain. The ALTIROC2 FPGA emulator described in the previous section can be used for
 5920 this purpose. Finally, the ALTIROC2 will be connected to the readout chain, in this case
 5921 a GBT chip can be used to interface the FELIX board and the ASIC. On a second stage, a
 5922 Phase-II FELIX board will be purchased for the integration and validation of lpGBT, since
 5923 this protocol is not currently supported in the Phase-I FELIX board. The DAQ demonstrator
 5924 roadmap is shown in Fig. 14.7.

5925 Furthermore, the DAQ demonstrator will be used to measure the different contributions to
 5926 the clock jitter at different stages (FELIX, lpGBT, FLEX, ALTIROC2). On the other hand, it
 5927 will be used to develop a calibration procedure close to the final design. Finally, the DAQ
 5928 demonstrator will be integrated in the full demonstrator set-up.

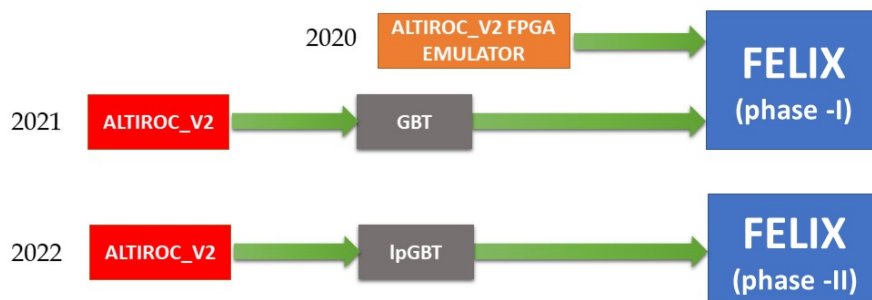


Figure 14.7: DAQ demonstrator roadmap. In 2020, an ALTIROC2 FPGA will be used to interface with FELIX. In 2021 an ALTIROC2 ASIC will be connected to FELIX using a GBT chip as interface. In 2021 a Phase-II FELIX board will interface the ALTIROC2 via lpGBT close to the final design.

5929 14.3.3 HGTD module

5930 The production of the HGTD modules will be used to validate the module assembly and
 5931 loading process (gluing, wire bonding and mounting) and quality control measurements
 5932 procedures used during the production.

5933 To get experience of this process, smaller bare modules will be assembled in house during
 5934 2019 using the ALTIROC1 ASICs and the existing 5×5 pads sensor. While for the test beam
 5935 purpose dedicated printed circuit boards have been developed and already used to test
 5936 the ASIC, it is also foreseen to develop a flex compatible with the ALTIROC1 read-out to
 5937 exercise the gluing and wire bonding of the bare module, as a first step of the validation of
 5938 the module assembly. Dedicated custom made readout boards will be used to validate these
 5939 modules, using the calibration signals and a beta source. These read-out boards might be
 5940 used on the demonstrator until the FELIX setup is operational.

5941 The bump bonding of the sensor to the ASIC will be outsourced in a company and require a
 5942 complete wafer for the under-bump-metalization process before the flip-chip. A specification
 5943 document has been prepared and is currently discussed with two companies in Germany
 5944 and China. Complete wafers will be available only after the production of ALTIROC2 and a
 5945 dedicated sensor production. The validation of the industrial bump bonding process will
 5946 be validated therefore only early Q3/2020. The possibility to produce the hybrids for the
 5947 demonstrator program in the HGTD institutes that have this capability in-house is also an
 5948 option. Between 5 to 10 bare HGTD modules are expected to be delivered by end of Q3
 5949 2020.

5950 Prototypes of the flex cable should also be produced, but the connector to the peripheral
 5951 board might still be not the final one.

5952 **14.4 Full demonstrator**

5953 The assembly of the demonstrator will start in Q4 2020. It will be made of :

- 5954 • The mechanical structure as used in the heater demonstrator, available by mid 2020.
- 5955 • Five to ten HGTD modules available by end of Q3 2020 and heater modules. A test of
5956 these modules after integration on the stave should be done using the custom made
5957 read-out board to qualify the modules.
- 5958 • At least one peripheral board able to read up to five HGTD modules connected through
5959 flex cables.
- 5960 • A peripheral electronics board and a FELIX board with its DAQ PC.
- 5961 • Prototypes of Low Voltage and High Voltage modules, with DCS, might be used but
5962 are not mandatory for this test.

5963 **14.5 Demonstrator tests**

5964 A period of about three months will be available before the first FDR. While intense elec-
5965 tronics calibration sequence tests will be performed, two options are investigated for the
5966 calibration sources : cosmic test bench with a precise trigger time measurement (although
5967 the rate might be insufficient) or a portable x-ray source (8 keV or 40 keV source) with a
5968 motorised stage to scan the stave.

5969 **14.6 Schedule and organisation**

5970 A dedicated WBS for the demonstrator is under construction and a tentative schedule of this
5971 demonstrator program is shown in Fig. 14.8. While the schedule for the heater demonstrator
5972 contains some contingency, the main risks for the full demonstrator rely on the availability of
5973 the modules in Q3 2020. This is strongly linked to the ASIC and sensor productions. From Q3
5974 2020, a weekly follow-up will be mandatory to fulfil this aggressive schedule. A dedicated
5975 working group will be set up in Q2 2019 with two coordinators, one mechanics/module
5976 oriented for the heater demonstrator and a second electronics/DAQ oriented. Beyond March
5977 2021, the demonstrator is expected to stay operational until the start of the production for
5978 additional tests.

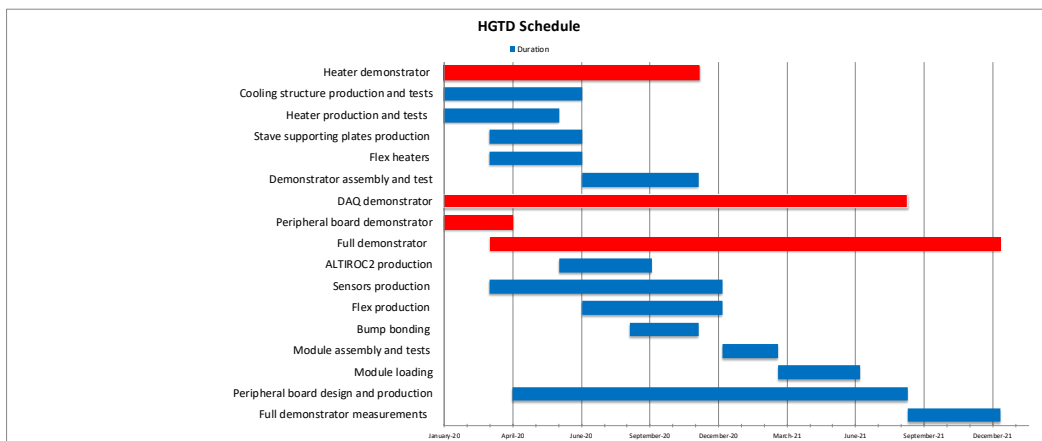


Figure 14.8: *UPDATE* Planning of the heater and full demonstrator from March 2019 to April 2021.

5979 **15 Project Organization, Costs, and Schedule**

5980 **This chapter will be extended in content (in particular in section 13.2-13.4 and the sched-**
5981 **ule re-worked/tuned assuming a new baseline that is to have only 1 layer per end-cap**
5982 **(instead of 2) in LS3 and the second one will be installed in the next YETS 2027. This**
5983 **will be done/completed in close collaboration with the ATLAS Project Management Of-**
5984 **ice (PMO), to be finished before the next TDR draft delivery . This chapter should also**
5985 **include a manpower table with needed and available resources.**

5986

5987 This chapter describes the overall organization of the HGTD project. Sec. 15.1 presents the
5988 way the project is organized and the management of the different activities, including a
5989 detailed breakdown for each component of the project. Sec. 15.2 discusses the schedule
5990 towards the detector completion. The foreseen available resources are discussed in Sec. 15.3.
5991 Finally, in Sec. 15.4 the risks involved with the project and the strategies to mitigate them are
5992 discussed.

5993 **15.1 Organization and management**

5994 **15.1.1 Upgrade organisation in ATLAS**

5995 The highest-level executive body in ATLAS is the Executive Board (EB), chaired by the
5996 Spokesperson with the Technical Coordinator (TC) as deputy chair. The overall steering and
5997 monitoring of the upgrade activities is delegated to the Upgrade Steering Committee (USC),
5998 which is a sub-committee of the EB, with an extended membership. The USC is chaired by
5999 the Upgrade Coordinator (UC). The review and approval of Upgrade Projects is steered by
6000 the UC and the USC, with approval of such projects by the EB, subject to endorsement by the
6001 Collaboration Board (CB). The UC also oversees and monitors the overall upgrade planning
6002 and schedules. The management of approved Upgrade Projects rests with the Upgrade
6003 Project Leader (UPL) of that Upgrade Project, acting together with the parent system's Project
6004 Leader and Institute Board chair. The Upgrade Coordinator should be well informed of the
6005 activities in the Upgrade Projects, and interacts regularly with the Upgrade Project Leaders
6006 to anticipate technical, schedule, resource, or other problems. The TC, supported by the
6007 Technical Coordination organization (TCn), is responsible for ensuring that all the upgrades
6008 can be successfully integrated in the ATLAS detector, that their installation schedules are

6009 compatible with shutdown schedules, and that there are adequate resources allocated for the
6010 installation and commissioning of the upgrade detectors. To this end the TC has organized
6011 an Upgrade Project Office (UPO) that provides technical support for the Upgrade Projects
6012 and the Upgrade Coordinator. Moreover the TC is responsible for the upgrade of all the
6013 common infrastructure needed for the upgrade program.

6014 The Review Office is an independent body embedded in Technical Coordination. In close
6015 collaboration with the UC, the TC, and the UPLs, the Review office develops and organizes
6016 technical reviews for the components of the upgrades following the ATLAS review strategy,
6017 comprising specifications, preliminary design, final design, and production readiness re-
6018 views.

6019 **15.1.2 HGTD organisation**

6020 The HGTD activity is currently managed by the ATLAS Liquid Argon unit but performed
6021 by Institutes not all belonging to the LAr unit yet so in addition to the LAr Institute Board
6022 the HGTD has also its own Institute Board. It started as an organized activity in summer
6023 2015 and this new sub-detector proposal was already part of the ATLAS Upgrade Scoping
6024 Document [68]. The HGTD Initial Design Report and Expression of Interest were approved
6025 by ATLAS and LHCC in 2017. The Technical Proposal was approved by LHCC in June 2018
6026 [69], with the recommendation to proceed to the Technical Design Report. Two Interim
6027 Upgrade Project Leaders (co-UPLs) have been chosen by the LAr management and endorsed
6028 by the LAr Institute Board. These two co-UPLs are part of the LAr steering group, and
6029 represent this HGTD project in the ATLAS Upgrade steering committee. They report to the
6030 LAr Project Leader and the UPLs, and they chair the HGTD steering group. Once the TDR is
6031 accepted, a search committee will be formed to identify the best candidates for UPL election.
6032 Either the LAr Institute or the HGTD Institute board will elect the UPL (one or two)

6033

6034 The project is organized in 8 working groups (WG):

- 6035 • **Sensors** : this WG is currently in charge of the R& D on sensors including irradiation
6036 tests with the aim of delivering the specifications of the final sensors. It works closely
6037 with the electronics WG as the expected performance relies strongly on the combined
6038 performance sensor+ASIC and with the testbeam WG. After the R& D Phase, it will
6039 have the charge of market survey and managing the production and QA tests.
- 6040 • **Electronics** this WG covers all electronics activities from the ASIC (design, specifica-
6041 tions, production and QA) to the Peripheral electronics boards (design, specifications,
6042 production and QA). It interacts with the sensors WG (for th ASIC specifications, High
6043 Voltage), the DAQ WG (data format, bandwidth) and the Module assembly (for the
6044 flex) and the Mechanics/assembly WP (CO2 cooling power, services).

