

Contents

3	0.1 Introduction	1
4	0.1.1 Historical retrospective	1
5	0.1.2 Actual challenges	2
6	0.1.3 Thesis composition	3
7	1 The Standard Model	5
8	1.1 General composition and key parameters	5
9	1.2 Classical fields and gauge invariance principle	8
10	1.3 Quantum electrodynamics	10
11	1.4 Electroweak theory and the Higgs mechanism	12
12	1.4.1 Electroweak gauge fields	12
13	1.4.2 Fermion sector	12
14	1.4.3 Higgs fields breaking the symmetry	13
15	1.4.4 Physical interpretation of gauge fields and parameters	14
16	1.5 Chromodynamics	15
17	2 The W boson	19
18	2.1 The motivation for the W mass measurement	19
19	2.2 Massive boson production at hadron colliders	20
20	2.2.1 Deep Inelastic scattering	21
21	2.2.2 The Drell-Yan process	23
22	2.3 Transverse momentum of massive vector bosons	26
23	3 The Large Hadron Collider	29
24	3.1 Introduction	29
25	3.2 The LHC running sequence	29
26	3.3 Special low pile-up run during LHC Run 2	33
27	4 The ATLAS detector	35
28	4.1 General description and layout	35
29	4.2 Coordinate system	35

30	4.3 Magnet system and magnetic field	36
31	4.4 Inner detector	36
32	4.5 Calorimeter system	38
33	4.5.1 Electromagnetic calorimeter	39
34	4.5.2 Hadronic calorimeter	42
35	4.6 Muon detectors	43
36	4.7 Forward detectors	45
37	4.8 Trigger system	47
38	5 Electromagnetic shower shapes correction in the electromagnetic calorimeter	49
39	5.1 Introduction	49
40	5.2 Shower shapes measurement and correction	52
41	5.2.1 Event selection	52
42	5.2.2 Data/MC discrepancies	52
43	5.2.3 The correction procedure	54
44	5.3 Results	55
45	5.4 Appendix: control plots	56
46	6 Event reconstruction	67
47	6.1 Charged particles track reconstruction	67
48	6.2 Determining the primary vertex of the event	68
49	6.3 Muon reconstruction and identification	70
50	6.3.1 Muon reconstruction	70
51	6.3.2 Muon identification	72
52	6.3.3 Muon isolation	73
53	6.4 Electron reconstruction and identification	73
54	6.4.1 Electron reconstruction	73
55	6.4.2 Electron identification	77
56	6.4.3 Electron isolation	78
57	6.5 Particle flow objects	79
58	7 MC samples and selection	81
59	7.1 Data samples	81
60	7.2 MC Signal and Background samples	81
61	7.3 Observables	81
62	7.3.1 Hadronic recoil	81
63	7.3.2 SET	81
64	7.3.3 SETUE	81
65	7.3.4 Transverse mass	81
66	7.3.5 Missing transverse energy	81

67	7.4 Event selection	81
68	7.4.1 Electrons	81
69	7.4.2 Muons	81
70	7.4.3 Truth	81
71	7.4.4 W boson	81
72	7.5 MC samples and backgrounds	81
73	8 MC samples and selection	83
74	8.1 Data and MC samples	83
75	8.1.1 MC samples and cross-sections	83
76	8.2 Z vertex reweighting	85
77	8.2.1 Multijet background	85
78	8.3 W analysis event selection and control plots	85
79	8.3.1 Event selection	90
80	8.3.2 $\sqrt{s} = 13$ TeV dataset control plots	94
81	8.3.3 $\sqrt{s} = 5$ TeV dataset control plots	102
82	9 W boson pT spectrum	109
83	9.1 Uncertainties propagation	109
84	9.2 Uncertainties breakdown plots	109
85	9.2.1 Unfolding	109
86	10 Hadronic recoil regression with deep neural networks	111
87	10.1 Deep neural networks	111
88	10.1.1 Gradient descent optimization	111
89	10.1.2 DNN structure and training	113
90	10.1.3 Batch normalization	115
91	10.2 HR regression	116
92	10.2.1 Input features and model	116
93	10.2.2 Results	117
94	10.3 Technical details	117
95	Bibliography	119

List of Figures

97	11	The list of particles that enters the Standard Model (SM)[sm_wiki].	7
98	12	The evolution of the SM running coupling constants [coupl_wiki].	8
99	13	The QEQ diagrams: Compton scattering, electron self-energy, photon self-energy.	11
100	14	Electroweak sector and the Weinberg rotation [coupl_wiki].	16
101	15	The formation of a jet [conf_wiki].	17
102	21	Next-to-leading order diagrams for W boson propagator containing contributions from heavy quarks and the Higgs boson.	19
103	22	W mass measurements and predictions	20
104	23	Examples of hard QCD scatterings.	21
105	24	The evolution of a PDF4LHC15 NNLO Hessian set from $Q^2 = 10^2$ GeV to $Q^2 = 10^4$ GeV using the DokshitzerGribovLipatovAltarelliParisi (DGLAP). Notice the increase in the sea quark density. The PDFs include one standard deviation uncertainty band.	23
106	25	Rapidity distribution for the vector bosons using MSTW2008 and CTEQ6.6 PDF sets for the centre-of-mass energies of 7 and 14 TeV [Halzen:2013bqa].	24
107	26	W and Z boson cross sections LHC at 7 TeV. ATLAS and CMS results, compared to NNLO predictions for various PDF sets [Mangano:2015ejw].	25
108	27	Parton contributions to the cross-sections of W^+ and W^- bosons for LHC and Tevatron cases [Martin:392675].	25
109	28	Kinematic distributions for W^\pm with corrections.	27
110	31	A Toroidal LHC ApparatuS (ATLAS) A hydrogen tank supplies LHC with protons [hydro].	30
111	32	Schematic depiction of the LHC ring.	31
112	33	Bunching at RF cavities	31
113	34	The crossing	32
114	35	Number of Interactions per bunch crossing in ATLAS Run 2 [run2lumi]. A little bump around $\mu \approx 2$ corresponds to special low pile-up runs.	34
115	41	ATLAS detector general layout	36
116	42	Geometry of ATLAS magnet windings	37
117	43	Fichier Gerber des modèles d CFR-34 et CFR-35.	38

125	44	ATLAS calorimeter general layout	39
126	45	ATLAS EM calorimeter layers	41
127	46	Accordion absorbers of the electromagnetic calorimeter (EMC)	42
128	47	ATLAS muon system	43
129	48	Cut views of the muon systems	45
130	49	ATLAS forward detectors	47
131	410	The scheme of ATLAS trigger systems	48
132	51	The schematic portrayal of EM shower development	49
133	52	Visualisations of the 7x11 calorimeter cluster	50
134	53	Shower shapes in the second layer of the electromagnetic calorimeter	50
135	54	Distribution of R_η in simulation (GEANT4) for electrons and jets [sshapes_simulation].	51
136	55	Data/MC Comparison for Calorimeter Shower Shapes of High Et Electrons [geant_corr].	52
137	56	R_η in the barrel and in the end-cap, Data vs MC	53
138	57	R_ϕ in the barrel and in the end-cap, Data vs MC	53
139	58	W_η^2 in the barrel and in the end-cap, Data vs MC	54
140	59	R_ϕ in the barrel and in the end-cap, Data vs MC	55
141	510	Schematic energy profile in ϕ dimension. Bremsstrahlung tails subtraction based on e^+ and e^- energy profiles.	55
143	511	R_η in the barrel and in the end-cap	56
144	512	R_ϕ in the barrel and in the end-cap, Data, MC, reweighted MC	57
145	513	W_η^2 in the barrel and in the end-cap, Data, MC, reweighted MC	57
146	514	Distributions integrated over pT (a) R_ϕ ; (b) R_η ; (c) $W_{\eta 2}$	58
147	515	Electron identification efficiency as a function of the electron pseudo-rapidity	58
148	516	Reta 2	59
149	517	Rphi in all eta slices	60
150	518	Reta 2	61
151	519	Reta 2	62
152	520	Reta 2	63
153	521	Reta 2	64
154	522	Reta 2	65
155	61	Impact parameters	67
156	62	Ambiguity solving.	69
157	63	Vertices	70
158	64	sagitta	71
159	65	muon combined	72
160	66	Electron path	74
161	67	topocluster	75
162	68	Supercluster	76

163	69	Isolation plot	78
164	610	pflow	80
165	81	Distributions for the 5 TeV low- μ dataset in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. The left and middle plots show the actual μ in a coarsely-binned and a finely-binned version. The right plot shows the number of reconstructed primary vertices N_{PV}	84
170	82	Distributions for the 13 TeV low- μ datasets taken in 2017 and 2018 in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. All distributions are (roughly) normalised to the same number of selected events in the 2017 dataset. The left and middle plots show the actual μ in a coarsely-binned and a finely-binned version. The right plot shows the number of reconstructed primary vertices N_{PV}	84
176	83	Distributions for the 5 TeV (left) and 13 TeV (right) low- μ dataset(s) in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. The distributions of the z-position of the primary vertex selected as the hard interaction are compared for the dataset(s) and the MC simulation before (“no $z_{\text{vtx}} \text{ rwgt}$ ”, blue, only 13 TeV) and after reweighting (black). For the 13 TeV data the 2017 and 2018 data are shown separately and all distributions are (roughly) normalised to the same number of selected events in the 2017 dataset.	90
183	84	$\Sigma \bar{E}_T$ distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	95
184	85	ΣE_T distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	96
185	86	\vec{E}_T^{miss} distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	97
186	87	Transverse mass distribution of the W boson in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	98
188	88	Lepton pseudorapidity distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	99
190	89	Lepton transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	100
192	810	W transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.	101
194	811	$\Sigma \bar{E}_T$ distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	102
195	812	ΣE_T distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	103
196	813	\vec{E}_T^{miss} distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	104
197	814	Transverse mass distribution of the W boson in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	105
199	815	Lepton pseudorapidity distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	106

201	816	Lepton transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	107
203	817	W transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.	108
205	101	A: The nodes perform a linear transformation of the inputs, then apply a non-linear activation function. B: The architecture of a deep neural network: neurons are arranged into layers [dnn1].	114
208	102	A model of the Deep Neural Network (DNN) used in the analysis.	117
209	103	learning curve.	118
210	104	The schematic portrayal of EM shower development	118

211

List of Tables

212	31	Energy and luminosity of the special low-mu runs.	33
213	41	ATLAS calorimeter in numbers	40
214	42	ATLAS muon spectrometer subsystems coverage and parameters	46
215	81	Monte Carlo samples at $\sqrt{s} = 13$ TeV. Given is a short description of the process, the ATLAS MC data set number (DSID), the names and version numbers of the MC generator(s), the used value of the higher order cross section times any branching and filter efficiencies ($\sigma \cdot BR \cdot \epsilon_{filter}$) with the theoretical uncertainty in percent (“th. unc.”), and finally the number of events analysed after skimming at derivation production (N_{evt}^{skim}) as well as the number of events originally processed and simulated (N_{evt}^{unskim}). In the case of $Z \rightarrow \ell\ell$ samples, the given $\epsilon_{filter} > 1$ is related to the fact, that the cross sections were calculated for $66 < m_{\ell\ell} < 116$ GeV, but the generated mass range is larger. The last section of $t\bar{t}$ samples refers to variation samples for systematics studies. The MC equivalent luminosity $N_{evt}^{unskim}/(\sigma \cdot BR \cdot \epsilon_{filter})$ is generally above 3fb^{-1} for signal and significant backgrounds, the exception are Powheg $W \rightarrow \tau\nu$ and $Z \rightarrow \tau\tau$ samples, that have about 0.45fb^{-1} only.	86

226	82	Alternative signal $Z \rightarrow \ell\ell$ Monte Carlo samples at $\sqrt{s} = 13\text{TeV}$ produced with SHERPA . General description of the table see Table 81. The samples are split into a long list of orthogonal slices based on “max(pTV,HT)” and filtered further into “b/c/light-jet” subcomponents. For the purpose of this analysis, the number of events in each slice is such that the samples are about 2fb^{-1} each (after application of a penalty factor for negative weight events) and an “inclusive sample” is restored after merging the slices.	87
232	83	Alternative signal $W \rightarrow \ell\nu$ Monte Carlo samples at $\sqrt{s} = 13\text{TeV}$ produced with SHERPA . See Table 82 for a description of the table. The samples are split into a long list of orthogonal slices based on “max(pTV,HT)” and filtered further into “b/c/light-jet” subcomponents. For the purpose of this analysis, the number of events in each slice is such that the samples are about 1fb^{-1} each (after application of a penalty factor for negative weight events) and an “inclusive sample” is restored after merging the slices.	88
238	84	Monte Carlo samples at $\sqrt{s} = 5\text{TeV}$. The table follows the same format as Table 81. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}} / (\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above 2.5fb^{-1} for signal and significant backgrounds, the exception are Powheg $W \rightarrow \tau\nu$ and $Z \rightarrow \tau\tau$ samples, that have about 0.20fb^{-1} and 0.45fb^{-1} only.	89
242	85	Analysis cut flow for $W^+ \rightarrow e^+\nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	91
244	86	Analysis cut flow for $W^+ \rightarrow e^+\nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	91
246	87	Analysis cut flow for $W^+ \rightarrow \mu^+\nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	92
248	88	Analysis cut flow for $W^+ \rightarrow \mu^+\nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	92
251	89	Analysis cut flow for $W^- \rightarrow e^-\nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	92
252	810	Analysis cut flow for $W^- \rightarrow e^-\nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	92
254	811	Analysis cut flow for $W^- \rightarrow \mu^-\nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	92
256	812	Analysis cut flow for $W^- \rightarrow \mu^-\nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.	93
258	813	Observed and Expected yield comparison for all signal selections.	93

259 **0.1 Introduction**

260 **0.1.1 Historical retrospective**

261 The reductionistic idea that all the countless variety of matter types that surrounds us could be in fact
262 brought to a combination of much fewer substances has been around at least since the time of Ancient
263 Greece. A thought that you can construct everything you see around out of one or few (e.g. fire, earth,
264 water and air) indivisible elements ($\alpha\tauομοζ$ in Greek) is simple, logical and therefore conceptually
265 attractive. Knowing all about these elements could potentially grant us profound understanding
266 of nature. But it wasn't before the XIX century when this idea has become something more than a
267 philosophical concept and obtained solid scientific evidence.

268 The composition of the periodic table of elements in 1860s [1] was a tremendous step forward, reducing
269 the number of elements to O(100). The elements of the periodic table resembled the ancient Greek
270 concept so much, that they were christened atoms. But the periodic character of the table and strong
271 correlation of atom position in the table with its chemical properties was insinuating on a certain inner
272 structure of the atoms, a possibility for them to be composed out of even smaller objects. The discovery
273 of isotopes in 1913 [2] left little room for other explanation.

274 Further evidences in favour of atomistic views kept coming in late XIX and early XX centuries from
275 theoretical and experimental sides. The molecular kinetic theory has been heavily criticized throughout
276 the XIX century, but the explanation of the Brownian motion [3] has secured its dominance from there
277 on lying a foundation for what is to become the statistical physics. Of particular importance was the
278 discovery of the first subatomic particle in 1897, which was called the electron [4].

279 Further studies of radioactive materials have allowed to compose a seemingly consistent understanding
280 of what matter is composed of. By the time of neutron discovery in 1932 [5] the list of what was called
281 elementary particles was reasonably short: an electron, a proton, and a neutron. It was still left to figure
282 out how these elements interact forming the known atoms, molecules and all the matter around. That
283 required additional efforts on the theoretical side, including resolving the inconsistencies between the
284 two new branches of physics supposed to describe the microworld and the fields, namely the quantum
285 theory and the field theory.

286 To move forward the physicist have made use of another source of elementary particles - the cosmic
287 rays. Cosmic rays contained particles of much higher energies comparing to the radioactive materials.
288 Cosmic ray experiments have led to the discovery of the first known antiparticle - the positron [6],
289 confirming the theoretical predictions by Dirac [7]. Further discoveries of the muon [8], pion [9], kaon
290 [10] and Λ_0 [11] have shown that the list of elementary particles was still far from being completed.
291 The second half of the XX century has pronounced a new era in particle physics with the extensive use
292 of particle accelerators. Accelerators have become the main experimental tool in the discovery of new
293 particles and investigation of their properties. Comparing to the cosmic rays, accelerators could offer
294 higher energies and better control over the experimental conditions. Thanks to these new tools by the
295 end of 1960s the number of newly discovered particles has exceeded one hundred and kept growing,

296 apparently taking away the reductionistic dream of having a reasonably small number of elementary
297 particles.
298 On the other hand, the properties of the newly discovered particles (sometimes called "the particle
299 zoo") had provided enough experimental data for theorists to make further assumptions. The particles,
300 if grouped by their properties, have formed patterns - a situation resembling the old story with the
301 atoms of the periodic table. This observation has allowed to assume the existence of even smaller
302 fundamental particles with a fractional charge that would make up all the visible hadrons. These
303 particles were eventually called quarks [12], [13]. By the late 1960s hypothesizing the existence of only
304 three quarks was enough to explain all the visible particles and successfully predict new ones [14].
305 Since then three more quarks were discovered and as of now all the experimental evidence suggests
306 that the quarks are truly fundamental particles being indivisible in the Ancient Greek sense.
307 At the same time serious theoretical efforts were taken in order to describe the interactions between
308 fundamental particles, taking into account the known fundamental forces. In the mid-1970s a theory
309 called The Standard Model was finalized. It included three out of four known fundamental forces
310 (excluding the gravity) and predicted a number of particles which were not discovered by that time.
311 All the key predictions of the theory were successfully confirmed by further experiments, making it a
312 dominant theory in particle physics. The theory was able to describe all the surrounding matter with
313 only 12 fundamental fermions (and their antiparticles) and 5 bosons. The SM is described in more
314 detail in the Chapter 1.
315 Theoretical efforts aimed to further simplify the list of fundamental particles are ongoing, but up to
316 the time of this thesis writing none of them were confirmed experimentally.

317 0.1.2 Actual challenges

318 The establishment of the Standard Model was a colossal step forward in understanding of the mi-
319 croworld physics. Nevertheless despite its great success and very good agreement with vast majority of
320 the experimental data there is a number inconsistencies and lacunae in the theory, which do not allow
321 to think of the SM as of the final theory. Here are most notable of these problematic questions:

- 322 1. A number of neutrino experiments have established that the neutrinos have a tiny though non-
323 zero mass. The minimal Standard Model assumes neutrinos to be massless and does not allow to
324 provide mass to the neutrinos.
 - 325 2. Astrophysical and cosmological evidences confirm the existence of the dark matter which does
326 not correspond to any of the SM particles.
 - 327 3. Cosmological observations show a substantial disproportion between observed matter and anti-
328 matter in favour of the former. The SM does not provide an explanation how such an imbalance
329 could have been formed.
-

330 4. The discovery of the gravitational waves in 2016 had confirmed the existence of the graviton - the
331 mediator of the gravitational force. The gravitational force is not represented in any way in the
332 SM.

333 5. No explanation is provided to the vastly different magnitude of the fundamental forces, i.e. why
334 the gravity is 10^{24} times weaker than the weak force.

335 In order to attack these and other problems numerous efforts have been taken to either modify
336 the SM or to replace it with a more fundamental theory, but so far none of these Beyond Standard
337 Model (BSM) theories were ever confirmed experimentally. The SM is still a source of most accurate
338 predictions for any physical process that involves elementary particle interactions. Description of the
339 BSM theories goes beyond the scope of current thesis.

340 The SM depends on the list of 18 free parameters (to be described in more detail in Chapter 1).
341 These parameters can not be calculated intrinsically and must be measured experimentally. The more
342 precisely we know the values of these parameters - the better is the accuracy of the SM prediction.
343 Precise knowledge of the SM input parameters can also give hints on where to look for a more
344 fundamental theory.

345 The LHC experiments have already contributed greatly by discovering the last missing piece of the SM,
346 the Higgs boson. This has ended the era of SM particle discoveries but at the same time started the era
347 of LHC precision measurements. The LHC experiments were capable to measure some parameters
348 of the SM for the first time (like the mass of the Higgs boson), but also could improve the existing
349 measurements, boosting the predictive power of the SM.

350 This thesis is a part of an ongoing effort at the ATLAS experiment to improve the precision of the W
351 boson mass, which is also among the SM free parameters. The mass of the W boson was first measured
352 at Large Electron-Positron (LEP) after its discovery in 1983. The precision of the measurement was
353 further improved by the experiments at Tevatron collider. The only LHC result performed so far was
354 published by ATLAS collaboration in 2018.

355 Hadron colliders are a challenging environment for the W boson-related measurements, the precision
356 is highly impacted by a number of factors one of them being the pile-up. Current analysis is based on
357 the data collected during two special LHC runs with low pile-up, taken in 2017 and 2018.

358 0.1.3 Thesis composition

359 The first chapter contains the description of the Standard Model, its constituents and input parameters.
360 Chapter 2 is dedicated to W boson and its properties. Chapter 3 tells about the Large Hadron Collider
361 (LHC) and its operations. ATLAS detector is described in Chapter 4. Chapter 5 is dedicated to the
362 description of the shower shapes reweighting. And so on and so forth...

365 “Potentielle citation sans aucun rapport avec le sujet”

³⁶⁶ — Personne inconnue, *contexte à déterminer*

The SM of particle physics is a quantum field theory that postulates the existence of three generations of quarks and leptons interacting through three fundamental forces: electromagnetic, weak and strong. From the mathematical point of view the SM is a gauge quantum field theory that has internal symmetries of the unitary product group $SU(3) \times SU(2)_L \times U(1)$. The fourth fundamental force, namely the gravity, is not included in the SM. Nevertheless, since the magnitude of the gravity interaction is negligible on the microscopic scale, it has little to no effect on the precision of the SM predictions. The model has 18^1 free input parameters - the physical constants that can not be predicted from within the theory and must be measured experimentally. Evidently, the SM predictions are based on these parameters, so the better we know them - the better we can predict how nature behaves on the micro level. The free parameters of the SM are briefly described in section 1.1

377

378 A comprehensive description of the quantum field theory formalism goes beyond the scope of
 379 current dissertation and can be found in the corresponding textbooks [1], [2], [3], [4], [5], [6]. In the
 380 following sections a brief overview of key SM features and constituent parts is provided.

381

382 1.1 General composition and key parameters

In this section I will describe the fields that enter the SM. Their existence and interactions result in the three fundamental forces that are taken into account by the theory. The quanta of these fields are also called fundamental particles and possess a number of properties like mass, charge (or charges), spin etc (see figure 11). The fundamental particles are divided into two groups based on their spin: particles with integer spin are called fermions and those with half-integer spin are bosons.

Let's start from the fermion sector. According to the Pauli exclusion principle[7] two fermions can not occupy the same quantum numbers. This in turn, has a consequence that the fermions must occupy a

¹There are SM extensions that take into account the non-zero neutrino mass. Then the model gets 7 additional parameters, so their total number reaches 25. Although current thesis only considers the SM where neutrinos are massless.

finite volume in space-time and as a result make up matter. Half of the fundamental fermions have colour charge and therefore take part in strong interaction - they are called quarks. The other six fermions do not have colour charge and are called leptons (from Greek " $\lambda\epsilonπτοσ$ " meaning "little", as they are lighter than the quarks of the same generation). Different types of quarks and leptons are also called flavours, so there are 6 flavours of quarks and 6 flavours of leptons.

For some reason which is yet unknown the twelve elementary fermions make three generations. Particles in the second and third generations have exactly the same charge and spin as the particles of the first generation, but are heavier and also unstable. Normally the particles of higher generations quickly decay down to their lighter kin of the first generation and can only be observed in cosmic rays and particle accelerators. That means all the matter that surrounds us consists of four fundamental fermions of the first generation²(the first column in Fig. 11).

The two quarks of the first generation are called up-quark and down-quark (or u-quark and d-quark for short). All the nuclei of the ordinary matter we see around are built out of these two types of quarks. Quarks are capable of interacting through all three SM forces: electromagnetic, weak and strong. Electrons, muons and tau-leptons are sensitive to electromagnetic and weak interaction, while neutrinos can interact (and therefore be detected) only through weak force. For this reason in particle physics the term "leptons" is sometimes used in a narrow sense referring to electron-like particles only. For all quarks and electron-like particles the antiparticles were observed as well as the corresponding annihilation phenomena. It is still not clear if neutrinos annihilate

From our experience we know that matter interacts with matter. But within the SM fermions do not interact with each other immediately. The interaction is mediated by boson-type particles. The SM includes five types of bosons: four vector bosons serving as force carriers for electromagnetic, weak and strong interactions, and a spinless Higgs boson whose role would be described in more detail in the corresponding subsection 1.4.1. The Higgs boson along with W and Z bosons are massive, while photons and gluons are massless.

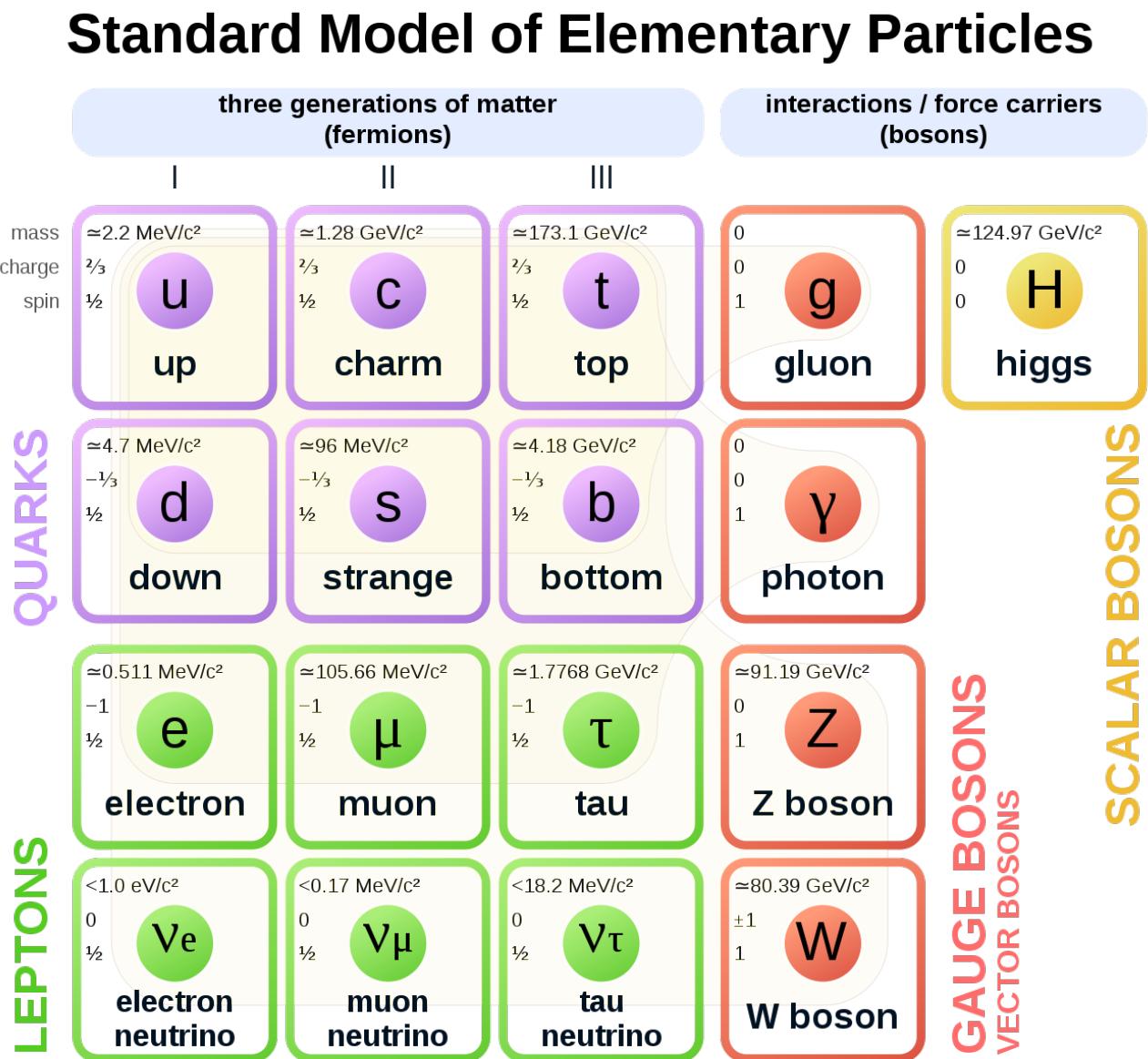
The masses of the fundamental particles make 12 out of 18 free parameters of the SM³.

As it was mentioned, bosons interact with fermions through fundamental interactions. The interaction depends on the charge of the interacting particles and on the type of the interaction itself. Each type of interaction has a coupling constant that defines the scale of the interaction. Hence two more parameters to the SM: the strong and electromagnetic coupling constants (the latter is also called the fine structure constant). Weak coupling constant is redundant since it can be obtained from other parameters. And the remaining four parameters are coming from the Cabibbo-Kobayashi-Maskawa (CKM), that contains information on the strength of the flavour-changing weak interaction. [8].

An important feature of the Quantum Field Theory (QFT) is that particles also interact with physical vacuum. For instance, a charged particle polarizes the physical vacuum, so the vacuum screens the

²Strictly speaking we already know that this is not completely true for the neutrinos, as they oscillate between the flavours due to their tiny mass. But in the SM neutrinos are assumed massless.

³The masses of W and Z bosons can be replaced by other parameters, e.g. weak mixing angle θ_W and Higgs potential vacuum expectation value (v. e. v.).

**Figure 11:** The list of particles that enters the SM[9].

425 charge of the particle[10].This interaction with virtual particles depends on the energy scale and so
 426 do the observed quantities like charge, mass etc. The SM is able to predict parameter evolution, so
 427 if the value of a certain input parameter q_0 is known at the energy Λ_0 then it is possible to predict
 428 its measurable value q at the energy Λ . This changing of physical parameters is an integral part of
 429 the QFT and is called *renormalisation* [2], [11]. In the picture 12 the dependence of the SM coupling
 430 constants on the energy is shown.

431

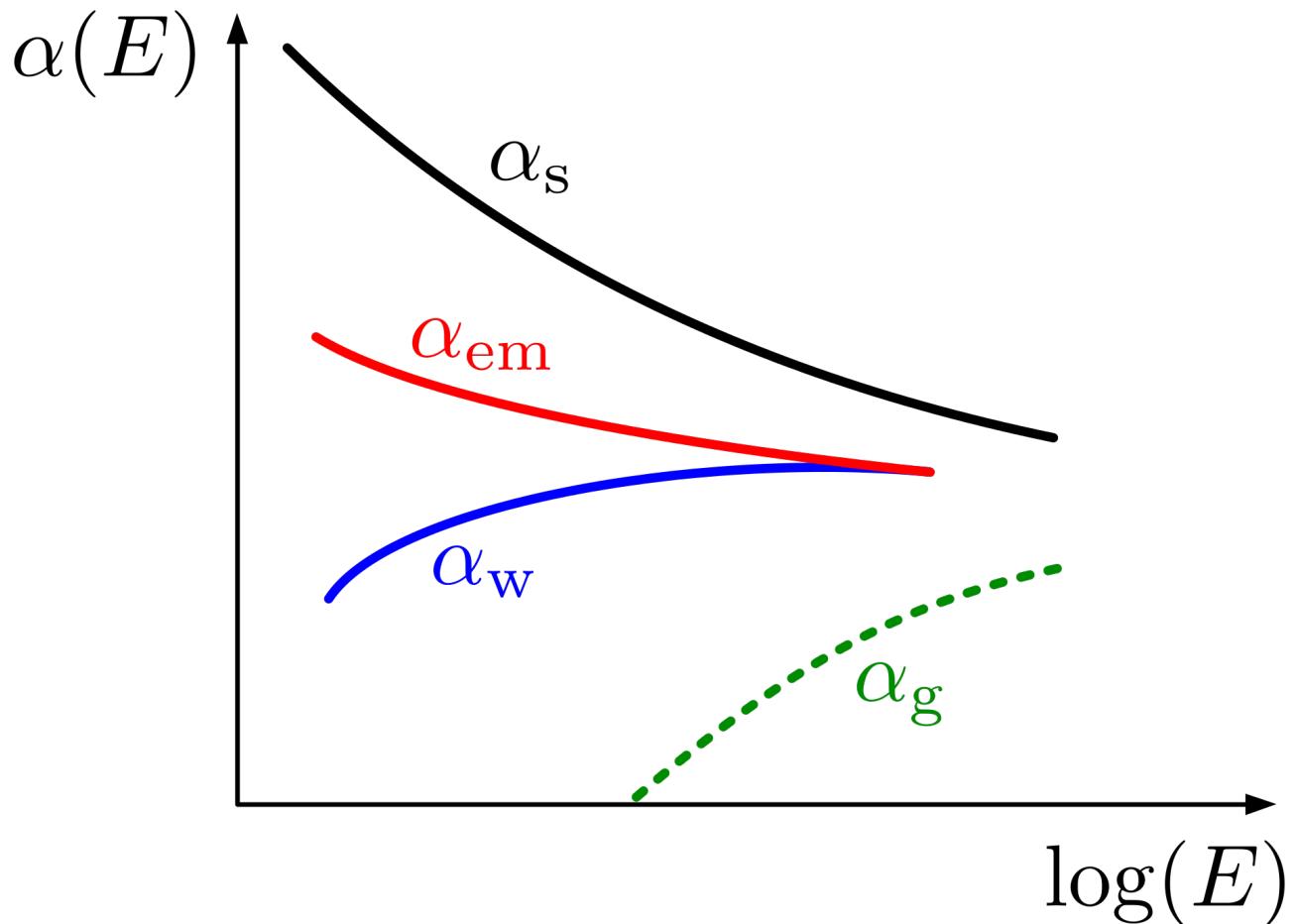


Figure 12: The evolution of the SM running coupling constants [12].

432 As we can see from picture 12 the strong coupling constant is getting smaller with the energy. This
 433 phenomena is called *the asymptotic freedom* [13], [14], [15].

434 **1.2 Classical fields and gauge invariance principle**

435 A consistent mathematical description of fields appears to be more challenging task compared to the
 436 description of physical objects that have definite size and shape. The derivation of Maxwell's equations

437 has been a great success and allowed to obtain the first equations of motion of relativistic fields. It has
 438 also subsequently led to understanding of special relativity [16], [17], [18]. Although for a more general
 439 case of fields other than electromagnetic it would be very useful to adopt a more systematic approach
 440 like that of Lagrangian or Hamiltonian in classical mechanics.

441 It has turned out that for the relativistic case Hamiltonian approach was not quite convenient, as
 442 the dedicated role of time over other degrees of freedom was in discord with relativistic space-time
 443 unification. However it was found possible to describe the fields within the Lagrangian approach. In
 444 the classic mechanics the action of a mechanical system of i mechanical objects is defined as:

$$S = \int L dt = \int \left(\sum_i T_i - U_i \right) dt,$$

445 where T_i and U_i are the kinetic and potential energies of the i^{th} object. Considering that by definition
 446 a field exists in every point of space-time, we need to define the Lagrangian density such that $L =$
 447 $\int \mathcal{L}(\phi, \partial_k \phi, \dot{\phi}) d^3x$, where ϕ is a field and $\partial_k \phi = \nabla \phi$ - the field gradient, $\partial_k = \frac{\partial}{\partial x^k}$, $k = 1, 2, 3$. Here and
 448 further Latin indices run through (1, 2, 3) and are used to denote spacial coordinates, while Greek
 449 indices denote space-time coordinates and run though (0, 1, 2, 3). So the action would look like:

$$S = \int L dt = \int \mathcal{L}(\phi, \partial_\mu \phi, \dot{\phi}) d^4x, \quad (1.1)$$

450 Now we may use the principle of least action to obtain the equations of motion using the Euler-
 451 Lagrange formalism. Let's check it with the example of electromagnetic fields. The Lagrangian density
 452 of electromagnetic fields in a vacuum can be written like:

$$S = -\frac{1}{4} \int F^{\mu\nu} F_{\mu\nu} d^4x. \quad (1.2)$$

453 Electromagnetic tensor can be defined in terms of electric and magnetic field intensities: $F_{i0} = -F_{0i} = E_i$,
 454 $F_{ij} = \epsilon_{ijk} H_k$, where ϵ_{ijk} - anti-symmetric Levi-Civita symbol. Alternatively $F_{\mu\nu}$ can be defined in terms
 455 of 4-potential A_μ :

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (1.3)$$

456 Now we can safely apply the variational principle and putting $\delta S = 0$ obtain the Maxwell equations in
 457 a vacuum:

$$\partial_\mu F_{\mu\nu} = 0. \quad (1.4)$$

458 Noticing the symmetries of the system and using the Noether's theorem[19] we can find the invariants
 459 of electromagnetic field. For example, translational symmetry in time and space ensures conservation
 460 of energy and momentum. Let's now consider a symmetry of a different kind. The field potential can
 461 be shifted by a gradient of an arbitrary function $\alpha = \alpha(x^\mu)$:

$$\begin{aligned} A_\mu(x) &\rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu \alpha(x) \\ F_{\mu\nu} &\rightarrow F'_{\mu\nu} = \partial_\mu(A_\nu(x) + \partial_\nu \alpha(x)) = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}. \end{aligned} \quad (1.5)$$

462 Let's now consider the electromagnetic theory in the presence of charges and currents:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + j^\mu A_\mu. \quad (1.6)$$

463 Now we have an interaction of a field potential A_μ with 4-current $j^\mu = (-\rho, j^i)$. It turns out to be a
464 general property of the field theories: the only form of interaction allowed is between a gauge field and
465 a current. After applying the gradient field transformation and the least action principle we can obtain
466 the corresponding conservation law:

$$\partial_\mu j^\mu = 0. \quad (1.7)$$

467 So this gradient symmetry[2] or as it is called more often gauge symmetry leads to the conservation of
468 electric current. If a theory is invariant under gauge transformations then it is called a gauge invariant
469 theory. As we have just seen electrodynamics is the simplest example of such a theory. Taking gauge
470 symmetries into consideration [20] has played a huge role in the development of the SM.

471 Gauge degree of freedom can be constrained in arbitrary way by applying additional conditions on the
472 gauge function. This is called fixing the gauge and becomes necessary after quantization. Any physical
473 result must be gauge-invariant, i.e. must not depend on the gauge.

474 1.3 Quantum electrodynamics

475 Quantum Electrodynamics (QED) is a theory of interaction between light and electrically charged
476 particles. Historically it was the first quantum field theory to reach good agreement between quantum
477 mechanics and special relativity. QED vacuum has zero expectation value. Nowadays it is considered
478 to be one of the most precise physical theories ever: theory predictions and experiment results agree up
479 to $O(10^{-8})$. It has also served as a model for composition of the subsequent parts of the SM, describing
480 other fundamental interactions.

481 Let's consider free Dirac field based Lagrangian:

$$\mathcal{L} = \bar{\psi}(x)(i\cancel{\partial} + m)\psi(x), \quad (1.8)$$

482 where ψ and $\bar{\psi}$ are Dirac wave function and its complex conjugate respectively, $\cancel{\partial} \equiv \gamma_\mu \partial^\mu$, γ_μ is one of
483 the four gamma-matrices and m is the mass of the Dirac field. Such a theory, though, would not be
484 physically consistent. This reflects the fact the quantum nature of spin and spinor fields have to be
485 treated as quantum fields. For instance, an attempt to calculate the energy of a Dirac field would lead
486 to a contradiction: the energy would not be positively defined, as some spinors would have negative
487 energies.

488 This Lagrangian has an internal symmetry to the U(1) transformation: $\psi \rightarrow e^{-i\alpha(x)}\psi$, $\bar{\psi} \rightarrow e^{i\alpha(x)}\bar{\psi}$.
489 According to Noether's theorem this symmetry implies current conservation: $j^\mu = \bar{\psi}\gamma^\mu\psi$. Now let's get
490 the combined Lagrangian of electromagnetic and Dirac fields, adding the interaction term:

$$\mathcal{L} = \mathcal{L}_{Dirac^{free}} + \mathcal{L}_{EM^{free}} + \mathcal{L}_{Interaction} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(x)(i\cancel{\partial} + m)\psi(x) - q\bar{\psi}\gamma^\mu A_\mu\psi, \quad (1.9)$$

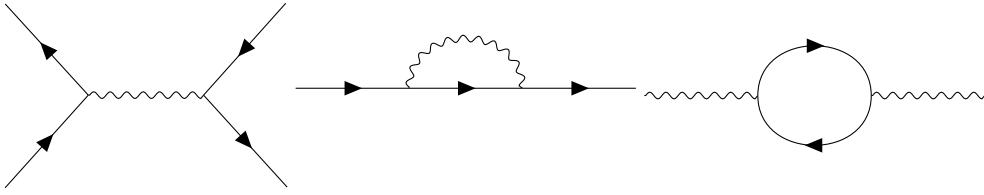


Figure 13: The QEQ diagrams: Compton scattering, electron self-energy, photon self-energy.

491 where q represents the elementary electric charge. This Lagrangian above is gauge invariant and can be
492 rewritten in a more convenient form:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(x)(iD + m)\psi(x), \quad (1.10)$$

493 where $D_\mu = \partial_\mu + iqA_\mu$ is a covariant derivative. If one considers space-time in the presence of a field as
494 curved, then A_μ would play a role of connectivity. It must be noted that values like m and q meaning
495 electron mass and charge⁴ are the SM input parameters mentioned in 1.1.

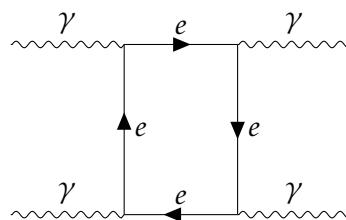
496 Further calculations are to be performed by the means of the quantum field theory formalism that
497 treats interaction terms like a perturbation to the free fields, making power series expansion in the
498 coupling constant. In the case of electrodynamics the coupling constant is quite small so good precision
499 is reached soon. Since the photons do not directly interact with other photons, QED allows only one
500 type of vertex - with two electron lines and one photon line.

501

502 Although the tree-level processes and diagrams were well understood by 1930th, the loop diagrams
503 were properly explained only by the end of the 1940th making it possible obtain numerical results
504 of the higher orders of power series expansion and achieve higher precision predictions for QED
505 processes[21], [10], [22], [23], [24], [25], [26], [27].

506 It must be noted that although direct photon-photon interaction is impossible, light-by-light scattering
507 is still possible through loops:

508



509

510 This process was theoretically described in 1936 [28] and experimentally observed 83 years after in
511 heavy ion collisions at the LHC [29].

⁴Charge of the electron is related to the electromagnetic coupling constant.

512 1.4 Electroweak theory and the Higgs mechanism

513 All the fermions of the standard model are subject to weak interaction, so its importance for physical
 514 processes can not be underestimated. At low energy weak interaction manifests itself mainly through
 515 flavour-changing decays like beta-decay and muon decay. The electroweak theory was created in the
 516 end of 1950s[11] [5] [30] thanks to numerous experimental results that allowed to shape its properties.
 517 The theory assumed that the electromagnetic and weak fundamental forces are actually manifestation
 518 of the same field that has a gauge symmetry $SU(2)_L \times U(1)$ with massive charged and neutral bosons.
 519 A few years later the structure of electroweak vacuum was explained along with the mechanism that
 520 has allowed the bosons to gain mass [31], [32]. Assuming this the Lagrangian of the electroweak theory
 521 must consist of three parts[33]:

- 522 • Gauge fields that would mediate the interaction.
- 523 • Fermions that interact with gauge fields
- 524 • A scalar Higgs field with non-zero vacuum energy that breaks the $SU(2)$ symmetry and couples
 525 to the fermions.

$$\mathcal{L}_{EW} = \mathcal{L}_{Gauge} + \mathcal{L}_{Higgs} + \mathcal{L}_{Fermions} \quad (1.11)$$

526 1.4.1 Electroweak gauge fields

527 As it was already pointed out before, knowing the symmetries of a physical system allows one to
 528 compose the gauge fields Lagrangian. The part with $U(1)$ symmetry would look like the electromagnetic
 529 field from 1.2 having the hypercharge Y , a vector potential B_μ and a gauge coupling g_1 . The $SU(2)$ field
 530 would have 3 vector components $W_\mu^{1,2,3}$, three isospin operators I_1, I_2, I_3 and a gauge coupling g_2 . We
 531 can pick the Pauli matrices σ^i as the representation of generators of the $SU(2)$ group, then the structure
 532 constants are ϵ_{abc} - Levi-Civita symbol.

$$\begin{aligned} \mathcal{L}_G &= -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W^{\mu\nu,a}B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \\ W_{\mu\nu}^a &= \partial_\mu W_\nu - \partial_\nu W_\mu + g_2 \epsilon_{abc} W_\mu^b W_\nu^c, \end{aligned} \quad (1.12)$$

533 where the term $g_2 \epsilon_{abc} W_\mu^b W_\nu^c$ appears due to the non-Abelian nature of the $SU(2)$ group (the generators
 534 don't commute).

535 1.4.2 Fermion sector

536 Each fundamental fermion generation expressed as left-handed doublets and right-handed singlets is a
 537 fundamental representation of the group $SU(2) \times U(1)$:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, (e_R), (\mu_R), (\tau_R), \quad (1.13)$$

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} b \\ t \end{pmatrix}_L, (u_R), (d_R), (c_R), (s_R), (t_R), (b_R). \quad (1.14)$$

538 Their quantum states are classified using the following quantum numbers: weak isospin I_3, I , weak
 539 hypercharge Y . Their electric charge can be obtained using the Gell-Mann-Nishijima relation:

$$Q = I_3 + \frac{Y}{2}. \quad (1.15)$$

540 The fermions are divided by their chirality: only the left-handed particles take part in weak
 541 interaction. The left-handed fermion fields of each lepton and quark generation j

$$\psi_j^L = \begin{pmatrix} \psi_{j+}^L \\ \psi_{j-}^L \end{pmatrix} \quad (1.16)$$

542 make SU(2) doublets, with indices $\sigma = \pm$, while the right-handed fermions can be written as singlets:

$$\psi_j^R = \psi_{j\sigma}^L. \quad (1.17)$$

543 Like in the electromagnetic case we can define the covariant derivative that would ensure the gauge
 544 invariance of the Lagrangian:

$$D_\mu = \partial_\mu - ig_2 I_a W_\mu^a + ig_1 \frac{Y}{2} B_\mu, \quad (1.18)$$

545 with $I_a \equiv \frac{\sigma_a}{2}$, then fermion Lagrangian takes the following form:

$$\mathcal{L}_{Fermions} = \sum_f \bar{\psi}_j^L i \gamma^\mu D_\mu \psi_j^L + \sum_{f,\sigma} \bar{\psi}_{f,\sigma}^R i \gamma^\mu D_\mu \psi_{f,\sigma}^R. \quad (1.19)$$

546 1.4.3 Higgs fields breaking the symmetry

547 The Higgs field is represented by single complex scalar doublet field $\Phi(x)$, that has 4 independent
 548 components. It spontaneously breaks the $SU(2) \times U(1)$ gauge symmetry, leaving the $U(1)_{EM}$ symmetry
 549 intact. The Higgs field doublet has the hypercharge $Y = 1$:

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}. \quad (1.20)$$

550 The Higgs field Lagrangian with non-zero vacuum expectation value:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^+ (D_\mu \Phi) - V(\Phi) + \mathcal{L}_{Yukawa}. \quad (1.21)$$

551 The gauge invariance of the Higgs Lagrangian is ensured in the traditional way by using the covariant
552 derivative:

$$D_\mu = \partial_\mu - ig_2 I_a W_\mu^a + i\frac{g_1}{2} B_\mu. \quad (1.22)$$

553 Higgs potential contains the mass term and quartic self-interaction:

$$V(\Phi) = -\mu^2 \Phi^+ \Phi + \frac{\lambda}{4} \partial_\mu (\Phi^+ \Phi)^2. \quad (1.23)$$

554 Valuum expectation value $\langle \Phi \rangle$ does not vanish:

$$\langle \Phi(x) \rangle = \frac{1}{\sqrt(2)} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v = \frac{2\mu}{\sqrt(\lambda)}. \quad (1.24)$$

555 Applying the unitarity gauge [34] we can constraint three out of four degrees of freedom of the Higgs
556 field and rewrite the Higgs doublet in the following way:

$$\Phi(x) = \frac{1}{2} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (1.25)$$

557 which leaves us with a physical real neutral scalar field $H(x)$ with

$$M_H = \sqrt(2)\mu. \quad (1.26)$$

558 This real field would couple to itself forming triple and quartic self-coupling vertices, to the gauge
559 fields through the covariant derivatives and to the charged fermions, giving them mass. Yukawa term
560 in Lagrangian the unitary gauge:

$$\mathcal{L}_{Yukawa} = - \sum_f m_f \bar{\psi}_f \psi_f - \sum_f \frac{m_f}{v} \bar{\psi}_f \psi_f H, \quad (1.27)$$

561 where

$$m_f = g_f \frac{v}{\sqrt(2)} = \sqrt(2) \frac{g_f}{g_2} M_W. \quad (1.28)$$

562 Higgs coupling constants to the corresponding fermion flavour are denoted as g_f . This relation
563 between the Higgs coupling and the mass of the W boson illustrates how much the SM parameters are
564 intertwined and particularly underlines the importance of the M_W measurement.

565 1.4.4 Physical interpretation of gauge fields and parameters

566 Higgs coupling to the gauge fields results in the following terms in the Lagrangian:

$$\frac{1}{2} \frac{g_2}{2} v (W_1^2 + W_2^2) + \frac{v^2}{4} (W_\mu^3, B_\mu) \begin{pmatrix} g_2^2 & g_1 g_2 \\ g_1 g_2 & g_1^2 \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \quad (1.29)$$

567 In order to get the physical meaning of this expression let us make a transition to the basis of physical
568 fields:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^+ \mp iW_\mu^-) \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \end{aligned} \quad (1.30)$$

569 where θ_W is called the weak mixing angle or the Weinberg angle. In the new basis expression 1.29 has
570 transparent physical sense:

$$M_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2}(A_\mu, Z_\mu) \begin{pmatrix} 0 & 0 \\ 0 & M_Z^2 \end{pmatrix} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix}, \quad (1.31)$$

571 with

$$\begin{aligned} M_W &= \frac{1}{2}g_2 v \\ M_Z &= \frac{1}{2}\sqrt{g_1^2 + g_2^2}v. \end{aligned} \quad (1.32)$$

572 The mixing angle θ_W also has a very clear physical meaning:

$$\cos\theta_W = \frac{g_2}{g_1^2 + g_2^2} = \frac{M_W}{M_Z}. \quad (1.33)$$

573 With A_μ having a sense of electromagnetic potential its coupling to the electron must have a physical
574 meaning of the electric charge $e = \sqrt{4\pi\alpha}$ we can express e in terms of gauge couplings:

$$e = \frac{g_1 g_2}{g_1^2 + g_2^2}, \quad g_2 = \frac{e}{\sin\theta_W}, \quad g_1 = \frac{e}{\cos\theta_W}. \quad (1.34)$$

575 Thus the demonstrated Weinberg rotation fully replaces the original parameters $g_1, g_2, \lambda, \mu^2, g_f$ by
576 another set of measurable values e, M_W, M_Z, M_H, m_f which are the input parameters of the SM.

577 1.5 Chromodynamics

578 The Quantum Chromodynamics (QCD) is a non-Abelian gauge theory that describes strong interaction.
579 The QCD is symmetric under unbroken SU(3) colour symmetry, so the interaction scheme is built in
580 the same way as electromagnetic and electroweak theories. To preserve the gauge invariance the gauge
581 field of gluons is introduced with 8 components, since SU(N) group has $\frac{N^2-1}{2}$ independent elements.
582 The gluons are massless vector bosons like the photons, although because of the non-Abelian nature of
583 the gauge group they couple not only to the fermions but also to the other gluons. The gauge invariant
584 QCD Lagrangian with kinetic term containing covariant derivative would look like:

$$\begin{aligned} \mathcal{L}_{QCD} &= -\frac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} + \bar{\psi}_a(i(\gamma^\mu D_\mu)^{ab} - m\delta^{ab})\psi_b, \\ F_{\mu\nu}^a &= \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_\mu^b A_\nu^c, \\ D_\mu &= \partial_\mu + ig_s A_\mu^a t_a. \end{aligned} \quad (1.35)$$

585 with ψ being the quark field, m is the mass of the quark, $a, b = 1, 2, \dots, 8$ are the colour indices, g_s is the
586 strong coupling constant, f^{abc} are the structure constants of the SU(3) group and t_a are the generators

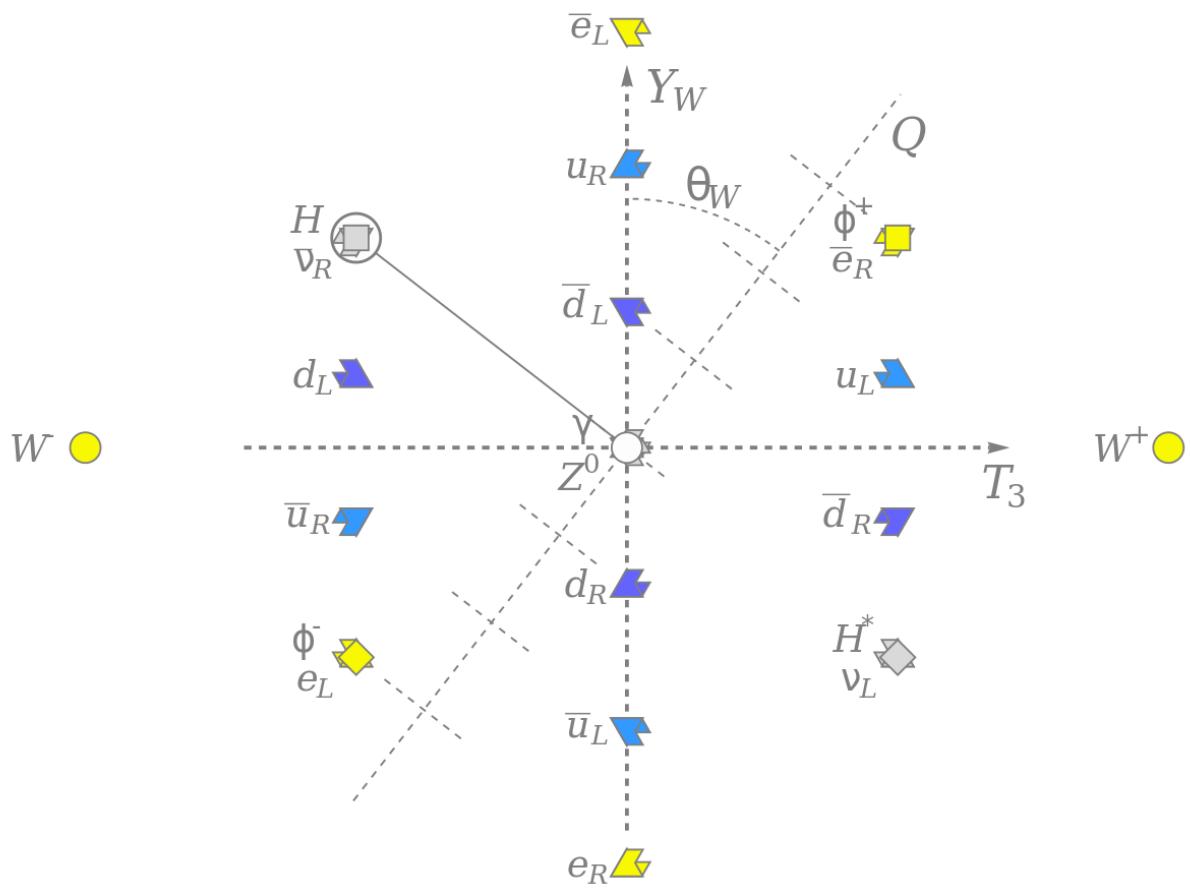


Figure 14: Electroweak sector and the Weinberg rotation [12].

587 of the SU(3) group.

588 As it was already mentioned in 1.3 quantitative calculations in QFT treat particle interaction as a
 589 perturbation to the free field theory. Coupling constant is considered to be a small parameter so every
 590 next power of the coupling constant is much smaller than the previous. Thanks to the asymptotic
 591 freedom α_s becomes small at higher energies and allows perturbative calculations. But at certain
 592 energy scale called $\Lambda_{QCD} \approx 200$ MeV, QCD becomes non-perturbative. It means we may no longer
 593 assume that interaction is a small perturbation of the free fields. This phenomena is known as the
 594 *colour confinement*.

595 Because of the colour confinement we can only observe colourless objects like baryons and mesons,
 596 but not quarks and gluons. If a high-energetic parton gets torn out of a hadron then it creates an
 597 avalanche-like process creating quark-antiquark pairs until fully hadronizes (see pic. 15) confining its
 598 colour. Such an avalanche is called a hadronic jet.

599

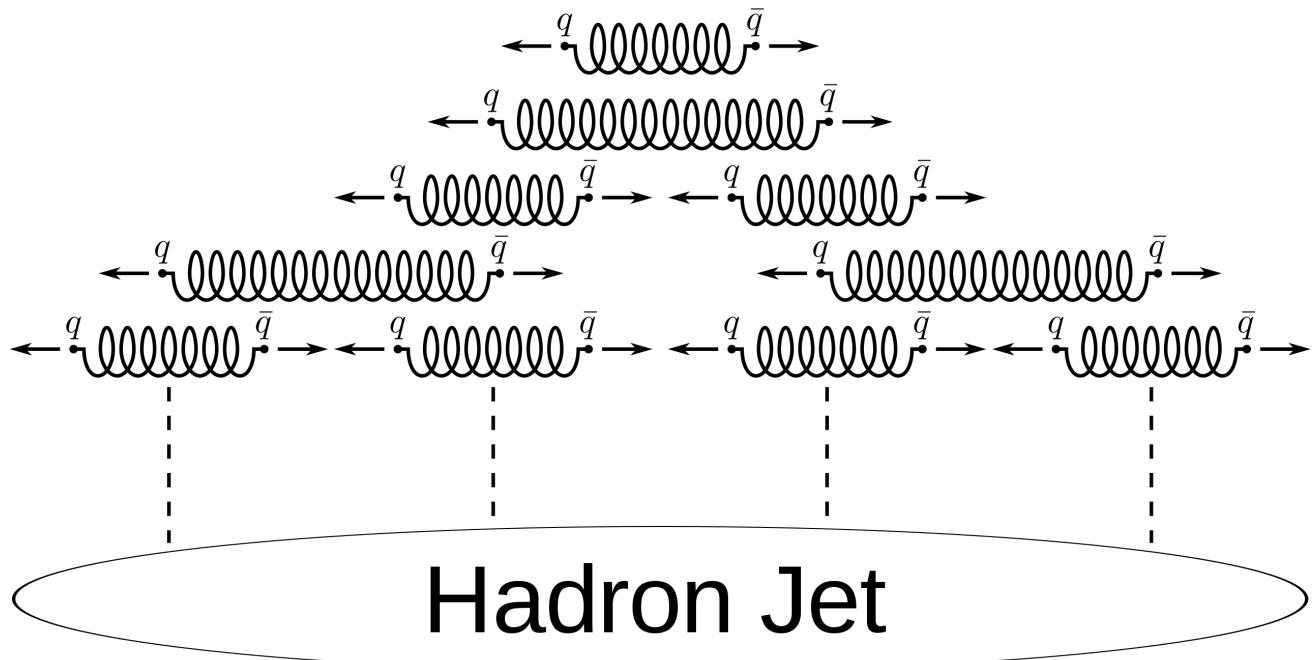


Figure 15: The formation of a jet [35].

600 Currently there is no viable physical theory that would describe QCD vacuum and low-energy
 601 behaviour of quarks and gluons. This also means that although nuclear forces are evidently residuals
 602 of the QCD interaction of partons within the baryons, there is no continuity between the QCD and
 603 nuclear physics. Confinement and low-energy QCD remain to be an unsolved problem of modern
 604 physics.

607 “Potentielle citation sans aucun rapport avec le sujet”

⁶⁰⁸ — Personne inconnue, *contexte à déterminer*

609 2.1 The motivation for the W mass measurement

610 Being one of the cornerstones of the SM, the W boson is tightly connected to the other parameters of
 611 the theory. In the leading order of the perturbation theory the W mass depends only on the electroweak
 612 parameters [1]:

$$M_W = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}} \frac{1}{\sin\theta_W}, \quad (2.1)$$

613 where G_F stands for the Fermi constant. The factor $\sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}} \approx 40$ GeV sets the lower bound for the
 614 possible W mass. Higher order corrections enter the equation in the following way:

$$M_W = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}} \frac{1}{\sin\theta_W} \frac{1}{1+\Delta r}, \quad (2.2)$$

where Δr contains the sum of all possible radiative corrections and depends also on other parameters of the SM, first of all on top quark and Higgs boson masses. The correction term is also sensitive to possible BSM effects. As it was mentioned in Chapter 1 the mass of the W boson is one of the input parameters of the SM, so the predictions of the theory directly depend on how precisely we know the value of the boson mass. On the other hand, we can theoretically constrain the value of the W boson mass assuming the already known values of the other SM parameters. Fig. 22 demonstrates that the uncertainty of the theoretical estimate for the W boson mass is about two times lower than that of the best available experimental measurement. This motivates the effort for a more precise experimental

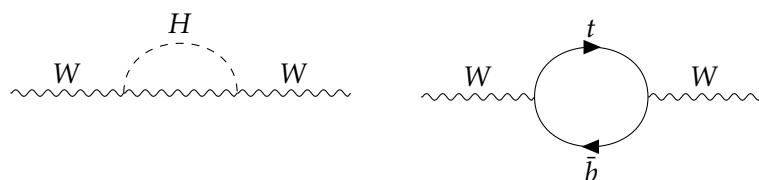
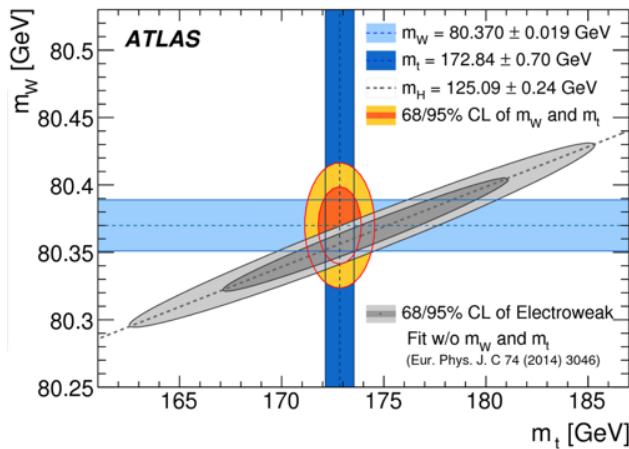
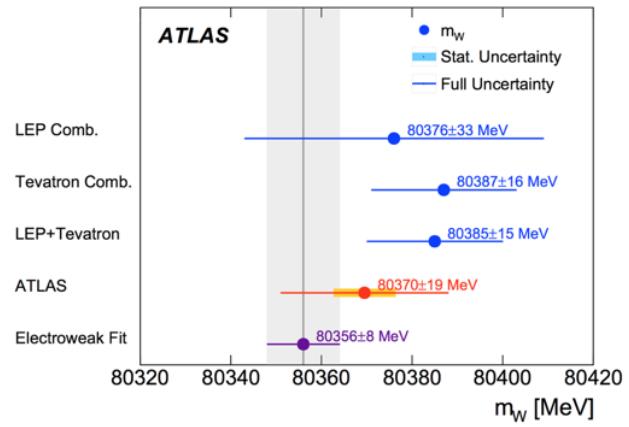


Figure 21: Next-to-leading order diagrams for W boson propagator containing contributions from heavy quarks and the Higgs boson.

623 measurement in order to test the consistency of the SM. Should the improved measurement reveal the inconsistency of the Standard Model - it would also allow to reveal viable BSM theories.



(a) W mass constraint [2]



(b) W mass measurements

Figure 22: W mass measurements and predictions

624

625 2.2 Massive boson production at hadron colliders

626 Hadron colliders provide a fruitful environment for the production and study of massive electroweak
627 bosons - all of them were discovered at hadron colliders. Hadron colliders allow to achieve much
628 higher centre-of-mass collision energy and luminosity comparing to their lepton counterparts. At the
629 same time precision measurements at hadron colliers demand much deeper theoretical understanding
630 of different aspects of the SM.
631 The main theoretical complication of the proton-proton colliders lies in the fact that in contrary to
632 leptons, protons are complex objects. This raises the following problems:

- 633 • A proton-proton collision is in general case a many-body problem. The absence of a consistent
634 theory for the QCD vacuum does not allow to describe the initial state of the proton constituents
635 in a consistent way.
- 636 • The initial energy of the whole proton is well known, but we don't know how this energy is
637 distributed between the proton constituents.
- 638 • We know that the proton consists of three valence quarks that have non-zero expectation value
639 and interact through gluons. In the course of these interactions all flavours of quarks (called sea
640 quarks) are appearing off-shell. The contribution of these sea quarks to the scattering cross-section
641 must also be taken in account.

642 In order to attack these problems and get accurate predictions for the proton-proton collisions it is
 643 necessary to take into account the asymptotic freedom that QCD demonstrates at short distances or
 644 high energies. At a certain energy scale of the momentum Q , transferred during the collision, we can
 645 assume that the interacting parts of the proton are asymptotically free and neglect the interaction
 646 with the rest of the proton. This is called *the factorization theorem*. The factorization occurs only if
 647 the transferred momentum $Q \gg \Lambda_{QCD}$ is large, and that is why these processes are called "hard". The
 648 physical conditions of the hard processes allow to use the perturbative QCD formalism, since at large
 649 energy scale the strong coupling constant α_s becomes small. Processes with lower energy scale of the
 650 transferred momentum are called "soft" and do not allow to use the perturbative QCD formalism. As it
 651 was mentioned in Section 1, a lot of things in the low-energy non-perturbative sector of the QCD are
 652 still unclear.
 653 The on-shell production of massive vector bosons occurs during the hard processes, however, precise
 654 measurements at hadron colliders require understanding of both hard and soft QCD regimes. It is
 655 common that the hard scattering of the proton constituents is accompanied by a soft scattering of the
 656 remaining proton parts. This forms what is called *underlying event* and also must be taken into account.

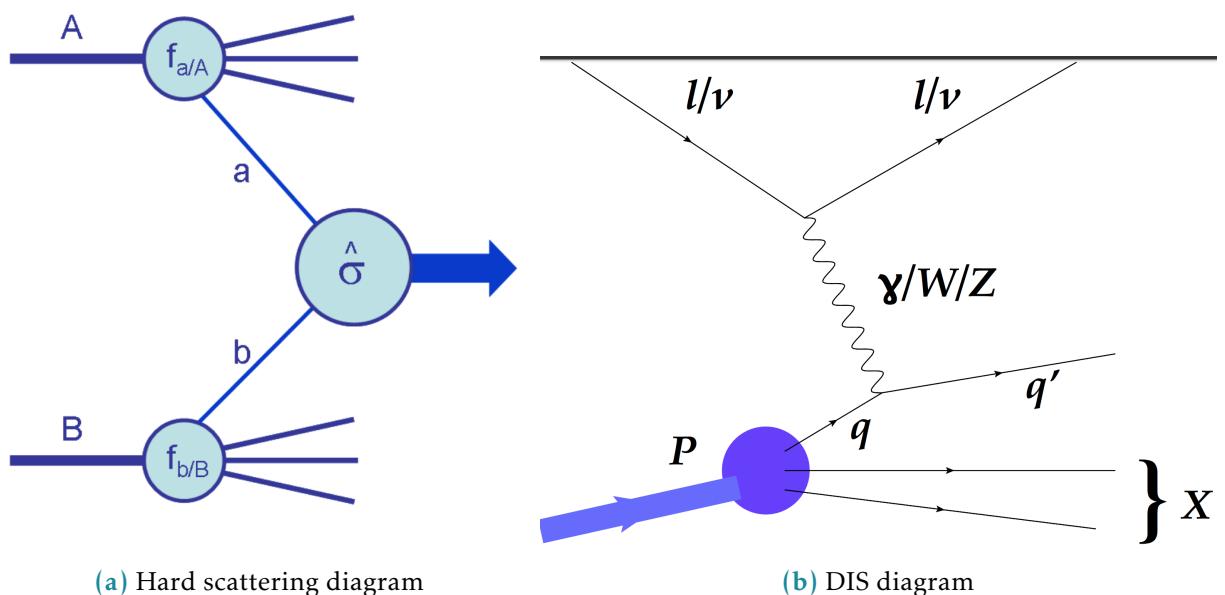


Figure 23: Examples of hard QCD scatterings.

657

658 2.2.1 Deep Inelastic scattering

659 In order to better illustrate the factorization approach let us first consider the lepton-hadron process
 660 called the Deep Inelastic Scattering (DIS). Historically it was the first experimental evidence for the
 661 complex structure of the proton and still serves an indispensable tool for the proton structure study.

662 Let's try to write a matrix element for a DIS process $e + A \rightarrow e + X$, exchanging a virtual photon with
 663 momentum q^μ :

$$|M|_{DIS}^2 = 4\pi M_N \frac{\alpha}{q^4} L_{\mu\nu} W_{hadron}^{\mu\nu}, \quad (2.3)$$

664 where $L_{\mu\nu}$ is the transverse lepton tensor, $q^\mu L_{\mu\nu} = q^\nu L_{\mu\nu} = 0$. The hadronic tensor $W_{\mu\nu}$ along with
 665 its normalization factor $4\pi M_N$ is unknown, but we can write it down in general form introducing
 666 longitudinal and transverse parts¹ [4]:

$$W_{\mu\nu} = F_1(x, Q^2) \left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) + F_2(x, Q^2) \frac{(p_\mu - q_\mu p \cdot q/q^2)(p_\nu - q_\nu p \cdot q/q^2)}{p \cdot q}, \quad (2.4)$$

667 with p_μ being the momentum of the hadron A, Q^2 is the exchange momentum, $x = \frac{Q^2}{2p \cdot q}$ and the form-
 668 factor functions $F_1(x, Q^2), F_2(x, Q^2)$ are unknown.

669 The cross-section of the DIS process can be measured experimentally, leaving the possibility to study
 670 the form-factor functions. It turned out that these functions do not depend (at least in the first
 671 approximation) on Q^2 [dis]. Further experiments have revealed that they form-factors depend only on
 672 the ratio x , as it was predicted before [5]. This type of behaviour was called the Bjorken scaling.
 673 These results have led to the idea of partons - point-like constituents of the proton [6]. Keeping in
 674 mind the idea of factorization we can assume the lepton only interacts with one of the partons. Then
 675 we can express the hadronic tensor $W_{\mu\nu}$ as a sum of all available partons:

$$W_{\mu\nu}(q_\mu, p_\nu) = \sum_a \int_x^1 \frac{d\xi}{\xi} f_{a/A}(\xi, \mu) H_{\mu\nu}^a(q_\mu, p_\nu, \mu, \alpha_s(\mu)) + NLO. \quad (2.5)$$

676 The functions $H_{\mu\nu}^a(q_\mu, p_\nu, \mu, \alpha_s(\mu))$ are called the hard scattering structure functions and only depend
 677 on parton type a , but not on hadron type A . These functions describe the high-energy behaviour and
 678 can be calculated in the framework of perturbative QCD. At the same time $f_{a/A}(\xi, \mu)$ is called Parton
 679 Density Function (PDF) and has a physical meaning of finding a parton of type a (gluon, u-quark,
 680 d-quark etc) in a hadron of type A (proton, neutron, meson) carrying the fraction of ξ of the hadron's
 681 momentum. These PDFs contain information on the momentum distribution of quarks and gluons
 682 within the hadron. This corresponds to the non-perturbative sector of the QCD which is beyond the
 683 reach of theoretical methods available so far. Note that they do not directly depend on the momentum
 684 Q^2 , but only on the energy scale μ .

685 The DGLAP equations show that once the PDFs are known at a certain energy scale μ they can be
 686 perturbatively extrapolated to a different energy scale [7], [8], [9], [10]. This means that the PDFs
 687 are universal - they can be measured experimentally at certain conditions in the course of the DIS
 688 (or any other) process and then used for numerical calculations of any other process (e.g. Drell-Yan
 689 (DY) process) at different conditions. Such a measurement allows a workaround - we may not be able
 690 to solve the many-body problem and perform non-perturbative calculations starting from the first
 691 principles, yet we still get a theoretical prediction with a good precision. Currently there exist a number

¹Given example assumes only electromagnetic interaction. For the more general electroweak case the tensor structure is more complicated and there are more than two scalar structure functions [3].

692 of different groups working on the PDF parametrizations and fits, constantly improving the fits using
 693 the data coming from hadron colliders. Using different PDF sets may give different results and also
 694 helps to estimate the systematic uncertainties implied by the PDFs.
 695 Historically the DIS experiments at HERA electron-proton collider have allowed to perform proton
 696 PDFs measurements with a good level of precision in the x region up to $x \sim 10^{-4}$ at high Q^2 . The HERA
 697 experiments operated until 2008, paving the path for precision predictions for the Drell-Yan process.
 698 Currently there are prospects for new experiments like Large Hadron Electron Collider (LHeC) that
 would involve DIS and further improve the PDF precision [11].

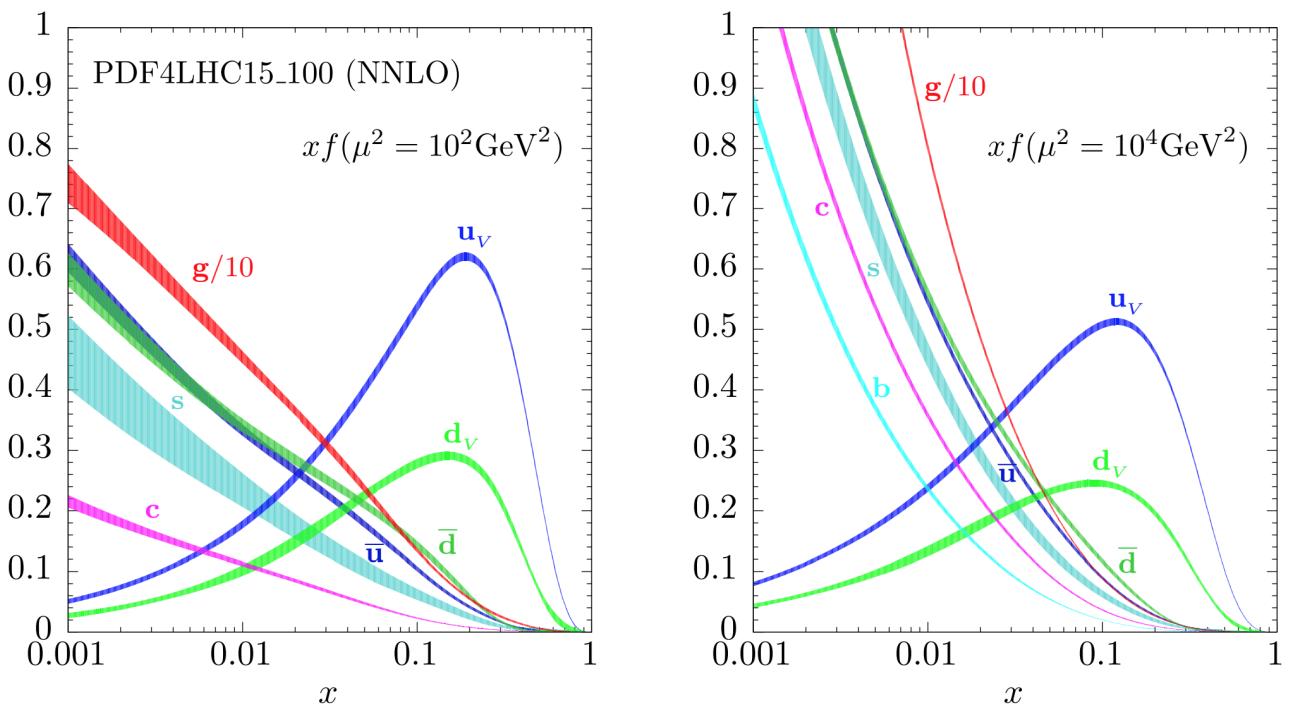


Figure 24: The evolution of a PDF4LHC15 NNLO Hessian set from $Q^2 = 10^2$ GeV to $Q^2 = 10^4$ GeV using the DGLAP. Notice the increase in the sea quark density. The PDFs include one standard deviation uncertainty band.

699

700 2.2.2 The Drell-Yan process

701 The DY process happens during the high-energy hadron-hadron scattering when quark and antiquark
 702 annihilate to form an electroweak boson [12]. It is postulated that DY cross-section in a proton-
 703 proton scattering $\sigma^{DY}(pp \rightarrow l^+l^- + X)$ can be expressed through the cross-sections of the corresponding
 704 parton-parton scattering cross-section convoluted with the PDFs of these partons:

$$\frac{d^2\sigma^{DY}}{dydM^2} = \sum_{a,b=q,\bar{q},g} \int_{\tau_1}^1 dx_1 f_a(x_1, \mu_F^2) \int_{\tau_2}^1 dx_2 f_b(x_2, \mu_F^2) \frac{d^2\hat{\sigma}_{ab}^{DY}}{dydM^2}(x_1, x_2, y, M^2, \mu_R^2, \mu_F^2). \quad (2.6)$$

In this equation $y = \frac{1}{2} \log \frac{E+p_z}{E-p_z}$ represents rapidity, M^2 is the invariant mass of the lepton pair, μ_F and μ_R are factorization and renormalisation scales correspondingly. Integration limits $\tau_{1,2} = \sqrt{\frac{Q^2}{s}} e^{\pm y}$ with s being the centre-of-mass energy of the two incoming protons. The partonic cross-sections can be in turn computed perturbatively as a series expansion in α_s [3]:

$$\frac{d^2\hat{\sigma}_{ab}^{DY}}{dydM^2}(x_1, x_2, y, M^2, \mu_R^2, \mu_F^2) = \sum_{n=0}^{\infty} \left(\frac{\alpha_s \mu_R^2}{2\pi} \right)^n \frac{d^2\hat{\sigma}_{ab}^{(n)DY}}{dydM^2}. \quad (2.7)$$

The exact sum of the expansion does not depend on the μ_F and μ_R parameters. However, finite-order calculations demand a specific choice for the two parameters. One of the common choices for the DY process is putting $\mu_F = \mu_R = M$, with M being the mass of the dilepton pair.

From equation 2.6 we can see that the rapidity distribution of the vector boson explicitly depends on the PDFs both in terms of flavour decomposition and in the sense of a particular PDF set. Figure 25 demonstrates different rapidity distributions for two centre-of-mass energies and two different PDF sets.

Let us consider partonic cross-sections, which can be constructed using an analogy from QED

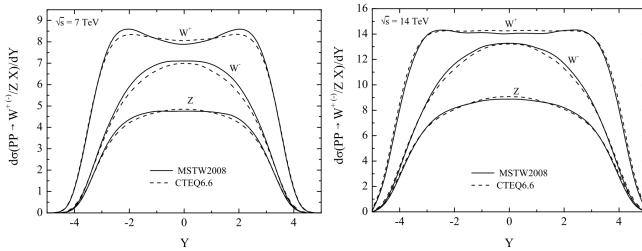


Figure 25: Rapidity distribution for the vector bosons using MSTW2008 and CTEQ6.6 PDF sets for the centre-of-mass energies of 7 and 14 TeV [13].

$e^+e^- \rightarrow \mu^+\mu^-$ - a flavour-changing process:

$$\hat{\sigma}(q\bar{q} \rightarrow e^+e^-) = \frac{4\pi\alpha_s^2}{3s} \frac{1}{N} Q_q^2. \quad (2.8)$$

Here Q_q^2 is the quark charge, $1/N$ stands for the colour factor and underlines the fact that quark and antiquark must have the matching colour in order to annihilate. In a similar way we can obtain the cross-section of the sub-processes of W and Z bosons production:

$$\begin{aligned} \hat{\sigma}^{q\bar{q}' \rightarrow W} &= \frac{\pi}{3} \sqrt{2} G_F M_W^2 |V_{qq'}|^2 \delta(s - M_W^2), \\ \hat{\sigma}^{q\bar{q}' \rightarrow Z} &= \frac{\pi}{3} \sqrt{2} G_F M_W^2 (v_q^2 + a_q^2) \delta(s - M_Z^2), \end{aligned} \quad (2.9)$$

where $V_{qq'}$ is the element of the CKM matrix, v_q (a_q) is a vector (axial vector) that couples the Z boson to the quarks. Figure 27 shows the contributions of different parton flavours into W^+ and W^- cross-sections. An assumption of narrow W resonance was used. The fact that the bosons with opposite charges are formed from different quarks makes a notable difference at the LHC experiments. Figure 26 contains the comparison of the results obtained at the LHC experiments with the NNLO theoretical predictions that use different PDF sets.

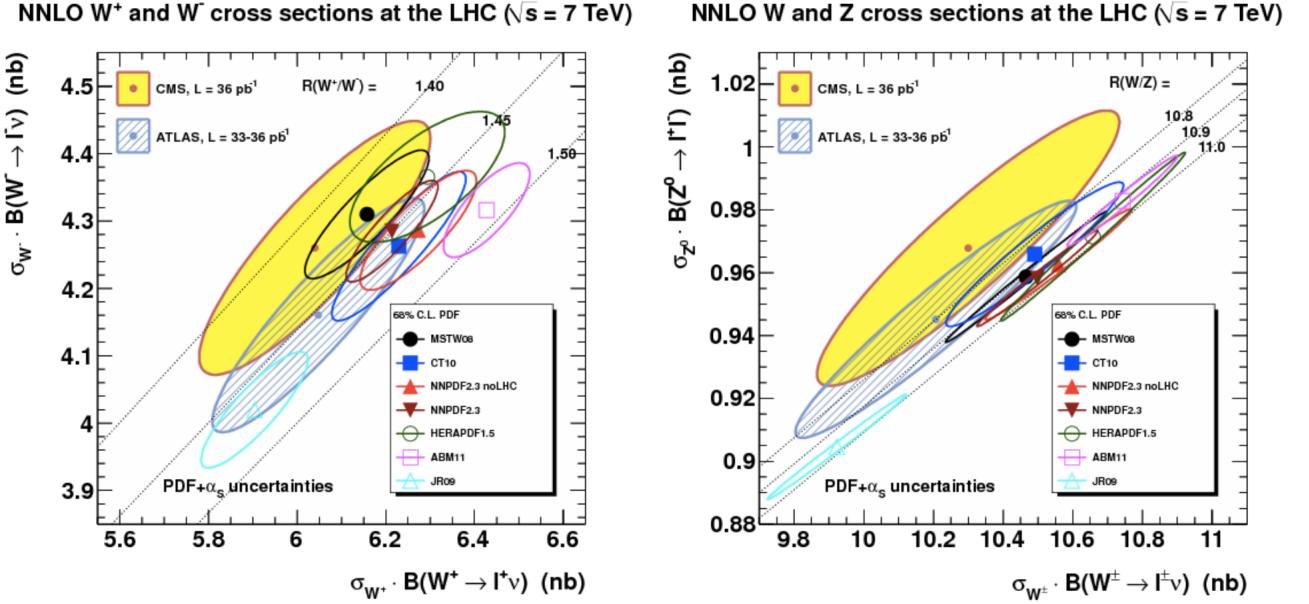


Figure 26: W and Z boson cross sections LHC at 7 TeV. ATLAS and CMS results, compared to NNLO predictions for various PDF sets [14].

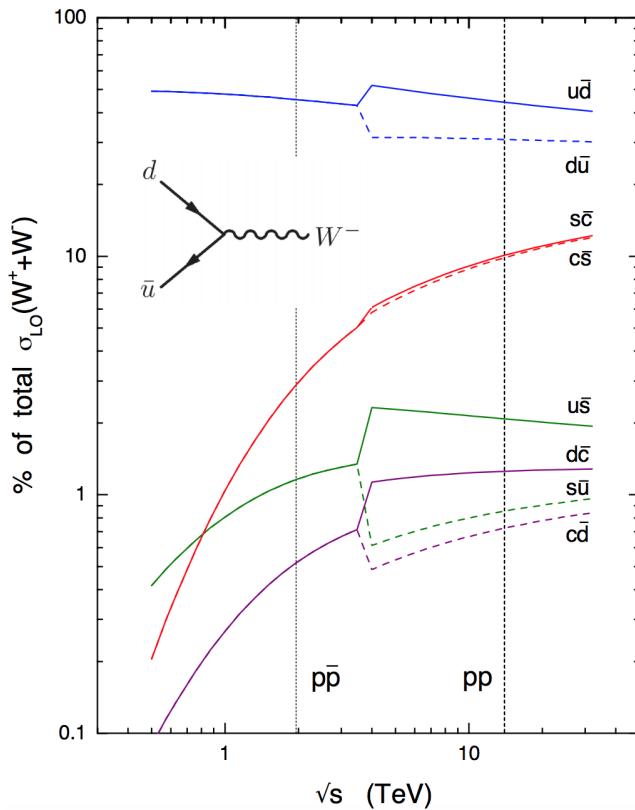


Figure 27: Parton contributions to the cross-sections of W^+ and W^- bosons for LHC and Tevatron cases [15].

727 2.3 Transverse momentum of massive vector bosons

728 The leading-order model of the DY process assumes the colliding partons to have their momentum
 729 perfectly collinear with the proton as a whole, which would mean that the vector boson p_T should
 730 peak at zero. However, most of the massive vector bosons produced in the DY process have a small
 731 yet non-zero transverse momentum $p_T \ll M_V$. This occurs due to a non-perturbative effect of partons
 732 moving within the colliding protons, having an intrinsic momentum of their own. This intrinsic
 733 momentum $\langle k_T \rangle \sim \Lambda_{QCD}$ is well parametrized using a Gaussian distribution with average value of
 734 500 [16] or 700 GeV [17], although there are ongoing efforts for a more sophisticated parametrization
 735 that would allow a better modelling of the lower part of vector boson p_T spectrum [18].
 736 The spectrum at higher values of p_T is determined by hard perturbative parton emission(s) like
 737 $q\bar{q} \rightarrow Vg$, $qg \rightarrow Vq$. The corresponding amplitudes can be conveniently expressed using Mandelstam
 738 variables:

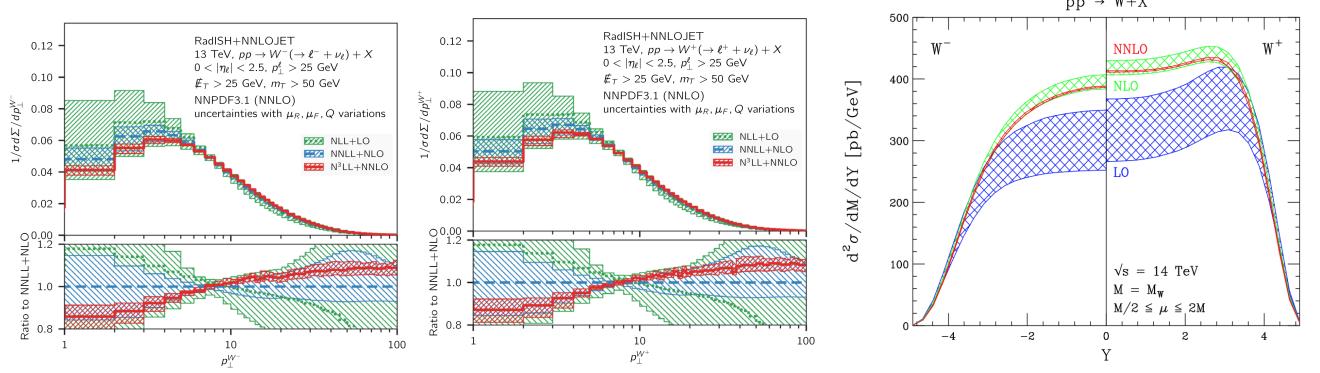
$$\begin{aligned}\sum |\mathcal{M}^{q\bar{q}' \rightarrow Wg}|^2 &= \alpha_s \sqrt{2\pi} G_F M_W^2 |V_{q\bar{q}'}|^2 \frac{8}{9} \frac{t^2 + u^2 + 2M_W^2 s}{tu}, \\ \sum |\mathcal{M}^{qg \rightarrow Wq'}|^2 &= \alpha_s \sqrt{2\pi} G_F M_W^2 |V_{q\bar{q}}|^2 \frac{1}{3} \frac{s^2 + u^2 + 2M_W^2 t}{-su},\end{aligned}\quad (2.10)$$

739 where the summation is performed over colours and spins in the final and initial states. Integrating
 740 these partonic matrix elements with the PDFs one can obtain the transverse momentum distribution
 741 $d\sigma/dp_T$. Further precision can be obtained by considering corrections from next-to-leading order
 742 processes $\sim O(\alpha_s^2)$ like $q\bar{q} \rightarrow Vgg$ - that would mainly affect the high p_T tail of the distribution.
 743 The matrix elements in 2.10 become singular when the emitted partons become soft or collinear to the
 744 initial-state partons - it is related to the poles at $u = 0$ and $t = 0$ in the denominator. Also for the NLO
 745 processes like $q\bar{q} \rightarrow Vgg$ a singularity arises if the two final-state gluons are collinear. This creates a
 746 problem for the calculation of the low- p_T part of the spectrum. Mathematically it is reflected in the
 747 appearance of different powers of logarithms like $\log M_W^2/p_T^2$ in all orders of cross-section expansion in
 748 α_s , which leads to divergences when p_T is small. This forces us to look for alternative approach that
 749 would take into account all the orders of the expansion.
 750 All-order resummation may be performed in a variety of approaches, one of the most popular is
 751 provided by parton showers. Its numerical implementation is available in a number of Monte-Carlo
 752 generators, PYTHIA, HERWIG and SHERPA are among the most used. It appears that for the case
 753 of soft and collinear gluon emission it is possible to factorize and exponentiate the logarithms in a
 754 *Sudakov form factor*, such that:

$$\begin{aligned}\frac{d\sigma}{dp_T^2} &= \sigma \frac{d}{dp_T^2} \exp\left\{-\frac{\alpha_s C_F}{2\pi} \log^2 \frac{M_W^2}{p_T^2}\right\}, \\ \exp\left\{-\frac{\alpha_s C_F}{2\pi} \log^2 \frac{M_W^2}{p_T^2}\right\} &= 1 - \frac{\alpha_s}{2\pi} C_F \ln^2 \frac{M_W^2}{p_T^2} + \frac{1}{2!} \left(\frac{\alpha_s}{2\pi}\right)^2 C_F^2 \ln^4 \frac{M_W^2}{p_T^2} - \frac{1}{3!} \left(\frac{\alpha_s}{2\pi}\right)^3 C_F^3 \ln^6 \frac{M_W^2}{p_T^2} + \dots\end{aligned}\quad (2.11)$$

755 The exponential $\exp\{G(\alpha_s, L)\}$, where $L = \log M_W^2/p_T^2$ is called the Sudakov form-factor. Its expansion
 756 by the powers of α_s defines the resummation accuracy: the term $\sim O(\alpha_s)$ is called the leading logarithm
 757 (LL), term with $\sim O(\alpha_s^2)$ is the next-to-leading logarithm (NLL) and so on.
 758 The cross-sections obtained with the resummation methods provide a good prediction for soft and

759 collinear emissions at low $p_T \ll M_W$. In order to get a combined cross-section for higher p_T region the
 760 resummed cross-section has to be *matched* with the fixed-order cross-sections of the corresponding power in α_s .



(a) W^- transverse momentum spectrum [19]. (b) W^+ transverse momentum spectrum [19].

(c) W^\pm rapidity distribution [20].

Figure 28: Kinematic distributions for W^\pm with corrections.

3

The Large Hadron Collider

762

763

764

“Potentielle citation sans aucun rapport avec le sujet”

765

— Personne inconnue, contexte à déterminer

766 3.1 Introduction

767 The study of elementary particles naturally demands a stable source of particles. At the dawn of
768 particle physics the two main sources were radioactive materials and the cosmic rays. However soon
769 researchers became in need of a more reliable source of particles in terms of particle energy, luminosity
770 and experimental repeatability. This has commenced the era of particle accelerators.

771 The first examples of particle accelerators were designed in late 1920s and early 1930s. Two different
772 designs emerged: linear and circular. The former accelerates particles via electric field during the
773 single pass through the machine, while the latter uses magnetic field to make accelerated particles go
774 in circles allowing to re-accelerate the same beam many times. On the other hand the circular design
775 comprises energy losses due to Bremsstrahlung radiation.

776 In the second half of the XX century the accelerators gradually got bigger and bigger in both size and
777 centre-of-mass energy of the accelerated particles. This has allowed to create an experimental basis for
778 the development of modern particle physics, notably the Standard Model.

779 Up to this day the biggest particle accelerator with the highest centre-of-mass energy is the Large
780 Hadron Collider (LHC). LHC is a circular collider that lies in a tunnel of 27 km under the French-Swiss
781 border next to Geneva [1]. In 2012 two biggest experiments of LHC have claimed the discovery of
782 the Higgs boson, the last elementary particle predicted by the Standard Model which was not yet
783 discovered by that time. [2], [3].

784 3.2 The LHC running sequence

785 It takes quite a journey for a proton to travel from a hydrogen tank (Fig. 31) into one of the LHC's
786 collision points. A resourceful system of pre-accelerators is necessary to make the proton beam ready
787 to get injected into one of the two LHC beam pipes. The LHC accelerator complex was not built from
788 scratch - it uses vast CERN infrastructure, that was built for the previous particle physics experiments.
789 After stripping the electrons off the atoms of hydrogen using the magnetic field the yielded protons get



Figure 31: ATLAS A hydrogen tank supplies LHC with protons [4].