- 6045 • **Luminosity and TDAQ** This WG covers the simulation studies and the specific hard-
6046 ware for the luminosity measurement and the TDAQ aspects (including the FELIX,
6047 and DCS). It make the interface with the ATLAS upgrade TDAQ and DCS project. A
6048 specific sub-group is in charge of studying and implementing the clock calibration
6049 (online and offline)
- 6050 • **Modules assembly and staves loading** This WG is in charge of defining the module
6051 assembly (bump bonding, gluing, flex specifications) specification, procedure and QA
6052 and the stave loading specification and QA.
- 6053 • **Test Beam** This WG is in charge of developing the needed tools for the testbeam
6054 (DAQ and hardware) and of the data analysis. It works closely with the sensors and
6055 electronics WG.
- 6056 • **Demonstrator** This WP is a transverse WP to all the other at the exception of Simu-
6057 lation/Performance. It will start its activity after the TDR delivery with the aim of
6058 building the demonstrator and validate the performance for the PDR of most of the
6059 components as described in chapter 12. This WG on long-term might take the charge
6060 of the commissioning of the final detector.
- 6061 • **Mechanics, assembly and installation** This WP is in charge of providing the specific-
6062 ations and building the vessels and cooling plates, the service definition and routine
6063 (with TC) an the CO2 cooling plant and N2 (with TC and CERN-DT). It should also
6064 design the tools needed for the assembly at surface and installation in the pit
- 6065 • **Simulation performance and physics** The role of this WG is to provide the most
6066 realistic simulation package and reconstruction tools (in interaction with the ATLAS
6067 Upgrade ITK simulation and performance and the Upgrade Physics group) to evaluate
6068 the performance on the object reconstruction and the impact on some physics channels.

6069 The coordinators of the Demonstrator WP will be appointed soon after the TDR submission.
6070 Each WG, coordinated by 2 to 3 co-coordinators, carry out several activities, as detailed in
6071 the current organization chart shown in Fig. 15.1. The level 2 activity coordinators will be
6072 appointed after TDR approval. All WP coordinators are members of the HGTD steering
6073 Group (SG) chaired by the two co-UPLs. HGTD general meetings are organized by the
6074 Interim UPLs and take place bi-monthly during 3-day mini-weeks. Topical meetings in each
6075 WG area are organized by the WG coordinators on a bi-weekly basis.

6076 The HGTD Institute board has one representative per institute with ex-officio the LAR PL
6077 and the UPL. During HGTD weeks joint steering group and Institute meetings are organized
6078 to discuss and endorse any strategic decision on detector layout, resource needs. The chair of
6079 the Institute Board is currently acting also as resource coordinator collecting and maintaining
6080 financial and manpower resource in close contact with all institutes. This has proven to be
6081 efficient during the R & D phase. A separate person will be appointed for this role after the

6082 TDR approval to help for the preparation of the MoU and the market survey (especially
 6083 sensors), working closely with the management team.

6084 The need of a technical coordinator after the TDR approval, or for the construction phase,
 6085 will be carefully evaluated, in particular if in the future organisation there will be only one
 6086 project leader instead of two PLs, as it was the case since the beginning of the R&D activities.
 6087 Depending on this choice a risk manager might also be appointed in order to develop and
 6088 maintain the risk register in coordination with the level-2 and 3 coordinators, and to track
 6089 and report any issue to the HGTD steering group.

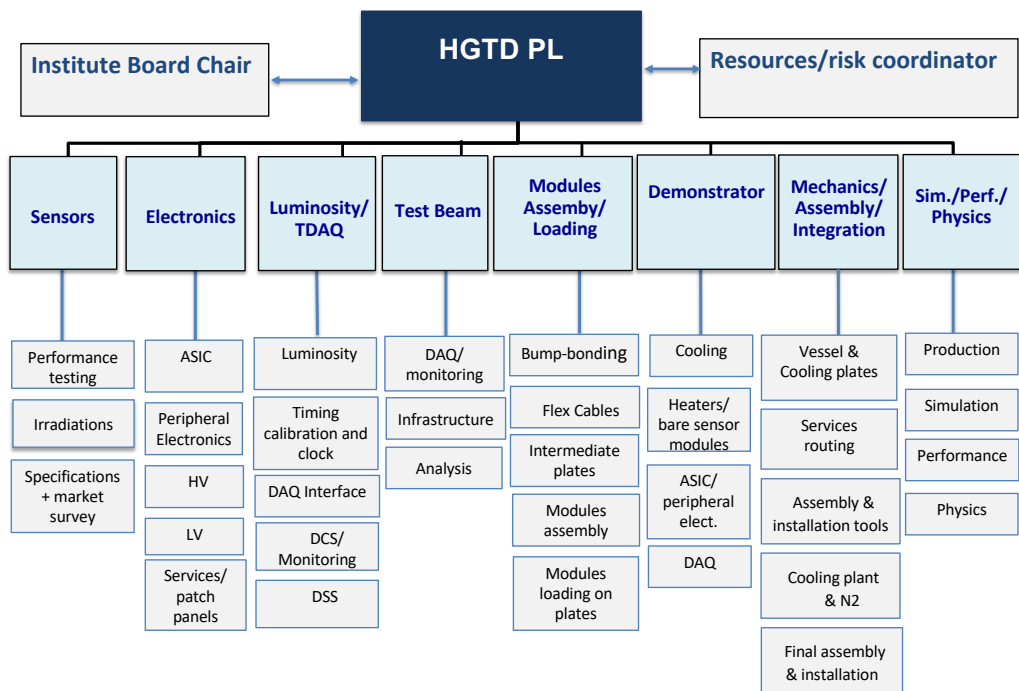


Figure 15.1: The HGTD organisation chart.

6090 The ongoing R&D is carried out by roughly 150 physicists, engineers and technicians from
 6091 24 ATLAS institutes from 12 Countries and 13 funding agencies, see Tab. 15.1. Those are
 6092 committed to carry out the R&D needed to mature the proposed detector and a sizable
 6093 fraction of the Institutes are already committed to the next steps of construction, Installa-
 6094 tion and Commissioning of HGTD. US groups have been quite involved in the R & D the
 6095 preparation of the TDR. After the delivery of the TDR some will stop their contribution
 6096 while the remaining US group will continue only the R& D phase up to early 2021 but will
 6097 not participate to the HGTD construction and part of the MoU. Sec. 15.1.2 summarizes the

Country /Funding agency	Institutes
Brazil	Sao Paulo Univ.
CERN	CERN
China (NSFC+MSTC)	IHEP, NJU, SDU, SJTU, USTC
France IN2P3	LAL, LPNHE, OMEGA, LPC
Germany BMBF	Mainz Univ., Giessen
JINR	JINR
Morocco	Univ. Hassan II
Slovenia	JSI
Spain	IFAE, CNM
Sweden	KTH
Taipei	AS, NTHU
USA (DOE+NSF+Univ. contributions)	BNL, SLAC, SMU, Ohio SU, UCSC UCSC, Iowa Univ., Stony Brook NY

Table 15.1: List of Countries, funding agencies and Institutes/Universities participating in the HGTD R&D. Some US groups will stop their contribution after TDR delivery while some others will extend their contribution up to the end of the R&D phase only.

6098 present involvement of the Institutes in the various R&D activities, planned until end of
2020.

R&D Activities/WG	Institutes
Sensors	BNL, CNM, CERN, Goettingen, IFAE, IHEP, JINR, JSI, NJU, USTC, SDU, SJTU, S. Paulo Univ., UCSC
Electronics	AS, Giessen, Hassan II Univ., IFAE, IHEP, Iowa Univ., JINR, KTH, LAL, LPC, NJU, NTHU, Omega, SDU, SLAC, SMU, Stony Brook NY, USTC
Luminosity and TDAQ	KTH, Ohio SU, UCSC, IHEP, Giessen, Iowa Univ.
Test beams	All institutes
Module assembly and staves Loading	BNL, IFAE, IHEP, JINR, LAL, LPNHE, Mainz Univ., USTC
Mechanics, assembly and Installation	CERN, IHEP, JINR, LAL, LPNHE
Computing-Software	AS, LAL, LPNHE, Mainz, NTHU, CERN, Hassan II Univ., KTH, SLAC, Taipei
Performance and Physics	All Institutes

6099

6100 15.1.3 Technical milestones

6101 All of the custom components used for the HGTD have to pass through a series of reviews
6102 before orders can be placed for procurement of parts and production of the deliverables, and
6103 before they can be used in the upgrade of the detector. These reviews are used to ascertain the
6104 quality and reliability of the components at various steps in the development and production

6105 process. They can also help to shorten the design phase, by enforcing in-depth presentations
6106 of the status at various stages. Reviews are conducted as (usually) half-day or full-day
6107 meetings of the group of people in charge of design and construction of the component with
6108 a team of reviewers. The review team is designated by the Upgrade Coordinator or by the
6109 Upgrade Review Office, and includes experts in the relevant technology, and, if applicable,
6110 users of the object to be reviewed or those interfacing other objects to it. This procedure is
6111 the ATLAS standard

6112 There are four main reviews for each custom component:

- 6113 • **Specifications Review (SPR)** This review is used to validate the specifications doc-
6114 ument, which describes the required functionality and performance of the device,
6115 its interfaces to other devices, tolerance to radiation, and reliability. The specified
6116 interfaces must be cross-checked for consistency with the corresponding component's
6117 specifications.
- 6118 • **Preliminary Design Review (PDR)** The PDR determines whether the design is sound
6119 and meets all requirements, including all interfaces to other components.
- 6120 • **Final Design Review (FDR)** The FDR is used to establish that the final prototype
6121 meets all requirements. Integrated tests with prototypes for the components the item
6122 interfaces directly to are required at this stage. A successful FDR gives the green light
6123 for a small pre-production.
- 6124 • **Production Readiness Review (PRR)** The results from pre-production are used to
6125 verify that larger scale production can be done with the acceptable yields, and that the
6126 quality control process is sufficiently thorough to filter out devices that will not meet
6127 the performance specification over the lifetime of ATLAS. After successful PRR, full
6128 production is launched.

6129 These reviews mark the transitions between different phases in each component's develop-
6130 ment and production schedule, and are thus used as key technical milestones in the overall
6131 project schedule.

6132 The co-coordinators of each WG are responsible for the preparation of the specifications
6133 and documentation, quality acceptance procedures, and material to be delivered to the
6134 reviews. Each individual component that will be built into the HGTD must have a written
6135 specification. The progress through the reviews is also used to monitor the progress of the
6136 project and to make sure it is on track. Production procurement, especially for large quantity
6137 items, will require a production plan and must follow procurement procedures required by
6138 the purchasing Institution. The CERN purchasing office will probably be in charge of large
6139 quantity items, to be CORE shared by several Funding Agencies.

6140 15.1.4 Deliverables and WBS

6141 The deliverables for the construction of the HGTD are organized in an hierarchical Product
6142 Break Down Structure (PBS), with a direct correspondence to the existing first 5 listed WG
6143 activities. The PBS organises the deliverables into 7 primary (L2) categories, with PBS
6144 numbering from 8.1 to 8.7, as described in Tab. 15.2. When appropriate PBS is further
6145 broken down into LV3 deliverables. At the lower levels, in particular in PBS item 8.6
6146 (Detector Assembly and QA on surface) and 8.7 (Detector Installation and Commissioning),
6147 the structure contains also the activities needed to build the deliverable. This part is referred
6148 as Work Breakdown Structure (WBS). All PBS items, except a few items that are considered
6149 WBS, have an associated CORE cost as described later in Sec. 15.3. The Involvement in
6150 CORE is discussed later in Sec. 15.3.

6151 15.2 Schedule and production schedule milestones

6152 The overall ATLAS installation schedule for Long Shutdown 3 provides constraints for the
6153 scheduling of the installation of the HGTD. In the ATLAS LS3 schedule the HGTD will be
6154 installed on the end-cap LAR calorimeter cryostat faces during one month for each end-cap,
6155 in June and July 2025 for the A and C side respectively. The closing of the calorimeter
6156 end-caps is scheduled approx. 6 months later, in January 2026. The design of the HGTD
6157 allows for a baseline schedule, detailed below, that assumes an installation of the two end-
6158 caps (HGTD-A and HGTD-C) with only one double sided layer per end-cap. The second
6159 double sided layer will be installed in the YETS 2027, in situ, sliding the front vessel cover
6160 along the beam pipe (properly protected) and installing the second layer in the shape of
6161 half circular disks inside the detector. The HGTD mechanical supports, external moderator,
6162 vessel, services, cooling station needed to operate the complete detector (with 2 layers per
6163 end-cap) need to be installed in LS3.