790 accelerated to the energy of 50 MeV by Linac 2¹ [5]. After that the beam gets into Proton Synchrotron
 791 Booster (PSB) to be accelerated to 1.4 GeV. The next link of the pre-acceleration chain is the Proton
 792 Synchrotron (PS) - a true veteran among CERN accelerators that first accelerated protons in 1959
 793 holding the world record in acceleration energy. Currently thanks to PSB and other modifications it
 794 can sustain proton beam intensity 1000 times larger than back in 1959. The PS accelerates the beam
 795 up to 25 GeV and conveys it further to the Super Proton Synchrotron (SPS) - the second-largest particle
 796 accelerator at CERN. Back in 1983 the massive electroweak bosons were discovered at the SPS but even
 797 now it serves as a main accelerator for a NA61/SHINE, NA62 and COMPASS experiments. The SPS
 798 raises the beam energy to 450 GeV and finally injects it into the LHC beam pipes (see Fig 32).

799

800 The LHC has inherited its 27 km tunnel from the predecessor, an electron-positron collider called
 801 Large Electron-Positron (LEP). However, all the LEP hardware has been replaced to sustain the
 802 conditions of the LHC beam. About 2/3 of the LHC circumference length is occupied by the dipole
 803 magnets that bend the trajectory of the proton beam to keep it within the pipe. These magnets use
 804 superconducting coils that conduct a current of 11080 amperes to produce a magnetic field of 8.3 tesla
 805 Proton acceleration is maintained by the radio-frequency (RF) cavities (Fig. 33a). Besides acceleration
 806 particles the RF cavities are also responsible for beam bunching i.e. separating the beam into a train of
 807 separated particle packs, each containing about 10^{11} protons. During LHC Run 2 the bunches were
 808 separated by 7 meters (25 ns) with a maximum of 2556 circulating bunches. The LHC has four crossing
 809 points, where the two beams are crossed in order to collide protons. Naturally, the particle detectors
 810 are installed at these four points. Before getting directed at the crossing point the beams get squeezed
 811 to make their cross-section as small as $16 \mu m^2$ (Fig 34a).

¹After Run 2 Linac 2 has been decommissioned to be succeeded by Linac 4.

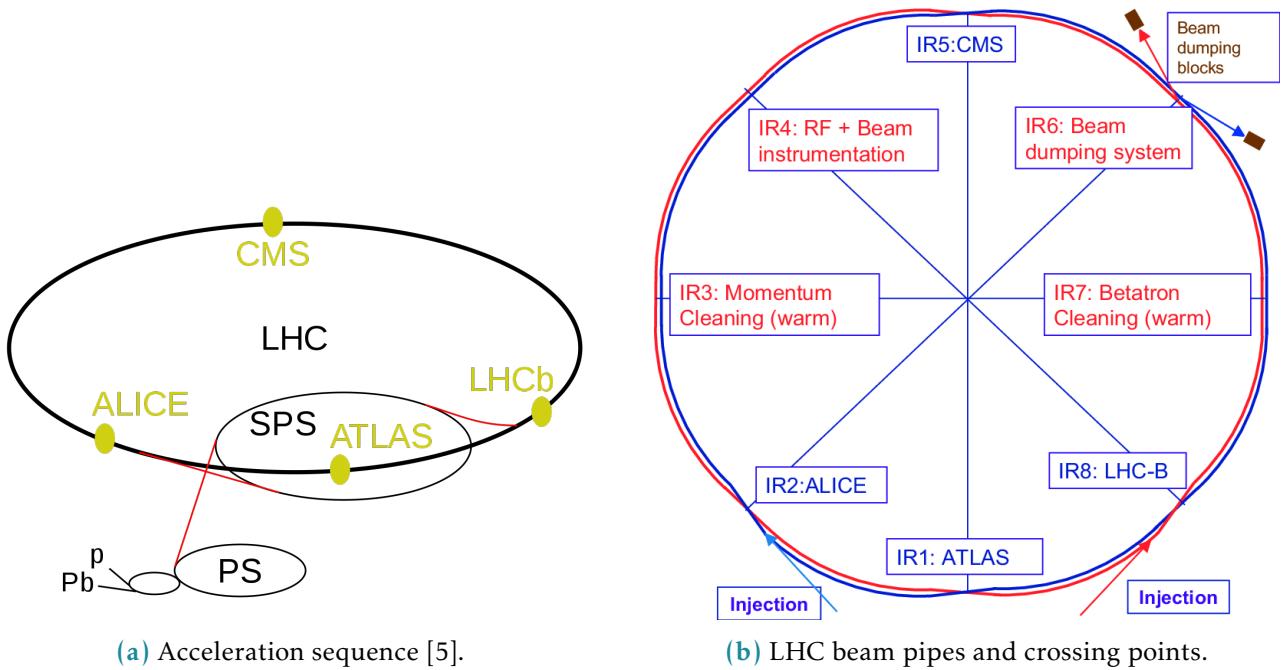


Figure 32: Schematic depiction of the LHC ring.

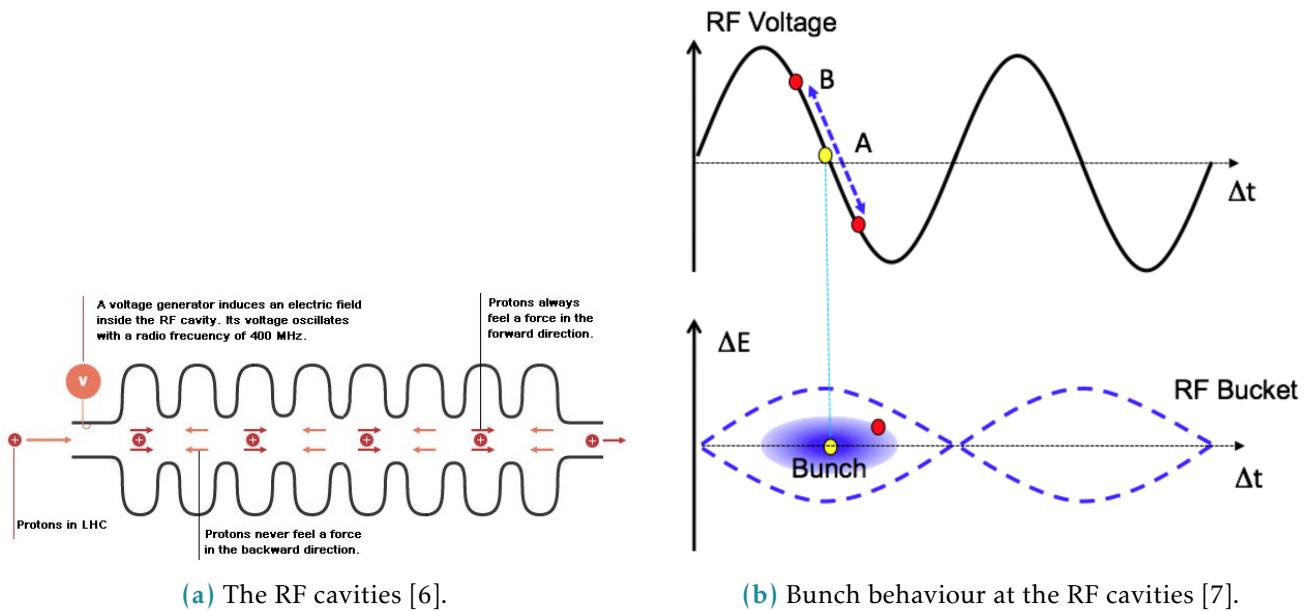
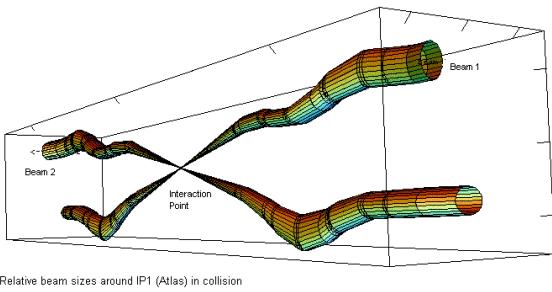
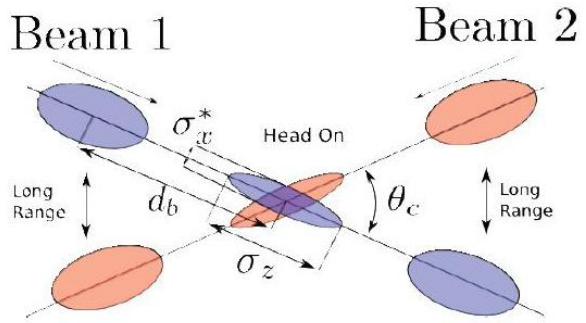


Figure 33: Bunching at RF cavities



(a) The two beams getting squeezed at the IP [8].



(b) Bunches at the collision point [9].

Figure 34: The crossing

812 In order to estimate the number of single proton-proton interactions in the crossing beams a value
 813 called instantaneous luminosity (simply called luminosity) is introduced. It is the proportionality
 814 factor between the number of events per second dR/dt and the cross-section σ_p :

$$\frac{dR}{dt} = \mathcal{L} \cdot \sigma_p.$$

815 For the case of head-on collisions the luminosity would equal to [10]:

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi \sigma_x \sigma_y}, \quad (3.1)$$

816 with N_1 and N_2 being the intensities of the two colliding beams, f is the revolution frequency, N_b -
 817 number of bunches per beam, σ_x, σ_y - r.m.s. beam widths in the corresponding dimensions, assuming
 818 that the bunches in both beams have the same size and Gaussian profiles.
 819

820 Head-on crossing of the beams would ensure maximal luminosity given the same beams, but on the
 821 other hand the measurement would suffer from unwanted beam-to-beam effects. To avoid it the beams
 822 at the LHC are crossed at an angle, which is called the crossing angle 34b. For the case of head-on
 823 collisions the luminosity gets a factor \mathcal{F} [10]:

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi \sigma_x \sigma_y} \cdot \mathcal{F}, \quad (3.2)$$

824 with geometric factor

$$\mathcal{F} = \frac{1}{\sqrt{1 + \left(\frac{\sigma_s}{\sigma_x} \frac{\theta_c}{2}\right)}},$$

825 where σ_s is the r.m.s. of the bunch length and θ_c is the crossing angle. Varying the parameters like
 826 beam intensity, bunch spacing, beam profile, crossing angle and others becomes a flexible tool for
 827 luminosity control. This comes in handy for different physics analysis, as some processes are rare
 828 and demand as much luminosity as possible (this is true, for example, for most of the Higgs studies),

Collision energy	Year	Integrated luminosity, pb^{-1}	Total uncertainty, %
5 TeV	2017	258	1.6
13 TeV	2017	148	2.1
13 TeV	2018	193	1.5

Table 31: Energy and luminosity of the special low- μ runs.

whereas the others suffer from high pile-up conditions. Instantaneous luminosity integrated over a period of time is called the integrated luminosity:

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t) dt, \quad (3.3)$$

and is directly related to the number of observed events $\mathcal{L}_{\text{int}} \cdot \sigma_p = N_{events}$. Precise measurement of the integrated luminosity is crucial for the LHC results and serves as one of the nuisance parameters for most of the analyses. A comprehensive overview on the luminosity determination at proton colliders can be found here [11]. Absolute luminosity measurements at the LHC are performed predominantly using the van-der-Meer (vdM) scan method [12], [13].

3.3 Special low pile-up run during LHC Run 2

During the Run 2 that lasted from 2015 to 2018 the ATLAS experiment has collected $146.9 fb^{-1}$ of data under different bunch crossing conditions. However the precise measurement of the W boson-related processes demands special conditions. High number of proton-proton collisions per bunch crossing leads to contamination of the final state signal with soft collisions products. This effect, known as pile-up, complicates object reconstruction and results in systematic uncertainties growth. For this reason two special runs with low number of interactions per bunch crossing have been performed by the LHC in 2017 and 2018 at the energies of 5 and 13 TeV.

The table ?? contains information on the data collected at ATLAS experiment during the special low pile-up run with $\langle \mu \rangle \approx 2$.

846

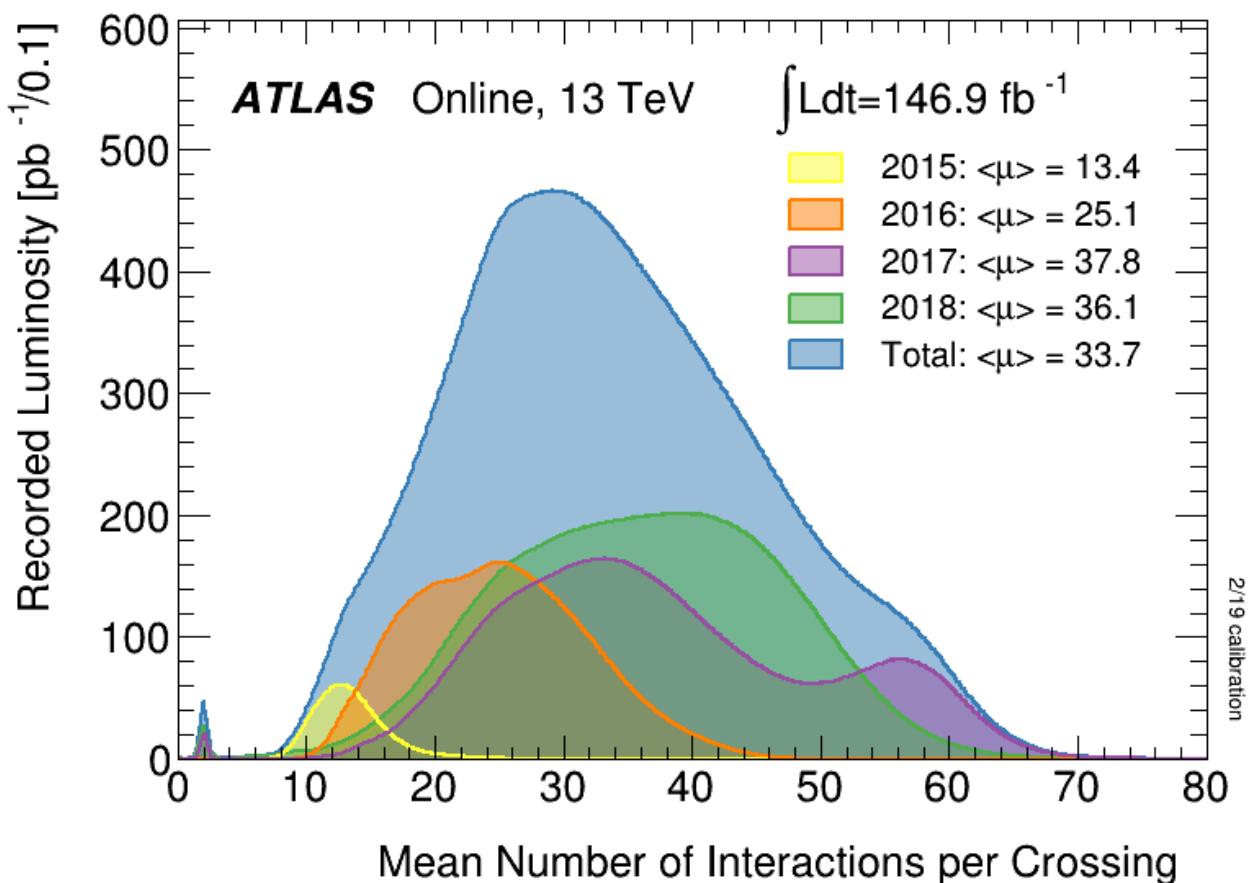


Figure 35: Number of Interactions per bunch crossing in ATLAS Run 2 [14]. A little bump around $\mu \approx 2$ corresponds to special low pile-up runs.

The ATLAS detector

849 “Potentielle citation sans aucun rapport avec le sujet”

— Personne inconnue, *contexte à déterminer*

851 4.1 General description and layout

ATLAS experiment is a multipurpose detector at the LHC built, along with its peer Compact Muon Solenoid (CMS), in order to probe the p-p, A-A and p-A collisions using the full LHC luminosity [1]. Being the largest (but not the heaviest) detector ever built for a collider experiment the ATLAS detector comprises 44m in length, 25m in height and weights 7000 tonnes.

The detector has a cylindrical shape and is an onion-like arrangement of several detector systems centered at the Interaction Point (IP) as shown in fig. 41. The sub-detectors operate in the magnetic field created by the solenoid and toroid magnets (ATLAS owes its name to the latter). Data acquisition and recording is controlled by the Trigger and Data Acquisition (TDAQ) systems, allowing eventually to lower the event rate to a value, acceptable for the data storage [2]. The named systems are described in more detail in this chapter.

862 4.2 Coordinate system

863 The ATLAS results often reference ATLAS coordinates briefly described in this subsection. The origin
 864 of the right-handed coordinate system is placed at the IP with z -axis directed along the beam direction.
 865 This, in turn, defines the transverse x - y plane with x axis pointing towards the center of the LHC ring
 866 and y axis directed upwards. All transverse observables like p_T and E_T are defined in this 2D plane.
 867 Besides the mentioned Cartesian coordinates the azimuthal angle ϕ is defined in the transverse plane
 868 around the beam axis. Polar angle θ is the elevation angle measured from the beam axis. The following
 869 metric quantities are also to be mentioned:

- Pseudorapidity $\eta = -\ln \tan(\theta/2)$,
 - Rapidity $y = 1/2 \ln [(E+p_z)/(E-p_z)]$
 - The distance between particles $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$

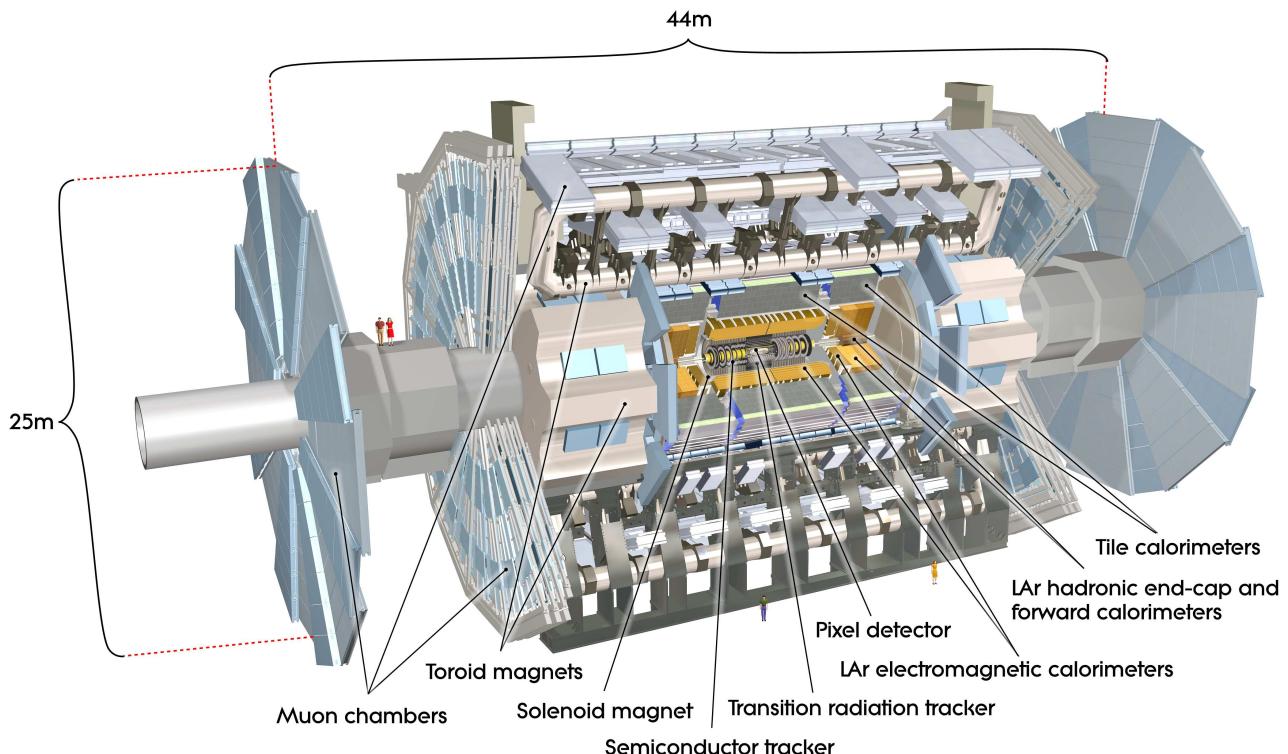


Figure 41: ATLAS detector general layout

873 4.3 Magnet system and magnetic field

874 ATLAS has a hybrid system of four superconducting magnets which has 22 m in diameter, 26 m in
 875 length and stores an energy of 1.6 GJ [3]. The windings of the magnets are schematically shown in fig.
 876 42. The four magnets that comprise the magnet system are the following:

- 877 • The central solenoid is aligned with the beam axis providing 2T axial magnetic field for the inner
 878 detector.
- 879 • A barrel toroid produces toroidal magnetic field of about 0.5T for the muon detectors in the
 880 barrel region.
- 881 • Two end-cap toroids produce toroidal magnetic field of approximately 1T for the muon detectors
 882 in the end-cap regions.

883 4.4 Inner detector

884 The ATLAS Inner Detector (ID) is designed to deliver pattern recognition, high-resolution momentum
 885 measurement [4],[5] along with primary and secondary vertex determination for charged particle tracks

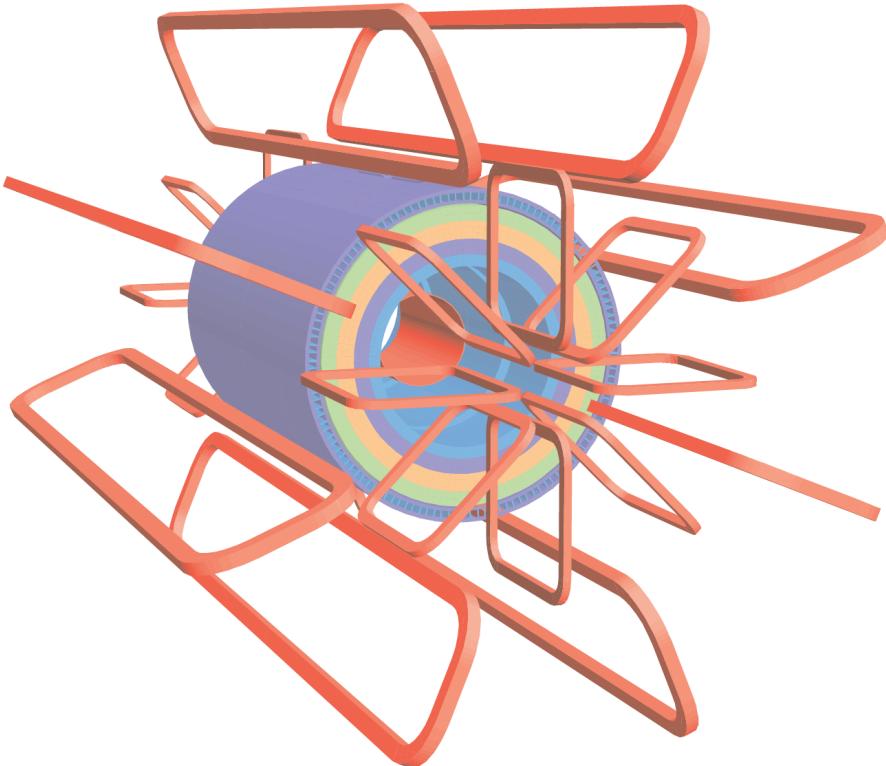


Figure 42: Geometry of ATLAS magnet windings

886 above a designated p_T threshold of 0.5 GeV (in some cases being capable of going as low as 0.1 GeV)
887 within the pseudorapidity range $|\eta| < 2.5$. The inner detector provides reliable electron identification
888 in rapidity range of $|\eta| < 2.0$ for energies from 0.5 GeV to 150 GeV.

889 The ID layout is a result of the technical requirements: it is assembled in a cylindrical envelope of 3512
890 mm in length and 1150 mm in radius. It is surrounded by the magnetic field of 2T imposed by the
891 superconducting solenoid (see section 4.3).

892 Three independent sub-detectors complement each other in the inner detector (see fig. 104a (a)):

893 • Silicon pixel with 3 cylindrical layers for the barrel and 3 discs on each side for the end-cap. It
894 provides the highest granularity around the vertex region. Normally each track hits three pixel
895 layers. The pixel detector has about 80.4 million readout channels. Each of 1744 identical pixel
896 sensors has 47232 pixels and 46080 readout channels. About 90% of the pixels have the size of
897 $50 \times 400 \mu\text{m}^2$, the remaining pixels are a bit longer: $50 \times 600 \mu\text{m}^2$.

898 • Silicon microstrip layers (SCT) with 4 cylindrical layers and 9 discs on each side for the end-cap.
899 A track typically crosses the strip layers in four space points. SCT has approximately 6.3 millions
900 readout channels from its 15912 sensors. There are 768 active strips of 12 cm lenght and $80 \mu\text{m}$
901 width per sensor plus two bias potential strips on the sensor edges.

902 • Transition radiation tracker (TRT) with 73 straw planes in the barrel and 160 straw planes in

the end-cap. The TRT has around 351,000 readout channels and detects in average 36 hits per track. The straw tubes that make up the TRT module are 4 mm thick and 1.44 m long (0.37 m in the endcap) and made out of polyamide films reinforced with carbon fibers. The straws are filled with gas mixture of 70% Xe, 27% CO₂ and 3%O₂ and supplied with gilded tungsten anodes which are directly connected to the readout channels. The pixel and SCT sensors are highly radiation-proof and operate in the temperature range from -5°C to -10°C to minimize the radiation damage, while the TRT module operates at room temperature.

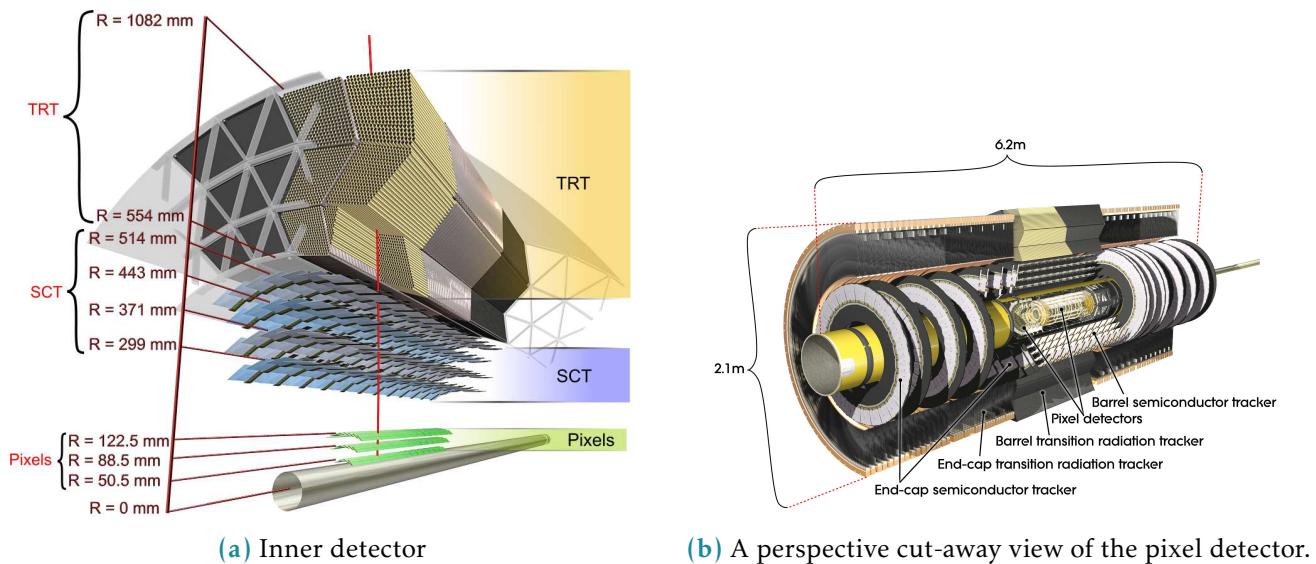


Figure 43: Fichier Gerber des modèles d CFR-34 et CFR-35.

910 4.5 Calorimeter system

911 The ATLAS calorimeter system covers the rapidity range within $|\eta| < 4.9$ and consists of several different
 912 detector systems. A rapidity region matched to the inner detector possesses fine granularity perfectly
 913 suited for high-precision measurements of photons and electrons. The remaining part's granularity is
 914 coarser but enough to perform jet reconstruction and measure E_T^{miss} . The view of ATLAS calorimeter is
 915 presented on fig. 44.

916 Besides measuring the energy of travelling particles calorimeters must also contain electromagnetic and
 917 hadronic showers, limiting their ability to go penetrate the calorimeter completely and get to the muon
 918 chambers. This provides a typical scale for size of the calorimeter modules: the EM calorimeter[6] is
 919 >22 radiation lengths (X_0) in the barrel and $>24X_0$ in the end-caps. The hadronic calorimeter has the
 920 thickness of 9.7 interaction lengths (λ) in the barrel and 10λ of in the endcap, which is enough to keep
 921 the leakage level below the typical muon background. This size also provides good resolution for the
 922 E_T^{miss} measuremet. The detailed description of the calorimeter system can be found in the table 41.

- 923 The tile calorimeter[7] uses scintillating tiles as active material alternated with steel absorbers. All the other calorimeter systems use liquid argon as an active medium with lead sampling.

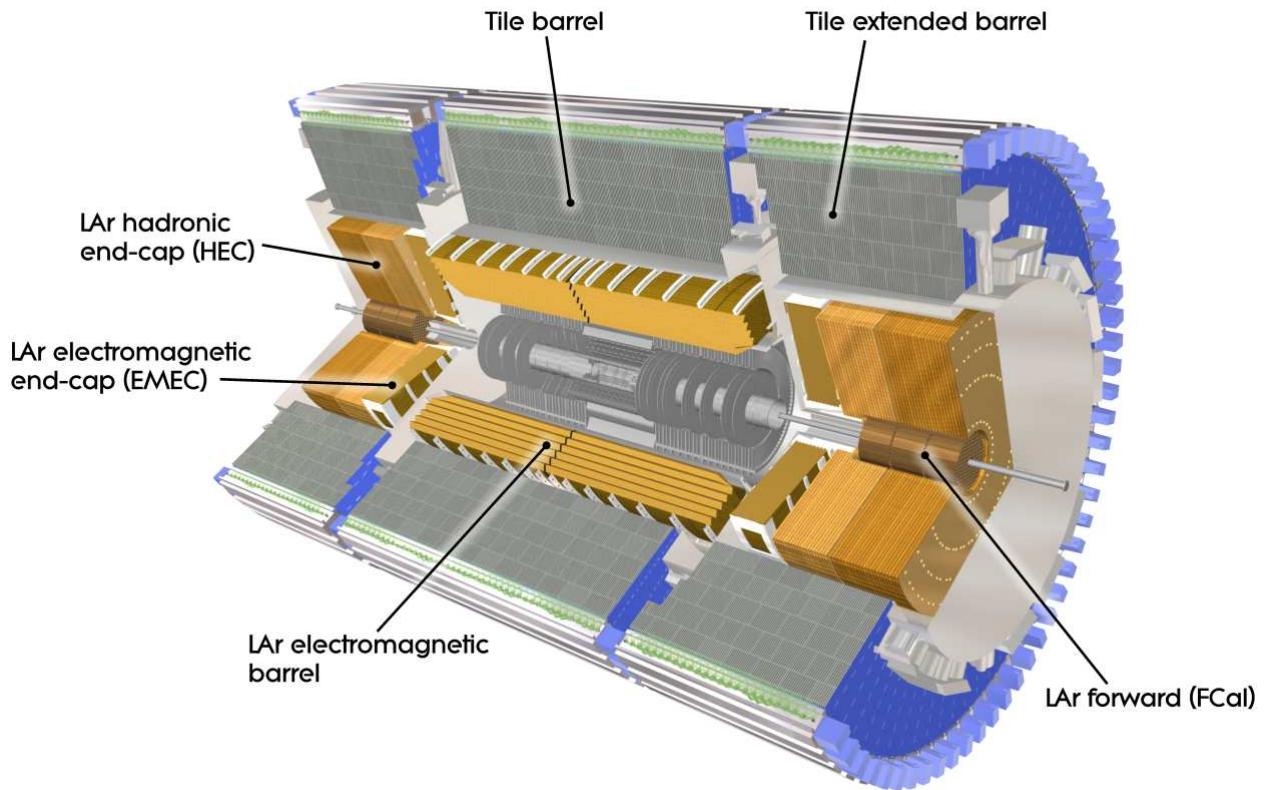


Figure 44: ATLAS calorimeter general layout

924

925 4.5.1 Electromagnetic calorimeter

- 926 The EMC has two submodules:

- 927 • EMC barrel detector.
928 • Electromagnetic end-cap calorimeter (EMEC) end-cap detector.

- 929 The EMC barrel module consists of two identical half-barrels 3.2 meters long with inner and outer radii
930 2.8 m and 4 m respectively. There is a 4mm gap at $z = 0$ between the half-barrels. The second crack is
931 situated between the barrel and the end-cap at $1.37 < |\eta| < 1.52$. The EMEC comprises of two pairs of
932 coaxial wheels of 63 cm thick having inner and outer radii of 330 mm and 2098 mm respectively. The
933 crack between the two wheels makes a third crack at $|\eta| = 2.5$. Both barrel and end-cap electromagnetic
934 calorimeters are designed to have an accordion-shaped absorbers made out of lead plates, coated in
935 stainless steel sheets. The readout electrodes are placed in the gaps between the absorbers. This type

	Barrel	End-cap			
EM Calorimeter					
Number of layers and $ \eta $ coverage					
Presampler	1	$ \eta < 1.52$	$1.5 < \eta < 1.8$		
Calorimeter	3	$ \eta < 1.35$	$1.375 < \eta < 1.5$		
Granularity $\Delta\eta \times \Delta\phi$ versus $ \eta $					
Presampler	0.025×0.1	$ \eta < 1.52$	$1.5 < \eta < 1.8$		
Calorimeter 1st layer	$0.025/8 \times 0.1$	$ \eta < 1.40$	$1.375 < \eta < 1.425$		
	0.025×0.025	$1.425 < \eta < 1.5$	$1.425 < \eta < 1.5$		
			$0.025/8 \times 0.1$		
			$1.5 < \eta < 1.8$		
			$0.025/6 \times 0.1$		
			$1.8 < \eta < 2.0$		
			$0.025/4 \times 0.1$		
Calorimeter 2nd layer	0.025×0.025	$ \eta < 1.40$	$1.375 < \eta < 1.425$		
	0.075×0.025	$1.4 < \eta < 1.475$	$1.425 < \eta < 2.5$		
			0.1×0.1		
Calorimeter 3rd layer	0.050×0.025	$ \eta < 1.35$	$1.5 < \eta < 2.5$		
Number of readout channels					
Presampler	7808	1536 (both sides)			
Calorimeter	101760	62208 (both sides)			
LAr hadronic end-cap					
$ \eta $ coverage		$1.5 < \eta < 3.2$			
Number of layers		4			
Granularity $\Delta\eta \times \Delta\phi$		0.1×0.1	$1.5 < \eta < 2.5$		
		0.2×0.2	$2.5 < \eta < 3.2$		
Readout channels		5632 (both sides)			
LAr forward calorimeter					
$ \eta $ coverage		$3.1 < \eta < 4.9$			
Number of layers		3			
Granularity $\Delta x \times \Delta y$		$F\text{Cal } 3.0 \times 2.6$	$3.15 < \eta < 4.30$		
		$F\text{Cal: } \sim\text{four times finer}$	$3.10 < \eta < 3.15$		
			$4.30 < \eta < 4.83$		
		$F\text{Cal2 } 3.3 \times 4.2$	$3.24 < \eta < 4.50$		
		$F\text{Cal2: } \sim\text{four times finer}$	$3.20 < \eta < 3.24$		
			$4.50 < \eta < 4.81$		
		$F\text{Cal3 } 5.4 \times 4.7$	$3.32 < \eta < 4.60$		
		$F\text{Cal3: } \sim\text{four times finer}$	$3.29 < \eta < 3.32$		
Readout channels		$4.60 < \eta < 4.75$			
		3524 (both sides)			
Scintillator tile calorimeter					
	Barrel	Extended barrel			
$ \eta $ coverage	$ \eta < 1.0$	$0.8 < \eta < 1.7$			
Number of layers	3	3			
Granularity $\Delta\eta \times \Delta\phi$		0.1×0.1	0.1×0.1		
		0.2×0.2	0.2×0.1		
Readout channels	5760	4092 (both sides)			

Table 41: ATLAS calorimeter in numbers

936 of geometry allows full coverage in ϕ without cracks together with fast extraction of the signal from
 937 both sides of the electrodes. The orientation of the accordion waves is axial in the barrel and radial in
 938 the end-caps (see fig. 44). These features of the calorimeter lead to virtually uniform performance in ϕ
 939 dimension.

940 Segmentation in η is very different in the layers of the calorimeter, but the second layer always has
 941 the finest granularity because the egamma particles are supposed to leave most of their energy in the
 942 second calorimeter layer. In order to correct for the energy losses upstream the barrel calorimeter is
 943 preceded by a thin LAr active layer of 11mm thick called a presampler. For more details on η coverage
 944 and granularity see table 41.

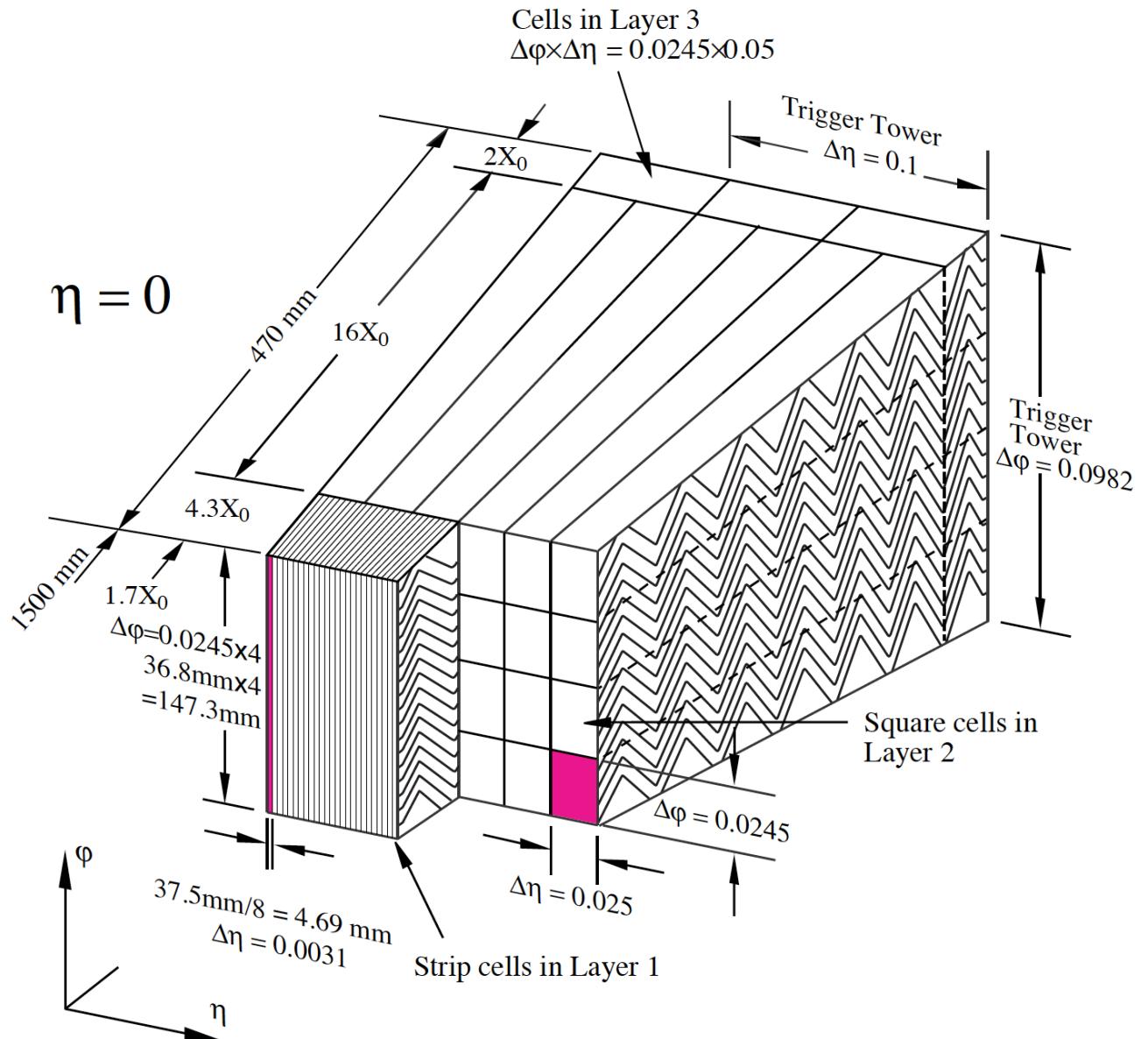
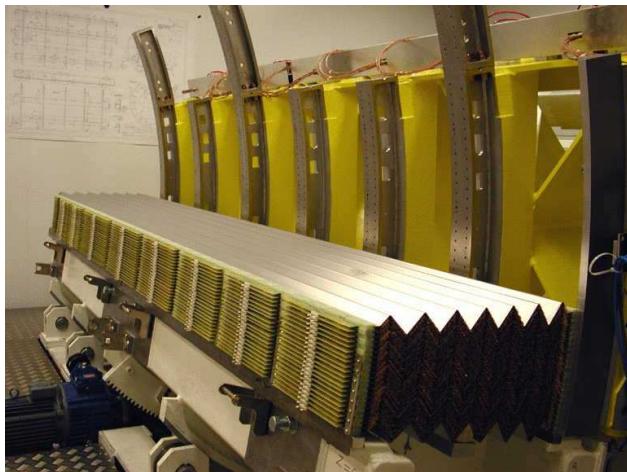


Figure 45: ATLAS EM calorimeter layers



(a) Barrel



(b) End-cap

Figure 46: Accordion absorbers of the EMC

945 4.5.2 Hadronic calorimeter

946 The hadronic calorimeter (HC) is combined of three submodules:

- 947 • HC scintilating tile detector, a steel sampled detector divided in turn into central barrel having
948 5.8 m in length and two extended barrels 2.6 m in length each. The extended barrels have inner
949 radii of 2.28 m and outer radii of 4.25 m. The tile calorimeter consists of three layers having
950 about 1.5, 4.1 and 1.8 interaction lengths λ in the barrel and 1.5, 2.6 and 3.3 λ s in the extended
951 barrel.
- 952 • Hadronic end-cap calorimeter (HEC) detector is a liquid argon calorimeter sampled with copper.
953 It has two pairs of independent wheels symmetrically located behind the EMEC called the frona
954 and the rear wheel. The wheels are cylindrical, their outer radius is 2030mm.
- 955 • Forward calorimeter (FCal) detector modules are located about 4.7 m from the IP and are
956 subjected to very high particle flux and radiation. It consists of three wheels 45 cm deep each.
957 The first one, FCal1 is sampled with copper intended for the measurement of electromagnetic
958 processes. The two other wheels FCal2 and FCal3 are sampled with tungsten and designed for
959 the hadronic showers measurement.

960 The number of the readout channels as well as the η coverage of every module and submodule is
961 described in the Table 41.

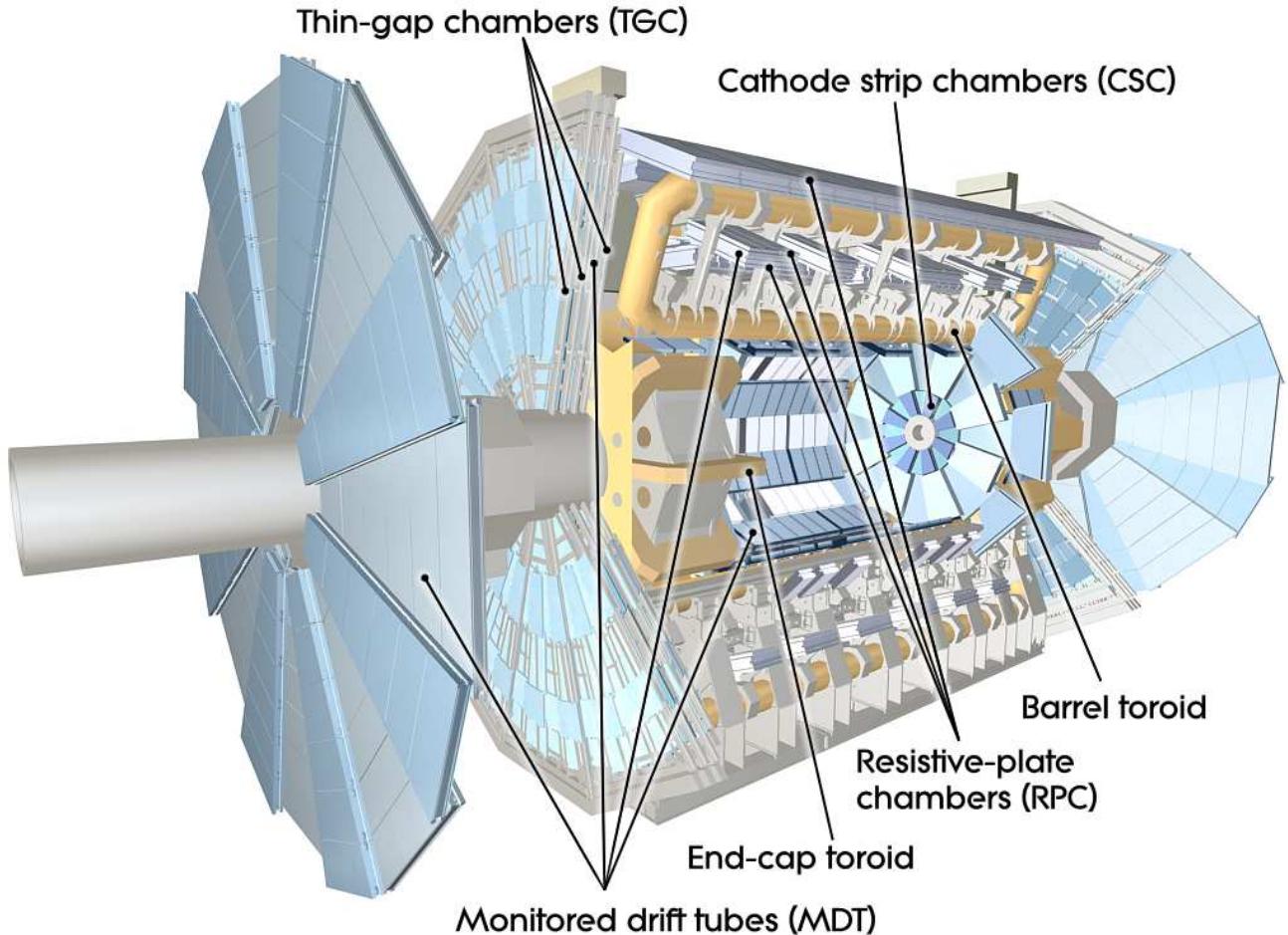


Figure 47: ATLAS muon system

4.6 Muon detectors

Most of the muons produced as a result of the p-p collisions are able to penetrate through calorimeters and make it to the muon detectors where their tracks are getting measured. The spectrometer provides high-precision measurement of the muon momenta in the rapidity range of $|\eta| < 2.7$ and approximate transverse momentum range of $3 \text{ GeV} < p_T < 3 \text{ TeV}$. The lower bound on the momentum is mainly due to energy losses in the calorimeter, while the upper bound is caused by the saggita bias coming from the tracking chambers alignment. The goal p_T resolution is about 10% for a 1 TeV muon track.

The muon tracks[8],[9] are bent by the torroidal magnets allowing to determine muon kinematic properties. The large barrel toroid covers the rapidity range of $|\eta| < 1.4$, while at $1.6 < |\eta| < 2.7$ the tracks are bent by the smaller end-cap magnets. The deflection in the transition region of $1.4 < |\eta| < 1.6$ is provided by the barrel and end-cap fields combined.

The general layout of the muon spectrometer is depicted on fig. 47, the parameters of the muon systems can be found in table ??.

Just like the rest of the detector systems the muon spectrometer is split into

975 the barrel and the end-cap parts.

976 The muon spectrometer possesses a fast triggering system able to trigger for muons in the rapidity
977 range of $|\eta| < 2.4$. It delivers the track information within a few tens of nanoseconds after the particle
978 passage which also allows to use it for the bunch-crossing identification. The trigger chambers measure
979 both η and ϕ coordinates of a track of which the former is in the bending plane and the latter is in the
980 non-bending plane.

981 There are two types of fast triggering detectors used in the muon spectrometer:

- 982 • The Resistive Plate Chambers (RPCs) is a gaseous electrode-plate detector filled with a $C_2H_2F_4/Iso-$
983 C_4H_{10}/SF_6 gas mixture (94.7/5/0.3). Two resistive plates of phenolic-melaminic plasctic lami-
984 nate are separated by insulating spacers of 2 mm thickness. The plates contain an electric field of
985 about 4.9 kV/mm such that the ionizing tracks cause avalanches towards the anode. The signal is
986 read out through the capacitive coupling of metalic strips, mounted to the resistive plates. The
987 RPCs have nominal operating voltage of 9.8 kV and provides an excellent time resolution of a
988 few ns with a supported local rate capability of 1000 Hz/cm^2
- 989 • Thin Gap Chambers (TGCs) are multi-wire proportional chambers with the wire-to-cathode
990 ditance of 1.4 mm and wire-to-wire distance of 1.8 mm and wire potential of 2900 V. The 2.8-mm
991 gas gap is filled with highly quenching gas mixture of CO_2 and $n-C_5H_{12}$ (55/45). Small distance
992 between the wires allows a very good time resolution of <25 ns in 99% of cases .

993 The precision-tracking chambers measure the coordinate of a track in the bending plane which is then
994 matched with the second coordinate, measured by the trigger chamber.

995 There are two types of precision tracking systems used:

- 996 • The Monitored Drift Tubes (MDTs) are pressurised drift tubes with a diameter of 29.970 mm
997 filled with Ar/CO_2 at 3 bar. Once the muon penetrates the tube it ionises the gas and the
998 ionised electrons are collected at the central tungsten-renium wire of $50 \mu\text{m}$ in diameter and at a
999 potential of 3080 V. This type of design carries several advantages: mechanical stiffness hence
1000 the alignment precision, reliability coming from the fact that a failure of a single tube would
1001 not cause malfunction of the others. MDTs counting rate is limited to 150 Hz/cm^2 which is not
1002 sufficient for the innermost layer in the forward region of $2.0 < |\eta| < 2.7$.
- 1003 • Cathode Strip Chambers (CSCs) are gas detectors filled with Ar/CO_2 in 80/20 proportion. The
1004 ionised electrons are collected at the wires which are oriented in the radial direction and operate
1005 at a potential of 1900 V. They are installed in the so-called Small Wheels and there are 16 CSCs
1006 on either side of the ATLAS detector. . The CSCs are able to provide a countng rate of 1000
1007 Hz/cm^2 which makes it a reasonable replacement for the MDTs in the region close to the beam.

1008 The precision-tracking chambers in the barrel are positioned between and on the coils of the supercon-
1009 ducting barrel thoroid magnet. They form three concentric cylindrical shells around the beam axis

at the approximate radii of 5 m, 7.5 m and 10 m. In the barrel region the RPCs were chosen for the fast triggering whereas the MDTs provide the precision tracking. The end-cap muon spectrometer is assembled in the form of large wheels perpendicular to the beam axis and located at distances about 7.4 m, 10.8 m, 14m and 21.5 m from the interaction point. The triggering in the end-cap is provided by the TGCs. Most of the precision tracking chambers are the MDTs similarly to the barrel, except for the forward region of $2.0 < |\eta| < 2.7$ where the CSCs are installed in the innermost tracking layer. The reason for that is their higher resistance to radiation and increased particle flow which becomes an issue if you get closer to the beam.
 Barrel and end-cap alignment is illustrated on fig. 48 which contains the side and transverse views of the muon spectrometer.

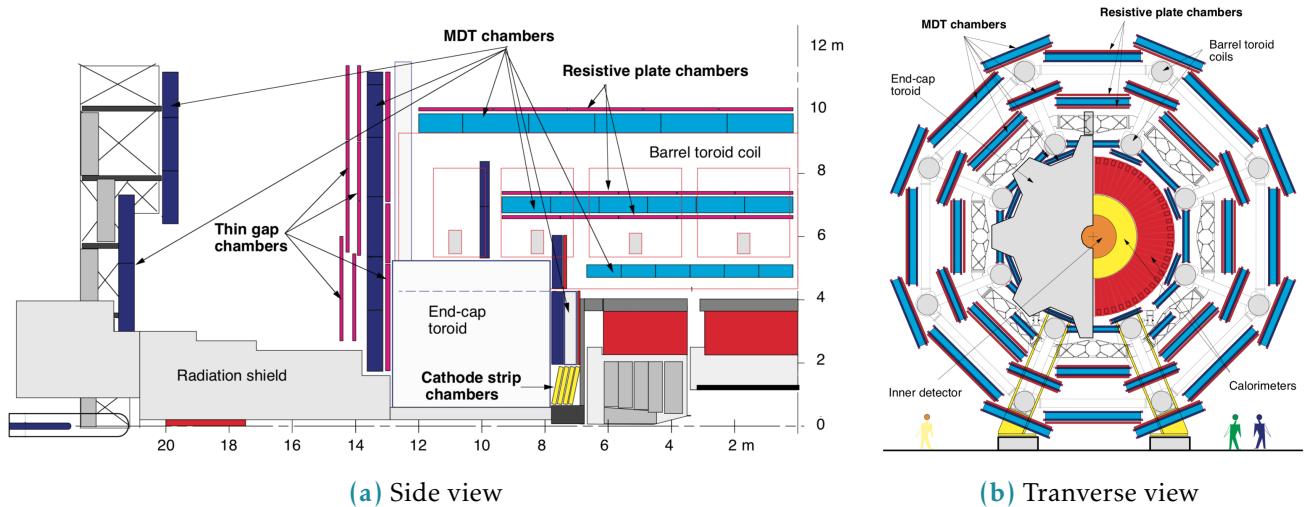


Figure 48: Cut views of the muon systems

1019

1020 4.7 Forward detectors

There are three detector systems that cover the ATLAS forward region (see fig. 49): LUminosity measurement using Cherenkov Integrating Detector (LUCID), Zero-Degree Calorimeter (ALFA) and Absolute Luminosity for ATLAS (ZDC). The measurement of luminosity is the main goal of the first two detectors and has fundamental importance: it provides the normalization scale for all the observed processes.
 LUCID[10], [11] is the main ATLAS relative luminosity monitor. The main purpose of the LUCID detector is to detect inelastic p-p scattering in the forward region measuring the integrated luminosity and performing online monitoring of the instantaneous luminosity and beam conditions with uncertainty of about few percent. It is symmetrically installed at ± 17 m from the interaction point and at a radial distance of about 10 cm from the beam line (resulting in $|\eta| \approx 5.8$). On each side four bundles of quartz fibers are used as a medium producing Cherenkov radiation directing the Cherenkov light into the 16

Monitored drift tubes	MDT
Coverage	$ \eta < 2.7$ (innermost layer: $ \eta < 2.0$)
Number of chambers	1088 (1050)
Number of channels	339 000 (354 000)
Function	Precision tracking
Cathode strip chambers	CSC
Coverage	$2.0 < \eta < 2.7$
Number of chambers	32
Number of channels	31 000
Function	Precision tracking
Resistive plate chambers	RPC
Coverage	$ \eta < 1.05$
Number of chambers	544 (606)
Number of channels	359 000 (373 000)
Function	Triggering, second coordinate
Thin gap chambers	TGC
Coverage	$1.05 < \eta < 2.7$
Number of chambers	3588
Number of channels	318 000
Function	Triggering, second coordinate

Table 42: ATLAS muon spectrometer subsystems coverage and parameters

1032 Photomultipliator Tubes (PMTs) placed outside the radiation shielding.

1033 The ALFA[12] detector is used to measure the absolute luminosity through elastic scattering at small
1034 angles. In order to perform such measurement we need to meet the following conditions:

1035

1036 • The beam has to be more parallel than normally. Special collider beam optics allowing high values
1037 of the amplitude function at the interaction point β^* together with reduced beam emittance.

1038 • To be sensitive to small angles the detectors have to be placed as far as possible from the
1039 interaction point and close to the beam. This is why the detectors are located inside the Roman
1040 pots at ± 240 from the interaction point. On each side there are two Roman pots separated by
1041 four meters.

1042 The Roman pot windows allow the elastically scattered protons reach the square scintillating fibres of
1043 0.5 mm width which are in turn connected to multi-anode PMTs through the light-guides. The detector
1044 provides a spacial resolution of $30 \mu\text{m}$ and allows to measure absolute luminosity with uncertainty of
1045 1.7% for the Run 2[13].

1046 ZDC are used to detect forward neutrons at $|\eta| > 8.3$ in heavy-ion collisions, which in turn allows to
1047 determine the centrality of such collisions. The detector is installed at ± 140 m from the interaction

1048 point. Every ZDC arm consists of 4 modules: one electromagnetic and three hadronic. These modules
 1049 are quartz rods shielded by the tungsten plates and connected to the PMTs via the light-guides allowing
 1050 to measure incending particle energy and position. The EM module has a better position resolution
 1051 mapping each of 96 quartz rods into a single pixel, while the hadronic modules map a bundle of four
 1052 rods into a pixel. Only one of the three hadronic modules per arm provide position-sensing rods and
 only the arm at -140 m has the position-sensing EM module.

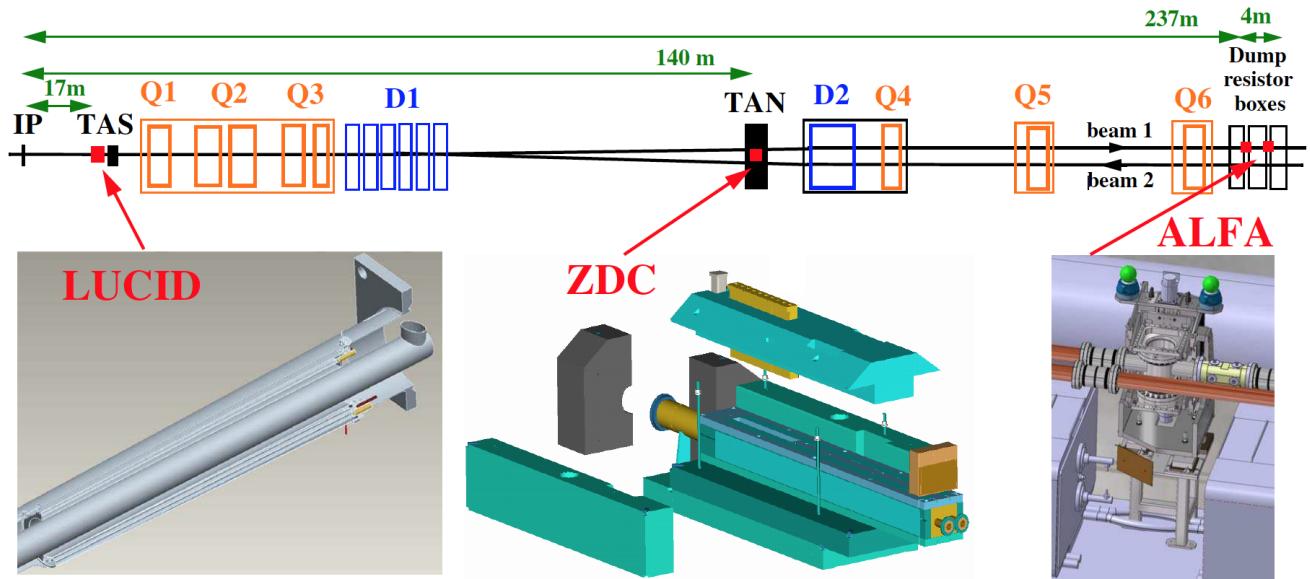


Figure 49: ATLAS forward detectors

1053

1054 4.8 Trigger system

1055 Considering that the bunch crossing rate at LHC is about 40 MHz and that ATLAS detector has over
 1056 one million read-out channels it would never be possible to store all the raw data without significant
 1057 preselection that would decrease the data rate. The selection criteria are picked to retain and store only
 1058 the events which might be interesting for the LHC physics. The preselection and storage is conducted
 1059 with the help of TDAQ systems.

1060 The trigger system has three distinct levels: L1, L2 and the event filter, the two latter levels are also
 1061 called High-level Trigger (HLT). Each next level refines the decisions made before and, if necessary,
 1062 applies additional selection, further lowering the event rate. The data acquisition system receives and
 1063 buffers the event data from the readout electronics at the L1 trigger accept rate which for Run 2 is
 1064 about 100 kHz [14]. The HLT then lowers the rate down to 1.5 kHz which is then stored for the offline
 1065 analysis.

1066 The L1 trigger looks for muons, electrons, photons, jets at hadrons from τ -lepton decays with high
 1067 transverse momentum, large missing and total transverse energy. The muons of interest are identified

1068 using the muon spectrometer trigger system described in section 4.6. The rest of the particles are
 1069 selected using the information from all the calorimeters with reduced granularity. During the Run 2 an
 1070 intermediate L1Topo trigger was also added allowing to combine the information from the spectrometer
 1071 and calorimeter and extend possible trigger selections. Results from these triggers gets processed by
 1072 the central trigger processor which implements the trigger menu made up of different combinations of
 1073 trigger selections. The decision latency for the L1 trigger must not exceed $2.5 \mu\text{s}$ after the corresponding
 1074 bunch crossing.

1075 For every selected event the L1 defines one or more regions called Region of Interest (RoI) which
 1076 include the η and ϕ coordinates of these regions for their subsequent use by the HLT. The L2 selection
 1077 is seeded RoI and uses full granularity and precision along with other detector data available. The
 trigger block diagram is presented in fig. 410.

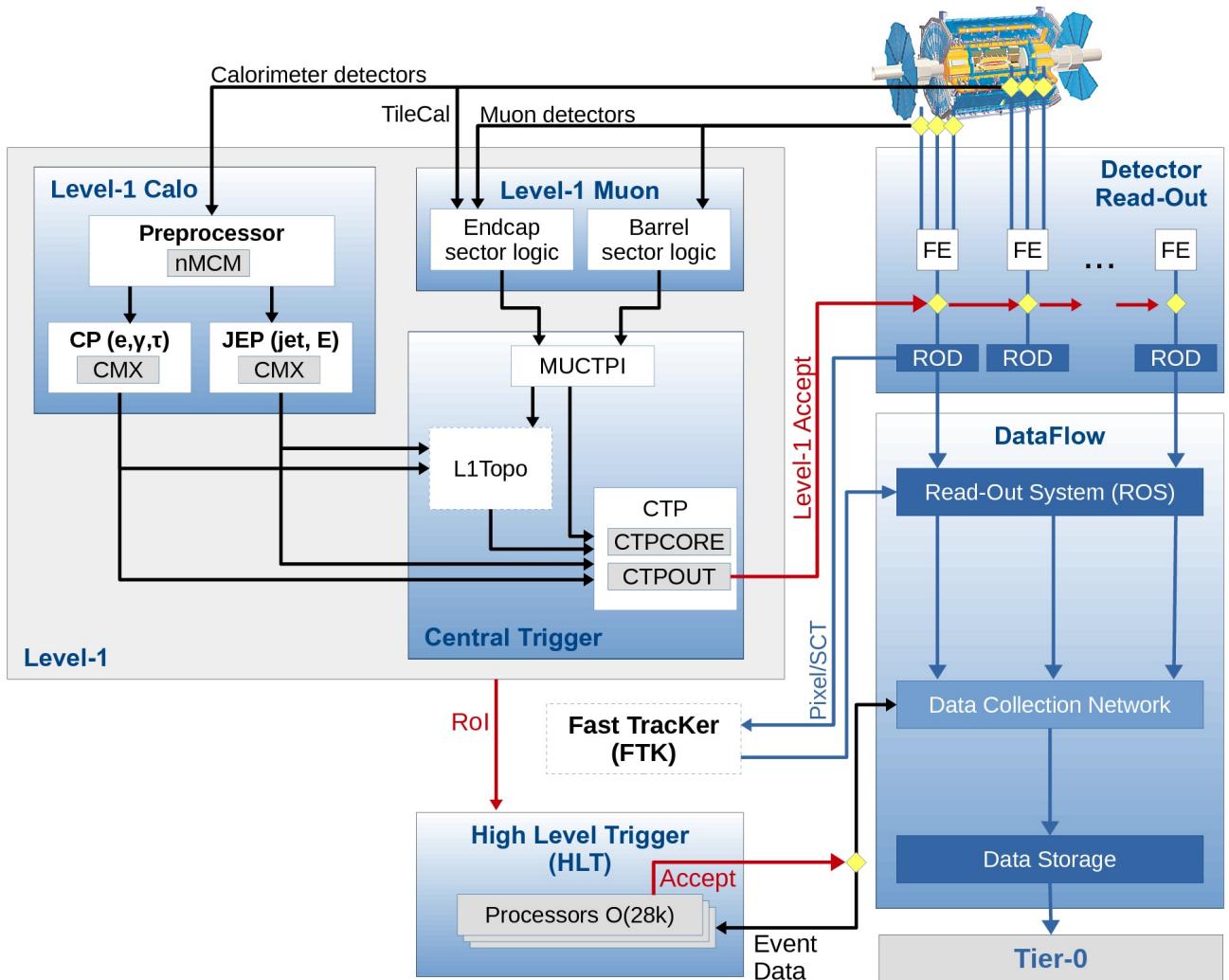


Figure 410: The scheme of ATLAS trigger systems

5

Electromagnetic shower shapes correction in the electromagnetic calorimeter

“Potentielle citation sans aucun rapport avec le sujet”

— Personne inconnue, contexte à déterminer

5.1 Introduction

The design and functionality of the ATLAS electromagnetic calorimeter was described in 4.5.1. Let's consider a bit more in detail the physical processes happening in the EMC. It order to measure particle's energy within the calorimeter we must make the particle to loose its entire energy within the calorimeter. For the electrons and photons with energies over few MeV (which is the case for the ATLAS experiment) the primary energy loss mechanism lies in bremsstrahlung radiation and pair creation). The two processes complete each other, so when a high-energy electron or photon gets into the calorimeter, it creates an avalanche-like processus called the electromagnetic shower when a bremsstrahlung-radiated photons create more electron-positron pairs which in turn radiate more bremsstrahlung photons and so on and so forth (see fig. 104.)

The longitudinal and transverse development of the shower depends on the type of the initial particle

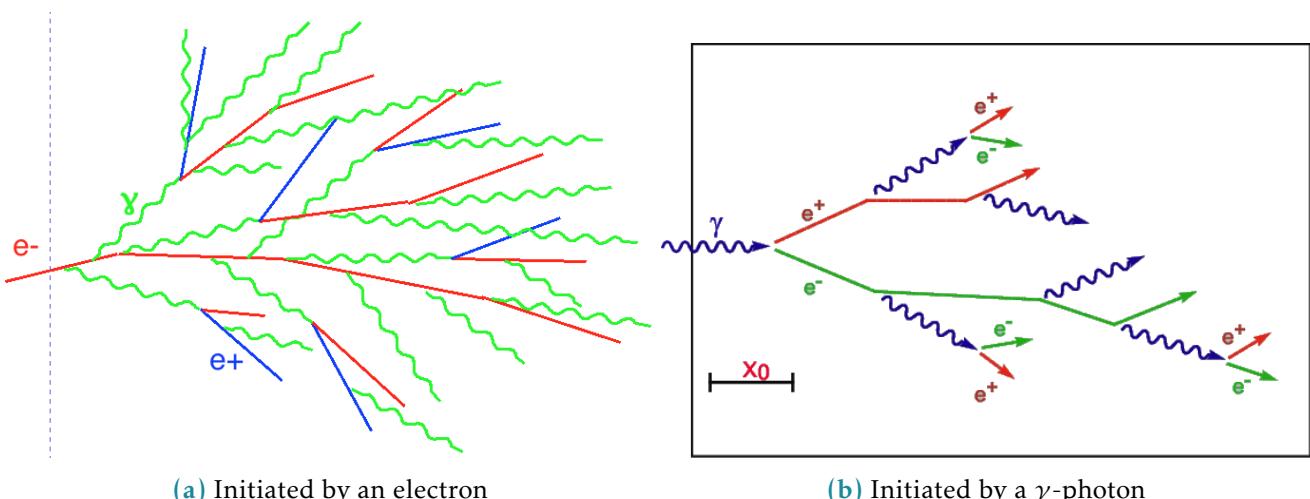
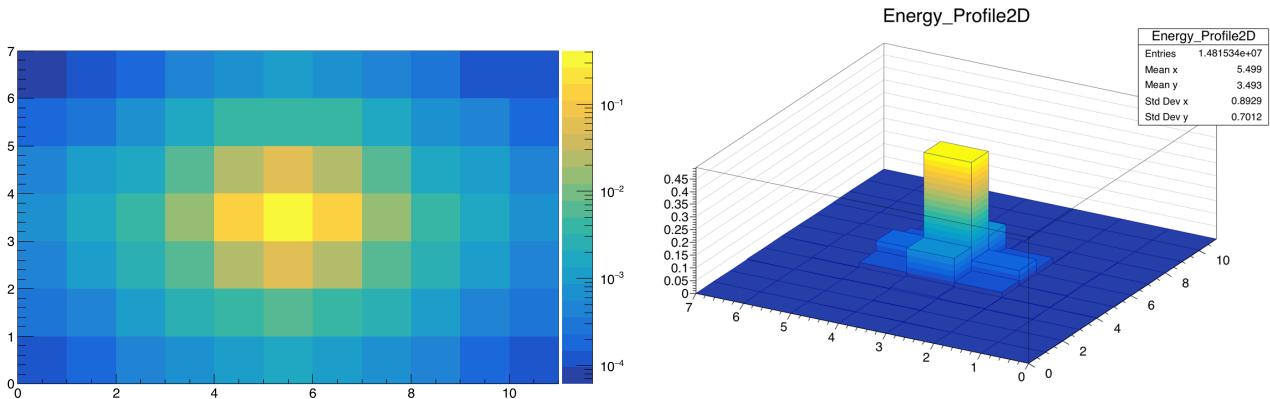


Figure 51: The schematic portrayal of EM shower development

and on its energy. The energy is well measured by the calorimeter, but identifying the particle still remains a challenging task. The transverse granularity of the ATLAS calorimeter allows to resolve the

1096 energy distribution within the electromagnetic shower in the transverse plane. This information can
 1097 later be used for particle identification.
 1098 When an e/ γ particle hits the calorimeter its footprint in the second layer of the calorimeter is visible as
 1099 a cluster of calorimeter cells centered at the central cell having the most energy deposited (sometimes
 1100 referred to as "the hottest cell"). Roughly 90% of shower energy is contained in the core 3x3 cells. We
 1101 have considered a cluster of 7x11 ($\eta \times \phi$) cells, which is schematically depicted on fig. 52a.

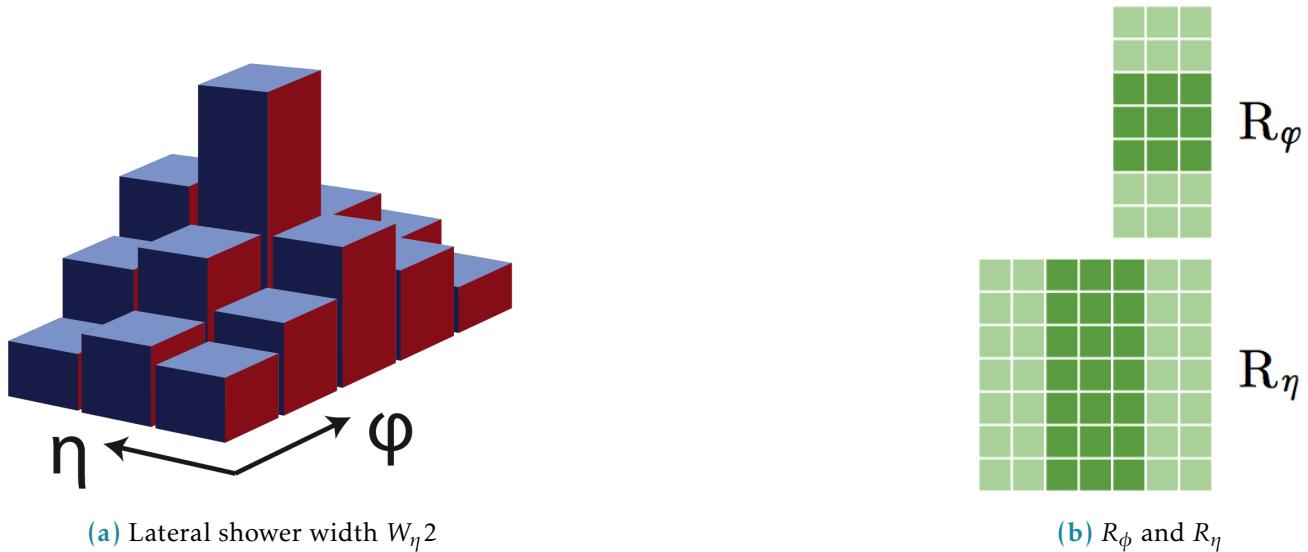


(a) Energy profile of a window of 7x11 cells in the 2nd calorimeter layer (logarithmic scale)

(b) 2D profile of the cluster

Figure 52: Visualisations of the 7x11 calorimeter cluster

1102 In order to characterise the energy distribution within the shower profile a number of observables
 1103 called shower shapes are used. They are then used as an input for particle identification MVA algorithm.
 1104 Current study focuses on the second layer of the calorimeter for which there are three shower shape
 1105 observables described below [1]:



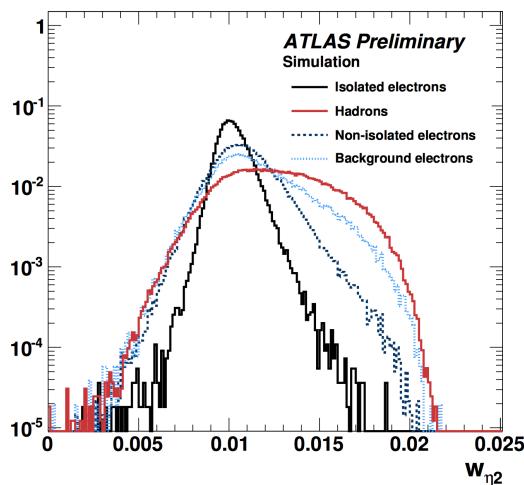
(a) Lateral shower width $W_{\eta}/2$

(b) R_ϕ and R_η

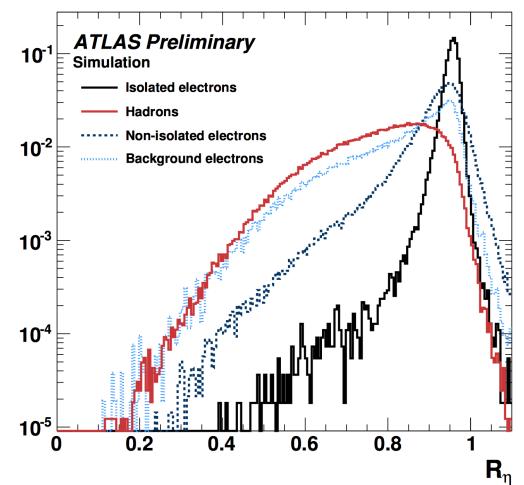
Figure 53: Shower shapes in the second layer of the electromagnetic calorimeter

- Lateral shower width $W_{\eta 2} = \sqrt{\sum(E_i \eta_i^2) - (\sum(E_i \eta_i)/\sum(E_i))^2}$ calculated within a window of 3x5 cells.
- R_ϕ - ratio of the energy in 3x3 cells over the energy in 3x7 cells centered around the hottest cell.
- R_η - ratio of the energy in 3x7 cells over the energy in 7x7 cells centered around the hottest cell.

The shower shapes distributions for different types of particles is shown in fig. ?? - although the distributions overlap, combining the shower shapes information with the inputs from other detectors allow to identify the particle.



(a) $W_{\eta 2}$ distribution simulation



(b) R_η distribution simulation

Figure 54: Distribution of R_η in simulation (GEANT4) for electrons and jets [2].

Figure ?? shows how R_η distribution is different in jets, signal electrons and background electrons. Background electrons denote non-prompt electrons which are not originated from primary vertex.

The shower shapes appear to be extremely sensitive to the detector material modelling. A simplification in the geometry of the EMCAL absorber geometry in GEANT4 9.2 (a layered structure of the accordion was represented as a homogenous material) has lead to visible discrepancies in the shower shapes between the data and MC. This was corrected in GEANT4 9.4 significantly improving the agreement, although not eliminating it completely (see fig. 55). The origin for the remaining discrepancy is not clear.

Disagreement in shower shapes between the data and MC leads to discrepancies in particle ID which are later fixed using η - and p_T -dependent scale factors. Correction of the shower shapes aims to get the scale factors closer to unity, reducing the corresponding systematic uncertainties and improving the precision of the measurements with electrons in the final states.

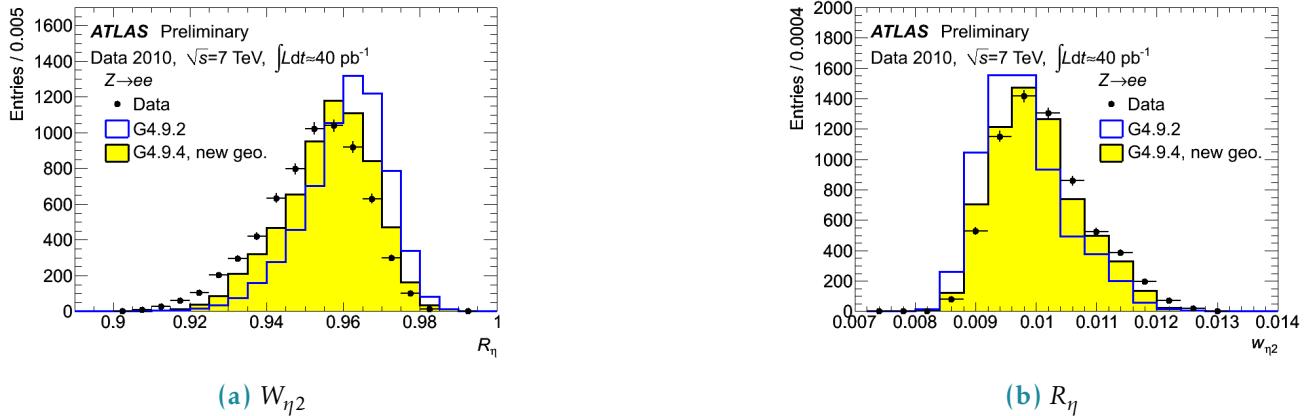


Figure 55: Data/MC Comparison for Calorimeter Shower Shapes of High Et Electrons [3].

1127 5.2 Shower shapes measurement and correction

1128 5.2.1 Event selection

1129 For this study we have considered electrons from the $Z \rightarrow ee$ decay. A set of recommended single
 1130 electron triggers was used (HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0,
 1131 HLT_e140_lhloose_nod0, HLT_e300_etcut). Each event was required to have 2 electrons at least one
 1132 of which has $p_T > 25$ GeV. In order to suppress the background both electrons had to pass gradient
 1133 isolation. Z invariant mass cut was applied with a window of 80 – 120GeV. To avoid identification bias
 1134 from triggering the tag and probe approach was used with only probe electrons taken into consideration
 1135 [4]. The electron cluster in the second calorimeter layer was required to contain information from 77
 1136 calorimeter cells. No pile-up reweighting has been applied. Datasets of 264786295 events in data (2017
 1137 proton-proton collisions) and 79340000 events in MC (Powheg+Pythia8) were used.

1138 5.2.2 Data/MC discrepancies

1139 Our consideration begins with the energy deposit of an electron in the second layer of the calorimeter.
 1140 A window of 7 cells in η and 11 cells in ϕ is centered around the cell with the highest energy.

1141 Shower shapes were considered in 14 η bins in the range between $|\eta| = (0, 2.4)$ in order to investigate
 1142 how the discrepancy depends on η .

1143 The η -dependent shower shapes in data are wider than the MC and show a larger discrepancy in the
 1144 endcap ($|\eta| = (1.52, 2.4)$). For ϕ dimension the situation is the opposite: MC is wider than the data and
 1145 the barrel ($|\eta| = (0, 1.52)$) shows larger discrepancy. Figures ??, ??, ?? contain examples of shower shapes
 1146 in different eta bins.

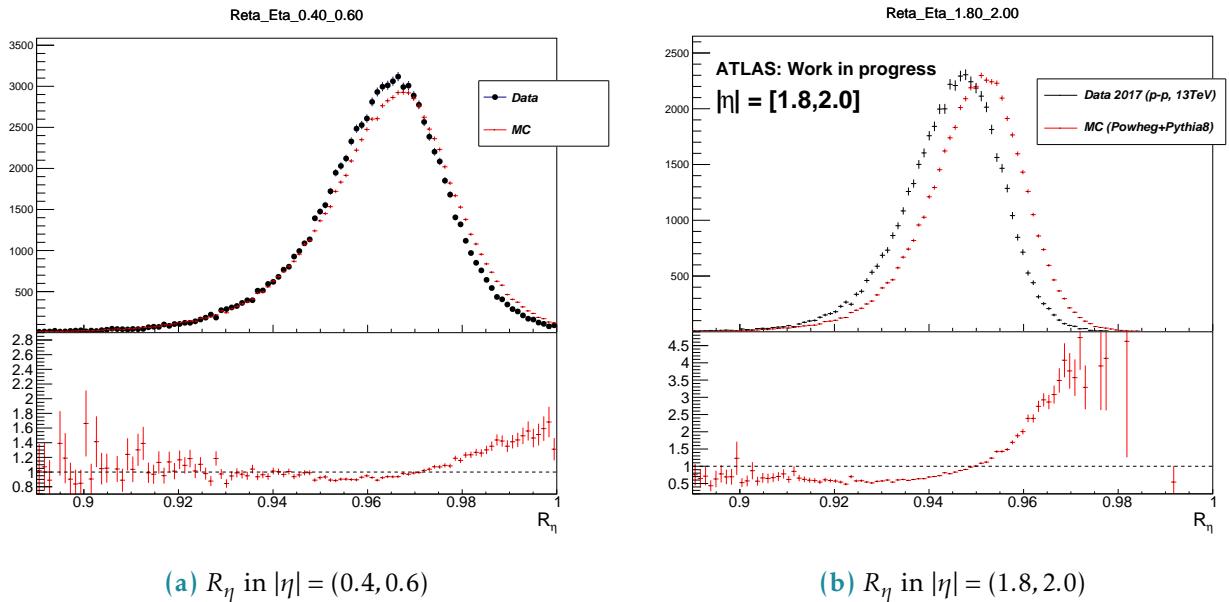


Figure 56: R_η in the barrel and in the end-cap, Data vs MC

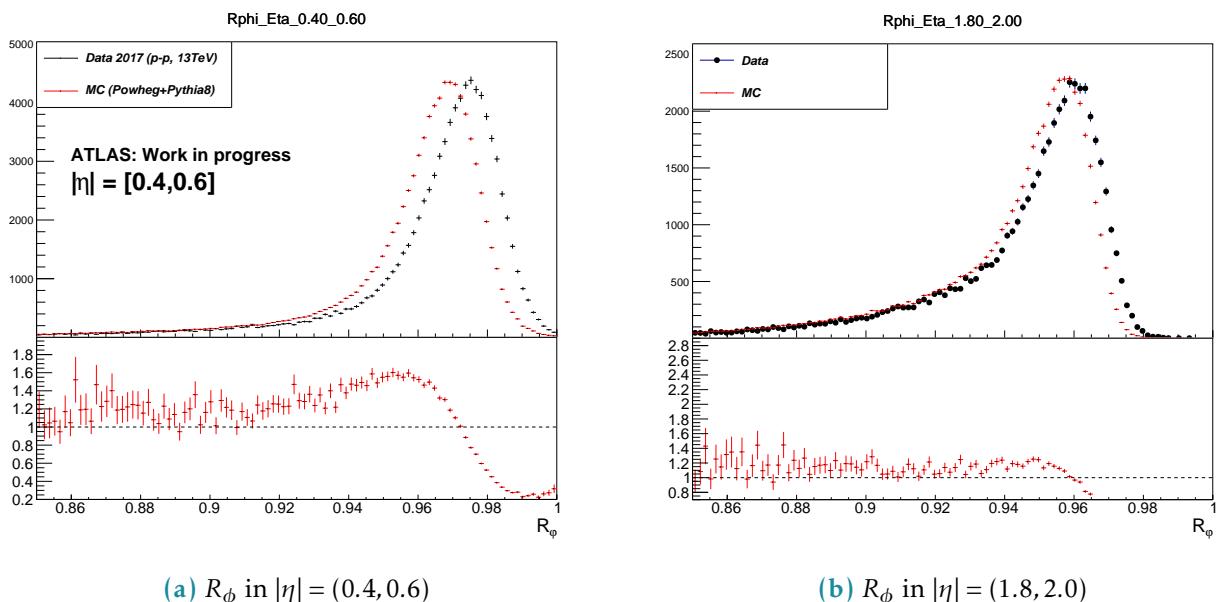


Figure 57: R_ϕ in the barrel and in the end-cap, Data vs MC

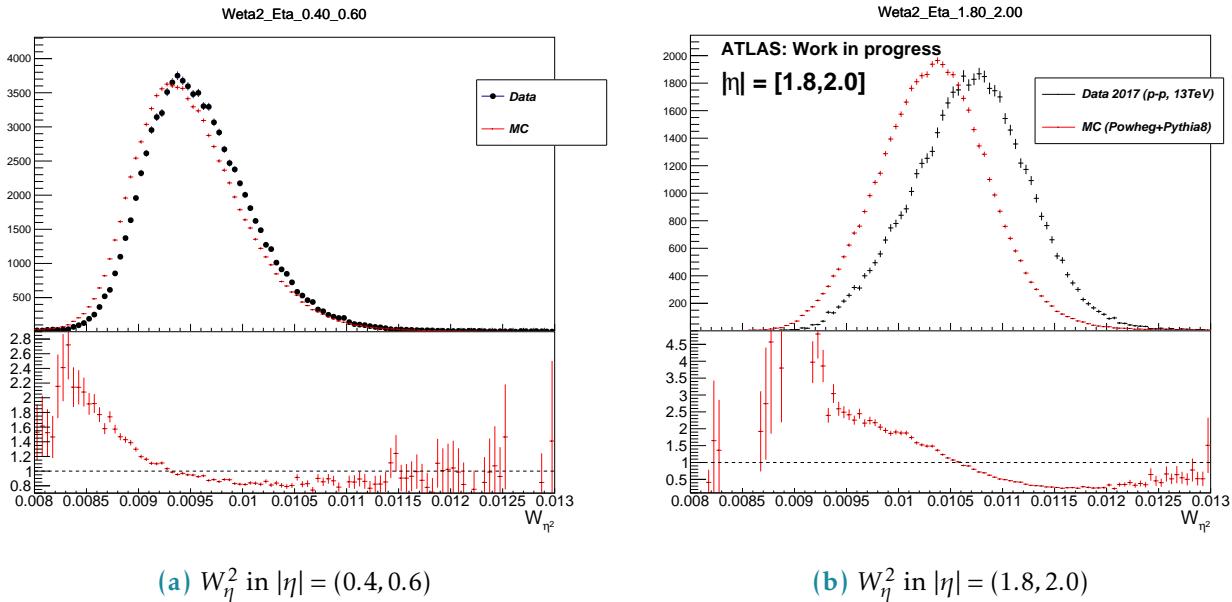


Figure 58: W_η^2 in the barrel and in the end-cap, Data vs MC

1147 5.2.3 The correction procedure

1148 The correction matrix

1149 The correction procedure is based on the redistribution of energy between the cluster cells in MC so
 1150 that the distribution becomes consistent with the data. For every η bin a correction matrix is derived in
 1151 the following way:

$$M_i^{Correction} = \frac{E_i^{Data}}{\sum E^{Data}} - \frac{E_i^{MC}}{\sum E^{MC}}$$

1152 $\sum_i M_i^{Correction} = 0$, $i = 1..77$.

1153 E_i^{Data} , E_i^{MC} - matrix elements of the averaged energy profiles. The correction is then applied to the
 1154 electron cluster cells on event-by-event basis:

$$E_i^{Reweighted} = E_i^{Non-reweighted} (1 + M_i^{Correction}).$$

1155 This redistributes the energy among the cells keeping the total energy exactly the same.

1156 Bremsstrahlung tails

1157 The magnetic field directed along the ϕ dimension leads to a significant asymmetry in energy deposits
 1158 for electrons and positrons (figure ??).

1159 Considering the fact that the reweighting is intended to correct for the data/MC discrepancies
 1160 themselves and not for the bremsstrahlung effect it makes sense to develop the bremsstrahlung-free
 1161 correction function based on e^+ and e^- correction matrices. The principle is schematically explained on

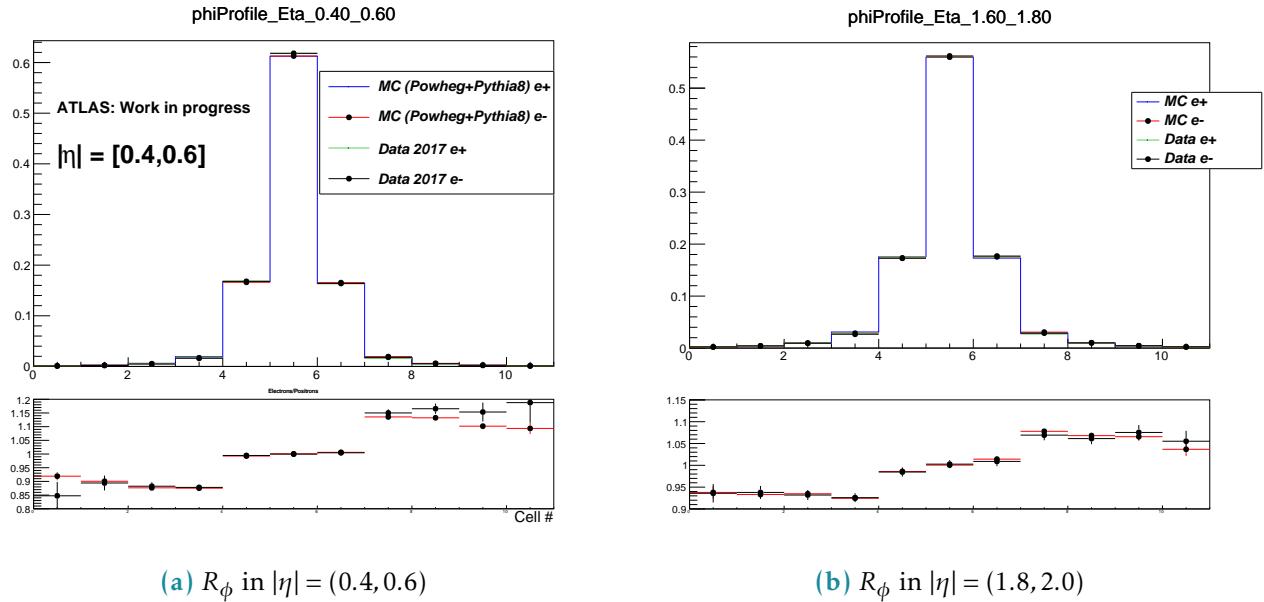


Figure 59: R_ϕ in the barrel and in the end-cap, Data vs MC

1162 figure 510.

Good agreement of data and MC description of e^+ and e^- asymmetry gives a hint that the material

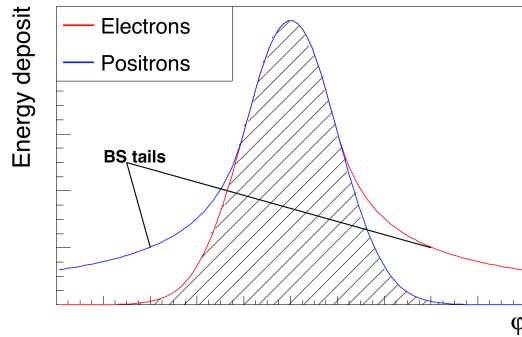


Figure 510: Schematic energy profile in ϕ dimension. Bremsstrahlung tails subtraction based on e^+ and e^- energy profiles.

1163

1164 mismodelling cannot be the main source of the data/MC disagreement.

1165

1166 5.3 Results

1167 Figures 511, 512, 513 show the effect of the correction. The shower shapes in MC become very close to
1168 the data, correcting a significant discrepancy.

1169 Figures 514 contain shower shapes vs p_T integrated over η . They demonstrate that the correction does

1170 not depend on the p_T which allows to expect the decreased systematic uncertainties for p_T regions
 1171 distant from 40 – 50 GeV.
 1172 Finally, figure 515 shows the effect of the correction on electron ID efficiency. We can see a visible
 1173 improvement, notably in the endcap region. Nevertheless the barrel region shows little improvement.
 1174 It can be explained by the fact that electron ID MVA relies on many variables while only a number of
 1175 them were corrected during current study.
 1176 The proposed method is getting integrated into ATLAS Athena framework as an option and is planned
 1177 to be used as a baseline for Run 3.