6164 There are three main schedule phases for HGTD:

- 6165 • **2018-2020** R&D
- 6166 • **2021-2024** Construction
- 6167 • **2025-2026** Integration, installation and commissioning

6168 For defining the schedule of the HGTD Project, a detailed bottom-up planning of activities
6169 has been performed. The schedule comprises the tasks that need to be undertaken between
6170 now and the completion of the project, and their dependencies, i.e. lists of tasks that have to
6171 be finished before a new task can begin. As the project moves forward, tasks will be broken
6172 down into sub-tasks of shorter duration for project tracking, so that delays can be spotted
6173 early and preventive actions can be taken if the project goes off track. Each sub-project

PBS/WBS	Deliverable
8.1	Sensors
8.2	Electronics
8.2.1	ASIC
8.2.2	Peripheral Electronics
8.2.3	High Voltage power supplies and crates
8.2.4	Low Voltage power supplies and crates
8.3	Luminosity/TDAQ (*)
8.3.1	Luminosity boards
8.3.2	DCS and interlocks
8.4	Module assembly and staves Loading
8.4.1	Bump-bonding ASIC/Sensor
8.4.2	Flex cables
8.4.3	Modules assembly
8.4.4	Modules loading on staves/plates
8.5	Mechanics, Services and Infrastructure
8.5.1	HGTD hermetic vessel
8.5.2	Moderator (**)
8.5.3	On detector cooling/support plate
8.5.4	CO ₂ /water Cooling and N ₂ systems
8.5.5	Tools for assembly and installation
8.5.6	Services (cables, connectors, fibres, pipes)
8.5.7	Patch panels
8.6	Detector Assembly and QA on surface
8.6.1	Test bench for detector certification
8.6.2	Components assembly on cooling plates (1 layer/EC) for LS3
8.6.3	Final integration inside 2 vessels (1 layer/EC) for LS3
8.6.4	Assembly of components for second layer/EC (for YETS 2027)
8.7	Detector Installation and commissioning
8.7.1	Services, patch panels and cooling installation
8.7.2	Back-end electronics installation in USA15
8.7.3	External moderator installation
8.7.4	Detector installation (w/ 1 layer/EC) and connectivity
8.7.5	Global commissioning in LS3
8.7.6	Install in-situ second layer/EC (in YETS2027)
8.7.7	Global commissioning in YETS2027

Table 15.2: Product Breakdown Structure (PBS) and Work Breakdown Structure (WBS) of the HGTD down to level 3. The PBS indicates the deliverables, to be assigned to a CORE value in the Memorandum of Understanding (MoU). The Work Breakdown Structure (WBS) is seeded by the PBS and includes the tasks required to produce the deliverables, those are mentioned explicitly when appropriate. (*) TDAQ related deliverables are not included in HGTD PBS and the corresponding costs are accounted separately in the HGTD CORE table. After TDR approval it should be added in TDAQ MoU. (**) The Core costs associated to the moderator (located inside and outside the HGTD vessel) is part of the ATLAS ITK common items.

6174 schedule includes an overview part that summarise the schedule in distinct phases. The start
6175 points and end points of these phases are delimited by appropriate high level milestones:

- 6176 • start of the design phase: SPR;
- 6177 • start of the prototyping phase: PDR;
- 6178 • start of the pre-production phase: FDR;
- 6179 • start of the production phase: PRR;
- 6180 • end of the production phase: Construction Completed;
- 6181 • end of the installation and commissioning phase: Installation Completed.

6182 The main production schedule milestones are included in the overall HGTD schedule,
6183 presented in Fig. 15.2 and Tab. 15.3. They use the PBS/WBS structure, detailed to level 3 as
6184 described in Tab. 15.2.

6185 The schedule takes into account realistic quantities for each component, accounting for
6186 inefficiencies in all the production steps until the final assembled detector. When possible
6187 activities are taking place in parallel, for those that need to be done in a sequential way it is
6188 given a time window of at least a few months before the start of the next activity.

6189 The schedule assumes a production model for the main components as follows.....
6190**DETAIL HERE THE PRODUCTION MODEL OF MODULES, SENSORS, ASICS, (THE**
6191 **5% PRE-PRODUCTION WILL GO IN THE DETECTOR) ; 10790 PRODUCED MOD-**
6192 **ULES.VS> 7984 NOMINAL QUANTITY (80% YIELD) IN MOD ASSEMBLY; ASICS:**
6193 **2x10790/.80 = 26950.VS. NOMINAL SICS QUANTITY=15968 , ASSUME 2 MODULES**
6194 **PER HV CHANNEL AT THE START AND 1 DOUBLE SIDED LAYER/EC INSTALLED**
6195 **IN LS3 AND THE SECOND ONE IN YETS 2027.....**

6197 The main schedule Milestones are detailed below.

6199 **Sensors (item 8.1)**

6200 ..add text with tuned dates....

6202 **Electronics (item 8.2)**

6203 ..add text with tuned dates....

6205 **Luminosity/TDAQ (item 8.3)**

6206 ..add text with tuned dates....

6207

6208 **Modules assembly and staves loading(item 8.4)**

6209 ..add text with tuned dates....

6210

6211 **Mechanics, Services and Infrastructure (item 8.5)**

6212 ..add text with tuned dates....

6213

6214 **Detector Assembly and QA on surface (item 8.5)**

6215 ..add text with tuned dates....

6216

6217 **Detector Installation and commissioning (item 8.7)**

6218 ..add text with tuned dates....

6219

6220 The schedules for the L2/L3 deliverables are defined by the sub-project coordinators and
6221 approved by the SG. It is the responsibility of sub-project coordinators to plan, implement,
6222 execute, and track the progress of their project against the baseline schedule for their re-
6223 spective deliverables. They report on the progress to the SG. It is the responsibility of the
6224 HGTD UPL to ensure that a comprehensive schedule is developed, to seek the necessary
6225 review process to baseline the schedule, to oversee the progress and take necessary corrective
6226 actions to ensure that the project remains on schedule, and to propose changes to the baseline
6227 as required.

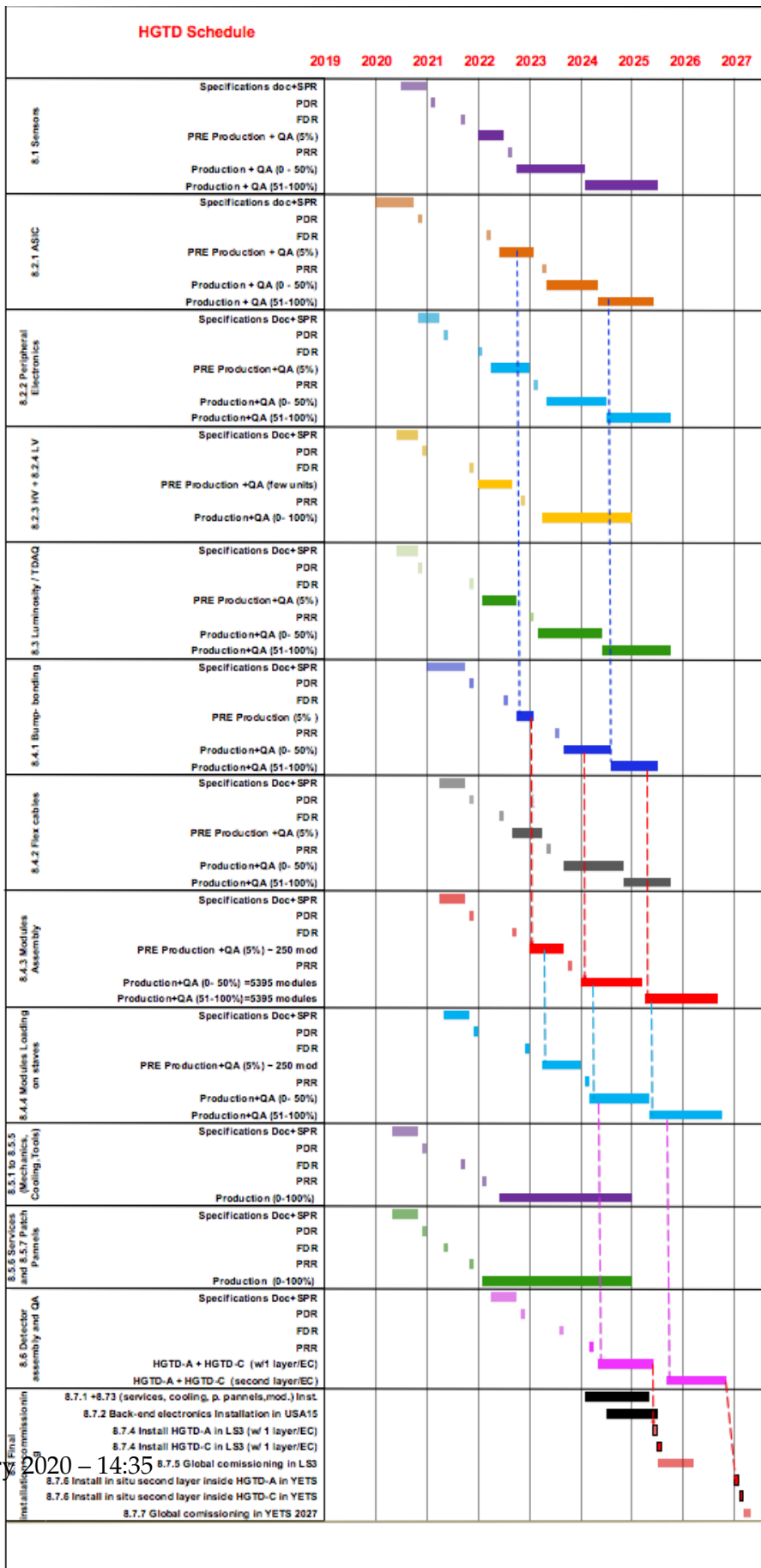
PBS-Deliverable	Milestone	Start	End
8.1-Sensors	Specifications doc. +SPR	1 Jul 20	30 Dec 20
	PDR	Q1 21	Q1 21
	FDR	Q3 21	Q3 21
	Pre-production	1 Jan 22	30 Jun 22
	PRR	Q3 22	Q3 22
	Production (0-50%)	1 Oct 22	30 Jan 24
	Production (51-100%)	1 Feb 24	30 Jun 25
8.2-Electronics 8.2.1-ASIC	Specifications doc.+SPR	1 Jan 20	30 Sep 20
	PDR	Q4 20	Q4 20
	FDR	Q2 22	Q2 22
	Pre-production	1 Jun 22	30 Jun 23
	PRR	Q3 23	Q3 23
	Production (0-50%)	1 May 23	30 Apr 24
	Production (51-100%)	1 May 24	30 May 25
8.2.2-Peripheral Electronics	Specifications doc.+SPR	1 Nov 20	30 March 21
	PDR	Q2 21	Q2 21
	FDR	Q1 22	Q1 22
	Pre-production	1 April 22	31 Dec 22
	PRR	Q1 23	Q1 23
	Production (0-50%)	1 May 23	30 June 24
	Production (51-100%)	1 July 24	30 Sep 25
8.2.3 and 8.2.4 (HV+LV in USA15/UX15)	Specifications doc.+SPR	1 Jun 20	30 Oct 20
	PDR	Q4 20	Q4 20
	FDR	Q4 21	Q4 21
	Pre-production	1 Jan 22	1 Sep 22
	PRR	Q4 22	Q4 22
	Production (0-100%)	1 April 23	30 Dec 24
	8.3-Luminosity/TDAQ 8.3.1 and 8.3.2 (Lumi boards+DCS/interlocks)	Specifications doc.+SPR	1 Jun 20
PDR		Q4 20	Q4 20
FDR		Q4 21	Q4 21
Pre-production		1 Feb 22	1 Oct 22
PRR		Q1 23	Q1 23
Production (0-50%)		1 Mar 23	30 May 24
Production (51-100%)		1 Jun 24	30 Sep 25
8.4-Module assembly+loading in staves 8.4.1- Bump-bonding	Specifications doc.+SPR	1 Jan 21	30 Sep 21
	PDR	Q4 21	Q4 21
	FDR	Q4 22	Q4 22
	Pre-production	1 Oct 22	30 Jan 23
	PRR	Q3 23	Q3 23
	Production (0-50%)	1 Sep 23	30 Jul 24
	Production (51-100%)	1 Aug 24	30 Jun 25
8.4.2-Flex cables	Specifications doc.+SPR	1 April 21	30 Sep 21
	PDR	Q4 21	Q4 21
	FDR	Q2 22	Q2 22
	Pre-production	1 Sep 22	30 Mar 23
	PRR	Q2 23	Q2 23
	Production (0-50%)	1 Sep 23	30 Oct 24
	Production (51-100%)	1 Nov 24	30 Sep 25
8.4.3-Modules assembly	Specifications doc.+SPR	1 Apr 21	30 Sep 21
	PDR	Q4 21	Q4 21
	FDR	Q3 22	Q3 22
	Pre-production	1 Jan 23	30 Aug 23
	PRR	Q4 23	Q4 23
	Production (0-50%)	1 Jan 24	15 Mar 25

	Production (51-100%)	1 Apr 25	1 Sep 26
8.4.4-Modules loading on staves	Specifications doc.+SPR	1 April 21	30 Oct 21
	PDR	Q4 21	Q4 21
	FDR	Q4 22	Q4 22
	Pre-production	1 April 23	30 Dec 23
	PRR	Q1 24	Q1 24
	Production (0-50%)	1 Mar 24	30 April 25
	Production (51-100%)	1 May 25	30 Sep 26
8.5-Mechanics, Services, Infrastructure	Specifications doc.+SPR	1 May 20	30 Oct 20
	PDR	Q4 20	Q4 20
	FDR	Q3 21	Q3 21
	PRR	Q2 22	Q2 22
	Production (0-100%)	1 Jun 22 ?	30 Dec 24
8.5.6 and 8.5.7 (services+patch panels)	Specifications doc.+SPR	1 May 20	30 Oct 20
	PDR	Q4 20	Q4 20
	FDR	Q2 21	Q2 21
	PRR	Q4 21	Q4 21
	Production (0-100%)	1 Feb 22 ?	30 Dec 24
8.6 Detector Assembly and QA on surface	Specifications doc.+SPR	1 Apr 22	30 Sep 22
	PDR	Q4 22	Q4 22
	FDR	Q3 23	Q3 23
	PRR	Q1 24	Q1 24
	Production w/ 1L/EC(0-50%)	1 May 24	30 May 25
	Production (51-100%)	1 Sep 25	30 Oct 26
8.7 Installation and commissioning			
8.7.1 and 8.7.3 (Services,p. panels,cool.,mod.)	Installation+QA (0-100%)	30 Jan 24	30 Apr 25
8.7.2 Back-end electronics inst. in USA15	Installation+QA (0-100%)	1 Jul 24	30 Jun 25
8.7.4 HGTD-A (w/ 1 layer)	Installation	2 Jun 25	2 Jul 25
8.7.4 HGTD-C (w/ 1 layer)	Installation	3 Jul 25	1 Aug 25
8.7.5 Commissioning in LS3 (w/ 1L/EC)	Commissioning	3 Jul 25	10 Mar 26
8.7.6 HGTD-A (w/ layer 2) in YETS27	Install in situ layer 2	1 Jan 27	30 Jan 27
8.7.6 HGTD-C (w/ layer 2) in YETS27	Install in situ layer 2	1 Feb 2027	2 Mar 27
8.7.7 Commissioning in YETS27 (w/ 2L/EC)	Commissioning	1 Mar 27	30 May 27

Table 15.3: Schedule of the main HGTD deliverables IN LS3 AND YETS 2027, including the planned reviews (SPR, PDR, FDR, PRR), Pre-production and Production, assuming as baseline only 1 double sided layer per end-cap and the second one to be installed in YETS 2027. **TUNE NEW SCHEDULE, ASSUMING AS BASELINE 1 LAYER/EC IN LS3 AND SECOND LAYER/EC IN YETS 2027.**

6228 15.3 Costs

6229 The cost of deliverables for ATLAS projects are expressed as their CORE cost. CORE is
6230 defined as the sum of the material value of each component that makes up the deliverable.
6231 The cost of generic infrastructure, prototypes, and spare components are all excluded by
6232 definition from the CORE costing, as is the cost of existing manpower, such as labour or
6233 travel for personnel employed by HGTD Institutions. Specialized infrastructure, such as
6234 custom-designed tooling, is included in CORE. For items bought in industry, the material
6235 value is simply their selling price, and depending on how the vendor calculated this price,



1st January 2020 – 14:35

275

Figure 15.2: Chart showing the main HGTD of the main HGTD deliverables, including the planned reviews (PDR, FDR, PRR), Pre-production and Production. **OPTIMIZE THIS NEW SCHEDULE MADE WITH ONLY 1 DOUBLE SIDED LAYER/EC IN LS3 AND THE SECOND LAYER/EC IN**

6236 it includes some unknown fraction of labour cost at the company. This type of labour
6237 cost is included in CORE. The CORE cost of a project does not represent its full cost, and
6238 Institutions participating in HGTD have to request funds to cover both CORE and non-CORE
6239 expenditure from their Funding Agencies, in a ratio that varies from country to country.

6240 Each HGTD PBS item, described in Tab. 15.2, has an associated CORE cost that is defined
6241 as the sum of the material value of all components making up the deliverable. A Work
6242 Breakdown Structure (WBS) is seeded by the PBS and adds to it the tasks that need to be
6243 performed to design, prototype, produce, assemble, and install the deliverables. The HGTD
6244 CORE cost has been estimated in a bottom-up approach. The cost estimates are detailed
6245 with individual elements for most items. The numbers of items have been calculated based
6246 on the layouts and schemes presented in this document.

6247 The Yield model used in the cost estimates accounts for failure and loss during the production
6248 phase, up to and including the installation of items in the ATLAS cavern. In contrast, spares
6249 account for failure and loss during the operations phase, i.e. from the beginning of Run 4
6250 onwards for items installed in the cavern during LS3. Yield is supported by upgrade funds
6251 and counts toward CORE cost, while spares are supported by maintenance and operations
6252 (M&O) funds and do not count as CORE.

6253 For each significant production step, the yield was estimated based on past production
6254 experience with similar or equivalent items or extrapolating from prototypes experience
6255 and a summary of the production model is shown in ???. It assumes that the total of the
6256 HGTD pre-production components of 5% quantity is of good quality and will be used in
6257 the detector. It is also considered that each HV channel will be reading 2 modules (2x2x225
6258 channels), at least in the initial phase when the irradiation levels are smaller.

6259 **ADD HERE A TABLE SUMMARIZING THE YIELD MODEL OF THE MAJOR COM-**
6260 **ONENTS (ASICS,SENSORS,MODULES) AND FEW LINES OF TEXT TO DESCRIBE**
6261 **THIS NEW TABLE.**

6262 The cost estimates for each item are quoted in CHF, using the exchange rates of 1Euro =
6263 1.085 CHF, 1 USD = 0.986 CHF and 100 JPY = 0.942 CHF, as in the other ATLAS phase II
6264 TDRs. The core estimates are based on existing contracts (ASICs), quotes from Industry
6265 (sensors, FPGAs, Flex cables,...), extrapolation from other ATLAS Upgrade Phase-II TDRs
6266 with already signed Memorandum of Understanding (MoU) that are using the same or
6267 similar type of components (power supplies, cables, cooling station used in the ITk).

6268 A summary of the HGTD core cost¹, detailed to the PBS level 3 is presented in Tab. 15.4, with
6269 a total of 9716 KCHF for the HGTD and 995 KCHF for the HGTD-TDAQ related costs.

6270 The costs for the planned replacement of the HGTD inner ring during the HL-LHC half life
6271 time should be accounted in the future (M&O) funds. Assuming that the outer radius of

¹ The item "Detector readout, data flow, and network" is considered a TDAQ deliverable but cannot be included in the TDAQ TDR until the HGTD TDR has been reviewed by LHCC/UCG.

6272 the inner ring will be kept at 320 mm the costs are estimated to be approximately 30% of
6273 the items: 8.1 (sensors), 8.2.1 (ASIC), and 8.4 (module assembly and loading on staves). The
6274 overall costs of maintenance and rolling replacement of hardware after its installation are
6275 also foreseen as part of the future HGTD maintenance and operation budget. The HGTD
6276 resources coordinator will review the HGTD costing in close collaboration with the experts
6277 of the different deliverables and the ATLAS resources coordinator. The cost of the project
6278 is expected to be covered by the Institutions participating in the HGTD Phase-II upgrade
6279 project, with their respective Funding Agencies. The details of responsibility and sharing
6280 among Institutes will be defined in an MoU to be prepared after the TDR approval. A
6281 preliminary survey of Institutes interests and resources indicate that a substantial fraction
6282 of the money and manpower resources required is already covered and with a balanced
6283 sharing by Institutes among the various PBS/WBS deliverables that are needed to construct
6284 and operate HGTD.

6285 15.4 Risk management

6286 **THIS SUB-CHAPTER STILL NEEDS TO BE COMPLETED WITH A RISK MANAGE-**
6287 **MENT ANALYSIS IN TERMS OF COST, SCHEDULE AND PERFORMANCE.**

6288

6289 Many of the technical choices in the HGTD concept were made already at the time of the
6290 Expression of Interest and Technical Proposal for the best compromise between performance
6291 and cost. The severe space constraints (in z , r), high radiation levels (still evolving with
6292 the finalization of the ITk services and supports layout) and the limited time available to
6293 implement the project in the LS3 shutdown have been seriously considered in the optimized
6294 layout presented in this TDR. Several risks are identified:

6295 • **Performance degradation** due to possible further increase in the expected radiation
6296 levels. This increase, not yet stabilized, is caused by the increasing amount of ITk
6297 services and supports in the patch panel PP2 region, in front of HGTD. The baseline
6298 layout has the transition radius between the replaceable HGTD inner ring and the
6299 permanent outer ring at $r = 320$ mm. Three possible scenarios are envisaged and
6300 should be decided once the final radiation levels are available with the ITk realistic
6301 services/supports and the performance of the irradiated HGTD sensor+ASIC system
6302 are validated with real size prototypes:

6303 – Increase the inner ring outer radius by few cm but bringing the extra modules
6304 inside the inner ring with unchanged 20% sensors overlap. This scenario does
6305 not increase the amount of modules in the detector, nor the CORE costs.

6306 – Increase the inner ring outer radius by few cm but the increased area will have
6307 80% overlapped sensors, to insure 3/hits/track in all the inner ring region. This

6308 scenario, if justified, should be carefully planned, since it will increase the amount
6309 of modules in the detector and amount of data to be transmitted to the peripheral
6310 electronics and to the outside of the detector.

6311 – Another possible action that can be cumulative to any of the 2 previous options is
6312 to replace twice the inner ring during the life time of the HL-LHC and not only
6313 one time as baselined in this project. The costs associated to each replacement
6314 of the inner ring (sensors, ASICs, flex cables) should be accounted in the future
6315 M&O resources.

6316 • **Resources shortage**, in particular experts manpower and Institutes from USA that
6317 today are actively involved in the R&D but will probably disappear for the construction
6318 phase. This scenario may lead to delays in the project that has already a late start
6319 compared to the other ATLAS Phase-II upgrade projects. An active action is on-going
6320 to attract new Institutes from ATLAS that have expertise in the relevant areas of the
6321 HGTD project and may bring the needed CORE and manpower expertise needed for
6322 the construction and later operation/maintenance of the HGTD.

6323 • **Delay scenarios**. The possible delay in the HGTD schedule is probably the main risk
6324 of this project, due to the late start compared to all the other ATLAS Phase-II upgrade
6325 projects. The high modularity of the detector, constructed from $2\text{ cm} \times 4\text{ cm}$ modules
6326 assembled into staves/intermediate plates that are later screwed into $1/2$ circular
6327 cooling disks allows several working scenarios in case of delays or lack of resources to
6328 bring a complete detector in Summer 2025 or later:

6329 – A priority should be to complete during the LS3 shutdown the construction and
6330 installation of all the services, patch panels, cooling station, moderator, empty
6331 vessel, to be locate in UX15 cavern.

6332 – Install all the crates and it's equipment in USA15 (power supplies, Luminos-
6333 ity/DCS/TDAQ boards, computers).

6334 – In summer 2025 install on the two end-cap calorimeter cryostat faces all the
6335 available $1/2$ circular pieces of HGTD inside the two HGTD vessels. In the
6336 present schedule the final assembly and integration of the detector components
6337 (modules/staves, peripheral electronics) into $1/2$ circular cooling disks is planned
6338 to be done at CERN in a sequential way for the A and C sides for a duration of 9
6339 months each. In case of cumulative delays the assembly of the 2 end-caps could
6340 be done in parallel, by duplicating the tools and Institutes manpower based at
6341 CERN for this operation.