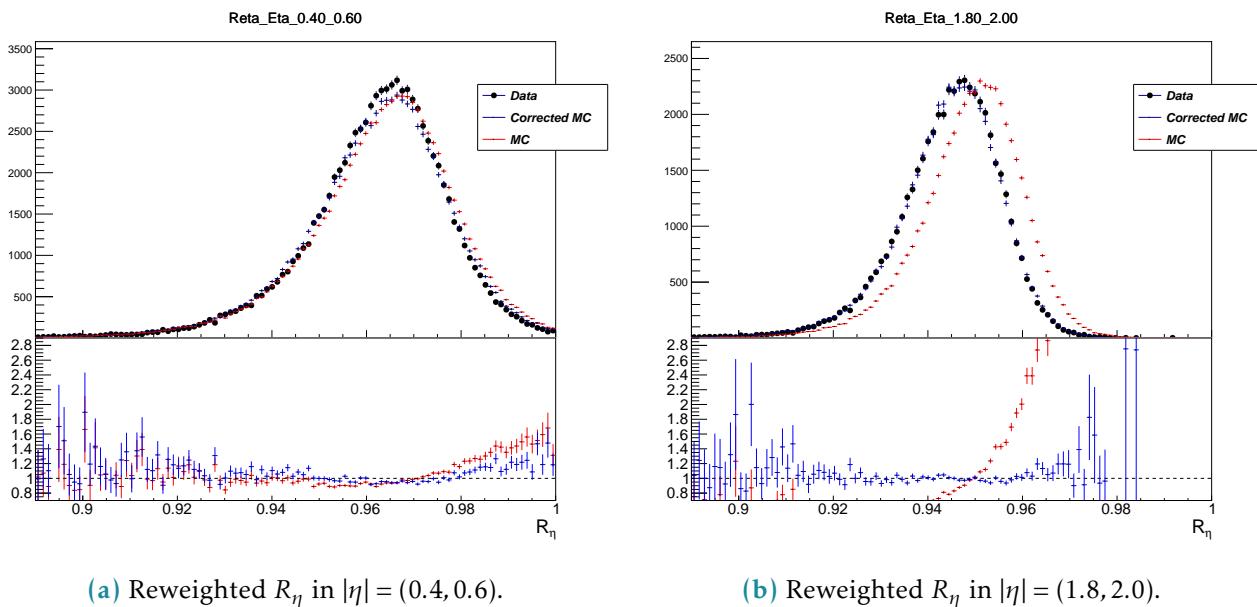


Figure 511: R_η in the barrel and in the end-cap

1178 **5.4 Appendix: control plots**

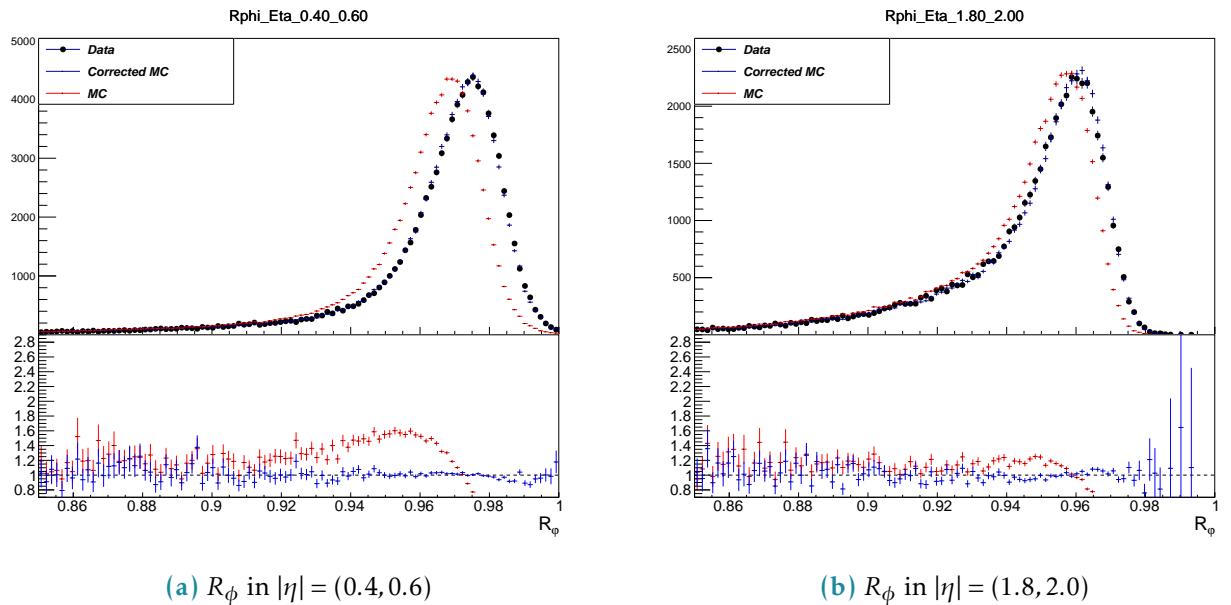


Figure 512: R_ϕ in the barrel and in the end-cap, Data, MC, reweighted MC

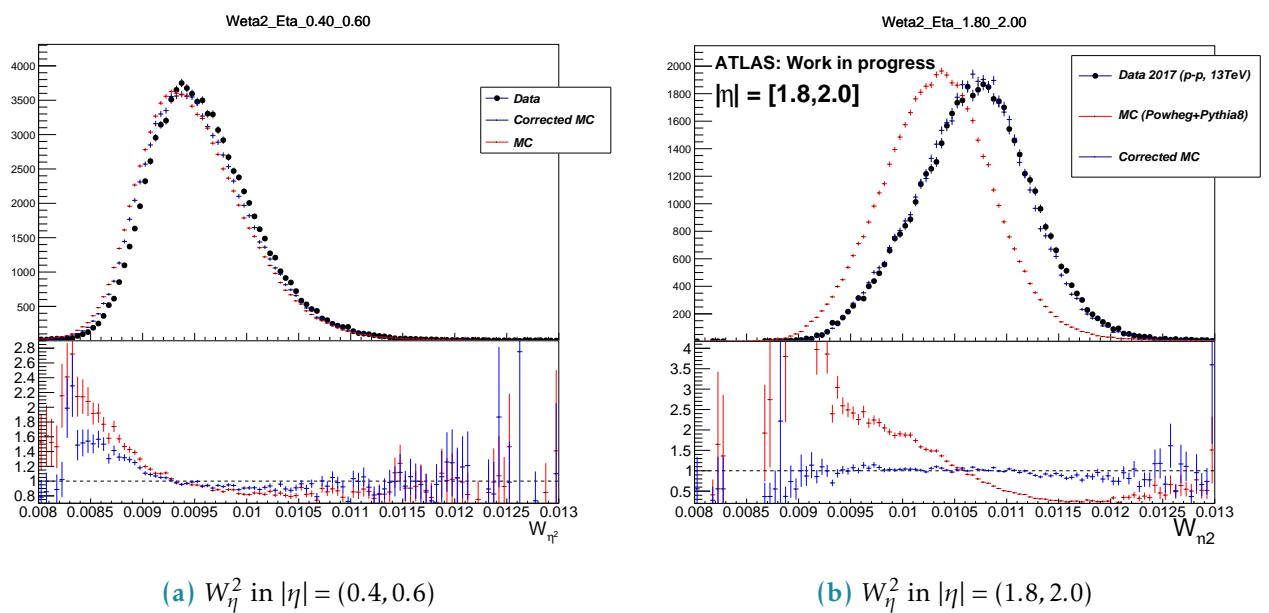


Figure 513: W_η^2 in the barrel and in the end-cap, Data, MC, reweighted MC

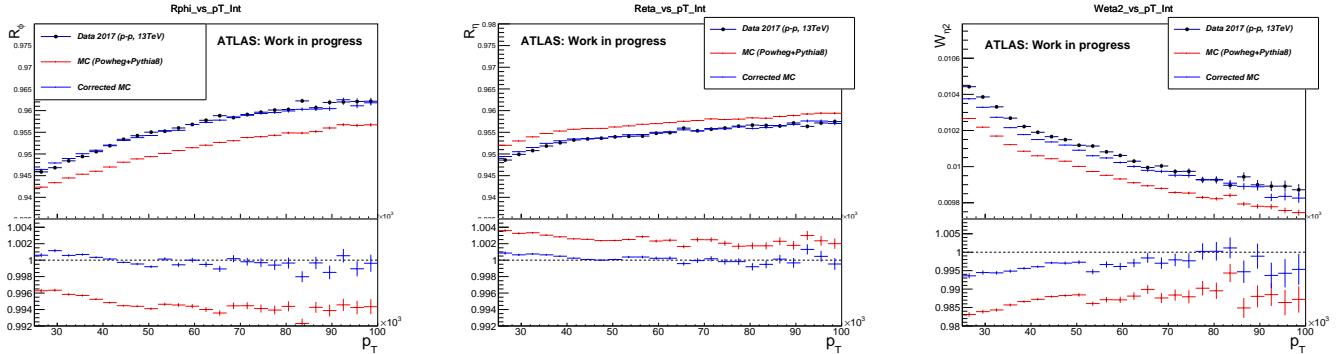


Figure 514: Distributions integrated over p_T (a) R_ϕ ; (b) R_η ; (c) $W_{\eta 2}$.

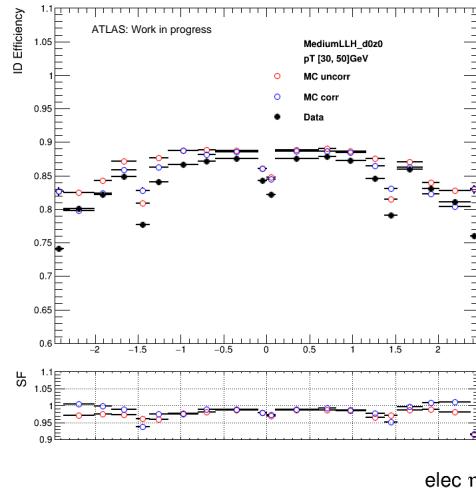


Figure 515: Electron identification efficiency as a function of the electron pseudo-rapidity

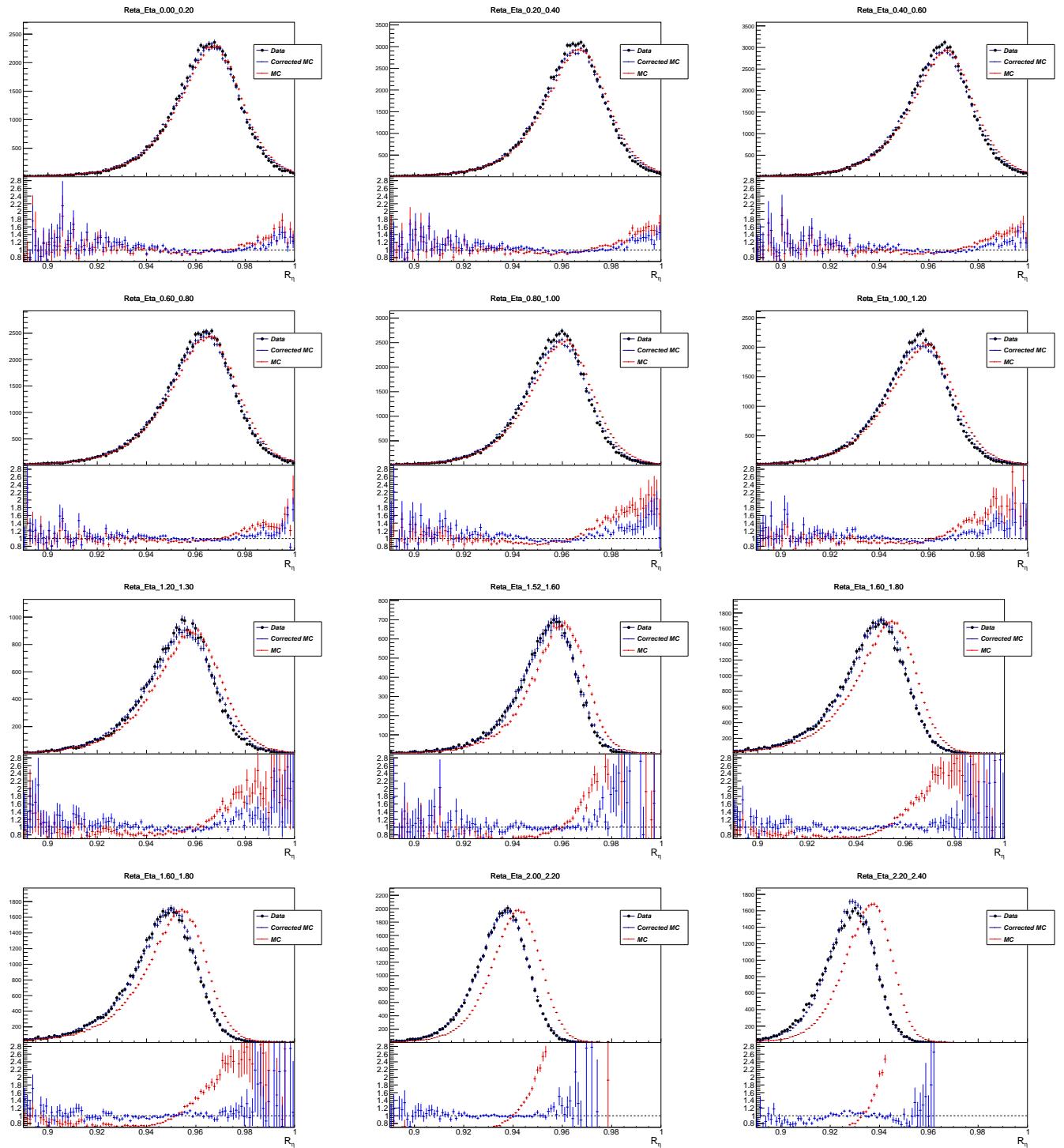


Figure 516: Reta 2

Mesure de la masse du boson W avec le détecteur ATLAS au LHC

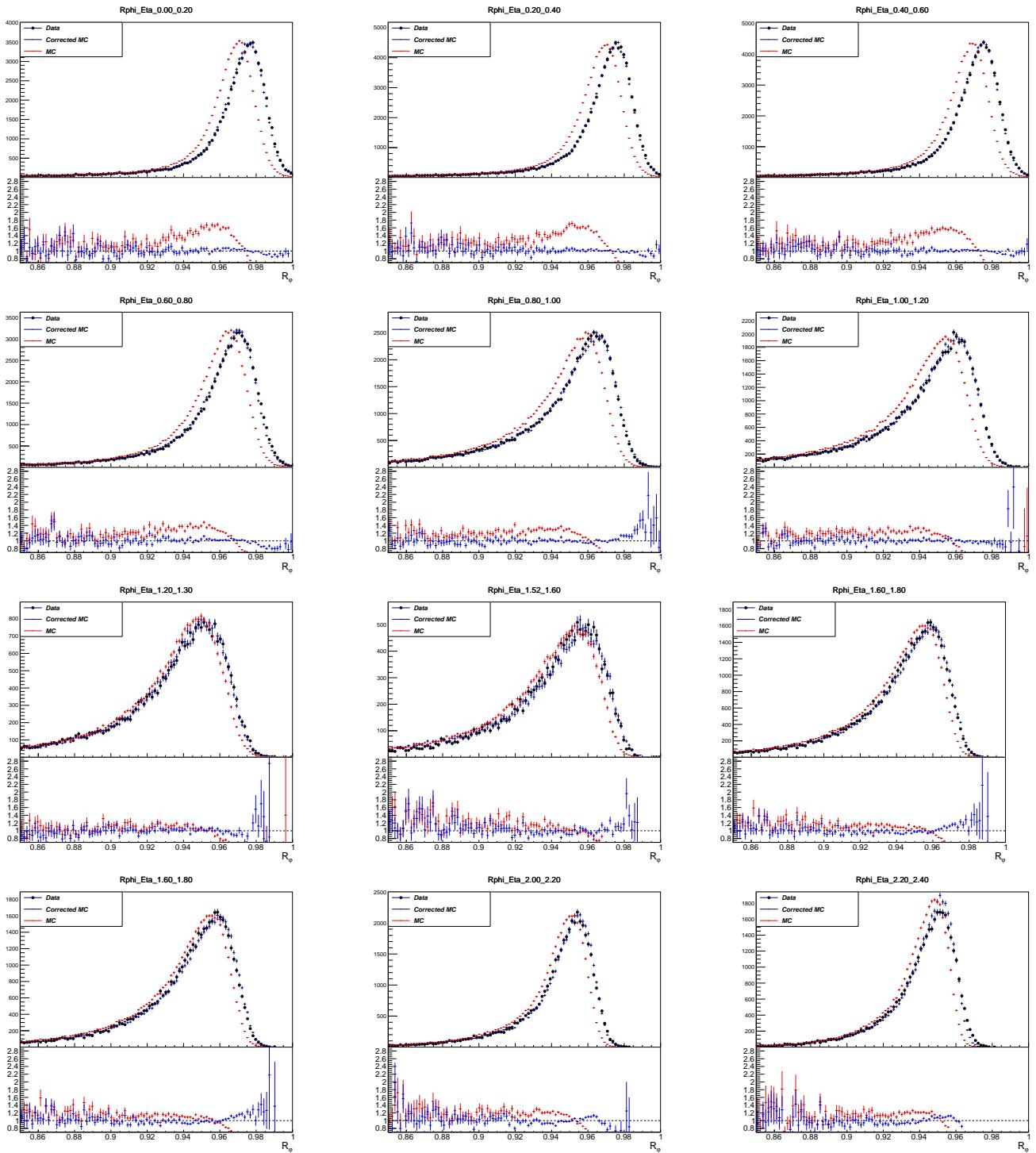


Figure 517: Rphi in all eta slices

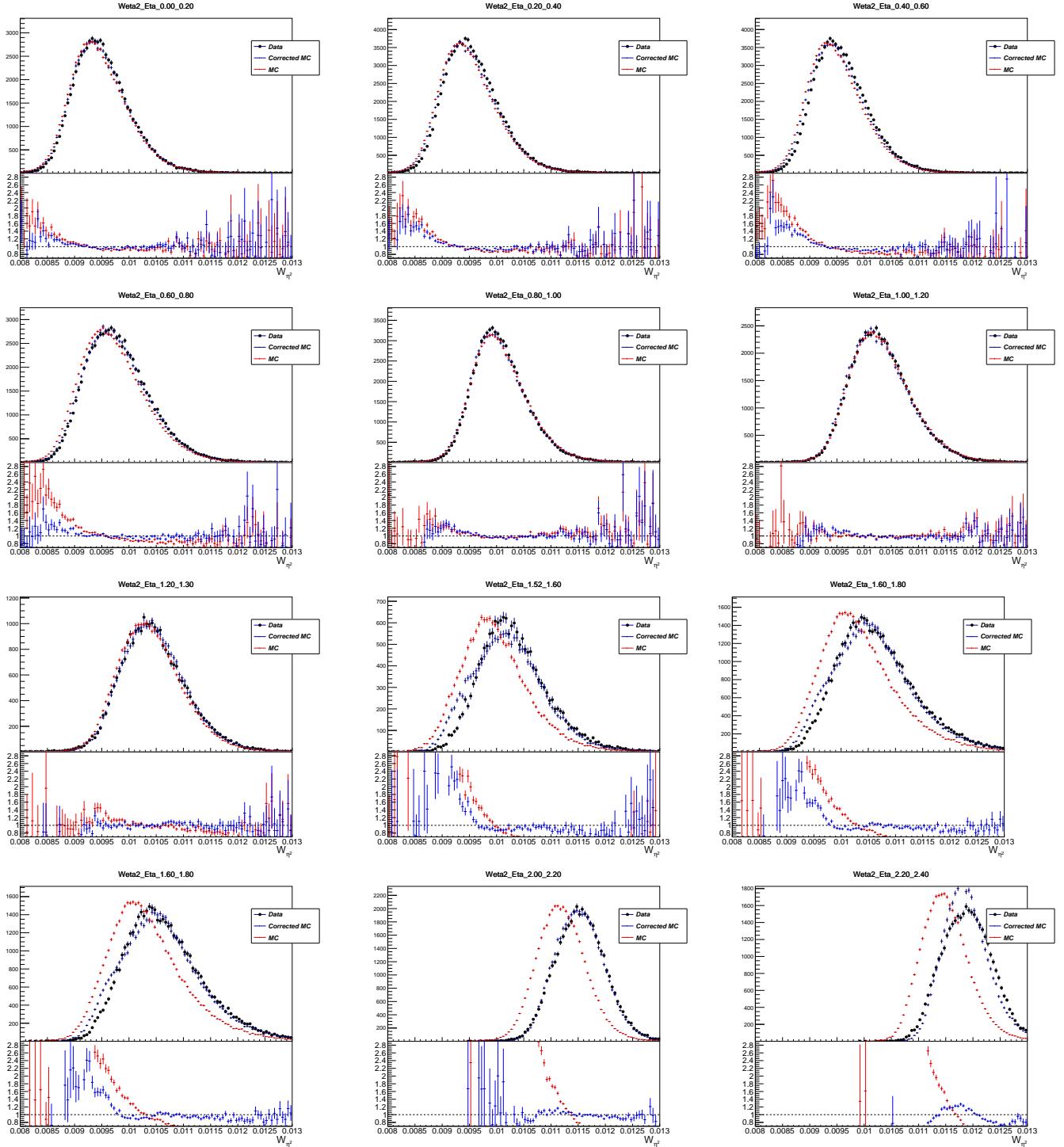


Figure 518: Reta 2

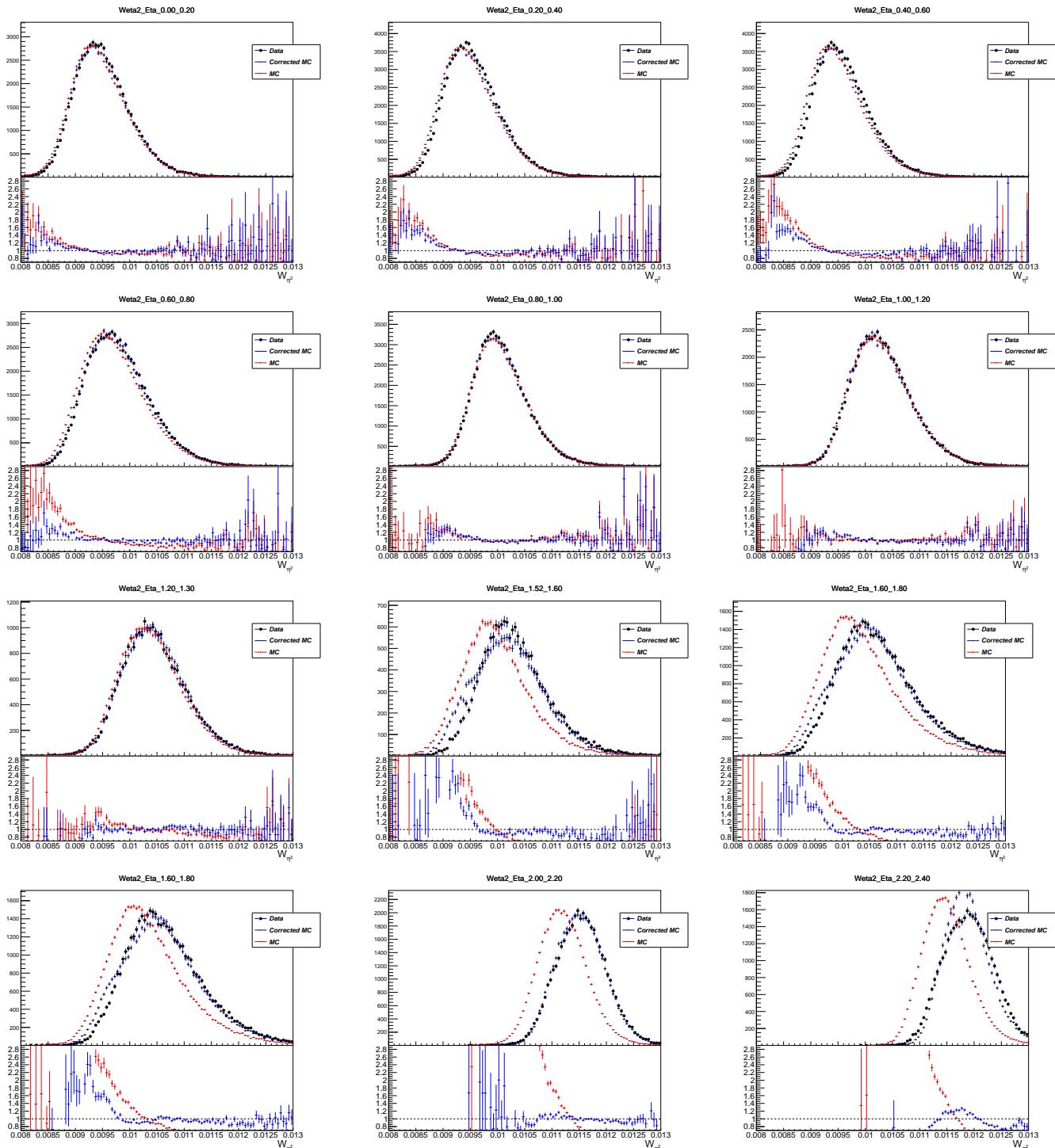


Figure 519: Reta 2

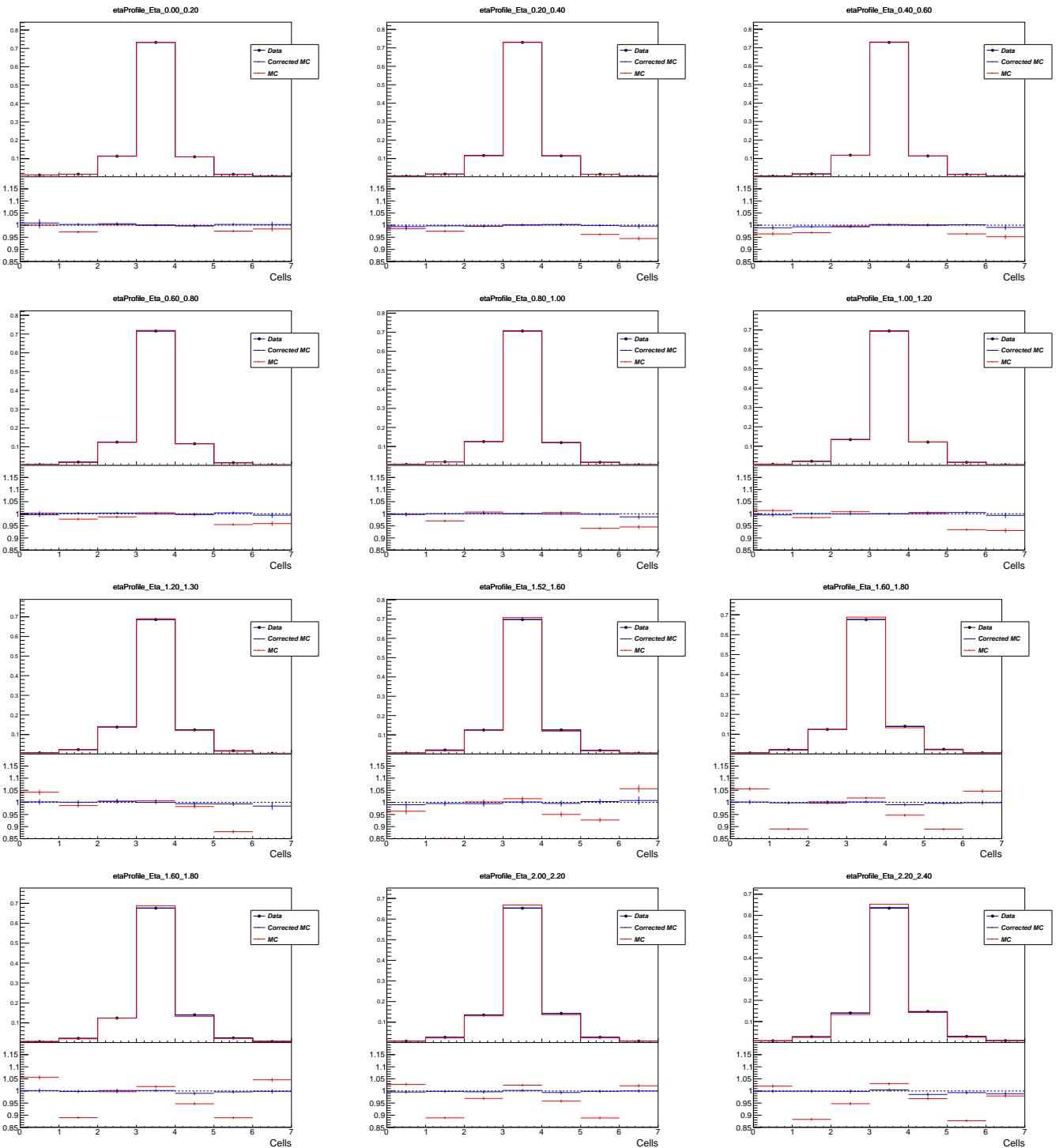


Figure 520: Reta 2

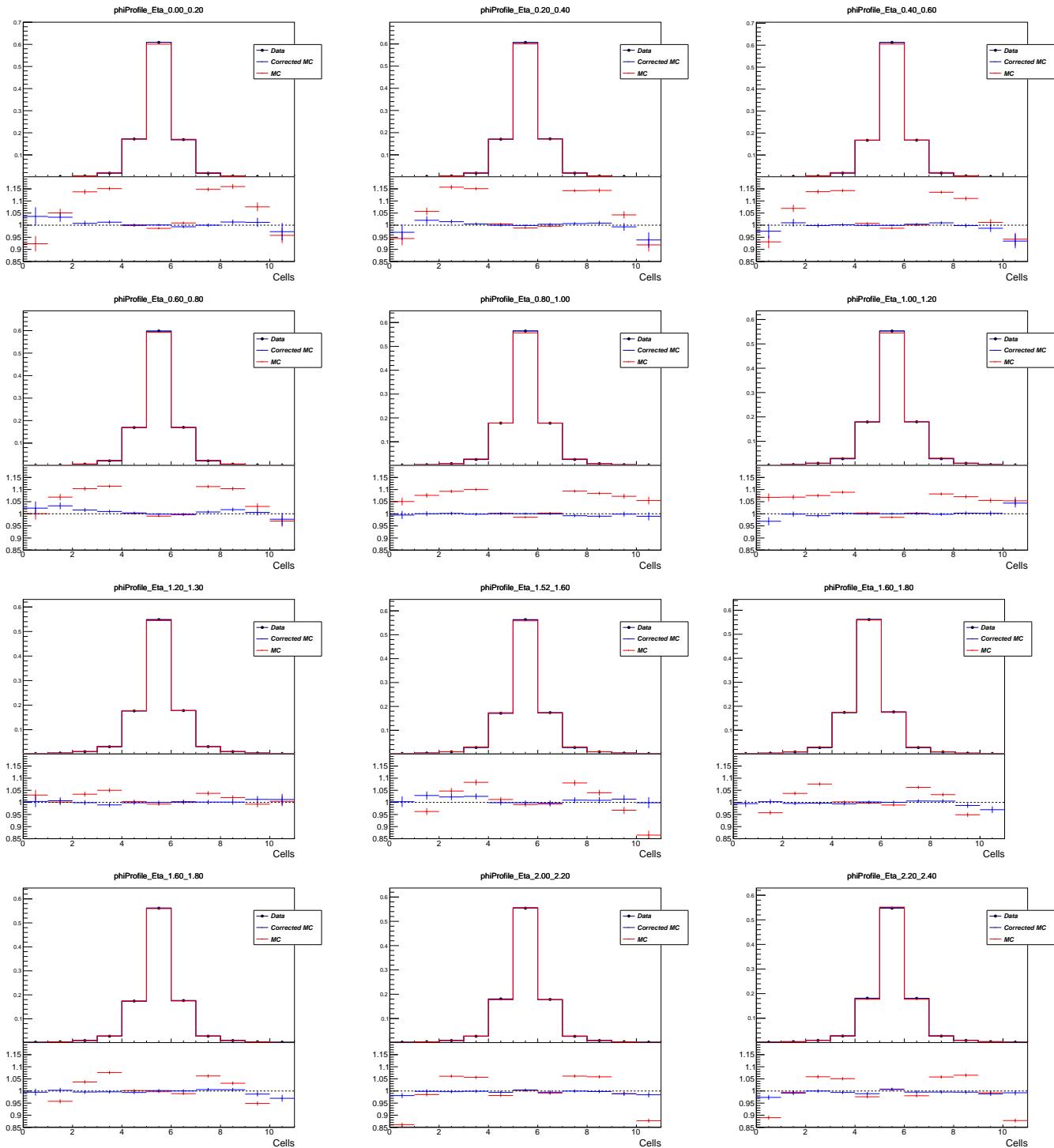


Figure 521: Reta 2

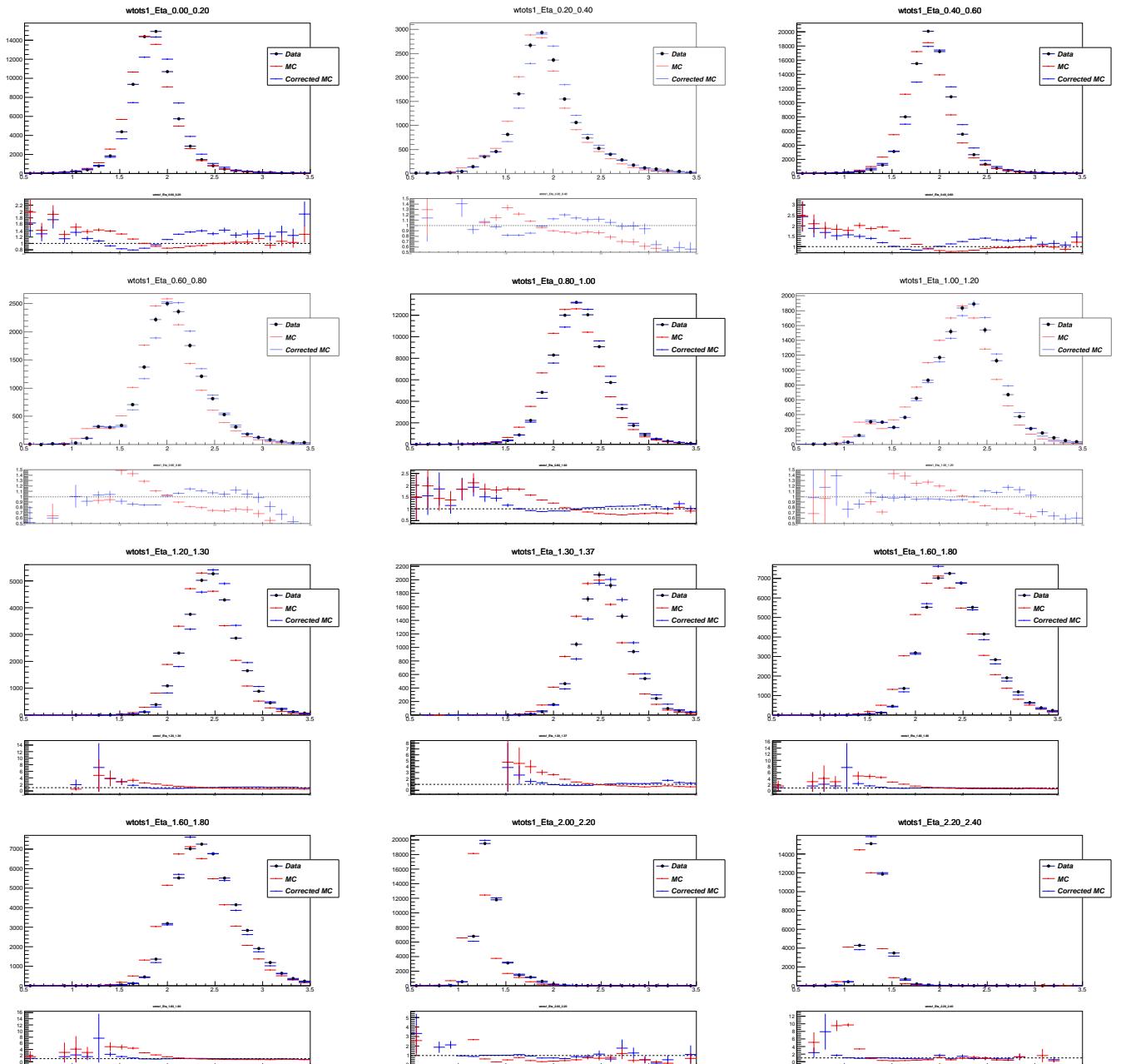


Figure 522: Reta 2

6

Event reconstruction

1179

1180

1181

“Potentielle citation sans aucun rapport avec le sujet”

1182

— Personne inconnue, contexte à déterminer

1183 6.1 Charged particles track reconstruction

1184 The track q is formed based on the information from the ID and contains five parameters: $q =$
1185 $(d_0, z_0, \phi, \theta, q/p)$, where d_0 is the distance from the track to the Z axis (transverse impact parameter), z_0
1186 is the Z coordinate of the perpendicular dropped from the track onto the Z axis (longitudinal impact
1187 parameter) (see fig. 61), ϕ and θ are the azimuthal and polar angles correspondingly and q/p is the
1188 charge to momentum ratio of the particle. The process of the track reconstruction is the same for lepton
1189 and charged hadron candidates.

To form tracks using the detector response information the following steps are performed [2]:

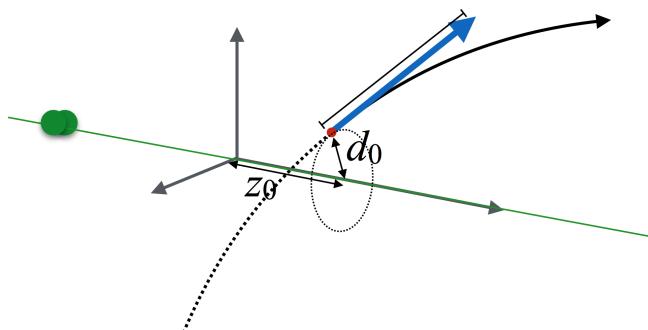


Figure 61: Impact parameters z_0 and d_0 [1].

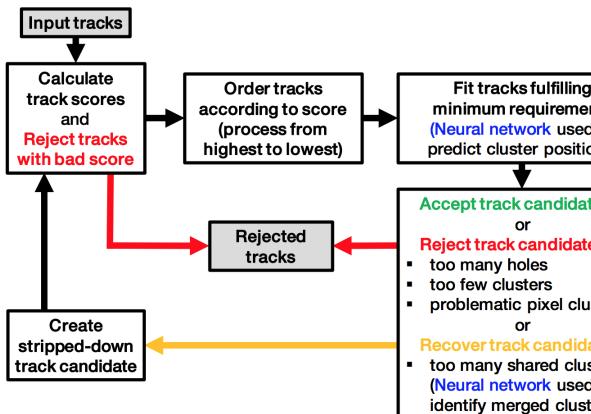
1190

- 1191 • **Clustering** single hits in the pixel and SCT detectors. Neighbouring hits are combined to form a
1192 single cluster, clusters are then transformed into *space points* that have having 3D coordinates. A
1193 cluster may be identified as a single-particle cluster or as merged cluster, created by two or more
1194 particles. Identification of a cluster as a merged one and separation of energy deposits between
1195 the particles (possible only for two particles) is performed by means of a Neural Network (NN)
1196 algorithm.

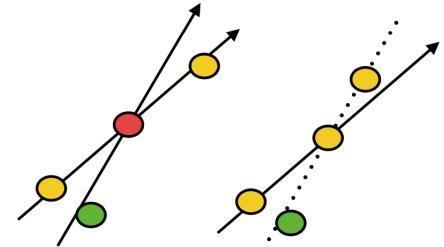
- **Forming seeds** out of the space points. To form a seed three space-points originating from unique layers of the silicon detectors (pixel or SCT) are used. All possible combinations of seeds are formed at this stage. For every seed a crude estimate of the track parameters is performed.
- **Track candidates** are formed out of the seeds by extending them within the silicon sub-detectors following the most likely path. Combinatorial Kalman filter [3] is used to build the track candidates. The purity of the seeds depends significantly on the sub-detector that recorded the corresponding space-points. SCT-only seeds are considered the most reliable, followed the seeds that origin only from the pixel detector space-points, and the least reliable are the seeds originating from both of these sub-detectors - that determines the order of seed consideration when composing track candidates.
Some fraction of the seeds that meet the necessary requirements become track candidates, the rest are discarded. A seed may be used for more than one track candidate if more than one space-point extension exists on the same layer.
- **Ambiguity solving** is the next step necessary to eliminate incorrectly assigned space-points or resolve conflicting track candidates that have and overlapping space-point. At this stage the track candidates are assigned a *track score*. The track score depends on the number of clusters associated to the track and which sub-detector these clusters originate from, the existence of holes (the absence of a cluster associated to a detector layer crossed by the track), the quality of the χ^2 fit of the track and track momentum.
The tracks are ordered by their track score and consequently fed to the ambiguity resolving sequence. A truck must pass a number of kinematic cuts, impact parameters cuts, number of holes, number of clusters and shared clusters cuts, otherwise the track candidate is rejected. If a track candidate has no shared clusters with other candidates it is accepted after that. If there are merged clusters then it is up to the NN to either accept the track, reject it or eliminate a space-point and recycle the updated track candidate (see Fig. 62a).
- **TRT extension** means matching of the track, composed using the information from silicon sub-detectors to the trace in the TRT tracker. This allows to improve momentum measurement benefiting from extended track length.
- Final high-resolution **track fit** is performed using all available information. Position and uncertainty of each cluster are determined by an additional NN allowing for more precise track parameters. The curvature of the particle track also serves for charge sign identification.

6.2 Determining the primary vertex of the event

Primary vertex determination is crucial for physics analyses for many reasons. One of them is the necessity to separate particles originating from hard events from pile-up. Another reason is to keep



(a) Track ambiguity resolver algorithm.



(b) Tracks sharing space-points.

Figure 62: Ambiguity solving.

1231 track of the decay chain and make difference between prompt and non-prompt particles. Flavour
 1232 tagging, background suppression and decay reconstruction also rely heavily on the primary vertex
 1233 determination.
 1234 After reconstructing the tracks of individual particles the obtained information is used to reconstruct
 1235 the Primary Vertex (PV) of the event [4]. The procedure relies on the reconstructed tracks and goes as
 1236 follows:

- 1237 • A seed from the first vertex is selected. The transverse position of the seed is taken as a centre of
 1238 the beam spot. The z-coordinate is the mode of z_0 coordinate of the tracks.
- 1239 • Using the seed and the available tracks an iterative fit is performed in order to find the best
 1240 position for the PV. In each iteration the tracks that are less compatible with the vertex are
 1241 down-weighted and the vertex position gets recomputed. With every iteration the spread in the
 1242 weight increases, separating track set into compatible tracks that mostly determine the vertex
 1243 position and incompatible tracks that have little weight and therefore very little influence on the
 1244 track position.
- 1245 • After the fit is done compatible tracks remain assigned to the vertex, while incompatible tracks
 1246 are removed from it. These incompatible tracks can be used in the determination of a different
 1247 vertex.
- 1248 • The procedure is repeated with the remaining tracks of the event.

1249 For the upcoming Run 3 of the LHC certain improvements and modifications are foreseen [5].

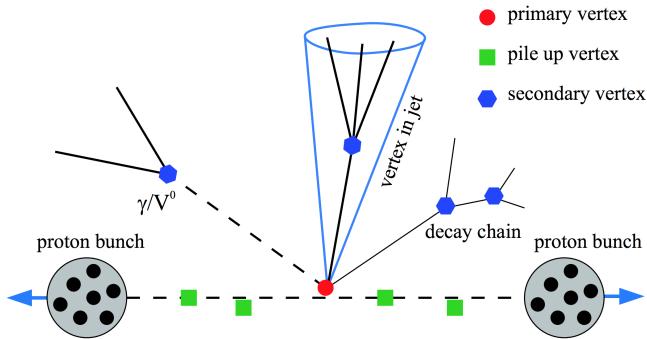


Figure 63: Primary, secondary and pile-up vertices [6].

1250 6.3 Muon reconstruction and identification

1251 Muon reconstruction relies primarily on the information from the ID (the muon track) and the Muon
 1252 Spectrometer (MS), sometimes also using additional information from the calorimeter. At the first
 1253 stage muon is independently reconstructed in the tracker and in the spectrometer, and then they are
 1254 combined to compose a muon track used in the physics analyses [7]. Track reconstruction is described
 1255 in subsection 6.1.

1256 6.3.1 Muon reconstruction

1257 Muon reconstruction on the muon spectrometer begins with a search for hit patterns in each muon
 1258 chamber and forming of the segments. Using the Hough transform [8] the hits in each MDT chamber
 1259 and nearby trigger chamber are aligned on trajectories in the bending plane. The orthogonal coordinate
 1260 is measured with RPC and TGC detectors. A separate combinatorial search is conducted in the CSC
 1261 detectors in ϕ and η detector planes.

1262 Then the track candidates are built by fitting hits from different layers. This algorithm starts a
 1263 combinatorial search first using the segments from the middle layers as seeds, as there are more trigger
 1264 hits in the middle layer. The search is later extended to include the segments from other layers as seeds.
 1265 Segment selection criteria are based on hit multiplicity and fit quality. The segments are matched using
 1266 their relative positions and angles. In all the regions, except barrel-endcap transition region, at least
 1267 two matching segments are needed to build a track (one segment is enough in the transition region).
 1268 A single segment can be used by two or more track candidates. An overlap removal algorithm decides
 1269 to which track should a segment belong or shares a segment between two tracks. A global χ^2 fit is used
 1270 to fit all the hits associated to every track. If the χ^2 fit meets the designated criteria then the track is
 1271 accepted. If a hit impair the χ^2 fit significantly, then this hit may be removed and the fit is repeated.
 1272 On the other hand, new hits may be recovered if they fit the track candidate trajectory.
 1273 Accurate fitting of the track trajectory is extremely important for the measurement of muon momentum.
 1274 A quantity called *sagitta* is measured by the MS (see Fig. 64). Knowing the length L and the sagitta S

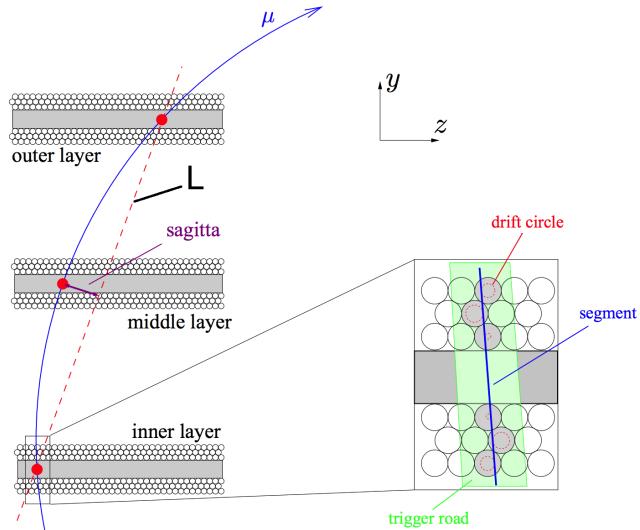


Figure 64: Sagitta used for the determination of the muon momentum [9].