6342 – In case of an incomplete detector at the start of HL-LHC, one should prioritize
6343 full instrumented $1/2$ disk pieces, with an equal number of $1/2$ disks for the A
6344 and C side. For example, one could start with only one layer per end-cap in the
6345 first year(s).

6346 – Install the missing 1/2 disk HGTD pieces at the next medium-short shutdown,
6347 when the end-cap calorimeters are opened to give access for the usual Tile or LAr
6348 barrel calorimeter electronics maintenance. The HGTD installation procedure and
6349 respective tools will be carefully designed to allow to complete the 1/2 circular
6350 detector disks installation even in the presence of the beam pipe. Only the vessel
6351 front cover made of 1 piece of circular shape, to improve thermal insulation, has
6352 to be installed without the beam pipe in place. The vessel front cover should
6353 be able to move on the installed beam pipe to allow a posteriori the completion
6354 of the 1/2 circular detector pieces and access during future HGTD maintenance
6355 periods.

PBS	Item	Core Cost (kCHF)
8.1	Sensors	2275
8.2	Electronics	3199
8.2.1	ASIC	833
8.2.2	Peripheral Electronics	767
8.2.3	High Voltage power supplies and crates	532
8.2.4	Low Voltage power supplies and crates	299
8.3	Luminosity/TDAQ (*)	395
8.3.1	Luminosity boards	315
8.3.2	DCS and interlocks	80
8.4	Module assembly and staves loading	1483
8.4.1	Bump-Bonding ASIC/Sensor	450
8.4.2	Flex cables	547
8.4.3	Modules assembly (incl tools)	386
8.4.4	Modules loading on staves/plates (incl tools)	100
8.5	Mechanics, Services and Infrastructure	2264
8.5.1	Vessel (including feedthroughs)	160
8.5.2	Moderator (**)	
8.5.3	On detector cooling/support plate	180
8.5.4	CO ₂ /water Cooling and N ₂ systems	1237
8.5.5	Tools for final assembly and installation	100
8.5.6	Services (cables, connectors, fibres)	526
8.5.7	Patch panels (w/ water cooling)	61
8.6	Detector Assembly and QA on surface	100
8.6.1	Test bench for detector certification	100
8.6.2	Assembly of components on cooling plates	-
8.6.3	Disks assembly	-
8.6.4	Final assembly inside vessel	-
8.7	Detector Installation and commissioning	-
8.7.1	Services and patch panels installation	-
8.7.2	Back-end elect. installation in USA15	-
8.7.3	Outer moderator part installation	-
8.7.4	Detector installation and connectivity	-
8.7.5	Detector commissioning	-
Total HGTD(kCHF)		9716
TDAQ(*)	Felix boards+LTI boards, emulator,...	995
Total w/ TDAQ		10711

Table 15.4: Estimated Core cost of the HGTD (in kCHF). The total cost is given with and without the costs of the TDAQ. It assumes that the total of the HGTD pre-production components corresponding to approximately 5% of the total needed production is of good quality and will be used in the detector. It is also considered that each HV channel will be reading 2 modules (2x2x225 channels), at least in the initial phase when the irradiation levels are smaller. (*) TDAQ related costs are estimated separately by TDAQ and should be moved to TDAQ CORE, after HGTD TDR approval. (**) The moderator core costs are accounted in the ATLAS ITK common items.

6356 Acknowledgements

6357 We thank CERN for the very successful operation of the LHC, as well as the support staff
6358 from our institutions without whom ATLAS could not be operated efficiently.

6359 We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia;
6360 BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil;
6361 NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China;
6362 COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF
6363 and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF,
6364 HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benozzi
6365 Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands;
6366 RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of
6367 Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and
6368 MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation,
6369 Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK,
6370 Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition,
6371 individual groups and members have received support from BCKDF, CANARIE, CRC and
6372 Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie
6373 Actions, European Union; Investissements d' Avenir Labex and Idex, ANR, France; DFG and
6374 AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-
6375 ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat
6376 de Catalunya, Spain; The Royal Society and Leverhulme Trust, United Kingdom.

6377 The crucial computing support from all WLCG partners is acknowledged gratefully, in
6378 particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark,
6379 Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1
6380 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities
6381 worldwide and large non-WLCG resource providers. Major contributors of computing re-
6382 sources are listed in Ref. [70]. Many thanks to Nicolo Cartiglia for providing the FBK UFSD3
6383 LGADs (production funded by INFN) and performing the Weightfield2 simulations.

A Expected Energy Spectra

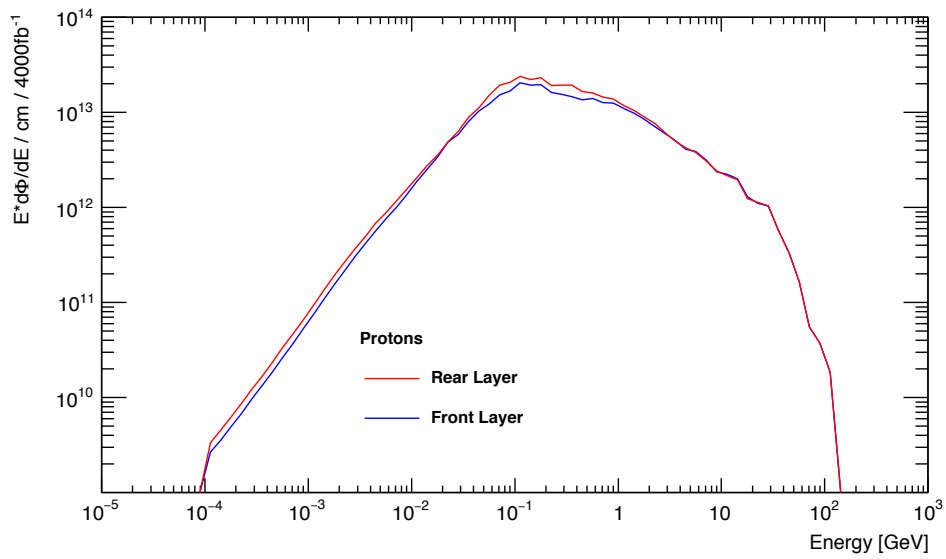


Figure A.1: Proton spectra averaged over the rear (outermost) and front (innermost) silicon layer of the HGTD. The uncertainties are of the order of 5%

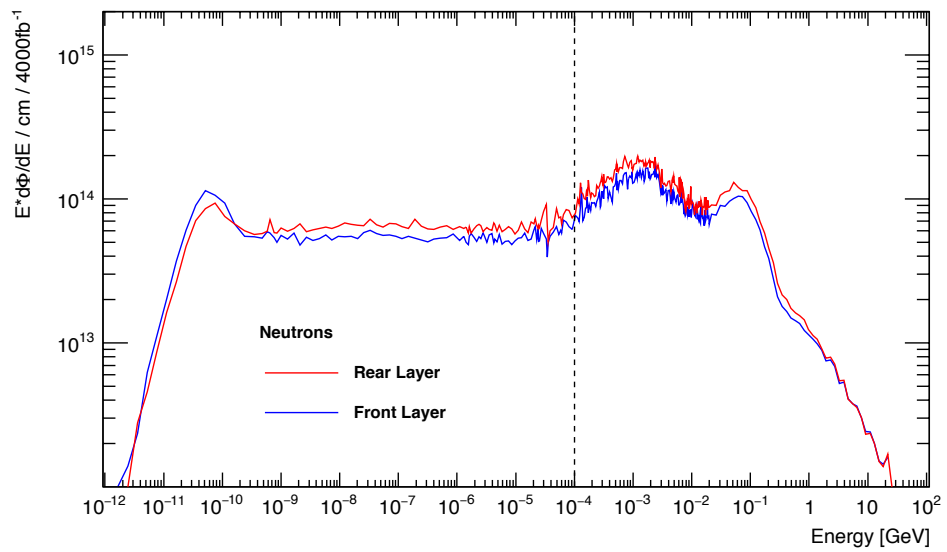


Figure A.2: Neutron spectra averaged over the rear (outermost) and front (innermost) silicon layer of the HGTD. The uncertainties are of the order of 5%. The fluctuations between 1 keV and 10 MeV are due to resonance.

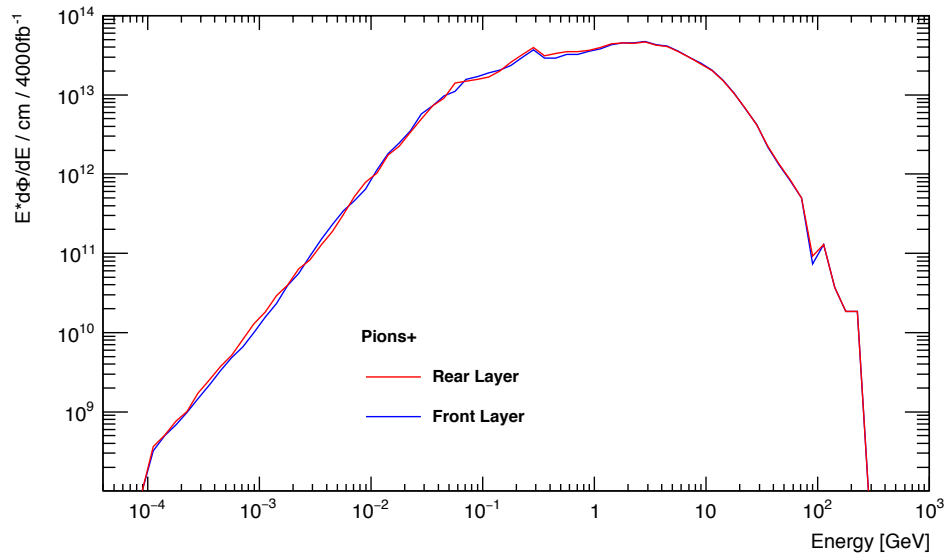


Figure A.3: Pion spectra averaged over the rear (outermost) and front (innermost) silicon layer of the HGTD. The uncertainties are of the order of 5%

6385 **B Technical Drawings**

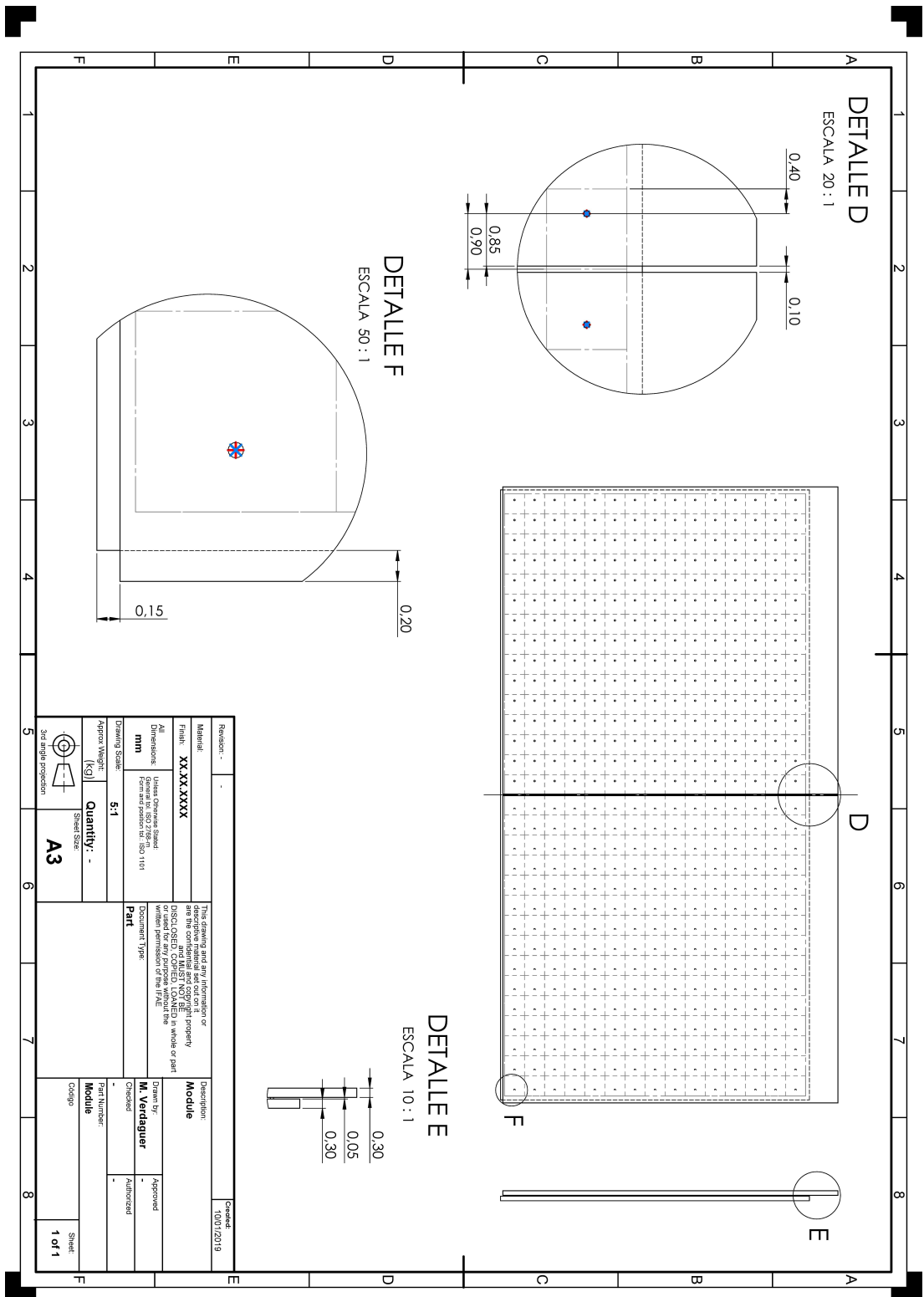


Figure B.1: Sketch of the bare module (sensor and ASIC). Distances are in millimeters. The bump pads on the sensor are shifted by 250 μm on each side of the sensor, to allow a 100 μm separation between the ASICs.

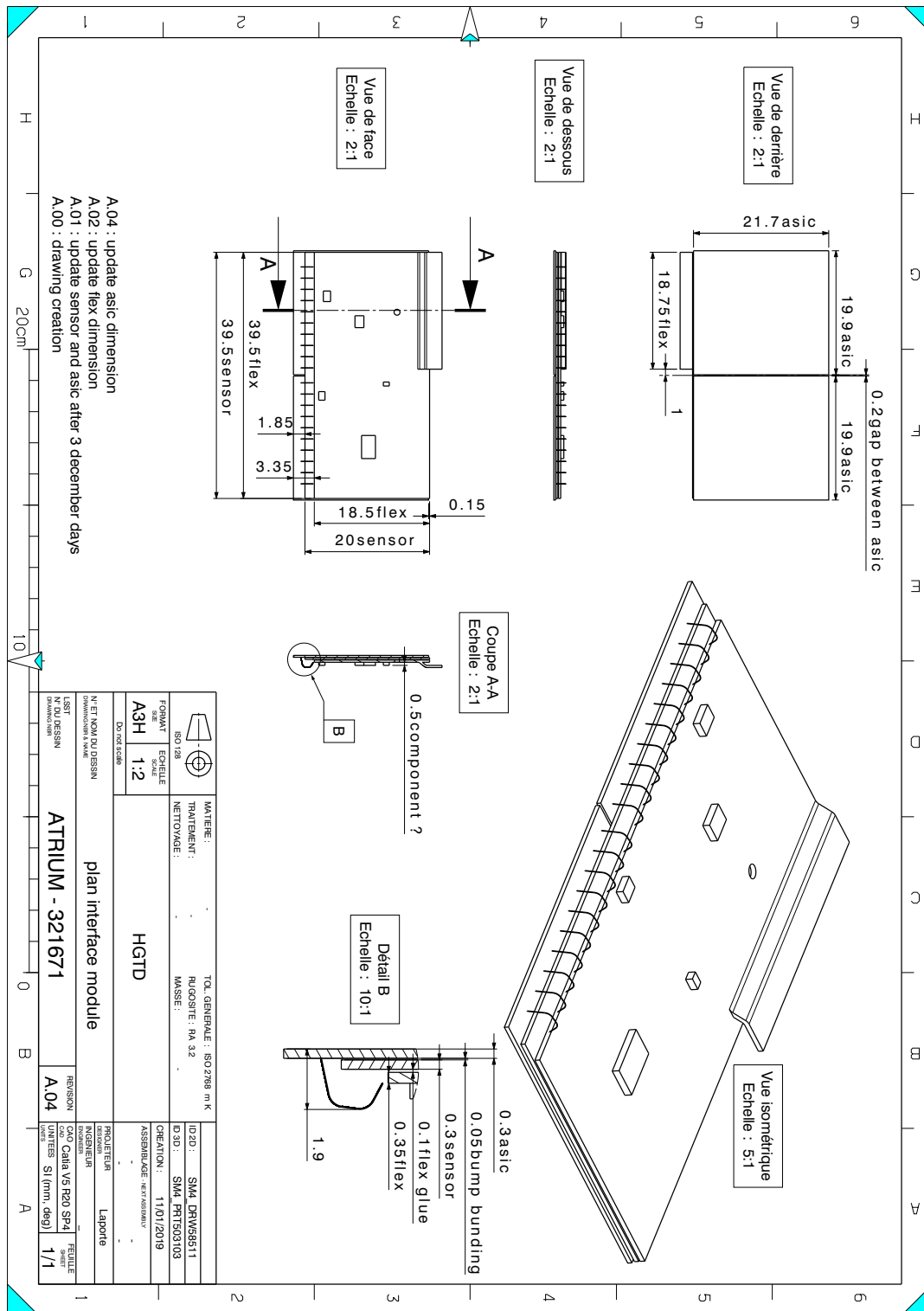


Figure B.2: Sketch of the module with the sensor, the ASICs and the Flex cable. Distances are in millimeters. Dimensions of the different components are visible, including the bumps pads, the glue and the wire-bonds.

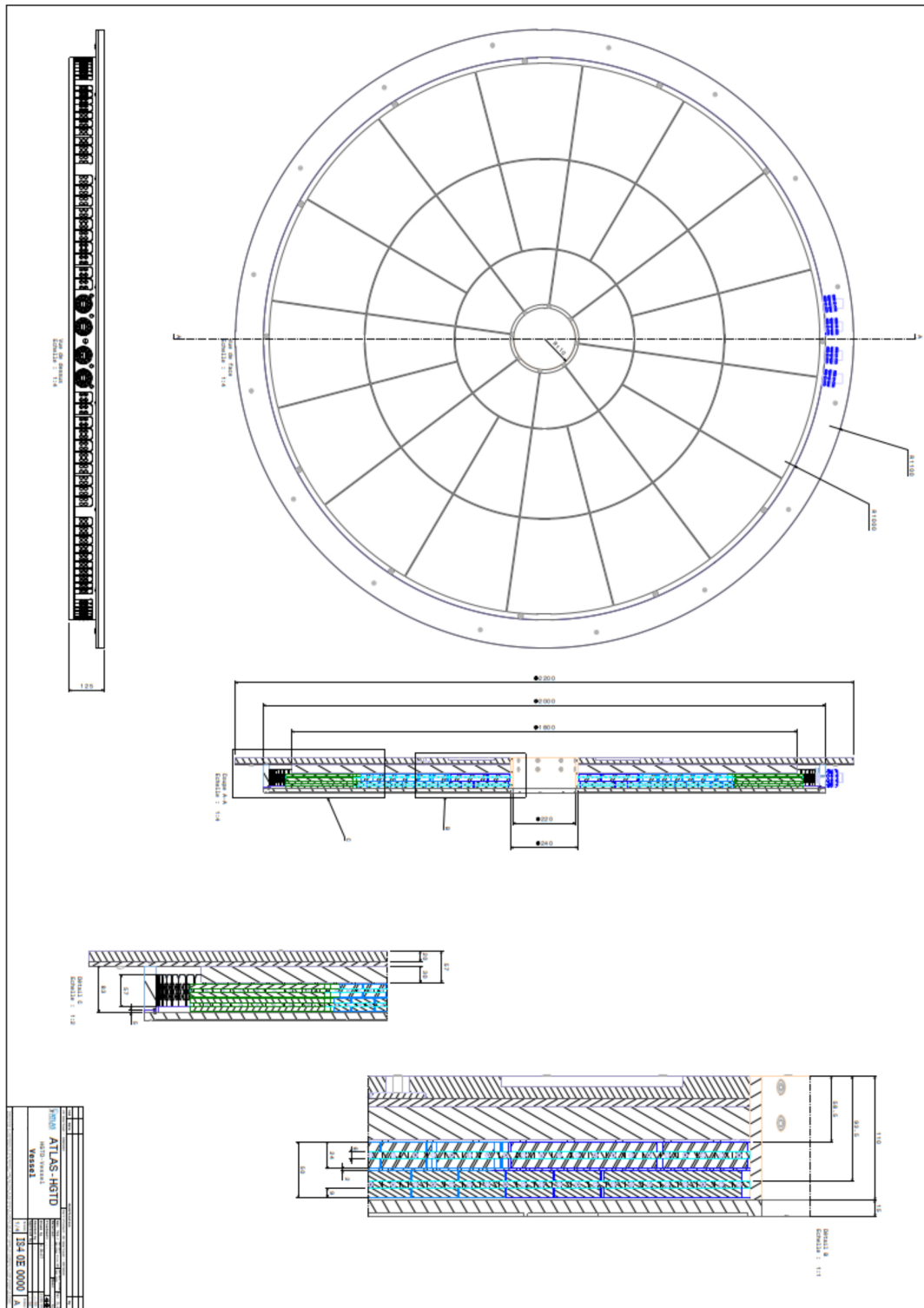


Figure B.3: top: View of the HGTD vessel front cover and feedthroughs region. The three bottom plots show an r - z view of the HGTD components inside the vessel, including a zoom of the inner radius and outer radius regions.

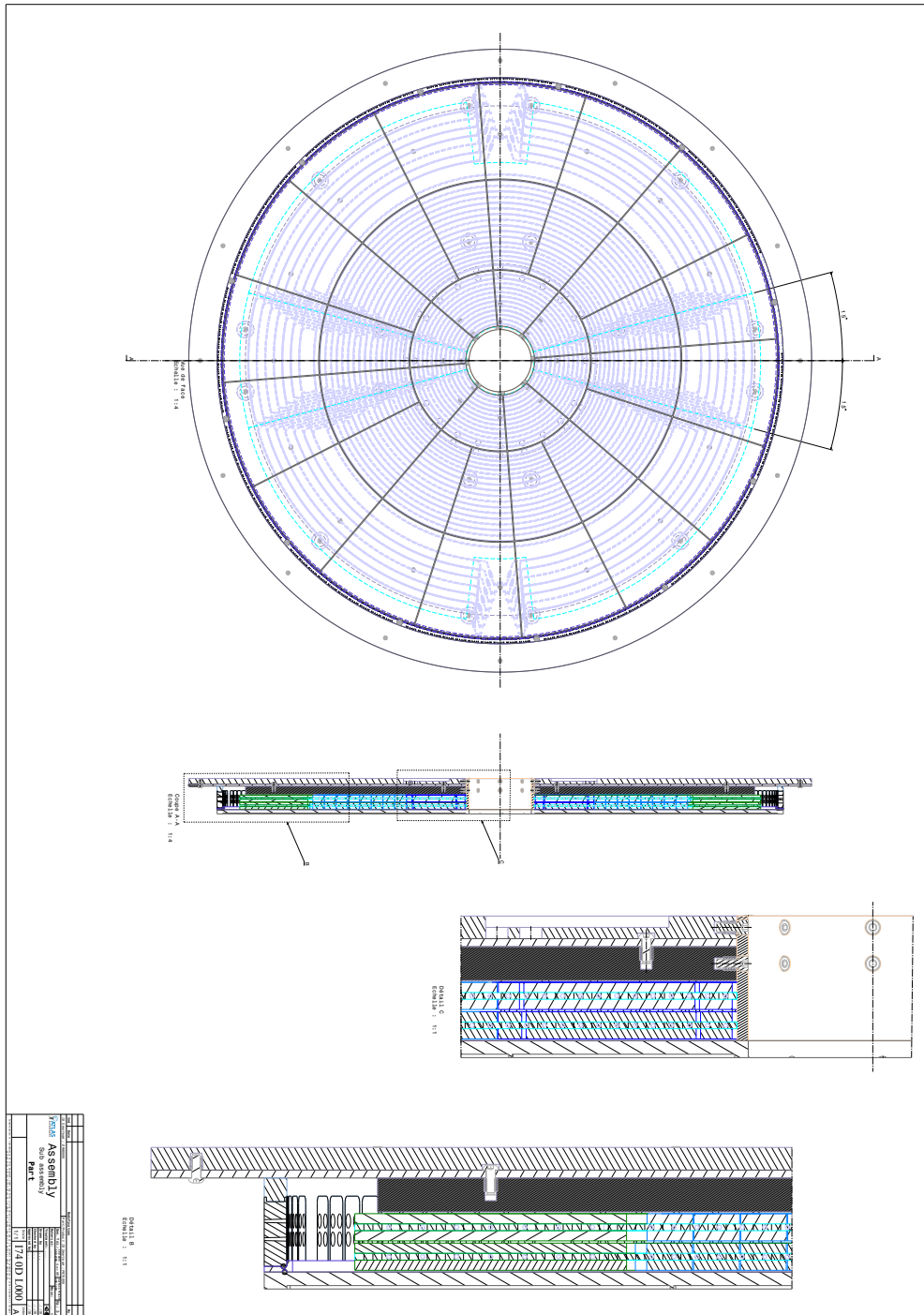


Figure B.4: View of the front and rear HGTD cooling disks inside the vessel, rotated by 15° in ϕ with respect to each other. The vessel front cover was put semi-transparent to allow to see the cooling disks. The three bottom drawings show r - z views of the HGTD components inside the vessel, including zooms of the inner radius and outer radius regions.