1275 we can determine the momentum:

$$p = \frac{BL^2}{8S}, \quad (6.1)$$

1276 where B is the magnetic field strength.

1277 After the muon gets reconstructed in every detector system separately, the obtained information is
1278 combined to form a reconstructed muon object. Depending on the detectors used for the combined
1279 reconstruction there are *four types of muons* defined (see Fig. 65):

- 1280 • **Combined (CB) muon** is formed from a global refit of the tracks reconstructed independently
1281 in the ID and in the MS. During this global refit the hits from both detectors are used and also
1282 new hits may be added. Normally the outside-in pattern is used, when MS track is extrapolated
1283 inwards to match ID track. Inverse inside-out procedure is used as a complementary approach.
- 1284 • **Segment-tagged (ST) muon** is a particle with an ID track that was extrapolated to the MS and
1285 associated with at least one local track segment in the MDT or CSC chambers. Normally these
1286 are muons with low p_T or their trajectory crosses regions with reduced MS acceptance.
- 1287 • **Calorimeter-tagged (CT) muon** has a valid ID track that can be associated to an energy deposit
1288 in the calorimeter compatible with minimum-ionizing particle. The CT muons have the lowest
1289 purity among the muon types although they provide acceptance where the MS coverage may be
1290 absent, like the very central region with $|\eta| \leq 0.1$ for $15 < p_T < 100$ GeV.
- 1291 • **Extrapolated (ME) muon** (standalone muon) trajectory is reconstructed base only on the MS
1292 track and a loose requirement to match the IP. ME muons allow to extend the muon acceptance
1293 to the region which is not covered by the ID, namely $2.5 < |\eta| < 2.7$.

1294 In case of overlap between different muon types the preference is given to CB muons, then to ST and then to CT muons. ME muons overlaps are resolved based on the MS track quality.

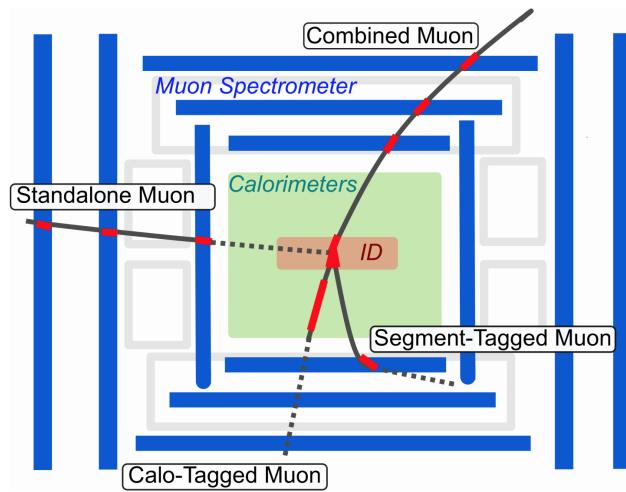


Figure 65: Four types of reconstructed muons.

1295

1296 6.3.2 Muon identification

1297 Muon identification is a set of measures to ensure that the registered particle has indeed the character-
 1298 istics of a muon and to identify the mechanism of its production. Muons created in the course of decay
 1299 of a short-lived particle (e.g. a massive boson) are called *prompt muons*, while those originating from
 1300 hadron or tau decays are called *non-prompt*. Muon identification plays an important role in background
 1301 suppression and guaranteeing a robust momentum measurement.

1302 Muons that are created during the in-flight decay of the charged hadrons in the ID usually have a
 1303 distinctive "kink" topology in their reconstructed track. This results in a decreased quality of the
 1304 resulting track fit and the incompatibility between the results of momentum measurement in the ID
 1305 and MS. Muons originating from W boson decays are called *signal*, while those coming from the hadron
 1306 decays are called *background*. For CB muons the three main identification variables are the following:

- 1307 • *q/p significance* is defined as $\frac{|(q/p)_{ID} - (q/p)_{MS}|}{\sqrt{\sigma^2(q/p)_{ID} + \sigma^2(q/p)_{MS}}}$ - an absolute difference between *q/p* measured in
 1308 the two detectors over the combined uncertainty.
- 1309 • Relative transverse momentum difference $\rho = \frac{|p_T^{ID} - p_T^{MS}|}{p_T^{combined}}$.
- 1310 • Normalized χ^2 fit of the combined track.

1311 Robust momentum measurement is ensured by specific requirements to the number of hits in the ID
 1312 and MS. A number of muon identification selections (working points) is developed to address specific
 1313 analyses.

1314 6.3.3 Muon isolation

1315 Isolated muons are a defining signature of massive boson decays. In the decays of W, Z and Higgs
1316 bosons muons are created separated from the rest of the particles. Quantitative measurement of
1317 detector activity around a muon candidate is called *muon isolation* and serves as an invaluable tool
1318 for background suppression. Muon isolation is assessed through two observables: one is track-based,
1319 another is calorimeter-based.

1320 The track-based observable $p_T^{varcone30}$ is defined as a scalar sum of all the particles with $p_T > 1$ GeV in a
1321 cone $\Delta R = \min(10\text{GeV}/p_T^\mu, 0.3)$ around the muon with transverse momentum p_T^μ excluding the proper
1322 track of the muon. The p_T dependence helps this definition to perform better for the muons created in
1323 the decay of the particles with high transverse momentum.

1324 The calorimeter-based isolation observable $E_T^{topocone20}$ is defined as the sum of the transverse energy of
1325 all the topological clusters in a cone of a size $\Delta R = 0.2$ around the muon after subtracting the proper
1326 muon energy deposit and correcting for the pile-up effects.

1327 Isolation criteria are normally defined using the relative isolation variables, using the ratio of $p_T^{varcone30}$
1328 and $E_T^{topocone20}$ to the transverse momentum. A number of working points exist, each having a certain
1329 requirements for one or both of the isolation variables.

1330 6.4 Electron reconstruction and identification**1331 6.4.1 Electron reconstruction**

1332 Electron reconstruction starts with two separate parts: track reconstruction in the ID and cluster
1333 reconstruction in the calorimeter, which are then matched to each other in order to make an electron
1334 candidate [10]. During Run 2 two algorithms were used for the cluster reconstruction, both of them are
1335 described below.

1336

1337 Sliding window

1338 The EMC is divided into a grid of 200x256 towers in $\eta \times \phi$ plane, each tower having a size of $\Delta\eta \times \Delta\phi =$
1339 0.025×0.025 , reproducing the granularity of the second layer in the EMC. Energy deposits in all
1340 available calorimeter layers (first, second and third layers of the EMC in the region $|\eta| < 2.47$ and the
1341 presampler in the region $|\eta| < 1.8$) are approximately calibrated at the EM scale and summed up for
1342 each tower. If the cumulative energy deposit in a certain tower exceeds 2.5 GeV then this tower is used
1343 as a seed. Then for every seed a sliding window algorithm of size 3×5 is used [11], forming a cluster
1344 around every seed.

1345 It happens that two seed-cluster candidates are found in close proximity. When their towers overlap

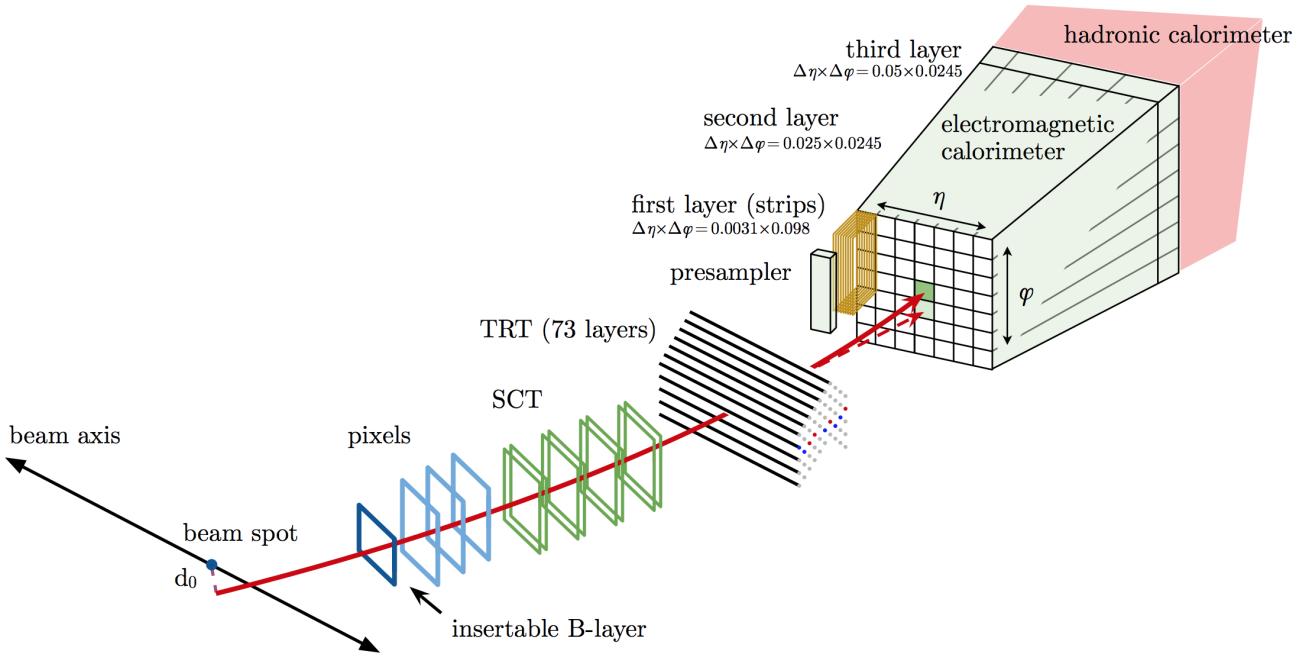


Figure 66: The path of an electron through the detector is shown by solid red line. The dashed red line denotes the trajectory of a photon, produced as a Bremsstrahlung radiation in the TRT.

1346 within an area of $\eta \times \phi = 5 \times 9$ in units of 0.025×0.025 the two clusters are considered overlapping. In
1347 this case two options are possible:

- 1348 • If the transverse energies of the two clusters are more than 10% different then the cluster with
1349 higher E_T is retained.
1350 • If the difference in the transverse energies is within 10% then the cluster with higher value of the
1351 E_T in the central tower is kept.

1352 After the overlap is resolved the duplicate cluster gets removed.
1353

1354 Topocluster reconstruction

1355 The algorithm for topocluster reconstruction [12], [13] starts with composing proto-clusters in the
1356 calorimeter using the noise threshold:

$$\zeta_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{noise,cell}^{EM}}, \quad (6.2)$$

1357 where E_{cell}^{EM} is the cell energy at the EM scale and $\sigma_{noise,cell}^{EM}$ is the expected cell noise. The latter
1358 comprises of the electronic noise and pile-up noise estimate base on the expected instantaneous

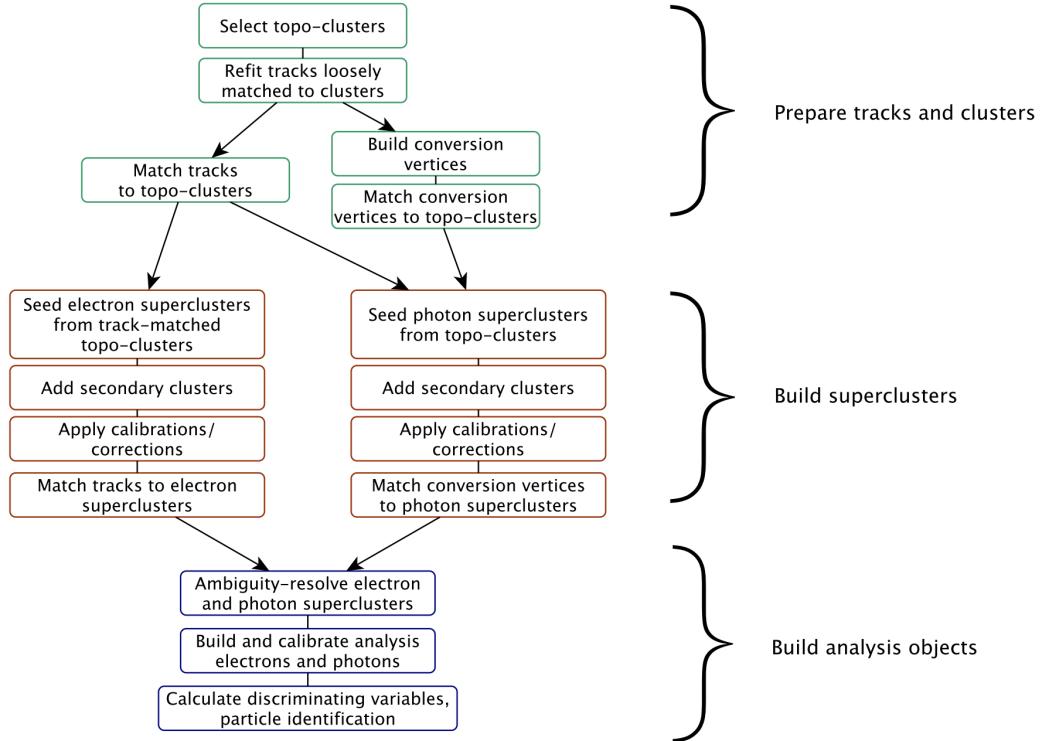


Figure 67: The algorithm scheme for topocluster reconstruction.

luminosity. The proto-cluster is formed around a cell with $|\zeta_{cell}^{EM}| \geq 4$. Then the neighbouring cell that pass the requirement of $|\zeta_{cell}^{EM}| \geq 2$ are added to the proto-cluster and serve as a seed for the next iteration, collecting all of its neighbours to the proto-cluster. If the two proto-clusters share a cell with $|\zeta_{cell}^{EM}| \geq 2$ then these proto-clusters are merged together. At every iteration neighbouring cells are added to the cluster even if they don't make the threshold. Proto-clusters with two local maxima are split into two clusters. For the proto-cluster to be considered as EM topocluster it must have at least 50% of its energy being contained in the EMC. At the stage of track reconstruction the tracks are first extended and fitted with the global χ^2 fitter using the pion hypothesis [14]. If it fails, then a more complicated pattern reconstruction algorithm based on Kalman filter is used [15]. This algorithm uses the electron hypothesis and allows up to 30% energy loss at each material surface. Then the tracks are loosely matched to the EM clusters if they meet one of the following criteria:

- The tracks extrapolated to the second layer of the EMC are consistent in ϕ and η (matching in η is not required for TRT-only tracks).
- The extrapolated tracks are consistent in ϕ (with a bit tighter requirements) and η after rescaling the track momentum to cluster momentum.

1374 **Track-cluster matching**

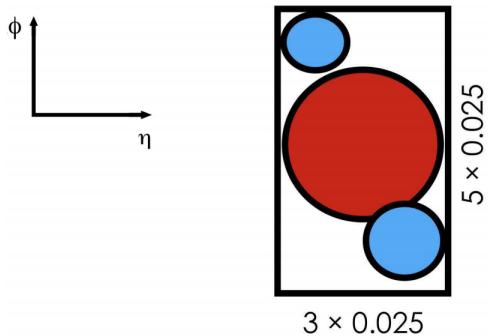
1375 Matching in ϕ coordinate assumes charge asymmetry to account for different direction of possible
 1376 Bremsstrahlung radiation for positive and negative particles. Then the loosely matched tracks that have
 1377 at least four silicon hits are refitted using the optimized Gaussian-sum filter (GSF) [16], that allows to
 1378 better take into account the energy losses in solid material.
 1379 After the track is fitted with the GSF algorithm the final matching with the cluster is performed using
 1380 tighter matching requirements between the track and the cluster barycentre. If matching criteria are
 1381 met with two or more tracks then an ambiguity resolving algorithm is used. This algorithm takes into
 1382 account a number of parameters like the distance between the cluster barycentre and the track in ϕ
 1383 and η , number of hits in the silicon detector and in the innermost silicon layer, association to photon
 1384 conversion vertex, E/p ration and p_T . This allows to rule out converted photons as electron candidates
 1385 and also helps to maintain high photon reconstruction efficiency. After track-cluster matching the
 1386 electron cluster is extended around the seed to 3×7 in the barrel region or 5×5 in the end-cap region
 1387 by adding one row of the cells on each side.

1388

1389 **Supercluster reconstruction**

All e^\pm, γ :

Add all clusters within 3×5 window
around seed cluster.



Electrons only:

Seed, secondary cluster
match the same track.

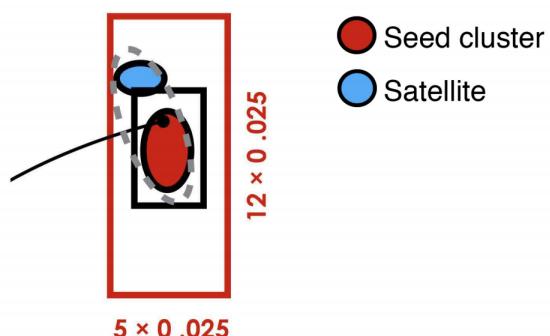


Figure 68: Supercluster reconstruction for electrons. Seed clusters are shown in red, satellite clusters in blue.

1390 The composition of electron supercluster is performed in two stages: first, the candidate EM topoclus-
 1391 ter is tested to be used as a seed for the supercluster. In the second stage the nearby EM topoclusters
 1392 can be identified as satellite clusters, emerging from Bremsstrahlung radiation or topocluster splitting.
 1393 First the EM topoclusters are sorted by their E_T in descending order. For the cluster to be considered
 1394 a seed it must have the $E_T > 1$ GeV, must be matched to a track with at least four hits in the silicon

1395 detectors and should not be assigned as a satellite cluster to any other seed. If these requirements
 1396 are met then the algorithm described in Fig. 68 is started. First, all topoclusters within a window of
 1397 $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre are added as satellite cluster, as they most
 1398 probably represent secondary EM showers coming from the same initial electron. Also, if a cluster
 1399 within $\Delta\eta \times \Delta\phi = 0.125 \times 0.3$ window around the seed cluster barycentre share the "best-matched" track
 1400 with the seed cluster - it is also added as a satellite. Finally the energy of the reconstructed cluster
 1401 must be calibrated. The calibration is performed using the multivariate technique based on data and
 1402 MC samples using $Z \rightarrow ee$ events [17], [18]. The shower shapes and other discriminating variables are
 1403 computed at this stage.

1404 **6.4.2 Electron identification**

1405 Prompt electrons in the central region of the ATLAS detector ($|\eta| < 2.47$) are selected using a likelihood-
 1406 based (LH) identification. The LH uses a number of inputs from ID and calorimeter detectors, as well
 1407 as combined information from both detectors (see Table 1 in [10]). The probability density functions
 1408 (pdfs) for the likelihoods of Run 2 were obtained using the simulated events.
 1409 The electron LH is based on the products of n pdfs P for signal L_S and background L_B :

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{S(B)}^i(x_i), \quad (6.3)$$

1410 where \mathbf{x} is the vector of the LH input parameters, P_S^i and P_B^i are the pdf values for parameter i at value x_i
 1411 for signal and background respectively. The LH operates at a number of working points, the higher the
 1412 likelihood - the lower is the efficiency. For example, the efficiencies for identifying a prompt electron
 1413 with $E_T = 40$ GeV for Loose, Medium and Tight working points are 93%, 88% and 80% respectively.
 1414 Prompt electrons are assumed to be the signal, while background includes the jets that mimic the
 1415 prompt electrons, electrons from photon conversions and non-prompt electrons from hadron decays.
 1416 For each electron candidate a discriminant d_L is composed:

$$d_L = \frac{L_S}{L_S + L_B}, \quad (6.4)$$

1417 that defines the electron likelihood identification. This discriminant d_L has a sharp peak at unity for
 1418 the signal and at zero for the background, which is not very convenient for picking working points.
 1419 That is why the discriminant distribution is transformed using the inverse sigmoid function:

$$d'_L = -\tau^{-1} \ln(d_L^{-1} - 1), \quad (6.5)$$

1420 where $\tau = 15$. Each operating point is assigned with a d'_L value - if a discriminant exceeds this value for
 1421 a given electron then this electron is considered signal.
 1422 There are two advantages of using likelihood-based approach comparing to selection-criteria-based
 1423 ("cut-based") identification:

- 1424 • The drawback of a cut-based approach is that if an electron fails to pass one of the cuts - it is
 1425 definitely removed from the selection, while in the LH approach it is still possible for this electron
 1426 to pass the selection thanks to other parameters. This quality promotes the selection efficiency.
- 1427 • In case of a significant overlap in signal and background distribution of a certain parameter using
 1428 it in a cut-based identification would entail large losses in efficiency. In the likelihood-based
 1429 identification this parameter may be added without penalty.
- 1430 The likelihood input parameters were obtained from the simulated events, which means that real
 1431 distributions in data may differ due to various mismodelling effects. These effects must be corrected
 1432 in order to get the accurate and efficient identification. Mismodelling may depend on coordinates or
 1433 energy. Chapter 5 of this dissertation is devoted to correction of electromagnetic shower shapes in the
 1434 calorimeter, which are among the likelihood input parameters.

1435 **6.4.3 Electron isolation**

1436 Electron isolation plays a very important role in background suppression in physics analyses. Since
 1437 electrons are reconstructed using the information from two different detectors - two different isolation
 1438 definitions are possible, track-based and calorimeter-based. Let's first consider calorimeter-based
 1439 isolation.

As depicted in Fig. 69 the raw isolation energy $E_{isol}^{T,raw}$ includes the energy of all the topoclusters,

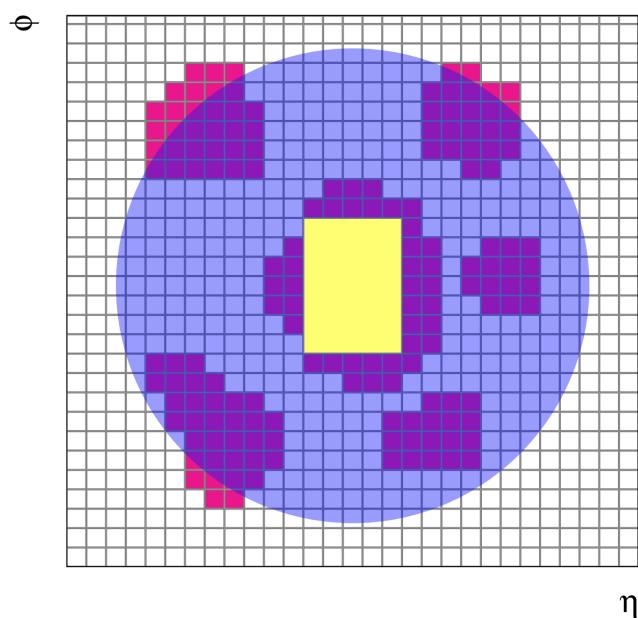


Figure 69: The isolation cone is centred at the candidate electron. All topological clusters, shown in red, are included in the raw isolation variable. The 5×7 cells included into core subtraction method are marked in yellow.

1440 barycentres of which fall within the isolation radius ΔR . It also includes core energy of the electron
 1441 candidate $E_{isol}^{T,core}$ which comprises the 5×7 cells within the area of $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$. The fixed
 1442 size of the core ensures simplicity and stability, although it may happen that the topocluster is larger
 1443 than the size of the core resulting in attributing the proper energy of the electron to the outside activity.
 1444 This leakage effect is corrected for using no pile-up simulated events, parametrizing the leakage with a
 1445 Crystal Ball function as a function of the transverse energy $E_{T,leakage} = E_{T,leakage}(E_T)$.
 1446 Another effect that must be corrected for is the pile-up and underlying event contribution. This
 1447 contribution is estimated from the ambient energy density [19]. This implies the calculation of
 1448 the median energy density $\rho_{median}(\eta)$ - a rapidity-dependent estimate of jet densities for every event.
 1449 Then the pile-up correction can be evaluated in the following way:
 1450

$$E_{T,pile-up}(\eta) = \rho_{median}(\eta) \times (\pi \Delta R^2 - A_{core}), \quad (6.6)$$

1451 where ΔR is the radius of the isolation cone, and A_{core} is the area of the subtracted signal core. Finally
 1452 the calorimeter isolation variable may be defined as follows:
 1453

$$E_{T,cone}^{isol} = E_{T,raw}^{isol} - E_{T,core} - E_{T,leakage} - E_{T,pile-up}. \quad (6.7)$$

1453 The track-based isolation includes all tracks with $p_T > 1$ GeV within a fiducial region of the ID that
 1454 satisfy basic track quality requirements. Pile-up is mitigated by requiring that $|z_0 \sin \theta| < 3$ mm, to
 1455 ensure that the track points at the primary vertex. The track-based isolation is composed of all the
 1456 tracks that fall within the radius ΔR excluding the candidate electron track.
 1457 The own contribution of the candidate track into the isolation must also include possible Bremsstrahlung
 1458 radiation emitted by the candidate electron. For that reason the tracks are extrapolated to the second
 1459 layer of the EMC and if they fall within a window of $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$ around the cluster position.
 1460 The resulting variable is called p_T^{isol} .
 1461 The track-based isolation allows to use variable-size cone, making the cone smaller for boosted particles.
 1462 The cone size for the $p_{T,var}^{isol}$ would be:
 1463

$$\Delta R = \min\left(\frac{10\text{GeV}}{p_T[\text{GeV}]}, R_{max}\right), \quad (6.8)$$

1463 where R_{max} is the maximum cone size and may vary depending on the analysis needs, typically between
 1464 0.2 and 0.4.
 1465

6.5 Particle flow objects

1466 The measurement of hadronic objects and particle showers remains a complicated task due to large
 1467 variety of particle types and properties they posses and because of the large energy/momentum span of
 1468 the measured objects. For the low-energy charged particles the ID shows better momentum resolution
 1469 and angular resolution. On the other hand, the calorimeter shows better performance at high energy
 1470

and is also capable of detecting neutral particles. The idea behind the Particle Flow (PF) algorithm [20] is to combine the information from the two detectors to obtain the best result possible. To properly take into account every particle it has to be ensured that every particle detected in both detectors is counted only once. This means that for a charged particle its deposit in the calorimeter must be found and subtracted. The Particle Flow Object (PFO) reconstruction process is schematically presented in Fig. 610. The process starts with getting *tight* tracks from the ID, meaning these tracks must have at

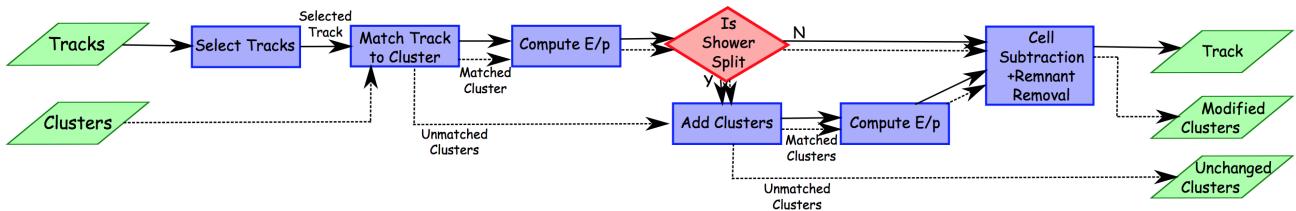


Figure 610: The algorithm scheme for particle flow object reconstruction.

least nine hits in the silicon detectors and no holes in the pixel detector. The tracks must have $|\eta| < 2.5$ and $0.5 < p_T < 40$ GeV, corresponding to the kinematic region where tracks offer better resolution than the calorimeter. The tracks associated to leptons are removed. The calorimeter topoclusters reconstructed like it was described in section 6.4.1 and calibrated using the EM scale are matched to the tracks based on their spacial position and measured momentum. First the ranked based on a distance metric:

$$\Delta R' = \sqrt{\left(\frac{\Delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2}, \quad (6.9)$$

where $\Delta\phi$ and $\Delta\eta$ are the angular distances between the topocluster barycentres and the track, σ_ϕ and σ_η are uncertainties in topocluster width. Preliminary matching is reached by requiring that $E^{clus}/p^{trk} > 0.1$, where E^{clus} is the cluster energy and p^{trk} is the track momentum. It often happens, that energy deposit of a particle is split between two (most often) or more clusters. Then a split shower recovery procedure is initiated, looking for matching clusters in the radius of $\Delta R = 0.2$ around the track extrapolated to the second layer of the EMC. Then it is estimated if the energy of the track and the energy of the associated topocluster is consistent. If it is the case then the topoclusters matched to the tracks are removed. Eventually two particle collections are obtained: a collection of charged particle flow objects (cPFOs) each with an associated track and neutral particle flow objects (nPFOs) with a calorimeter deposit. The former must also match the primary vertex, having $|z_0 \times \sin\theta| < 2$ mm.

7

MC samples and selection

1493

1494

1495 “Potentielle citation sans aucun rapport avec le sujet”

1496 — Personne inconnue, contexte à déterminer

1497 7.1 Data samples

1498 7.2 MC Signal and Background samples

1499 7.3 Observables

1500 7.3.1 Hadronic recoil

1501 7.3.2 SET

1502 7.3.3 SETUE

1503 7.3.4 Transverse mass

1504 7.3.5 Missing transverse energy

1505 7.4 Event selection

1506 7.4.1 Electrons

1507 7.4.2 Muons

1508 7.4.3 Truth

1509 7.4.4 W boson

1510 7.5 MC samples and backgrounds

8

MC samples and selection

1513 *“Potentielle citation sans aucun rapport avec le sujet”*

1514 — Personne inconnue, contexte à déterminer

1515 8.1 Data and MC samples

1516 The data and MC samples for this study were collected under special beam conditions that ensure low
1517 pile-up. The data samples were collected in three runs:

- 1518 • $\sqrt{s} = 5.02\text{TeV}$ data taken in November 2017, ATLAS data period M, preliminary calibrated
1519 luminosity 256.827 pb^{-1} with an uncertainty of $\pm 1.6\%$
- 1520 • $\sqrt{s} = 13\text{TeV}$ data taken in November 2017, ATLAS data period N, preliminary online luminosity
1521 146.6 pb^{-1}
- 1522 • $\sqrt{s} = 13\text{TeV}$ data taken in June 2018, ATLAS data period G4+J, preliminary online luminosity
1523 193.2 pb^{-1}

1524 The runs of November 2017 and the run of June 2018 had the same bunch spacing of 25 ns, but a
1525 different filling scheme. The two main differences from the high- μ data collection are the following:

- 1526 • In order to optimize topo-cluster response for the Hadronic Recoil (HR) lower topo-cluster
1527 thresholds were used.
- 1528 • Single e and μ triggers with significantly lower thresholds and looser identification criteria are
1529 run without prescale, most notably `HLT_e15_1hloose_nod0_L1EM12` and `HLT_mu14`.

1530 At the beginning of 5 TeV fills the pile-up reached $\mu \backslash 5$, slowly descending to $\mu \backslash 1$ by the end of the run.

1531 In the case of 13 TeV the luminosity was levelled at $\mu = 2$ in the course of the run. The corresponding
1532 distributions for μ and N_{PV} for the 5 TeV and 13 TeV runs are shown in Fig. 81 and Fig. 82.

1533 8.1.1 MC samples and cross-sections

1534 Signal and background processes are modelled using fully simulated and reconstructed using Monte-
1535 Carlo (MC) samples, specifically tuned for the special run conditions, namely the pileup conditions,

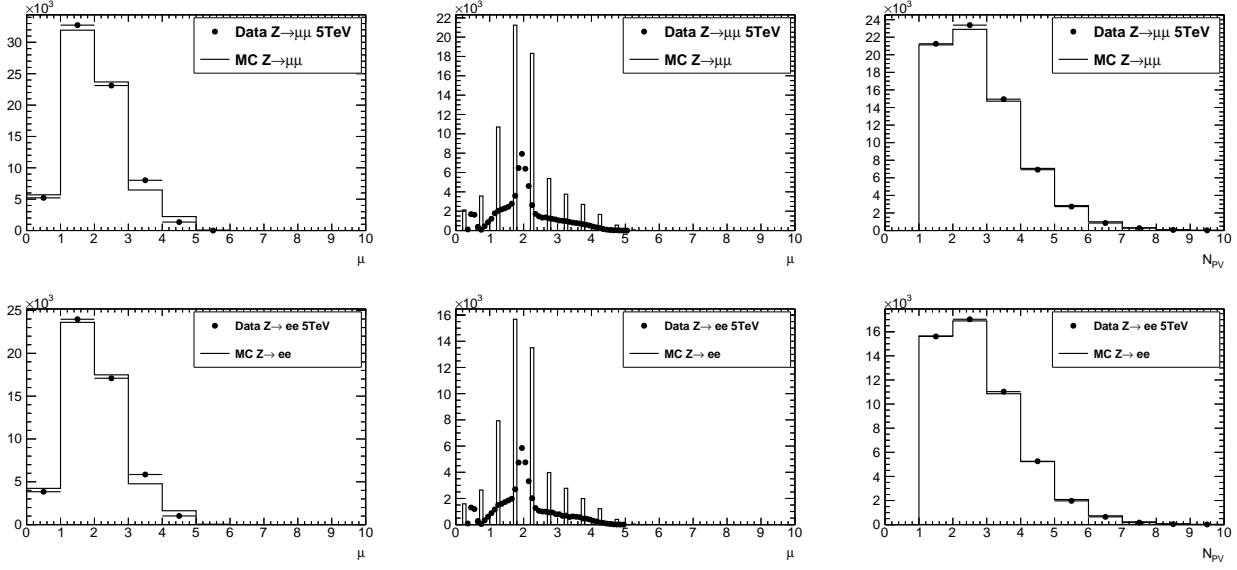


Figure 81: Distributions for the 5 TeV low- μ dataset in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. The left and middle plots show the actual μ in a coarsely-binned and a finely-binned version. The right plot shows the number of reconstructed primary vertices N_{PV} .

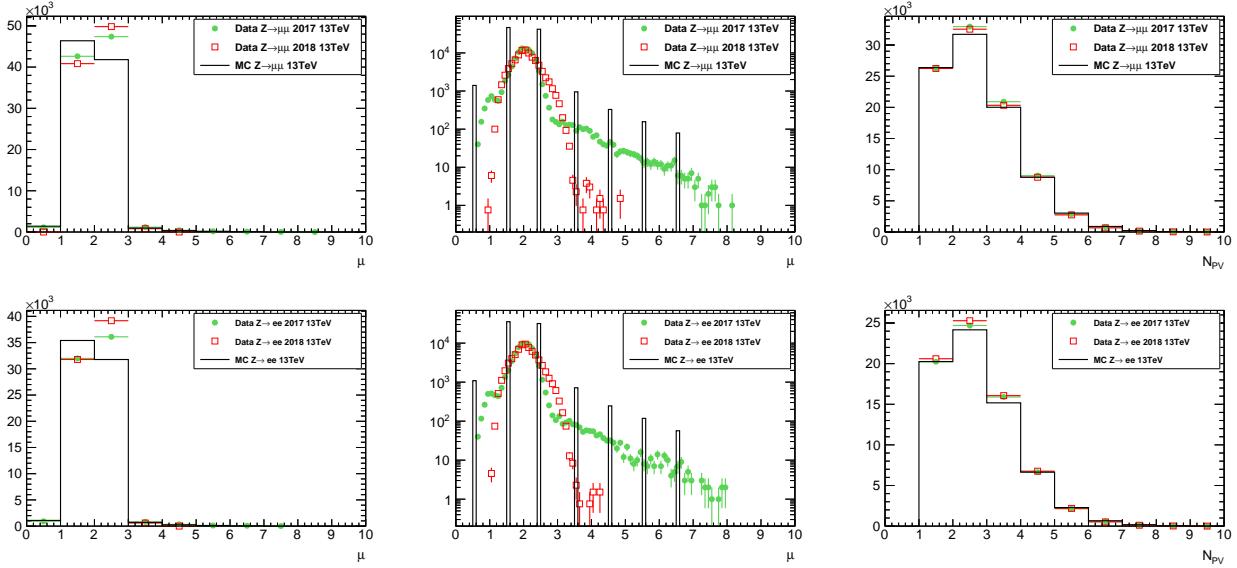


Figure 82: Distributions for the 13 TeV low- μ datasets taken in 2017 and 2018 in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. All distributions are (roughly) normalised to the same number of selected events in the 2017 dataset. The left and middle plots show the actual μ in a coarsely-binned and a finely-binned version. The right plot shows the number of reconstructed primary vertices N_{PV} .

1536 lower topo-cluster noise thresholds and adapter trigger menu. No pileup reweighting is performed.
1537 The information on the simulated samples and their properties is given in Tables 81, 82, 83, 84. The
1538 predicted event counts are normalized to the cross-sections quoted in the table.
1539 The primary signal event samples for W and Z production are obtained using PowHEG [1, 2, 3, 4] event
1540 generator with CT10 PDF, linked with PYTHIA8 [5] with AZNLO tune [6]. PowHEG+PYTHIA88 samples
1541 are interfaced to PHOTOS++ [7] for final state QED effects simulation.
1542 A set of alternative samples at $\sqrt{s} = 13\text{TeV}$ was prepared with SHERPA2.2.2 [8] using the NNPDF3.0
1543 PDFs and merging $V + 0, 1, 2$ at NLO accuracy with $V + 3, 4$ at LO accuracy with the MEPS@NLO scheme.
1544 A similar set for $\sqrt{s} = 5\text{TeV}$ was prepared with SHERPA2.2.5 with a setup similar to 13 TeV samples.
1545 Pileup is modelled by overlaying simulated soft events over the original hard-scattering event. These
1546 soft events were modelled using PYTHIA8 with NNPDF2.3LO set of PDFs [9] and the A3 tune [10].
1547 The W and Z processes samples are normalized to NNLO calculations performed using the DYTURBO,
1548 an optimised version of DYNNLO [11, 12] using the MMHT2014nnlo PDF set [13]. Corresponding
1549 numerical values were taken from the corresponding ATLAS publications of the 2015 data at 13 TeV [14]
1550 and 5.02 TeV [15] are presented in Table 81 for 13 TeV and Table 84 for 5 TeV. The uncertainties on
1551 those cross-sections arise from the choice of PDF set, from factorization and renormalisation scale
1552 dependence, and the strong coupling constant α_s uncertainty resulting in the total uncertainty estimate
1553 of 5%.

1554 Backgrounds from top-quark pair-production $t\bar{t}$ and single-top production (Wt , t-channel, s-
1555 channel) were generated with PowHEG+PYTHIA8. The 5 TeV $t\bar{t}$ cross section is taken as the top++
1556 prediction observed by CMS [16]. Di-boson combinations VV , $V = W, Z$ are generated with SHERPA in
1557 all decay channels with a requirement of having at least one real lepton in the final state.

1558 8.2 Z vertex reweighting

1559 The 5 TeV MC samples have been generated to be perfectly matched to the data. Although this is not
1560 the case for 13 TeV samples, which can be seen at Fig. 83. It is also seen from these plots that the
1561 2017 and 2018 data were collected at two different runs under different beam conditions. To avoid
1562 possible impact on the acceptance the MC samples were reweighted to the data using $Z \rightarrow ee$ and
1563 $Z \rightarrow \mu\mu$ selections.

1564 8.2.1 Multijet background

1565 8.3 W analysis event selection and control plots

Process	Data set	Generator	$\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.)	$N_{\text{evt}}^{\text{skim}} [10^6]$	$N_{\text{evt}}^{\text{unskim}} [10^6]$
$W^+ \rightarrow e^+ \nu$	361100	POWHEG+PYTHIA8	11.61 (5%)	40	40
$W^+ \rightarrow \mu^+ \nu$	361101	POWHEG+PYTHIA8	11.61 (5%)	40	40
$W^+ \rightarrow \tau^+ \nu$	361102	POWHEG+PYTHIA8	11.61 (5%)	0.28	5.0
$W^- \rightarrow e^- \bar{\nu}$	361103	POWHEG+PYTHIA8	8.630 (5%)	30	30
$W^- \rightarrow \mu^- \bar{\nu}$	361104	POWHEG+PYTHIA8	8.630 (5%)	29	29
$W^- \rightarrow \tau^- \bar{\nu}$	361105	POWHEG+PYTHIA8	8.630 (5%)	0.24	4.0
$Z \rightarrow ee$	361106	POWHEG+PYTHIA8	1.910×1.03 (5%)	10	10
$Z \rightarrow \mu\mu$	361107	POWHEG+PYTHIA8	1.910×1.025 (5%)	10	10
$Z \rightarrow \tau\tau$	361108	POWHEG+PYTHIA8	1.910×1.025 (5%)	0.12	1.0
$ZZ(q\bar{q}\ell\ell)$	363356	SHERPA 2.2.1	0.01556×0.141 (10%)	0.0064	0.010
$WZ(q\bar{q}\ell\ell)$	363358	SHERPA 2.2.1	0.003433 (10%)	0.0063	0.010
$WW(q\bar{q}\ell\nu)$	363359	SHERPA 2.2.1	0.02472 (10%)	0.0093	0.020
$WW(\ell\nu q\bar{q})$	363360	SHERPA 2.2.1	0.02472 (10%)	0.0093	0.020
$WZ(\ell\nu q\bar{q})$	363489	SHERPA 2.2.1	0.01142 (10%)	0.0047	0.010
$ZZ(4\ell)$	364250	SHERPA 2.2.2	0.001252 (10%)	0.0057	0.010
$WZ(3\ell\nu)$	364253	SHERPA 2.2.2	0.004583 (10%)	0.0062	0.010
$WW(2\ell 2\nu)$	364254	SHERPA 2.2.2	0.01250 (10%)	0.0073	0.010
$WZ(\ell 3\nu)$	364255	SHERPA 2.2.2	0.003235 (10%)	0.0050	0.010
Wt	410013	POWHEG+PYTHIA8	0.03582 (10%)	0.0037	0.010
$W\bar{t}$	410014	POWHEG+PYTHIA8	0.03399 (10%)	0.0037	0.010
$t\bar{t}$ (nominal)	410470	POWHEG+PYTHIA8	0.8318×0.544 (7%)	1.2	2.0
$t(t-\text{chan.}t)$	410642	POWHEG+PYTHIA8	0.03699 (10%)	0.016	0.030
$t(t-\text{chan.}\bar{t})$	410643	POWHEG+PYTHIA8	0.02217 (10%)	0.011	0.020
$t(s-\text{chan.}t)$	410644	POWHEG+PYTHIA8	0.002027 (10%)	0.0050	0.010
$t(s-\text{chan.}\bar{t})$	410645	POWHEG+PYTHIA8	0.001268 (10%)	0.0052	0.010
$t\bar{t}$ (syst.)	410480	POWHEG+PYTHIA8	0.8318×0.438 (7%)	0.85	1.5
$t\bar{t}$ (syst.)	410482	POWHEG+PYTHIA8	0.8318×0.105 (7%)	0.40	0.50
$t\bar{t}$ (syst.)	410557	POWHEG+PYTHIA8	0.8318×0.438 (7%)	0.85	1.5
$t\bar{t}$ (syst.)	410558	POWHEG+PYTHIA8	0.8318×0.105 (7%)	0.40	0.50

Table 81: Monte Carlo samples at $\sqrt{s} = 13\text{TeV}$. Given is a short description of the process, the ATLAS MC data set number (DSID), the names and version numbers of the MC generator(s), the used value of the higher order cross section times any branching and filter efficiencies ($\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$) with the theoretical uncertainty in percent (“th. unc.”), and finally the number of events analysed after skimming at derivation production ($N_{\text{evt}}^{\text{skim}}$) as well as the number of events originally processed and simulated ($N_{\text{evt}}^{\text{unskim}}$). In the case of $Z \rightarrow \ell\ell$ samples, the given $\epsilon_{\text{filter}} > 1$ is related to the fact, that the cross sections were calculated for $66 < m_{\ell\ell} < 116\text{GeV}$, but the generated mass range is larger. The last section of $t\bar{t}$ samples refers to variation samples for systematics studies. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}} / (\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above 3fb^{-1} for signal and significant backgrounds, the exception are Powheg $W \rightarrow \tau\nu$ and $Z \rightarrow \tau\tau$ samples, that have about 0.45fb^{-1} only.

Process	Data set	Generator	$\sigma \cdot BR \cdot \epsilon_{\text{filter}} [\text{nb}] (\text{th. unc.})$	$N_{\text{evt}}^{\text{skim}} [10^6]$	$N_{\text{evt}}^{\text{unskim}} [10^6]$
$Z \rightarrow \mu\mu$	364100	SHERPA 2.2.1	$1.932 \times 0.822 (5\%)$	8.0	8.0
$Z \rightarrow \mu\mu$	364101	SHERPA 2.2.1	$1.933 \times 0.114 (5\%)$	1.5	1.5
$Z \rightarrow \mu\mu$	364102	SHERPA 2.2.1	$1.932 \times 0.0660 (5\%)$	1.1	1.1
$Z \rightarrow \mu\mu$	364103	SHERPA 2.2.1	$0.1063 \times 0.690 (5\%)$	1.5	1.5
$Z \rightarrow \mu\mu$	364104	SHERPA 2.2.1	$0.1062 \times 0.200 (5\%)$	0.40	0.40
$Z \rightarrow \mu\mu$	364105	SHERPA 2.2.1	$0.1063 \times 0.114 (5\%)$	0.25	0.25
$Z \rightarrow \mu\mu$	364106	SHERPA 2.2.1	$0.03889 \times 0.593 (5\%)$	0.20	0.20
$Z \rightarrow \mu\mu$	364107	SHERPA 2.2.1	$0.03885 \times 0.235 (5\%)$	0.060	0.060
$Z \rightarrow \mu\mu$	364108	SHERPA 2.2.1	$0.03889 \times 0.156 (5\%)$	0.035	0.035
$Z \rightarrow \mu\mu$	364109	SHERPA 2.2.1	$0.008310 \times 0.561 (5\%)$	0.020	0.020
$Z \rightarrow \mu\mu$	364110	SHERPA 2.2.1	$0.008310 \times 0.266 (5\%)$	0.010	0.010
$Z \rightarrow \mu\mu$	364111	SHERPA 2.2.1	$0.008320 \times 0.177 (5\%)$	0.0050	0.0050
$Z \rightarrow \mu\mu$	364112	SHERPA 2.2.1	$0.001740 (5\%)$	0.0050	0.0050
$Z \rightarrow \mu\mu$	364113	SHERPA 2.2.1	$0.0001400 (5\%)$	0.0050	0.0050
$Z \rightarrow ee$	364114	SHERPA 2.2.1	$1.933 \times 0.821 (5\%)$	8.0	8.0
$Z \rightarrow ee$	364115	SHERPA 2.2.1	$1.932 \times 0.114 (5\%)$	1.5	1.5
$Z \rightarrow ee$	364116	SHERPA 2.2.1	$1.932 \times 0.0658 (5\%)$	1.1	1.1
$Z \rightarrow ee$	364117	SHERPA 2.2.1	$0.1080 \times 0.694 (5\%)$	1.5	1.5
$Z \rightarrow ee$	364118	SHERPA 2.2.1	$0.1077 \times 0.191 (5\%)$	0.40	0.40
$Z \rightarrow ee$	364119	SHERPA 2.2.1	$0.1078 \times 0.119 (5\%)$	0.25	0.25
$Z \rightarrow ee$	364120	SHERPA 2.2.1	$0.03964 \times 0.616 (5\%)$	0.20	0.20
$Z \rightarrow ee$	364121	SHERPA 2.2.1	$0.03967 \times 0.233 (5\%)$	0.060	0.060
$Z \rightarrow ee$	364122	SHERPA 2.2.1	$0.04068 \times 0.150 (5\%)$	0.035	0.035
$Z \rightarrow ee$	364123	SHERPA 2.2.1	$0.008460 \times 0.569 (5\%)$	0.020	0.020
$Z \rightarrow ee$	364124	SHERPA 2.2.1	$0.008450 \times 0.266 (5\%)$	0.010	0.010
$Z \rightarrow ee$	364125	SHERPA 2.2.1	$0.008470 \times 0.177 (5\%)$	0.0050	0.0050
$Z \rightarrow ee$	364126	SHERPA 2.2.1	$0.001760 (5\%)$	0.0050	0.0050
$Z \rightarrow ee$	364127	SHERPA 2.2.1	$0.0001451 (5\%)$	0.0050	0.0050

Table 82: Alternative signal $Z \rightarrow \ell\ell$ Monte Carlo samples at $\sqrt{s} = 13\text{TeV}$ produced with SHERPA. General description of the table see Table 81. The samples are split into a long list of orthogonal slices based on “max(pTV,HT)” and filtered further into “b/c/light-jet” subcomponents. For the purpose of this analysis, the number of events in each slice is such that the samples are about 2fb^{-1} each (after application of a penalty factor for negative weight events) and an “inclusive sample” is restored after merging the slices.