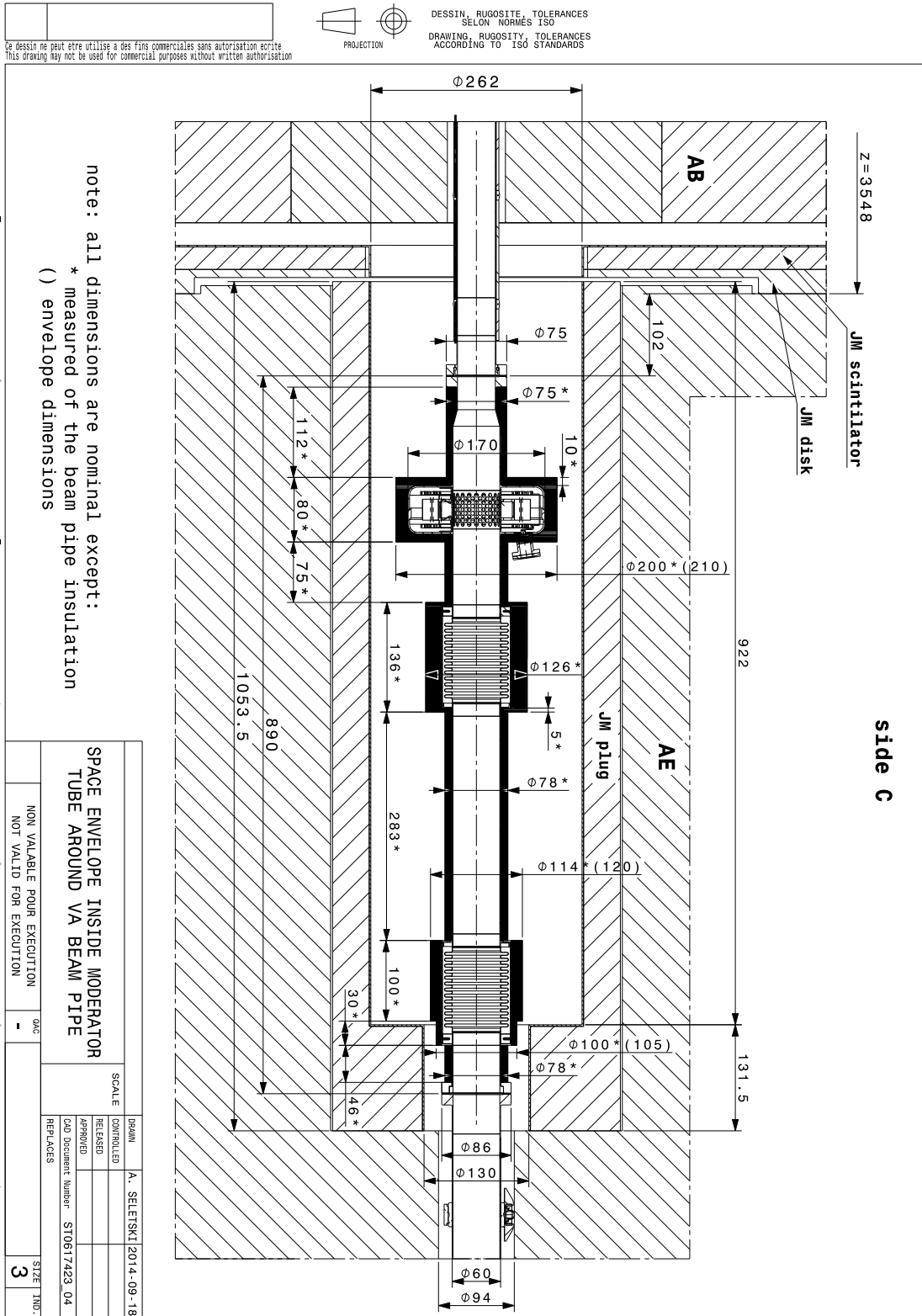


Figure B.6: Details of the space envelope around the beam pipe and moderator.

6386

Bibliography

- 6387 [1] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*,
6388 tech. rep. CERN-LHCC-2017-021, CERN, 2018, URL: [https://cds.cern.ch/
6389 record/2285585](https://cds.cern.ch/record/2285585) (cit. on p. 1).
- 6390 [2] F. Bordry et al., *Machine Parameters and Projected Luminosity Performance of Proposed
6391 Future Colliders at CERN*, tech. rep. CERN-ACC-2018-0037, CERN, 2018, URL: <https://cds.cern.ch/record/2645151> (cit. on p. 3).
- 6392 [3] RD50 collaboration, URL: <https://rd50.web.cern.ch/rd50> (cit. on pp. 5, 79,
6393 84).
- 6394 [4] G. Pellegrini et al., *Technology developments and first measurements of Low Gain Avalanche
6395 Detectors (LGAD) for high energy physics applications*, *Nucl. Instrum. Meth. A* **765** (2014) 12
6396 (cit. on pp. 5, 79).
- 6397 [5] ATLAS Collaboration, *Tracking Performance of the ATLAS Inner Tracker at the HL-LHC*,
6398 tech. rep., 2019, URL: <https://cds.cern.ch/record/2665904> (cit. on pp. 7, 10).
- 6399 [6] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250
6400 (cit. on p. 17).
- 6401 [7] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput.
6402 Phys. Commun.* **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 24).
- 6403 [8] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*,
6404 *JHEP* **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on p. 24).
- 6405 [9] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower
6406 simulations: the POWHEG method*, *JHEP* **11** (2007) 070, arXiv: 0709.2092 [hep-ph]
6407 (cit. on p. 24).
- 6408 [10] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO
6409 calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043,
6410 arXiv: 1002.2581 [hep-ph] (cit. on p. 24).
- 6411 [11] ATLAS Collaboration, *Expected Performance of the ATLAS Inner Tracker at the High-
6412 Luminosity LHC*, ATL-PHYS-PUB-2016-025, 2016, URL: [https://cds.cern.ch/
6413 record/2222304](https://cds.cern.ch/record/2222304) (cit. on p. 42).
- 6414 [12] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*,
6415 ATLAS-CONF-2014-018, 2014, URL: <https://cds.cern.ch/record/1700870>
6416 (cit. on p. 42).
- 6417

- 6418 [13] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with*
6419 *the ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-
6420 2015-027, 2015, URL: <https://cds.cern.ch/record/2037904> (cit. on p. 44).
- 6421 [14] A. Ferrari, P. R. Sala, A. Fasso and J. Ranft, *FLUKA: A multi-particle transport code*,
6422 tech. rep., 2005 (cit. on p. 48).
- 6423 [15] *Beam Background Generator*, 2010, URL: <http://bbgen.web.cern.ch/bbgen/>
6424 (cit. on p. 48).
- 6425 [16] R. Kwee-Hinzmann et al., *Machine-induced Background Simulation Studies for LHC Run*
6426 *1, Run 2 and HL-LHC*, tech. rep. CERN-ACC-2017-0025, CERN, 2017, URL: <https://cds.cern.ch/record/2261862> (cit. on p. 49).
- 6427 [17] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the*
6428 *ATLAS detector at the LHC*, *Eur. Phys. J.* **C76** (2016) 653, arXiv: 1608.03953 [hep-ex]
6429 (cit. on pp. 53, 195, 196, 200).
- 6430 [18] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in*
6431 *ATLAS*, *J. Inst.* **13** (2018) P07017 (cit. on pp. 53, 195).
- 6432 [19] ATLAS Collaboration, *Projections for measurements of Higgs boson cross sections, branching*
6433 *ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC*, tech. rep.,
6434 2018, URL: <https://cds.cern.ch/record/2652762> (cit. on p. 53).
- 6435 [20] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration,
6436 SLD Collaboration, LEP Electroweak Working Group, SLD Electroweak Group, SLD
6437 Heavy Flavour Group, *Precision electroweak measurements on the Z resonance*, *Phys. Rept.*
6438 **427** (2006) 257, arXiv: hep-ex/0509008 [hep-ex] (cit. on p. 54).
- 6439 [21] J. C. Collins and D. E. Soper, *Angular distribution of dileptons in high-energy hadron*
6440 *collisions*, *Phys. Rev. D* **16** (1977) 2219 (cit. on p. 54).
- 6441 [22] ATLAS Collaboration, *Precision measurement and interpretation of inclusive W^+ , W^-*
6442 *and Z/γ^* production cross sections with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) 367,
6443 arXiv: 1612.03016 [hep-ex] (cit. on pp. 54, 55).
- 6444 [23] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector*
6445 *using LHC Run 1 data*, *Eur. Phys. J.* **C74** (2014) 3071, arXiv: 1407.5063 [hep-ex]
6446 (cit. on p. 55).
- 6447 [24] P. A. M. Dirac, *Quantised singularities in the electromagnetic field*, *Proc. Roy. Soc. Lond.*
6448 **A133** (1931) 60 (cit. on p. 57).
- 6449 [25] G. 't Hooft, *Magnetic Monopoles in Unified Gauge Theories*, *Nucl. Phys.* **B79** (1974) 276,
6450 [291(1974)] (cit. on p. 57).
- 6451 [26] Y. M. Cho and D. Maison, *Monopoles in Weinberg-Salam model*, *Phys. Lett.* **B391** (1997) 360,
6452 arXiv: hep-th/9601028 [hep-th] (cit. on p. 57).
- 6453 [27] Y. M. Cho, K. Kimm and J. H. Yoon, *Mass of the Electroweak Monopole*, *Mod. Phys. Lett.*
6454 **A31** (2016) 1650053, arXiv: 1212.3885 [hep-ph] (cit. on p. 57).
- 6455