Process	Data set	Generator	$\sigma \cdot BR \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.)	$N_{\text{evt}}^{\text{skim}} [10^6]$	$N_{\text{evt}}^{\text{unskim}} [10^6]$
$W \rightarrow \mu\nu$	364156	SHERPA 2.2.1	18.58×0.825 (5%)	31	31
$W \rightarrow \mu\nu$	364157	SHERPA 2.2.1	18.57×0.131 (5%)	8.1	8.1
$W \rightarrow \mu\nu$	364158	SHERPA 2.2.1	18.57×0.0433 (5%)	2.6	2.6
$W \rightarrow \mu\nu$	364159	SHERPA 2.2.1	0.9173×0.674 (5%)	6.3	6.3
$W \rightarrow \mu\nu$	364160	SHERPA 2.2.1	0.9172×0.244 (5%)	2.1	2.1
$W \rightarrow \mu\nu$	364161	SHERPA 2.2.1	0.9163×0.0847 (5%)	0.23	0.23
$W \rightarrow \mu\nu$	364162	SHERPA 2.2.1	0.3296×0.600 (5%)	0.80	0.80
$W \rightarrow \mu\nu$	364163	SHERPA 2.2.1	0.3297×0.293 (5%)	0.27	0.27
$W \rightarrow \mu\nu$	364164	SHERPA 2.2.1	0.3295×0.111 (5%)	0.099	0.099
$W \rightarrow \mu\nu$	364165	SHERPA 2.2.1	0.06993×0.548 (5%)	0.068	0.068
$W \rightarrow \mu\nu$	364166	SHERPA 2.2.1	0.06995×0.320 (5%)	0.034	0.034
$W \rightarrow \mu\nu$	364167	SHERPA 2.2.1	0.06991×0.125 (5%)	0.014	0.014
$W \rightarrow \mu\nu$	364168	SHERPA 2.2.1	0.01456 (5%)	0.020	0.020
$W \rightarrow \mu\nu$	364169	SHERPA 2.2.1	0.001200 (5%)	0.004	0.004
$W \rightarrow e\nu$	364170	SHERPA 2.2.1	18.58×0.825 (5%)	31	31
$W \rightarrow e\nu$	364171	SHERPA 2.2.1	18.57×0.131 (5%)	8.3	8.3
$W \rightarrow e\nu$	364172	SHERPA 2.2.1	18.57×0.0448 (5%)	2.5	2.5
$W \rightarrow e\nu$	364173	SHERPA 2.2.1	0.9168×0.675 (5%)	6.4	6.4
$W \rightarrow e\nu$	364174	SHERPA 2.2.1	0.9176×0.244 (5%)	2.1	2.1
$W \rightarrow e\nu$	364175	SHERPA 2.2.1	0.9173×0.0851 (5%)	0.79	0.79
$W \rightarrow e\nu$	364176	SHERPA 2.2.1	0.3295×0.599 (5%)	0.76	0.76
$W \rightarrow e\nu$	364177	SHERPA 2.2.1	0.3297×0.288 (5%)	0.28	0.28
$W \rightarrow e\nu$	364178	SHERPA 2.2.1	0.3295×0.111 (5%)	0.10	0.10
$W \rightarrow e\nu$	364179	SHERPA 2.2.1	0.06993×0.548 (5%)	0.070	0.070
$W \rightarrow e\nu$	364180	SHERPA 2.2.1	0.06996×0.320 (5%)	0.034	0.034
$W \rightarrow e\nu$	364181	SHERPA 2.2.1	0.06994×0.137 (5%)	0.014	0.014
$W \rightarrow e\nu$	364182	SHERPA 2.2.1	0.01460 (5%)	0.020	0.020
$W \rightarrow e\nu$	364183	SHERPA 2.2.1	0.001200 (5%)	0.0050	0.0050

Table 83: Alternative signal $W \rightarrow \ell\nu$ Monte Carlo samples at $\sqrt{s} = 13\text{TeV}$ produced with SHERPA. See Table 82 for a description of the table. The samples are split into a long list of orthogonal slices based on “max(pTV,HT)” and filtered further into “b/c/light-jet” subcomponents. For the purpose of this analysis, the number of events in each slice is such that the samples are about 1fb^{-1} each (after application of a penalty factor for negative weight events) and an “inclusive sample” is restored after merging the slices.

Process	Data set	Generator	$\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.)	$N_{\text{evt}}^{\text{skim}} [10^6]$	$N_{\text{evt}}^{\text{unskim}} [10^6]$
$W^+ \rightarrow e^+ \nu$	361100	POWHEG+PYTHIA8	4.357 (5%)	11	11
$W^+ \rightarrow \mu^+ \nu$	361101	POWHEG+PYTHIA8	4.357 (5%)	11	11
$W^+ \rightarrow \tau^+ \nu$	361102	POWHEG+PYTHIA8	4.357 (5%)	0.065	0.94
$W^- \rightarrow e^- \bar{\nu}$	361103	POWHEG+PYTHIA8	2.902 (5%)	7.0	7.0
$W^- \rightarrow \mu^- \bar{\nu}$	361104	POWHEG+PYTHIA8	2.902 (5%)	7.0	7.0
$W^- \rightarrow \tau^- \bar{\nu}$	361105	POWHEG+PYTHIA8	2.902 (5%)	0.039	0.59
$Z \rightarrow ee$	361106	POWHEG+PYTHIA8	0.6600×1.025 (5%)	6.3	6.3
$Z \rightarrow \mu\mu$	361107	POWHEG+PYTHIA8	0.6600×1.025 (5%)	3.4	3.4
$Z \rightarrow \tau\tau$	361108	POWHEG+PYTHIA8	0.6600×1.025 (5%)	0.039	0.29
$Z \rightarrow ee$	364381	SHERPA 2.2.5	0.6600×1.12 (5%)	5.0	5.0
$Z \rightarrow \mu\mu$	364382	SHERPA 2.2.5	0.6600×1.12 (5%)	5.0	5.0
$Z \rightarrow \tau\tau$	364383	SHERPA 2.2.5	0.6600×1.12 (5%)	1.5	1.5
$W \rightarrow e\nu$	364384	SHERPA 2.2.5	7.259 (5%)	25	25
$W \rightarrow \mu\nu$	364385	SHERPA 2.2.5	7.259 (5%)	25	25
$W \rightarrow \tau\nu$	364386	SHERPA 2.2.5	7.259 (5%)	6.0	6.0
$ZZ(4\ell)$	361063	SHERPA 2.1	0.004624 (10%)	0.017	0.049
$WZ(\ell\ell\ell^-\nu\text{SF})$	361064	SHERPA 2.1	0.0005324 (10%)	0.0073	0.015
$WZ(\ell\ell\ell^-\nu\text{OF})$	361065	SHERPA 2.1	0.001041 (10%)	0.012	0.030
$WZ(\ell\ell\ell^+\nu\text{SF})$	361066	SHERPA 2.1	0.0008433 (10%)	0.010	0.020
$WZ(\ell\ell\ell^+\nu\text{OF})$	361067	SHERPA 2.1	0.001633 (10%)	0.016	0.039
$WW(2\ell 2\nu)$	361068	SHERPA 2.1	0.003356 (10%)	0.068	0.090
$WW(q\bar{q}\ell\nu)$	361091	SHERPA 2.1	0.006059 (10%)	0.078	0.15
$WW(\ell\nu q\bar{q})$	361092	SHERPA 2.1	0.006082 (10%)	0.14	0.26
$WZ(\ell\nu q\bar{q})$	361093	SHERPA 2.1	0.002503 (10%)	0.039	0.075
$WZ(q\bar{q}\ell\ell)$	361094	SHERPA 2.1	0.0007518 (10%)	0.017	0.025
$ZZ(q\bar{q}\ell\ell)$	361096	SHERPA 2.1	0.003789×0.148 (10%)	0.0070	0.010
$t\bar{t}$	410470	POWHEG+PYTHIA8	0.06890×0.544 (7%)	1.8	2.8
$t(s - \text{chan.} t)$	410644	POWHEG+PYTHIA8	0.0005400 (10%)	0.028	0.050
$t(s - \text{chan.} \bar{t})$	410645	POWHEG+PYTHIA8	0.0002751 (10%)	0.028	0.050
Wt	410646	POWHEG+PYTHIA8	0.002990 (10%)	0.018	0.050
$W\bar{t}$	410647	POWHEG+PYTHIA8	0.002983 (10%)	0.019	0.050
$t(t - \text{chan.} t)$	410658	POWHEG+PYTHIA8	0.005414 (10%)	0.028	0.050
$t(t - \text{chan.} \bar{t})$	410659	POWHEG+PYTHIA8	0.002682 (10%)	0.028	0.050

Table 84: Monte Carlo samples at $\sqrt{s} = 5\text{TeV}$. The table follows the same format as Table 81. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}} / (\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above 2.5fb^{-1} for signal and significant backgrounds, the exception are Powheg $W \rightarrow \tau\nu$ and $Z \rightarrow \tau\tau$ samples, that have about 0.20fb^{-1} and 0.45fb^{-1} only.

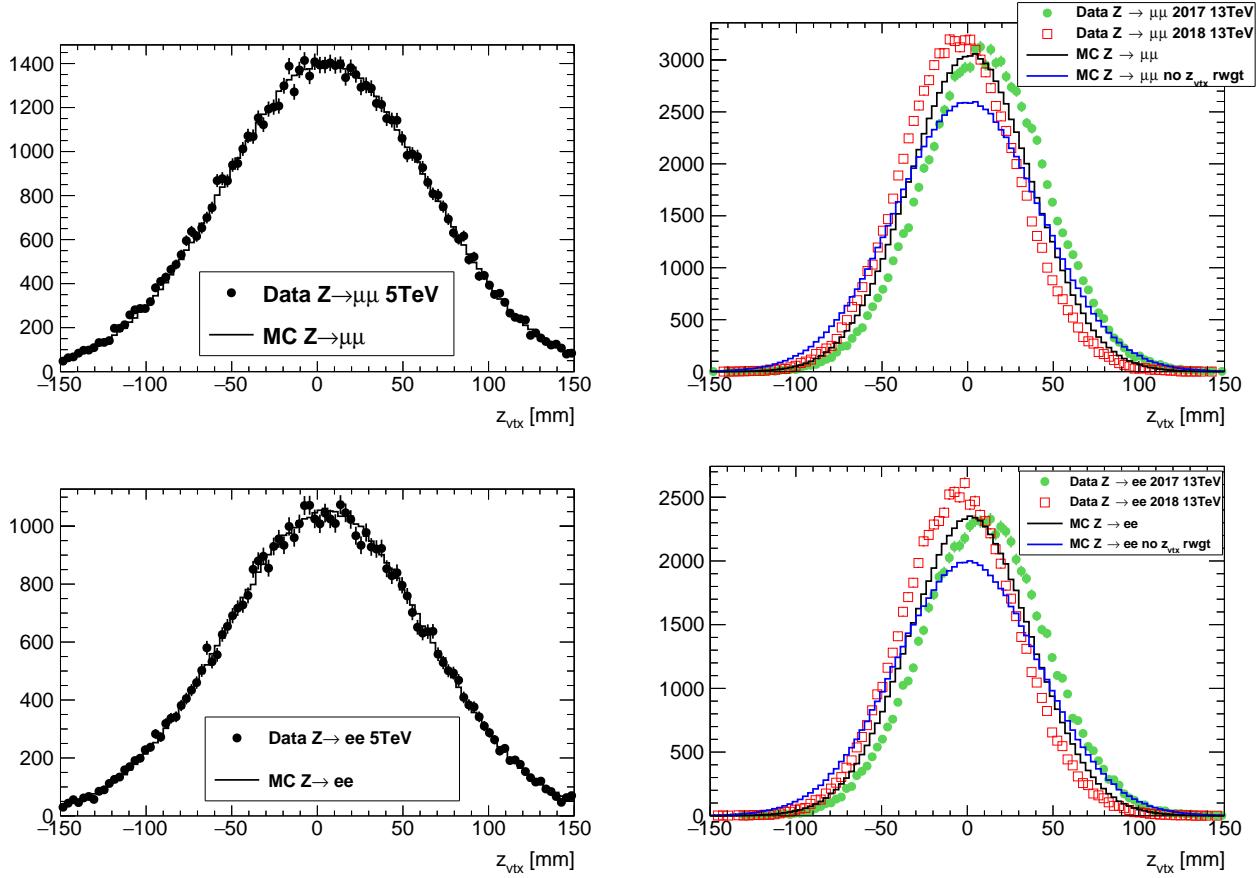


Figure 83: Distributions for the 5 TeV (left) and 13 TeV (right) low- μ dataset(s) in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. The distributions of the z -position of the primary vertex selected as the hard interaction are compared for the dataset(s) and the MC simulation before (“no $z_{\text{vtx}} \text{ rwgt}$ ”, blue, only 13 TeV) and after reweighting (black). For the 13 TeV data the 2017 and 2018 data are shown separately and all distributions are (roughly) normalised to the same number of selected events in the 2017 dataset.

1566 8.3.1 Event selection

1567 Both in case of 5 and 13 TeV events with $W \rightarrow \ell\nu$ candidate were selected base on a single-lepton
 1568 trigger requirement. The trigger for $W \rightarrow e\nu$ event candidate HLT_e15_1h1loose_nod0_L1EM12 require
 1569 at least one reconstructed electron with E_T larger than 15 GeV passing *loose* identification requirements.
 1570 Candidates for $W \rightarrow \mu\nu$ were triggered by HLT_mu14 trigger requiring one muon with E_T larger than
 1571 14 GeV.

1572 Events are required to contain exactly one lepton (muon or electron) candidate having $p_T > 25\text{GeV}$.
 1573 Electrons are required to have $|\eta| < 2.47$ excluding transition region $1.37 < |\eta| < 1.52$. Muons Events
 1574 with additional leptons of the same flavour with transverse momentum greater than 20 GeV satisfying
 1575 some ID criteria are discarded, to better reject the Z background. The ID point is *medium* for the muon

1576 channel, and *loose* for the electron channel. There is no requirement on the number of leptons with
 1577 different flavour than the channel under study.

1578 To suppress background, in particular from multijet processes, events are required to have E_T^{miss}
 1579 greater than 25 GeV. The W boson transverse mass m_T is demanded to be larger than 50 GeV. This
 1580 transverse mass is defined as follows:

$$m_T = \sqrt{2p_T^\nu p_T^l (1 - \cos \Delta\phi^\nu)} \quad (8.1)$$

1581 analogously to Eq. (??) with p_T^ν replaced by E_T^{miss} and ϕ^ν replaced by the azimuthal angle related to
 1582 the E_T^{miss} .

1583 In tables 85,87,89,811 $W^\pm \rightarrow \ell^\pm \nu$ signal selection event yields for the $\sqrt{s} = 5$ TeV low- μ dataset are
 1584 shown, and the corresponding numbers for the 13 TeV low- μ dataset are shown in tables 86,88,810,812.
 1585 A summary of observed compared to expected yields is then provided in table 813. The yields for
 1586 the signal and for the main electroweak and top backgrounds are obtained from the dedicated MC
 1587 samples described in Ref [Kretzschmar:2657141]. The events referred to as $W \rightarrow \ell\nu$ in the tables in
 1588 the plots are adding the W of opposite charge, same flavour, with respect to the signal, to the $W \rightarrow \tau\nu$
 1589 background events. The estimate of the multijet background, which contain contributions from fake
 1590 leptons produced in semi-leptonic decays of heavy quarks, in-flight pion decays, photon conversions,
 1591 etc, is done using a data-driven techniques described in Reference [Xu:2657146].

Cut	Data	Signal	$W^\pm \rightarrow \ell^\pm \nu$ BG	$Z \rightarrow \ell\ell$	Top	Diboson	Multijet
One electron	1993720	643610 \pm 260	32940 \pm 190	44338 \pm 71	1754.4 \pm 3.9	772.2 \pm 3.7	-
Electron trig matched	1907724	612940 \pm 250	30790 \pm 190	42100 \pm 69	1698.5 \pm 3.8	741.1 \pm 3.6	-
Isolation	1438941	610320 \pm 250	30590 \pm 190	41923 \pm 69	1663.6 \pm 3.8	722.5 \pm 3.6	-
$p_T^e > 25\text{GeV}$	720284	482240 \pm 220	14790 \pm 130	31955 \pm 53	1464.5 \pm 3.5	592.1 \pm 3.2	-
$E_T^{\text{miss}} > 25\text{GeV}$	440605	421510 \pm 210	9650 \pm 100	1336 \pm 20	1223 \pm 3.2	420.8 \pm 2.4	-
$m_T > 50\text{GeV}$	430620	417430 \pm 210	8800 \pm 96	1047 \pm 16	944.3 \pm 2.9	373.5 \pm 2.2	3030 \pm 550

Table 85: Analysis cut flow for $W^+ \rightarrow e^+ \nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Cut	Data	Signal	$W^\pm \rightarrow \ell^\pm \nu$ BG	$Z \rightarrow \ell\ell$	Top	Diboson	Multijet
One electron	7915023	1797340 \pm 390	92520 \pm 270	147490 \pm 140	63207 \pm 89	3069 \pm 63	-
Electron trig matched	7840239	1709140 \pm 380	86370 \pm 260	139760 \pm 140	61110 \pm 88	2967 \pm 62	-
Isolation	5413483	1698430 \pm 380	85560 \pm 260	138890 \pm 140	59834 \pm 87	2939 \pm 61	-
$p_T^e > 25\text{GeV}$	2452868	1342200 \pm 330	44450 \pm 190	106270 \pm 110	53811 \pm 82	2565 \pm 58	-
$E_T^{\text{miss}} > 25\text{GeV}$	1275513	1136520 \pm 310	28580 \pm 150	8313 \pm 46	45707 \pm 75	1990 \pm 53	-
$m_T > 50\text{GeV}$	1207776	1117560 \pm 310	24760 \pm 130	6443 \pm 36	34580 \pm 65	1718 \pm 50	28000 \pm 1800

Table 86: Analysis cut flow for $W^+ \rightarrow e^+ \nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Mesure de la masse du boson W avec le détecteur ATLAS au LHC

Cut	Data	Signal		$W^\pm \rightarrow \ell^\pm \nu$ BG		$Z \rightarrow \ell\ell$		Top		Diboson	Multijet
One muon	2434459	760980	\pm 280	35090	\pm 200	37015	\pm 82	2025.3	\pm 4.1	864.7	\pm 3.7
Muon trig matched	2353403	664100	\pm 260	30610	\pm 190	32554	\pm 76	1725.6	\pm 3.8	746.6	\pm 3.4
Isolation	1186616	659200	\pm 260	30400	\pm 190	32303	\pm 76	1574.6	\pm 3.7	710.1	\pm 3.3
$p_T^\mu > 25\text{GeV}$	632016	508270	\pm 230	13900	\pm 130	22556	\pm 57	1335.3	\pm 3.4	568.2	\pm 2.9
$E_T^{\text{miss}} > 25\text{GeV}$	470856	442600	\pm 210	8700	\pm 100	9959	\pm 31	1111.8	\pm 3	424.5	\pm 2.5
$m_T > 50\text{GeV}$	457053	438280	\pm 210	7879	\pm 97	9649	\pm 27	879.7	\pm 2.8	381.7	\pm 2.3

Table 87: Analysis cut flow for $W^+ \rightarrow \mu^+ \nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Cut	Data	Signal		$W^\pm \rightarrow \ell^\pm \nu$ BG		$Z \rightarrow \ell\ell$		Top		Diboson	Multijet
One muon	9570104	2100770	\pm 410	83110	\pm 270	2019400	\pm 2200	71602	\pm 94	3442	\pm 63
Muon trig matched	9382783	1840550	\pm 390	72820	\pm 250	1750400	\pm 2000	61519	\pm 87	2956	\pm 59
Isolation	3905612	1821750	\pm 380	71780	\pm 250	595700	\pm 1100	56849	\pm 84	2916	\pm 59
$p_T^\mu > 25\text{GeV}$	1930655	1393330	\pm 340	34470	\pm 170	170840	\pm 490	49338	\pm 78	2471	\pm 54
$E_T^{\text{miss}} > 25\text{GeV}$	1321407	1173860	\pm 310	21450	\pm 140	51090	\pm 180	41956	\pm 72	1930	\pm 49
$m_T > 50\text{GeV}$	1244892	1153800	\pm 310	18270	\pm 130	38304	\pm 81	32375	\pm 63	1705	\pm 44

Table 88: Analysis cut flow for $W^+ \rightarrow \mu^+ \nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Cut	Data	Signal		$W^\pm \rightarrow \ell^\pm \nu$ BG		$Z \rightarrow \ell\ell$		Top		Diboson	Multijet
One electron	1724472	374900	\pm 200	24150	\pm 160	41995	\pm 70	1590.5	\pm 2.9	684.8	\pm 4
Electron trig matched	1645694	359010	\pm 200	22070	\pm 160	39854	\pm 68	1539.9	\pm 2.9	655.7	\pm 3.9
Isolation	1176976	357660	\pm 200	21920	\pm 160	39686	\pm 68	1504.6	\pm 2.8	640.7	\pm 3.8
$p_T^e > 25\text{GeV}$	529183	302070	\pm 180	11920	\pm 110	30214	\pm 52	1330.8	\pm 2.6	532.9	\pm 3.5
$E_T^{\text{miss}} > 25\text{GeV}$	281957	266750	\pm 170	8084	\pm 90	1293	\pm 20	1112.5	\pm 2.4	380	\pm 3
$m_T > 50\text{GeV}$	274329	264540	\pm 170	7317	\pm 84	994	\pm 16	855.2	\pm 2.1	338.1	\pm 2.9

Table 89: Analysis cut flow for $W^- \rightarrow e^- \nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Cut	Data	Signal		$W^\pm \rightarrow \ell^\pm \nu$ BG		$Z \rightarrow \ell\ell$		Top		Diboson	Multijet
One electron	7471742	1323710	\pm 330	78230	\pm 230	140980	\pm 140	61951	\pm 86	3059	\pm 58
Electron trig matched	7402574	1267710	\pm 330	72240	\pm 230	133580	\pm 140	59950	\pm 85	2968	\pm 57
Isolation	4949352	1260540	\pm 330	71550	\pm 230	132740	\pm 140	58689	\pm 84	2937	\pm 57
$p_T^e > 25\text{GeV}$	2113364	1053510	\pm 300	39660	\pm 160	101350	\pm 110	52923	\pm 79	2544	\pm 53
$E_T^{\text{miss}} > 25\text{GeV}$	1008915	900640	\pm 280	25900	\pm 130	7954	\pm 45	45065	\pm 73	1962	\pm 48
$m_T > 50\text{GeV}$	949362	887810	\pm 270	22400	\pm 120	6052	\pm 35	34177	\pm 64	1695	\pm 44

Table 810: Analysis cut flow for $W^- \rightarrow e^- \nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Cut	Data	Signal		$W^\pm \rightarrow \ell^\pm \nu$ BG		$Z \rightarrow \ell\ell$		Top		Diboson	Multijet
One muon	2075709	440560	\pm 220	22510	\pm 170	34440	\pm 80	1835.6	\pm 3.1	751.5	\pm 3.3
Muon trig matched	2002955	383720	\pm 200	19640	\pm 160	30277	\pm 75	1561.6	\pm 2.9	648	\pm 3.1
Isolation	883078	381010	\pm 200	19450	\pm 160	30046	\pm 74	1411	\pm 2.7	616.9	\pm 2.9
$p_T^\mu > 25\text{GeV}$	426119	314370	\pm 180	9370	\pm 110	20749	\pm 56	1202.1	\pm 2.5	505	\pm 2.5
$E_T^{\text{miss}} > 25\text{GeV}$	298992	276060	\pm 170	5893	\pm 89	8716	\pm 29	1004.2	\pm 2.3	372.6	\pm 2
$m_T > 50\text{GeV}$	287870	273710	\pm 170	5158	\pm 82	8408	\pm 26	788.2	\pm 2	335.6	\pm 1.9

Table 811: Analysis cut flow for $W^- \rightarrow \mu^- \nu$ 5 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Cut	Data	Signal		$W^\pm \rightarrow \ell^\pm \nu$ BG		$Z \rightarrow \ell\ell$		Top		Diboson	Multijet
One muon	8773414	1518070	\pm 360	64930	\pm 230	2019900	\pm 2200	70580	\pm 90	3230	\pm 60
Muon trig matched	8597493	1322980	\pm 330	56520	\pm 210	1750300	\pm 2000	60579	\pm 84	2806	\pm 56
Isolation	3298569	1310310	\pm 330	55680	\pm 210	593700	\pm 1100	55949	\pm 80	2751	\pm 55
$p_T^\mu > 25\text{GeV}$	1561721	1069770	\pm 300	28230	\pm 150	166810	\pm 490	48544	\pm 75	2362	\pm 52
$E_T^{\text{miss}} > 25\text{GeV}$	1030406	910150	\pm 280	17380	\pm 120	47370	\pm 180	41259	\pm 69	1842	\pm 46
$m_T > 50\text{GeV}$	963568	896850	\pm 270	14710	\pm 110	34572	\pm 80	31772	\pm 61	1598	\pm 43
										9050	\pm 620

Table 812: Analysis cut flow for $W^- \rightarrow \mu^- \nu$ 13 TeV signal selection. The minimum lepton p_T required before the final p_T cut is 18 GeV.

Selection	Observed	Expected		
5TeV $W^+ \rightarrow e^+ \nu$	430620	431620	\pm 600	
5TeV $W^+ \rightarrow \mu^+ \nu$	457053	457790	\pm 300	
5TeV $W^- \rightarrow e^- \nu$	274329	276450	\pm 530	
5TeV $W^- \rightarrow \mu^- \nu$	287870	289160	\pm 250	
13TeV $W^+ \rightarrow e^+ \nu$	1207776	1213000	\pm 1800	
13TeV $W^+ \rightarrow \mu^+ \nu$	1244892	1253490	\pm 870	
13TeV $W^- \rightarrow e^- \nu$	949362	979500	\pm 2000	
13TeV $W^- \rightarrow \mu^- \nu$	963568	988560	\pm 690	

Table 813: Observed and Expected yield comparison for all signal selections.

1592 **8.3.2 $\sqrt{s} = 13 \text{ TeV}$ dataset control plots**

1593 Control plots for the 13 TeV low- μ dataset are provided here after applying all corrections described
1594 in section ??, and after applying the selection described above in this section. In each figure, the
1595 right(left)-hand column shows distributions for the W^+ (W^-) process. The top (bottom) row shows the
1596 muon (electron) decay channel. In the ratio panels, the grey band is the total systematic uncertainty,
1597 whilst the brown band adds the MC statistical uncertainty in quadrature on top of it. In regions
1598 of the distributions insensitive to the modelling of p_T^W there is generally good agreement between
1599 data and predictions. The bulk of the m_T distribution is a typical example of distribution that is
1600 mostly insensitive to the modeling of p_T^W . The u_T distribution is an exception, and it can therefore be
1601 concluded that the baseline simulation is not modeling p_T^W satisfactorily.

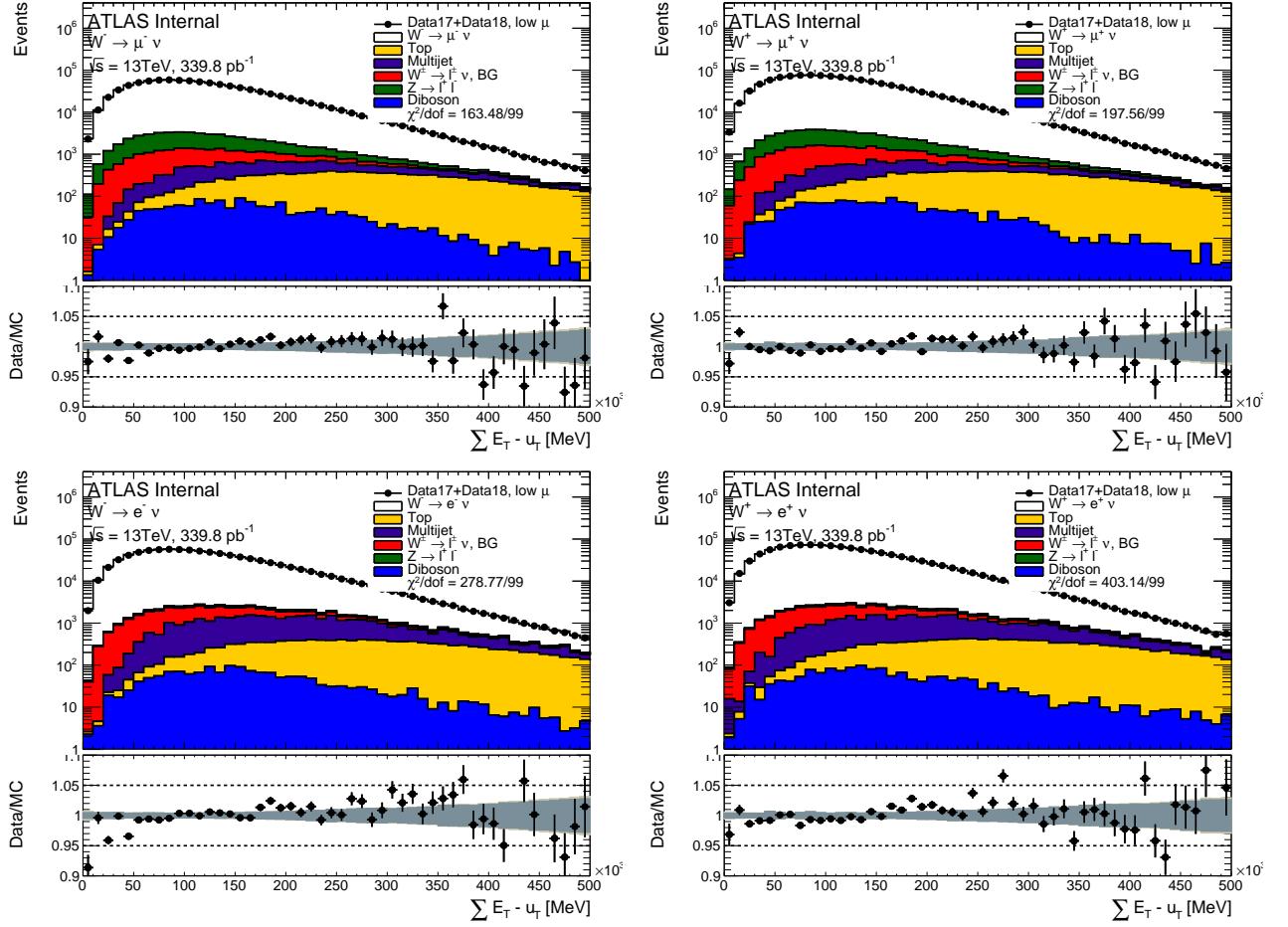


Figure 84: ΣE_T distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.

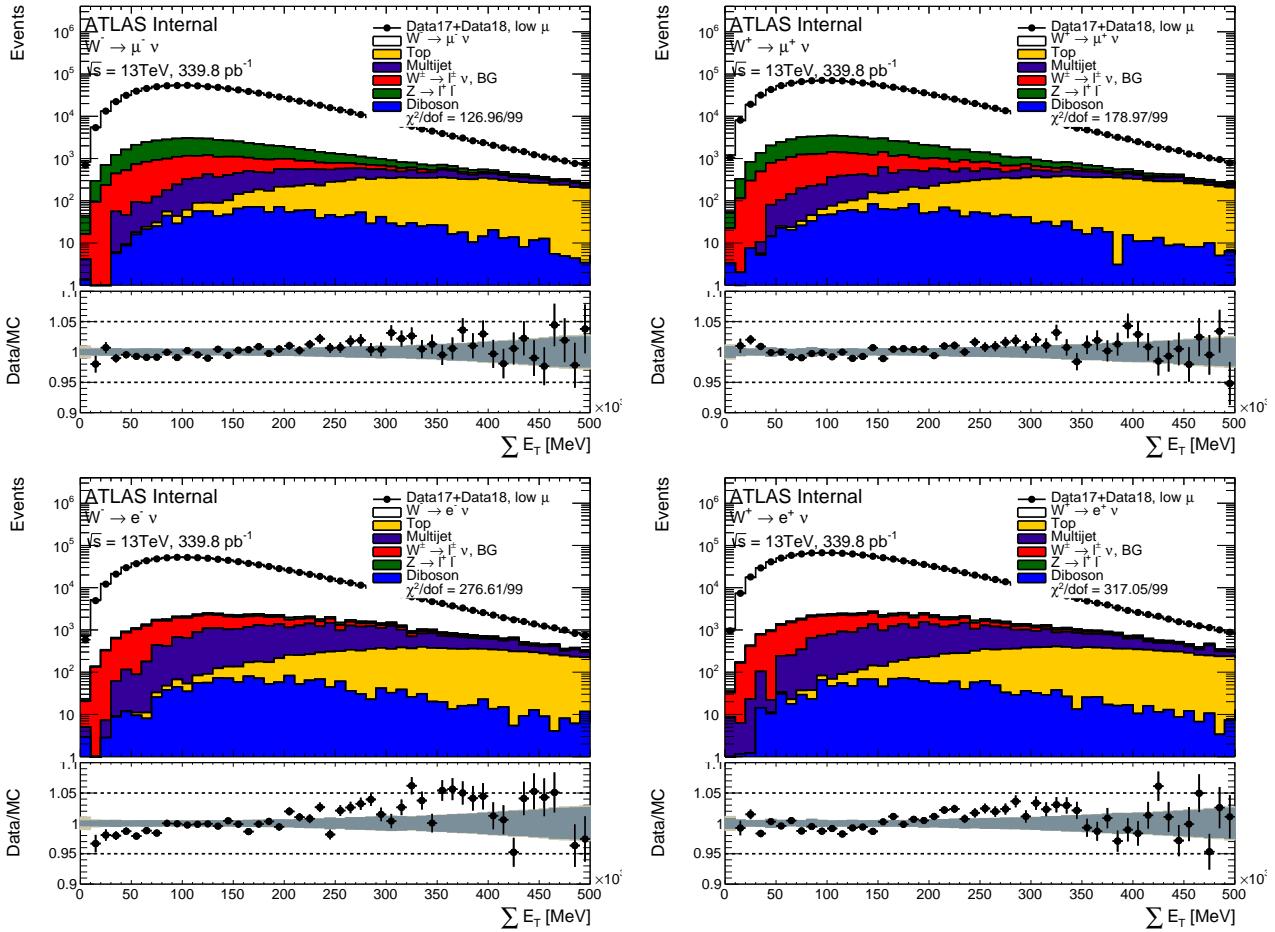


Figure 85: ΣE_T distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.

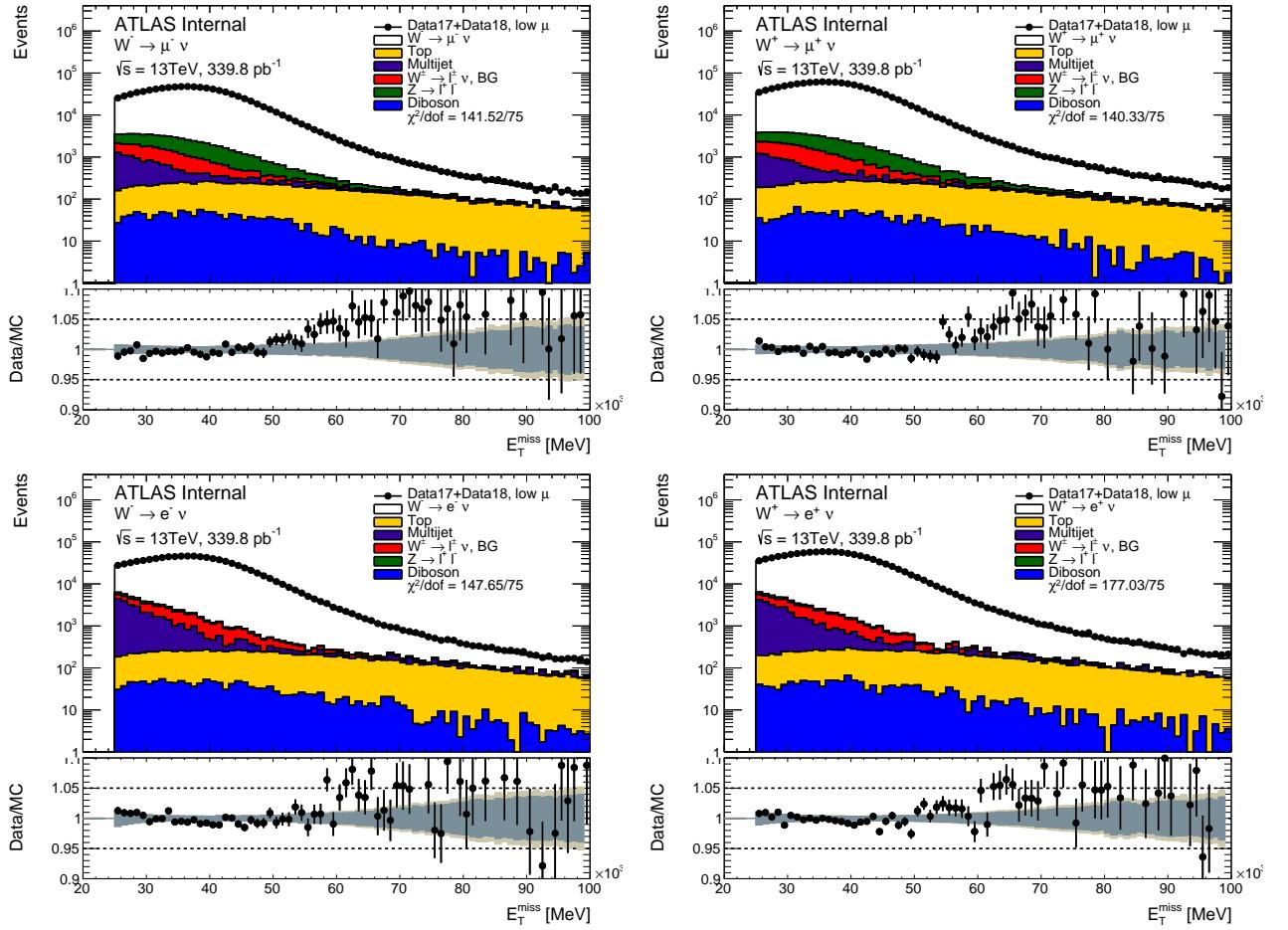


Figure 86: \vec{E}_T^{miss} distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.

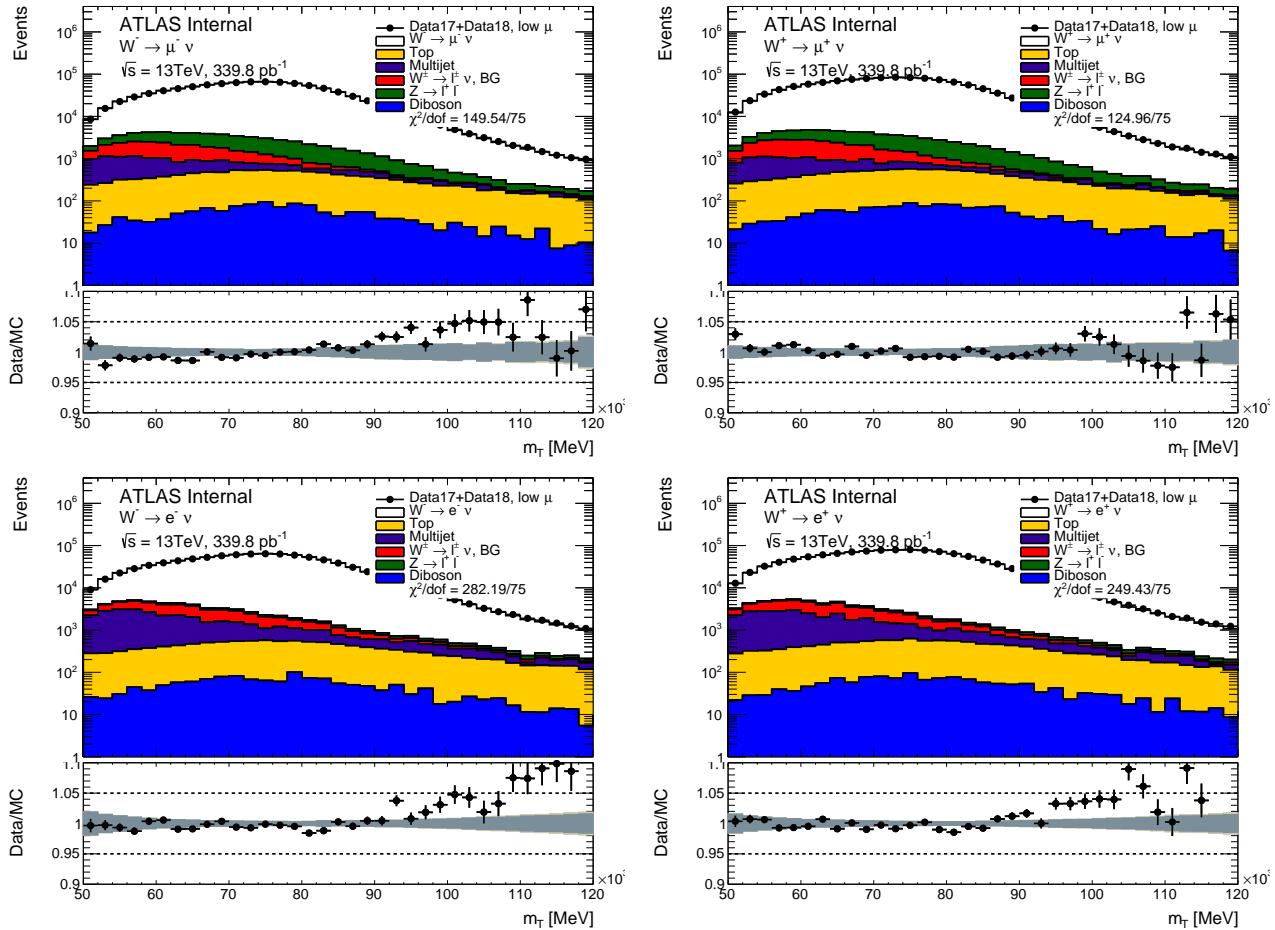


Figure 87: Transverse mass distribution of the W boson in the muon and electron channel for the $\sqrt{s} = 13\text{TeV}$ dataset.

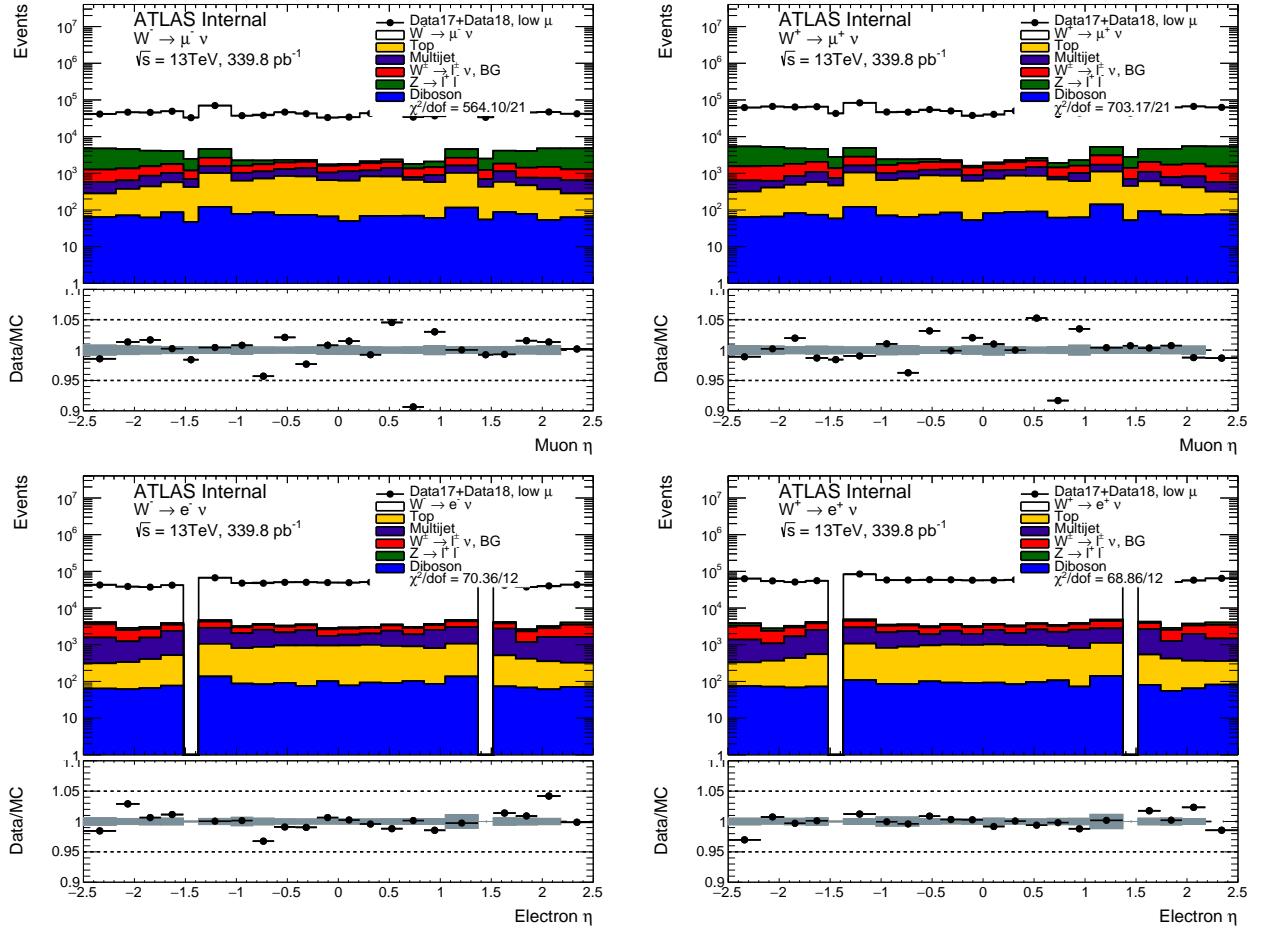


Figure 88: Lepton pseudorapidity distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.