- 6456 [28] ATLAS Collaboration, *Search for Massive Long-lived Highly Ionising Particles with the*
6457 *ATLAS Detector at the LHC*, **Phys. Lett. B** **698** (2011) 353, arXiv: 1102.0459 [hep-ex]
6458 (cit. on p. 57).
- 6459 [29] ATLAS Collaboration, *Search for magnetic monopoles in $\sqrt{s} = 7$ TeV pp collisions with*
6460 *the ATLAS detector*, **Phys. Rev. Lett.** **109** (2012) 261803, arXiv: 1207.6411 [hep-ex]
6461 (cit. on p. 57).
- 6462 [30] ATLAS Collaboration, *Search for magnetic monopoles and stable particles with high electric*
6463 *charges in 8 TeV pp collisions with the ATLAS detector*, **Phys. Rev. D** **93** (2016) 052009, URL:
6464 <https://link.aps.org/doi/10.1103/PhysRevD.93.052009> (cit. on p. 57).
- 6465 [31] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS*
6466 *Trigger and Data Acquisition System*, tech. rep. CERN-LHCC-2017-020 ; ATLAS-TDR-029,
6467 CERN, 2017, URL: <http://cds.cern.ch/record/2285584> (cit. on pp. 62, 187,
6468 189).
- 6469 [32] H. Sadrozinski, A. Seiden and N. Cartiglia, *4-Dimensional Tracking with Ultra-Fast*
6470 *Silicon Detectors*, Reports on Progress in Physics (2017), arXiv: 1704.08666, URL:
6471 <http://iopscience.iop.org/10.1088/1361-6633/aa94d3> (cit. on pp. 80,
6472 85, 95, 96).
- 6473 [33] G. Pellegrini et al., *Status of LGAD production at CNM*, 30th RD50 Workshop, Krakow,
6474 Poland, 2017, URL: [https://indico.cern.ch/event/637212/contributions/](https://indico.cern.ch/event/637212/contributions/2608652/attachments/1470919/2276240/pellegrini_rd50.pdf)
6475 [2608652/attachments/1470919/2276240/pellegrini_rd50.pdf](https://indico.cern.ch/event/637212/contributions/2608652/attachments/1470919/2276240/pellegrini_rd50.pdf) (cit. on
6476 p. 80).
- 6477 [34] CMS Collaboration, *Technical proposal for a MIP timing detector in the CMS Experiment*
6478 *Phase 2 upgrade*, tech. rep. CERN-LHCC-2017-027. LHCC-P-009, CERN, 2017, URL:
6479 <https://cds.cern.ch/record/2296612> (cit. on p. 82).
- 6480 [35] G. Kramberger et al., *Radiation effects in Low Gain Avalanche Detectors after hadron*
6481 *irradiations*, **JINST** **10** (2015) P07006 (cit. on pp. 84, 85).
- 6482 [36] G. Kramberger et al., *Radiation hardness of thin Low Gain Avalanche Detectors*, **Nucl.**
6483 **Instrum. Meth. A** **891** (2018) 68, ISSN: 0168-9002, arXiv: 1711.06003, URL: [http :](http://www.sciencedirect.com/science/article/pii/S0168900218301682)
6484 [//www.sciencedirect.com/science/article/pii/S0168900218301682](http://www.sciencedirect.com/science/article/pii/S0168900218301682)
6485 (cit. on pp. 84, 85).
- 6486 [37] N. Cartiglia et al., *Beam test results of a 16 ps timing system based on ultra-fast silicon detect-*
6487 *ors*, **Nucl. Instrum. Meth. A** **850** (2017) 83, arXiv: 1608.08681 [physics.ins-det]
6488 (cit. on pp. 85, 86, 95, 96, 118).
- 6489 [38] J. Lange et al., *Gain and time resolution of 45 μm thin Low Gain Avalanche Detectors*
6490 *before and after irradiation up to a fluence of 10^{15} $n_{\text{eq}}/\text{cm}^2$* , **JINST** **12** (2017) P05003, arXiv:
6491 [1703.09004](https://arxiv.org/abs/1703.09004) [physics.ins-det] (cit. on pp. 85, 95, 96).
- 6492 [39] Z. Galloway et al., *Properties of HPK UFSD after neutron irradiation up to $6e15$ n/cm^2* ,
6493 submitted to NIM A (2017), arXiv: 1707.04961 (cit. on pp. 85, 96, 102).

- 6494 [40] Y. Zhao et al., *Comparison of 35 and 50 μm thin HPK UFSD after neutron irradiation up to*
6495 $6 \times 10^{15} \text{ neq/cm}^2$, (2018), arXiv: 1803.02690 (cit. on pp. 85, 96, 102).
- 6496 [41] S. M. Mazza et al., *Proprieties of FBK UFSDs after neutron and proton irradiation up to*
6497 $6 * 10^{15} \text{ neq/cm}^2$, (2018), arXiv: 1804.05449 [physics.ins-det] (cit. on pp. 85,
6498 96).
- 6499 [42] L. Masetti et al., *Beam test measurements of Low Gain Avalanche Detector single pads and*
6500 *arrays for the ATLAS High Granularity Timing Detector*, JINST **13** (2018) P06017, arXiv:
6501 1804.00622 [physics.ins-det] (cit. on pp. 85, 86, 94–96, 98, 118).
- 6502 [43] F. Cenna et al., *Weightfield2: A fast simulator for silicon and diamond solid state detector*,
6503 Nucl. Instrum. Meth. **A796** (2015) 149 (cit. on p. 86).
- 6504 [44] Synopsys, URL: <https://www.synopsys.com/silicon/tcad/device-simulation/sentaurus-device.html> (cit. on p. 86).
- 6506 [45] Agapopoulou et al., *A prototype ASIC for the ATLAS High Granularity Timing Detector*,
6507 In preparation () (cit. on pp. 109, 131, 144).
- 6508 [46] ATLAS Collaboration, *ATLAS Trigger and Data Acquisition Phase-II Upgrade Technical*
6509 *Design Report*, tech. rep. ATL-COM-DAQ-2017-185, This version represents the draft
6510 for EB approval and LHCC release.: CERN, 2017, URL: <https://cds.cern.ch/record/2296879> (cit. on p. 110).
- 6512 [47] K. Wyllie et al., *A Gigabit Transceiver for Data Transmission in Future High Energy Physics*
6513 *Experiments*, Phys. Procedia **37** (2012) 1561, URL: <https://cds.cern.ch/record/2103391> (cit. on p. 112).
- 6515 [48] F. Faccio et al., ‘FEAST2: A Radiation and Magnetic Field Tolerant Point-of-Load Buck
6516 DC/DC Converter’, 2014 IEEE Radiation Effects Data Workshop (REDW), IEEE, 2014
6517 (cit. on p. 113).
- 6518 [49] *Radiation tolerance of the 130 nm TSMC technology*, URL: <https://asicsupport-community.web.cern.ch/t/radiation-tolerance-of-the-130nm-tsmc-technology/144> (cit. on p. 113).
- 6521 [50] F. Cenna et al., *Weightfield2: A fast simulator for silicon and diamond solid state detector*,
6522 Nucl. Instr. Meth. A **796** (2015) 149, URL: <http://www.sciencedirect.com/science/article/pii/S0168900215004842> (cit. on p. 119).
- 6524 [51] Cadence Sigrity and PowerSI Cadence Desgn Systems Inc, URL: https://www.cadence.com/content/dam/cadence-www/global/en_US/documents/tools/pcb-design-%20analysis/sigrity-powersi-ds.pdf (cit. on p. 153).
- 6527 [52] CAEN Electronic Instruments, *DT55xxE Desktop HV Power Supply*, URL: <https://www.caen.it/products/dt5521he/> (cit. on p. 153).
- 6529 [53] Tektronix, *TDR Impedance Measurements: A Foundation for Signal Integrity*, URL: http://www.tek.com/dl/55W_14601_2.pdf (cit. on p. 153).
- 6530

- 6531 [54] KC705 Evaluation Board. Xilinx Inc., URL: [https://www.xilinx.com/support/](https://www.xilinx.com/support/documentation/boards_and_kits/kc705/ug810_KC705_Eval_Bd.pdf)
6532 [documentation/boards_and_kits/kc705/ug810_KC705_Eval_Bd.pdf](https://www.xilinx.com/support/documentation/boards_and_kits/kc705/ug810_KC705_Eval_Bd.pdf)
6533 (cit. on p. 154).
- 6534 [55] G. Beck and G. Viehhauser, *Analytic model of thermal runaway in silicon detectors*, **Nucl.**
6535 **Instrum. Meth. A618 (2010) 131** (cit. on pp. 165, 167, 168).
- 6536 [56] L. Amaral et al., *The versatile link, a common project for super-LHC*, *Journal of Instrumentation* **4** (2009) P12003 (cit. on pp. 175, 178).
- 6537
- 6538 [57] J. Anderson et al., *FELIX: a PCIe based high-throughput approach for interfacing front-end*
6539 *and trigger electronics in the ATLAS Upgrade framework*, *Journal of Instrumentation* **11**
6540 (2016) C12023 (cit. on p. 187).
- 6541 [58] J. Troska, S. Biereigel, S. Kulis and E. Mendes, *Preliminary lpGBT clock characteriz-*
6542 *ation*, tech. rep., URL: [https://espace.cern.ch/HighPrecisionTiming/](https://espace.cern.ch/HighPrecisionTiming/Evaluations/Components%20Evaluations/LpGBT/20190118lpGBTJitterReport.pdf)
6543 [Evaluations/Components%20Evaluations/LpGBT/20190118lpGBTJitterReport.](https://espace.cern.ch/HighPrecisionTiming/Evaluations/Components%20Evaluations/LpGBT/20190118lpGBTJitterReport.pdf)
6544 [pdf](https://espace.cern.ch/HighPrecisionTiming/Evaluations/Components%20Evaluations/LpGBT/20190118lpGBTJitterReport.pdf) (cit. on p. 191).
- 6545 [59] K. Chen and H. C. W. Wu, *Evaluation of System Clock from TTC to Front-Ends via FELIX*,
6546 tech. rep., URL: [https://indico.cern.ch/event/676116/contributions/](https://indico.cern.ch/event/676116/contributions/2767727/attachments/1548939/2440064/20171030_KChen_Evaluation_of_System_Clock_from_TTC_to_Front-Ends_via_FELIX.pdf)
6547 [2767727/attachments/1548939/2440064/20171030_KChen_Evaluation_](https://indico.cern.ch/event/676116/contributions/2767727/attachments/1548939/2440064/20171030_KChen_Evaluation_of_System_Clock_from_TTC_to_Front-Ends_via_FELIX.pdf)
6548 [of_System_Clock_from_TTC_to_Front-Ends_via_FELIX.pdf](https://indico.cern.ch/event/676116/contributions/2767727/attachments/1548939/2440064/20171030_KChen_Evaluation_of_System_Clock_from_TTC_to_Front-Ends_via_FELIX.pdf) (cit. on
6549 p. 191).
- 6550 [60] T. Mastoridis, P. Baudrenghien and J. Molendijk, *Cavity voltage phase modulation to*
6551 *reduce the high-luminosity Large Hadron Collider rf power requirements*, **Phys. Rev. Accel.**
6552 **Beams 20 (10 2017) 101003**, URL: [https://link.aps.org/doi/10.1103/](https://link.aps.org/doi/10.1103/PhysRevAccelBeams.20.101003)
6553 [PhysRevAccelBeams.20.101003](https://link.aps.org/doi/10.1103/PhysRevAccelBeams.20.101003) (cit. on p. 192).
- 6554 [61] CMS Collaboration, *CMS Luminosity Measurements for the 2016 Data Taking Period*,
6555 (2017) (cit. on p. 195).
- 6556 [62] G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the*
6557 *ATLAS detector at the LHC*, **Eur. Phys. J. C73 (2013) 2518**, arXiv: 1302.4393 [hep-ex]
6558 (cit. on pp. 195, 198).
- 6559 [63] P. Grafstrom and W. Kozanecki, *Luminosity determination at proton colliders*, **Prog. Part.**
6560 **Nucl. Phys. 81 (2015) 97** (cit. on p. 196).
- 6561 [64] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*,
6562 tech. rep. CERN-LHCC-2017-005 ; ATLAS-TDR-025, CERN, 2017, URL: [https://cds.](https://cds.cern.ch/record/2257755)
6563 [cern.ch/record/2257755](https://cds.cern.ch/record/2257755) (cit. on pp. 215, 221, 249).
- 6564 [65] EDMS 1866761, tech. rep. (cit. on p. 248).
- 6565 [66] M. Huhtinen, *Method for estimating dose rates from induced radioactivity in complicated*
6566 *hadron accelerator geometries - Write-up of the FIASCO code*, tech. rep., 1998, URL: [https://cds.](https://cds.cern.ch/record/2046812)
6567 [cern.ch/record/2046812](https://cds.cern.ch/record/2046812) (cit. on p. 247).
- 6568 [67] EDMS 1765924, tech. rep. (cit. on p. 250).

- 6569 [68] *ATLAS Phase-II Upgrade Scoping Document*, tech. rep. CERN-LHCC-2015-020. LHCC-
6570 G-166, CERN, 2015, URL: <https://cds.cern.ch/record/2055248> (cit. on
6571 p. 264).
- 6572 [69] ATLAS Collaboration, *Technical Proposal: A High-Granularity Timing Detector for the*
6573 *ATLAS Phase-II Upgrade*, tech. rep. CERN-LHCC-2018-023, CERN, 2018, URL: <https://cds.cern.ch/record/2623663> (cit. on p. 264).
- 6574
- 6575 [70] ATLAS Collaboration, *ATLAS Computing Acknowledgements*, ATL-GEN-PUB-2016-002,
6576 URL: <https://cds.cern.ch/record/2202407> (cit. on p. 281).