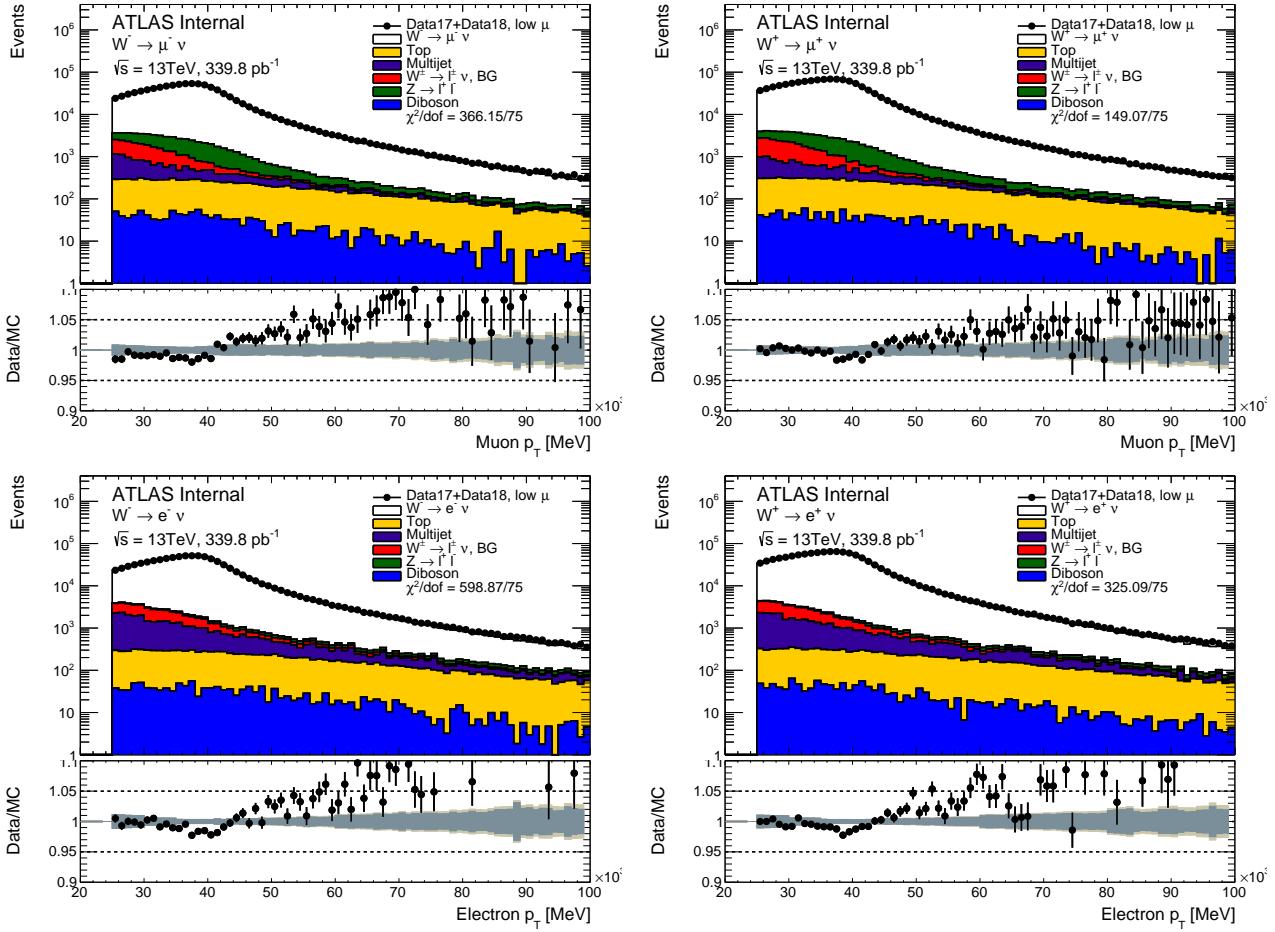


Figure 89: Lepton transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 13 \text{ TeV}$ dataset.

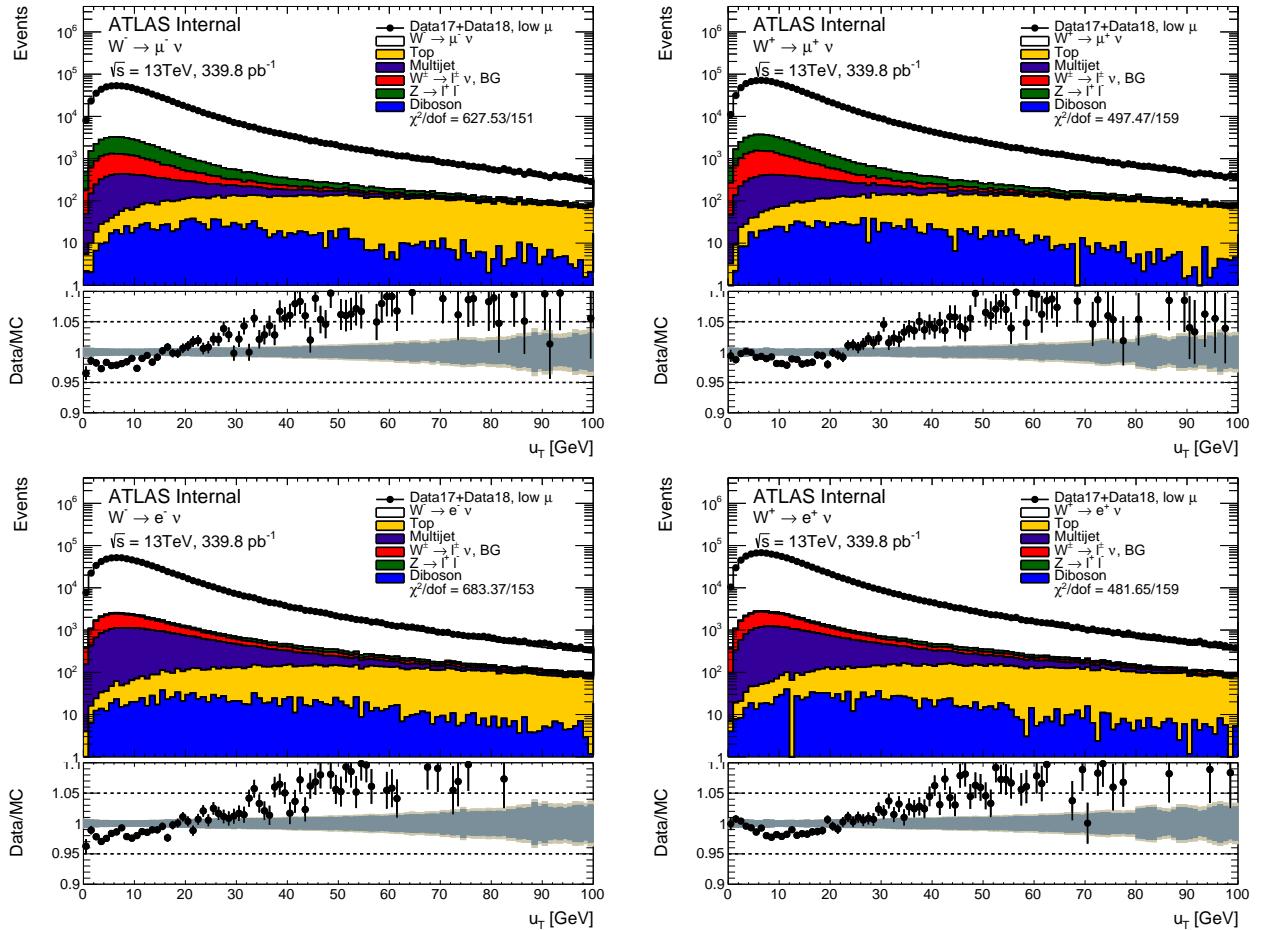


Figure 810: W transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 13\text{TeV}$ dataset.

1602 8.3.3 $\sqrt{s} = 5$ TeV dataset control plots

1603 Control plots for the 5 TeV low- μ dataset are provided here after applying all corrections described
 1604 in section ??, and after applying the selection described above in this section. In each figure, the
 1605 right(left)-hand column shows distributions for the W^+ (W^-) process. The top (bottom) row shows the
 1606 muon (electron) decay channel. In the ratio panels, the grey band is the total systematic uncertainty,
 1607 whilst the brown band adds the MC statistical uncertainty in quadrature on top of it. In regions of
 1608 the distributions insensitive to the modelling of p_T^W there is generally good agreement between data
 1609 and predictions. The bulk of the m_T distribution is a typical example of distribution that is mostly
 1610 insensitive to the modeling of p_T^W . Compared to the 13 TeV situation, the u_T distribution seems to
 1611 indicate that the baseline simulation models p_T^W more satisfactorily.

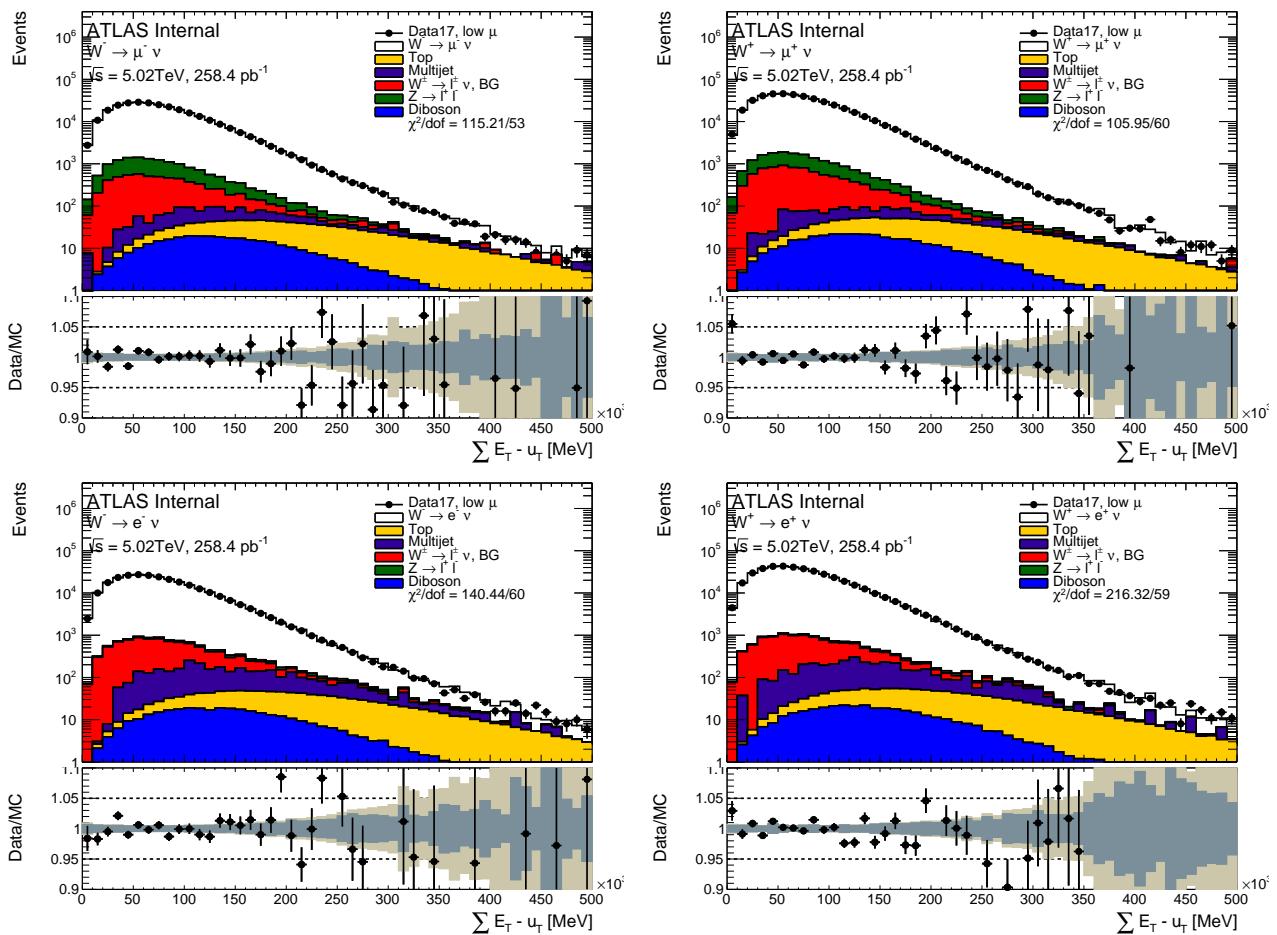


Figure 811: ΣE_T distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

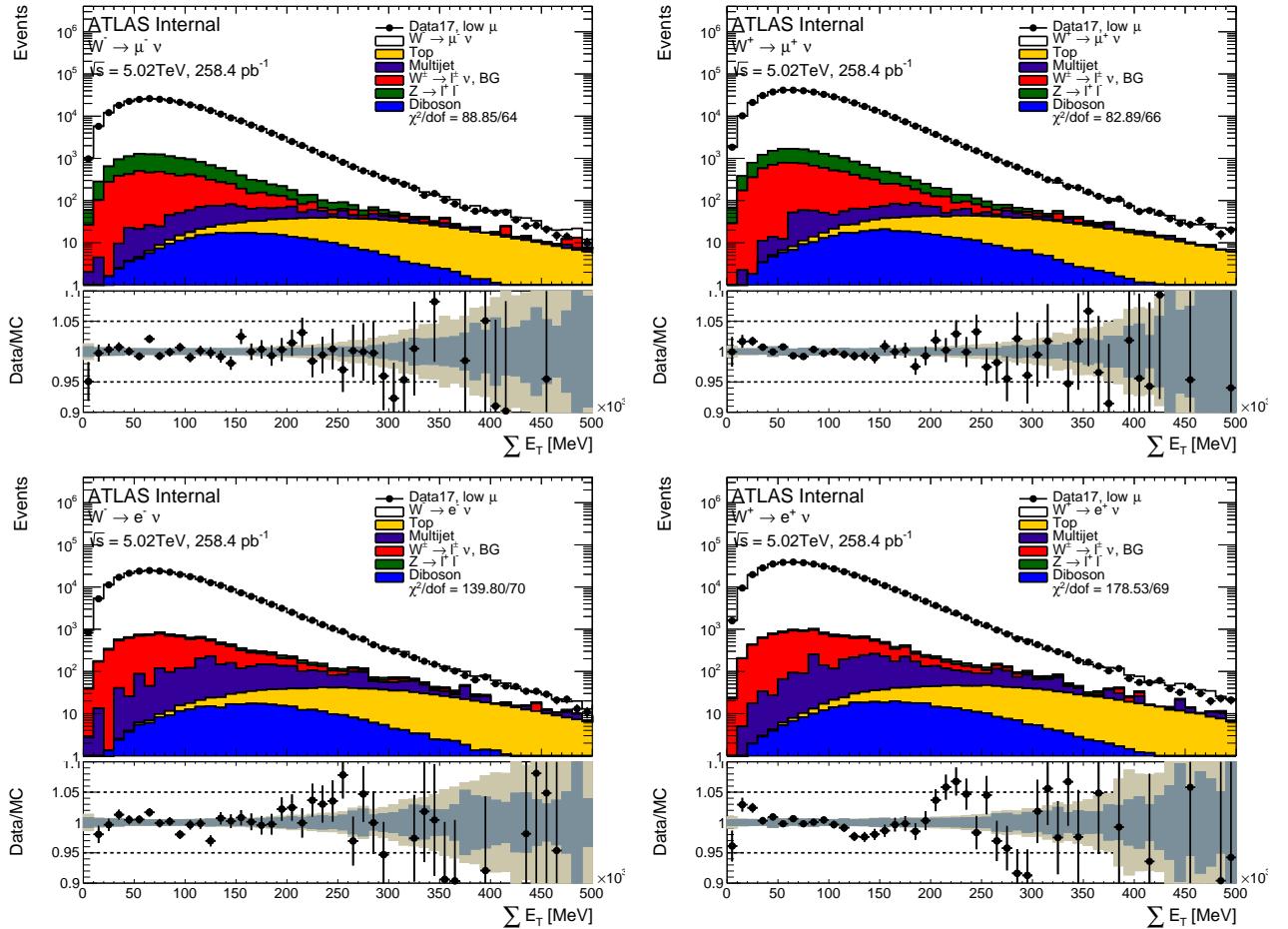


Figure 812: ΣE_T distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

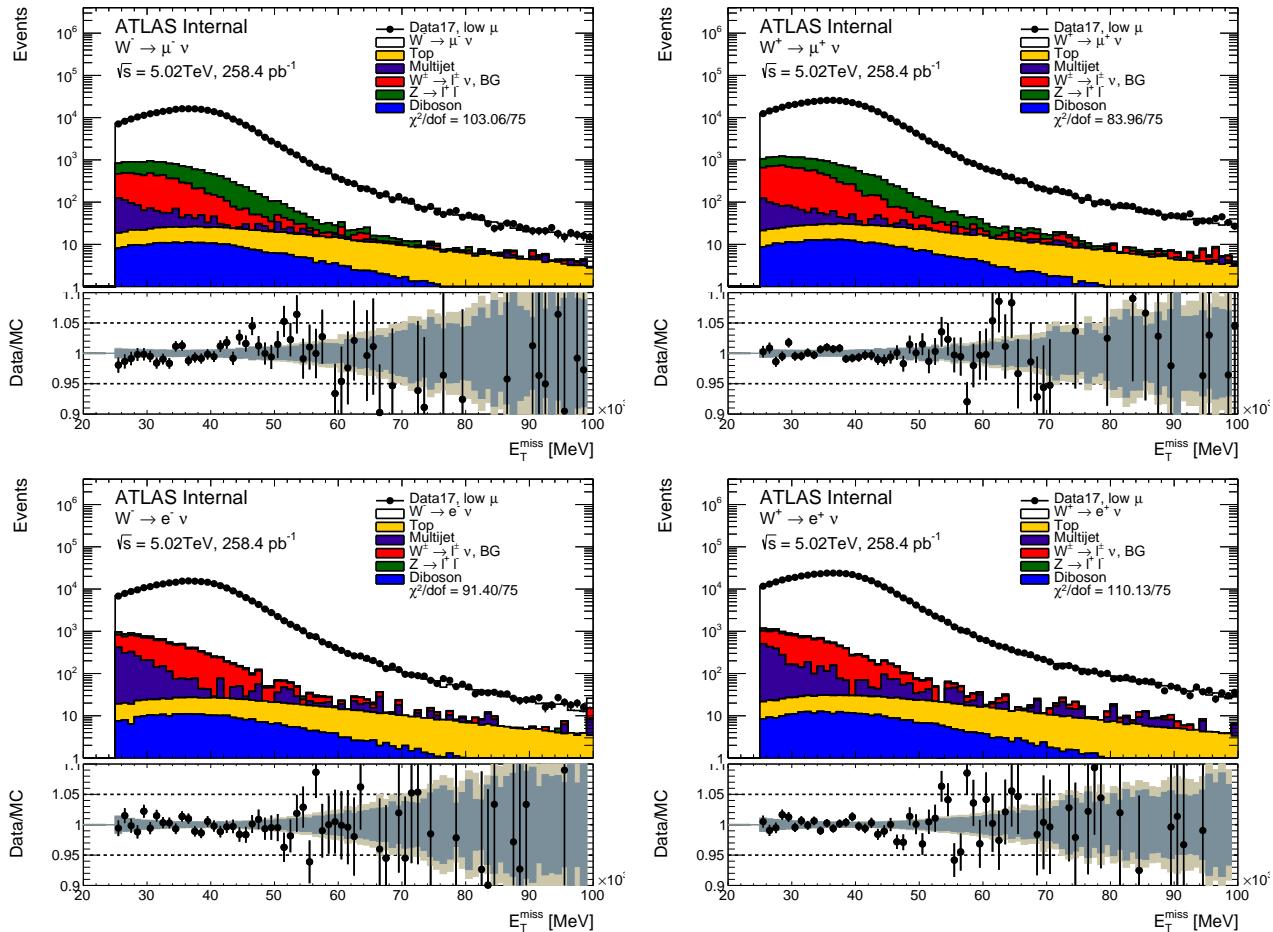


Figure 813: \vec{E}_T^{miss} distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

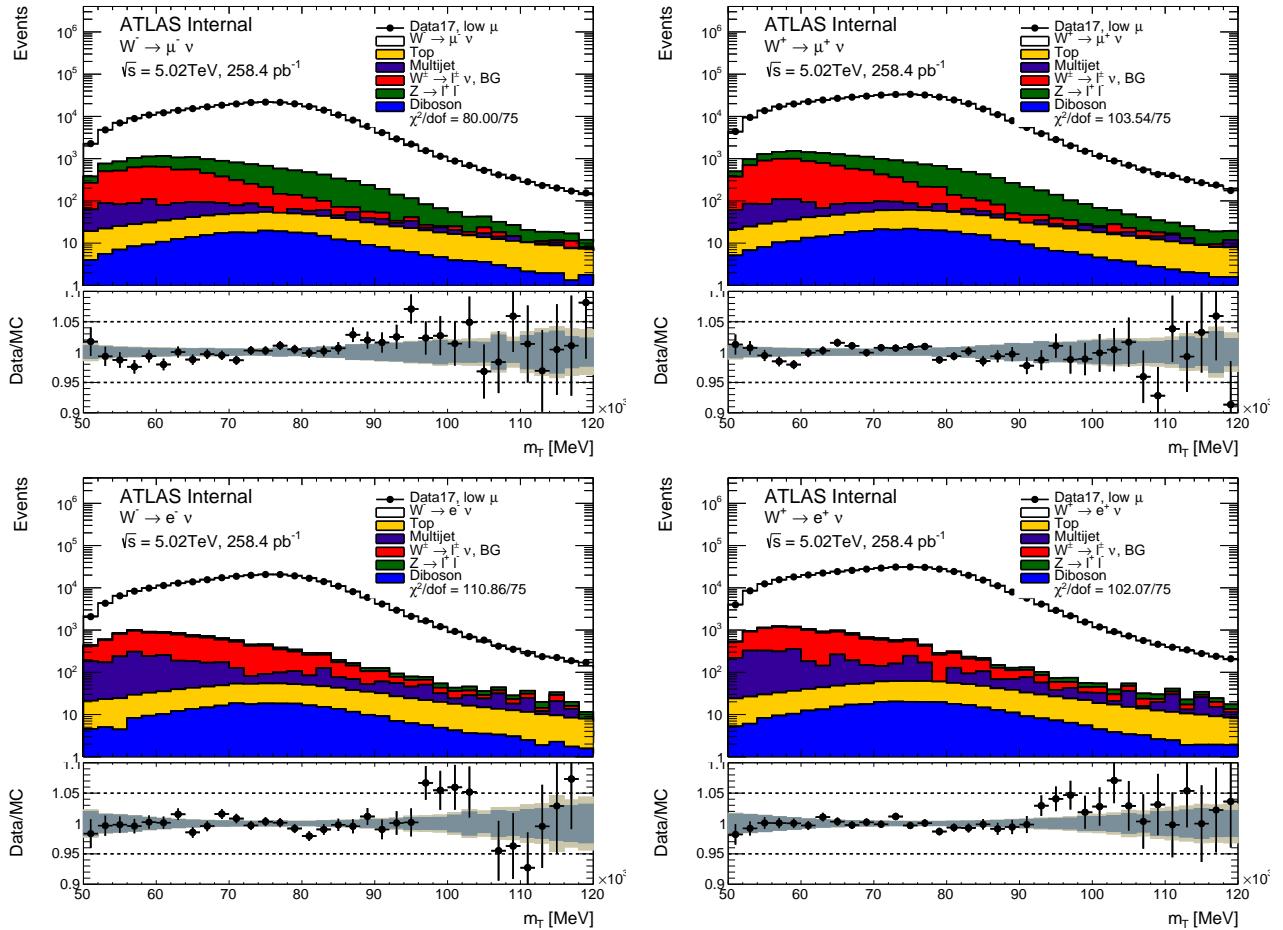


Figure 814: Transverse mass distribution of the W boson in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

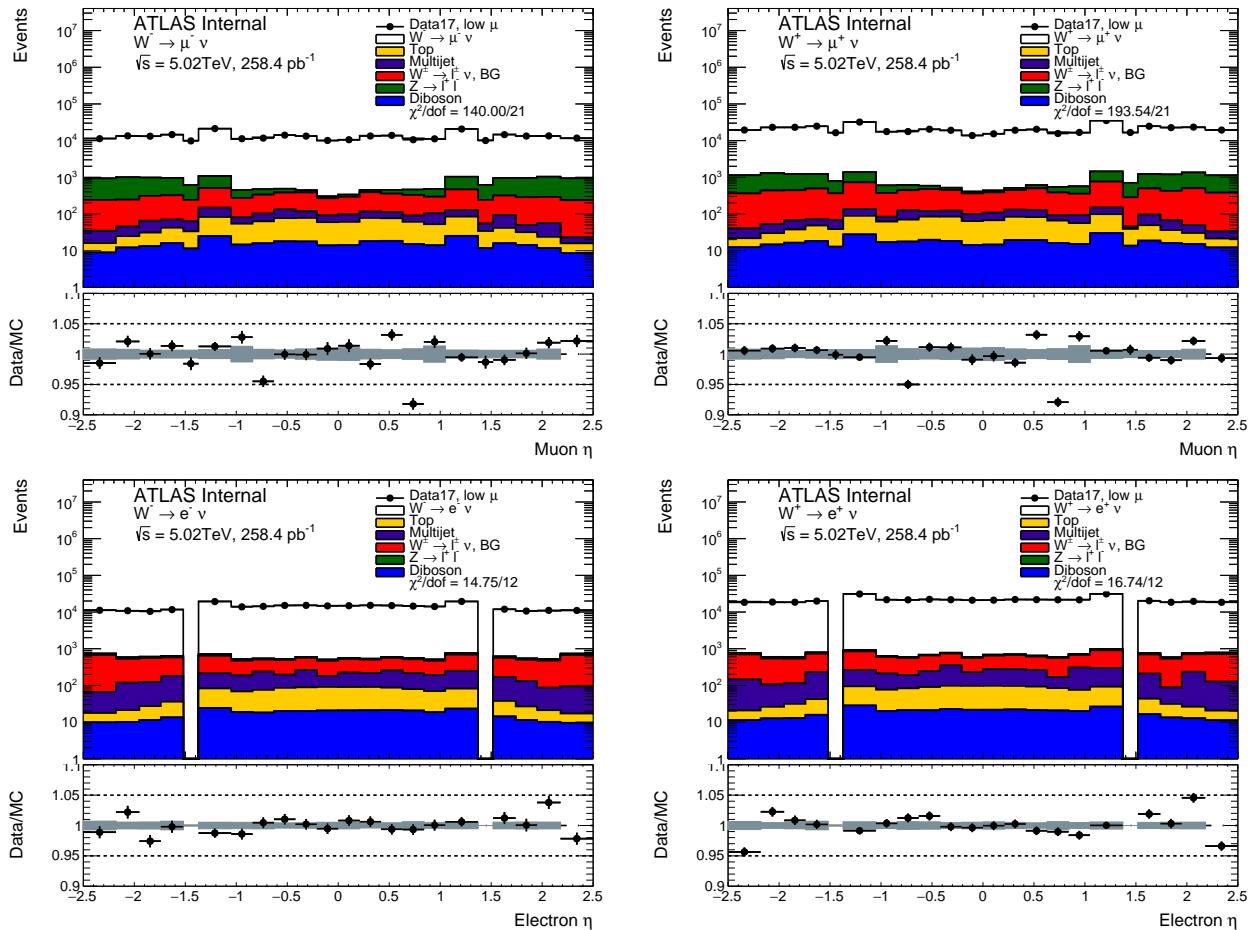


Figure 815: Lepton pseudorapidity distribution in the muon and electron channel for the $\sqrt{s} = 5\text{ TeV}$ dataset.

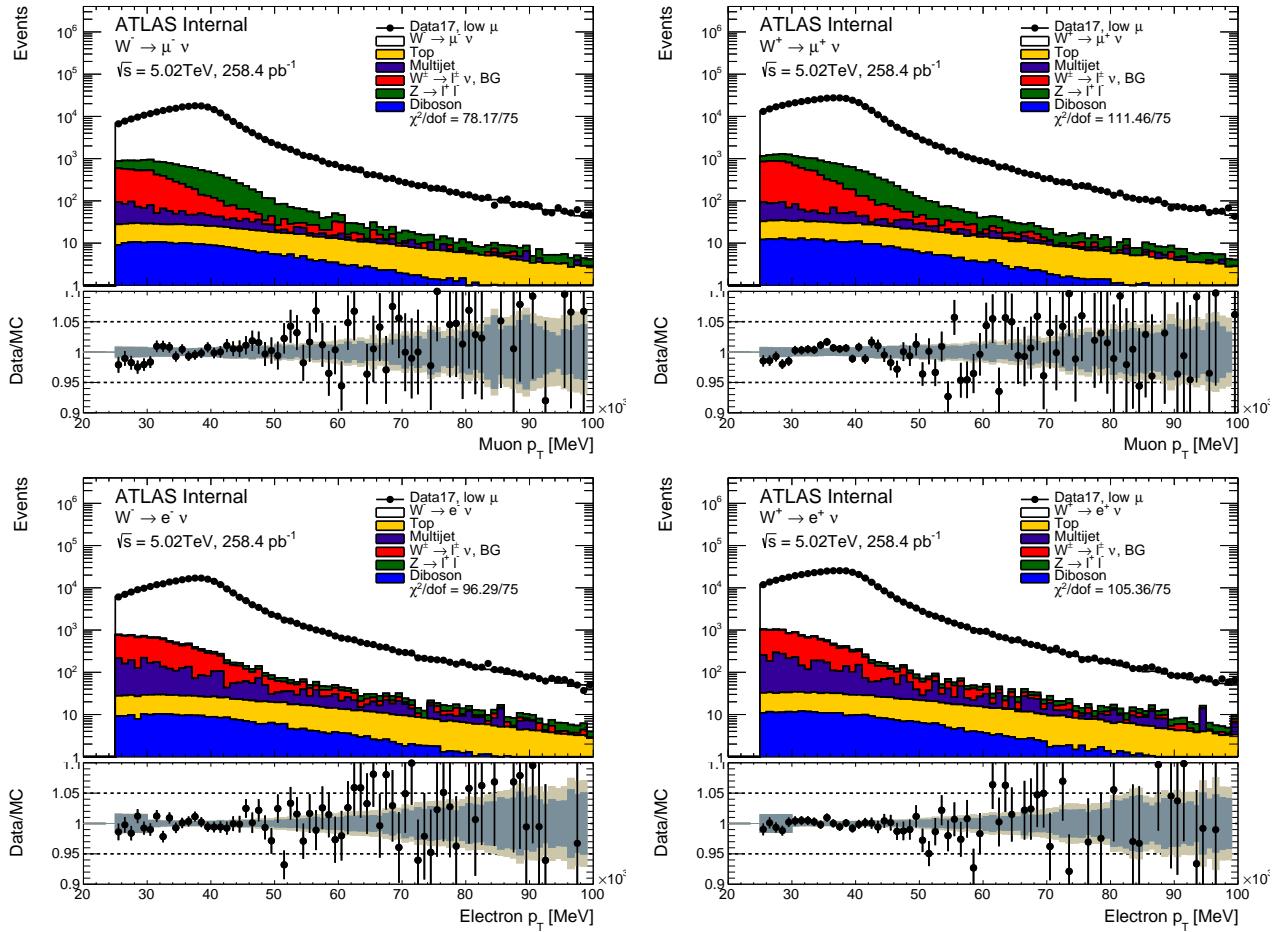


Figure 816: Lepton transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

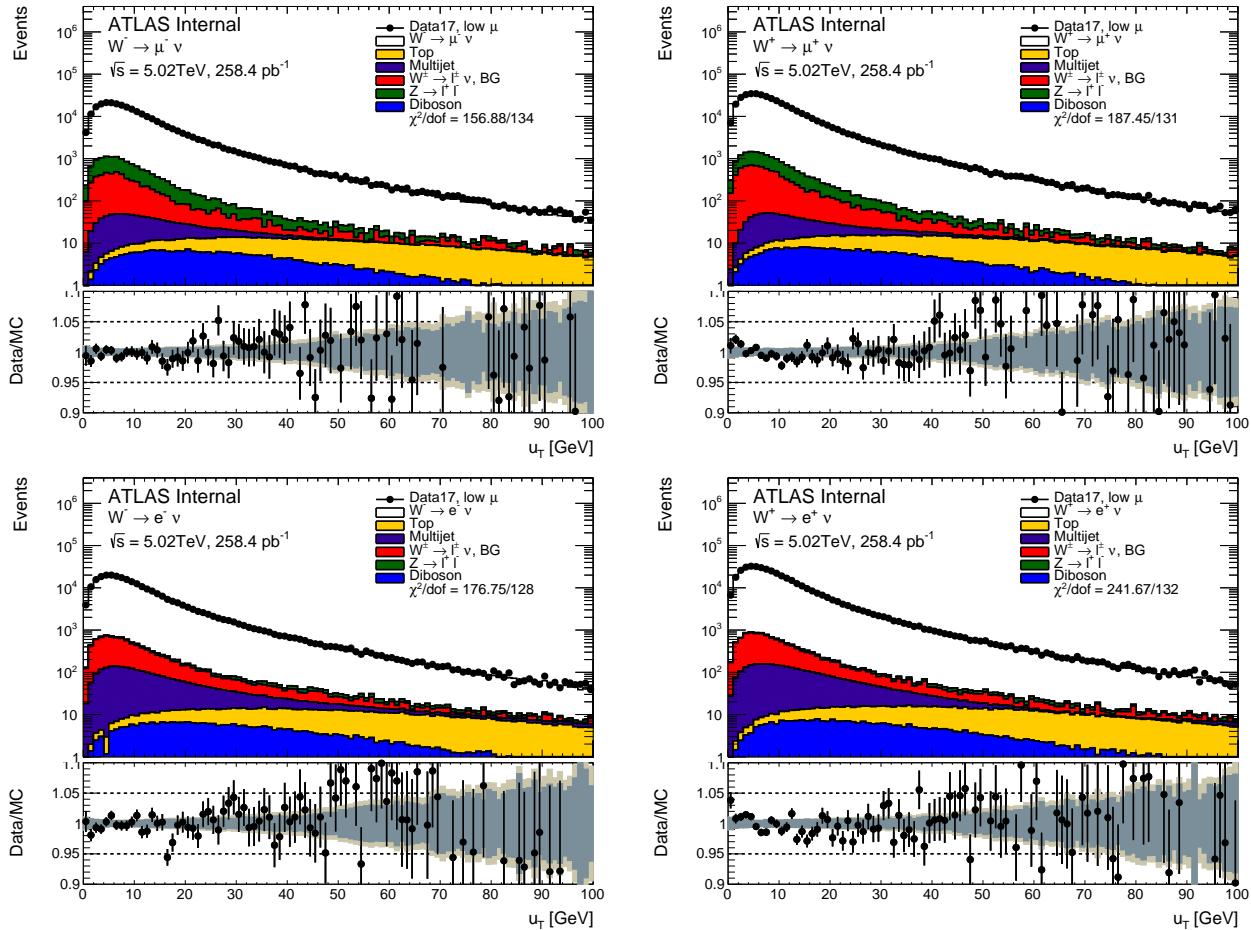


Figure 817: W transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 5\text{ TeV}$ dataset.

9

W boson pT spectrum

1612

1613

1614 "Potentielle citation sans aucun rapport avec le sujet"

1615 — Personne inconnue, contexte à déterminer

1616 In the recent years a significant progress was achieved in the field of big datasets analysis. There is a
1617 number of principles available for solving a wide variety of tasks. In this thesis

1618 9.1 Uncertainties propagation

1619 9.2 Uncertainties breakdown plots

1620 9.2.1 Unfolding

10 Hadronic recoil regression with deep neural networks

1621

1622

Hadronic recoil regression with deep neural networks

1623 In the recent years a significant progress was achieved in the field of big datasets analysis. There is a
1624 number of principles available for solving a wide variety of tasks. In this thesis a DNN was used for
1625 the regression of the 2-component hadronic recoil vector.

1626 10.1 Deep neural networks

1627 Normally a machine learning problem has a number of ingredients: a dataset \mathbf{X} , a set of parameters
1628 θ , a model $g(\theta)$ and a loss function $C(\mathbf{X})$ that tells us how well the model $g(\theta)$ describes the dataset.
1629 Finding the values of θ that would minimize the loss function we fit the model.

1630 10.1.1 Gradient descent optimization

1631 One of the most powerful and used class of methods in minimizing the loss function is called the
1632 *gradient descent*, [1] especially its sub-class, the stochastic gradient descent (SGD) [2], [3]. One of its
1633 modifications called ADAM [4] was used as an optimization algorithm in the work presented in this
1634 thesis.

1635 Let's assume that a loss function $E(\theta)$ may be estimated as a sum over n data points:

$$E(\theta) = \sum_{i=1}^n e_i(x_i, \theta). \quad (10.1)$$

1636 In the simplest case of the gradient descent (GD) algorithm we start looking for the values of parameters
1637 θ such that the sum of functions $\sum_{i=1}^n e_i$ is minimal. We start with a certain value θ_0 and then iteratively
1638 perform the following:

$$\begin{aligned} v_t &= \eta_t \nabla_{\theta} E(\theta_t), \\ \theta_{t+1} &= \theta_t - v_t, \end{aligned} \quad (10.2)$$

1639 where $\nabla_{\theta} E(\theta_t)$ is the gradient of $E(\theta)$ with respect to θ ; factor η_t is called the *learning rate* and defines
1640 the length of the step in the direction of θ performed with every iteration. Balancing learning rate
1641 is very important for learning process and convergence. A value too low can make the convergence
1642 "stuck" in the local minimum, it also increases the number of iterations. Picking a very high learning

1643 rate we risk to miss the minimum so the algorithm would never converge to a minimum. Also, if the
 1644 number of data points n is high, calculating the gradient is a costly task in terms of CPU time.
 1645 Some of the problems accompanying the use of GD are dealt with by using its modification - the SGD.
 1646 The idea is the following: instead of using all the available data points n at each iteration of the GD, we
 1647 split the data into k *minibatches*, each having M data points, such that $k = n/M$. Normally the size of
 1648 the batch is few hundreds of data points, to provide a certain degree of variance and incorporating
 1649 stochasticity. The transition to SGD algorithm is done in the following way:

$$\nabla_{\theta} E(\theta) = \sum_{i=1}^n \nabla_{\theta} e_i(x_i, \theta) \rightarrow \sum_{i \in B_l} \nabla_{\theta} e_i(x_i, \theta), \quad (10.3)$$

1650 where B_l is a set of data points belonging to a minibatch $l \in 1, \dots, n/M$. Now every next iteration of θ
 1651 parameters update is performed over a different batch, consecutively running over all the batches:

$$\begin{aligned} \nabla_{\theta} E^{EM}(\theta) &= \sum_{i \in B_l} \nabla_{\theta} e_i(x_i, \theta), \\ v_t &= \eta_t \nabla_{\theta} E^{EM}(\theta_t), \\ \theta_{t+1} &= \theta_t - v_t. \end{aligned} \quad (10.4)$$

1652 A full iteration over all the n/M batches is called an *epoch*. Now stochasticity prevents the gradient
 1653 algorithm from getting stuck in a local minimum. Also computing the gradient over fewer data point
 1654 notably decreases the CPU time spent.
 1655 The algorithm may be further improved, adding a "memory", that is to say making every next step t
 1656 dependent on the direction of the previous step $t-1$:

$$\begin{aligned} v_t &= \gamma v_{t-1} \eta_t \nabla_{\theta} E^{EM}(\theta_t), \\ \theta_{t+1} &= \theta_t - v_t. \end{aligned} \quad (10.5)$$

1657 Thanks to analogy from physics the parameter γ is called a *momentum*, having $0 \leq \gamma \leq 1$ [5], [6].
 1658 This parameter provides a certain "inertia" in the change of the direction of the gradient descent.
 1659 Introduction of the momentum helps for quicker convergence in the case of a slow but steady change
 1660 of a certain parameter during the gradient descent.
 1661 The convergence of the GD may be significantly improved if the learning rate could be different in
 1662 different directions, depending on the landscape of the parameter space θ : the steeper the gradient
 1663 in a certain direction - the smaller the corresponding step. The optimal step could be estimated by
 1664 obtaining the *Hessian matrix* in the vicinity of a point θ_0 , providing a description of the local curvature
 1665 in a multidimensional space. Although calculating Hessian matrix is complicated and slow-converging
 1666 process [7]. However, a number of methods use the second moment of the gradient to efficiently
 1667 estimate the optimal learning rate. One of such methods is called ADAM (ADAptive Momentum) [4],

1668 its iterative relations are the following:

$$\begin{aligned}
 g_t &= \nabla_{\theta} E(\theta_t) \\
 m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\
 s_t &= \beta_2 s_{t-1} + (1 - \beta_2) g_t^2 \\
 \hat{m}_t &= \frac{m_t}{1 - (\beta_1)^t} \\
 \hat{s}_t &= \frac{s_t}{1 - (\beta_2)^t} \\
 \theta_{t+1} &= \theta_t - \eta_t \frac{\hat{m}_t}{\sqrt{\hat{s}_t} + \epsilon}.
 \end{aligned} \tag{10.6}$$

1669 Here the parameters β_1 and β_2 set the memory lifetime for the first and second moment; η is the learning
 1670 rate and ϵ is a small regularization constant keeping the denominators from vanishing. Like in other
 1671 cases of the SGD here the iterations are performed batch-wise. Parameter s_t is linked to the variance of
 1672 the gradient size. This basically means that the learning rate is proportional to the first momentum of
 1673 the gradient and inverse proportional to its standard deviation.

1674 10.1.2 DNN structure and training

1675 A neural network is composed of single neurons, also called nodes, arranged in layers. The first layer is
 1676 called the input layer, the last one is called the output layer; all the layers in between are named hidden
 1677 layers (see Fig. 101).

1678 A single node i takes a vector of k input features $\mathbf{x} = (x_1, x_2, \dots, x_k)$ and produces a scalar input $a_i(\mathbf{x})$.
 1679 Function a_i may have a different form, although it normally can be decomposed into two steps. The
 1680 first step is a linear transformation of the inputs into a scalar value assigning each input a weight:

$$z^i = w_k^i \cdot x_k + b^i, \tag{10.7}$$

1681 where $\mathbf{w}^i = (w_1^i, w_2^i, \dots, w_k^i)$ is a set of k weights assigned to corresponding inputs. The weights \mathbf{w}^i are
 1682 specific to a neuron i , as well as the scalar bias b^i . The next step is where the non-linear function σ_i
 1683 comes into play: we can express the output function $a_i(\mathbf{x})$ as follows:

$$a_i(\mathbf{x}) = \sigma_i(z^i). \tag{10.8}$$

1684 There exists a number of options for the non-linear function σ ; in current thesis a tanh is used. When
 1685 the neurons are arranged in layers in a feed-forward neural network - the outputs from neurons of
 1686 the previous layer serve as inputs for the succeeding layers neurons. The universal approximation
 1687 theorem states, that a neural network with a single hidden layer can approximate any continuous
 1688 multiparametric function with arbitrary accuracy [8], [9]. However, in practice it is easier to reach the
 1689 possible precision having more hidden layers.

1690 So in terms of a DNN fitting the model means tuning the weights and biases (\mathbf{w}^i, b^i) in such a way
 1691 that a loss function applied to the new dataset would be minimal. It is reached through iterative
 1692 process called *training*, that involves the GD with an algorithm called *backpropagation* [10]. The
 1693 backpropagation algorithm allows to calculate the gradients and adjust the corresponding parameters

1694 in a very computation-efficient way.

Let us assume that there are L layers in the network $l = 1, \dots, L$, that w_{jk}^l and b_j^l are the weight of an

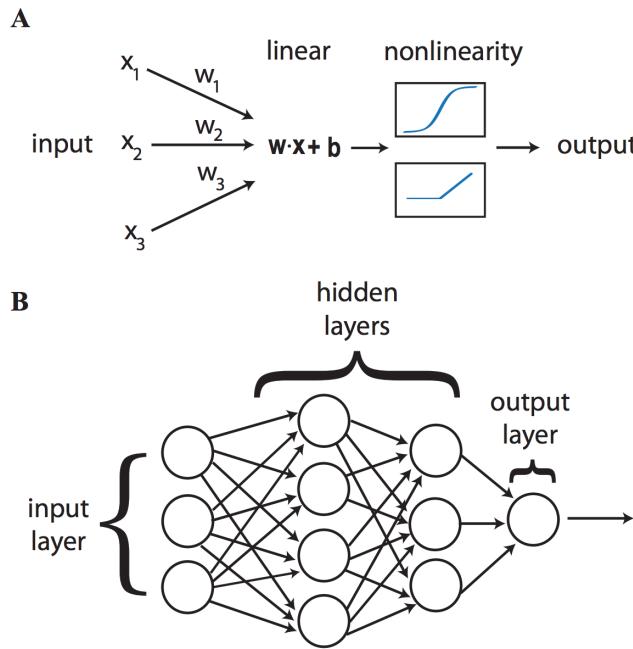


Figure 101: A: The nodes perform a linear transformation of the inputs, then apply a non-linear activation function. B: The architecture of a deep neural network: neurons are arranged into layers [11].

1695

1696 input parameter k and the bias for node k in layer l respectively. The layered structure of the neural
1697 network ensures that the inputs for the nodes in layer l depend only on the outputs of the nodes from
1698 layer $l - 1$, hence:

$$a_j^l = \sigma \left(\sum_k w_{jk}^l a_k^{l-1} + b_j^l \right) = \sigma(z_j^l), \quad (10.9)$$

1699 where the linear weighted sum is denoted as:

$$\sigma(z_j^l) = \sum_k w_{jk}^l a_k^{l-1} + b_j^l. \quad (10.10)$$

1700 The cost function E is computed from the output of the neural network, so it directly depends only on
1701 the values of a_j^L . Let us define the error Δ_j^L of the j -th node in the output (L -th) layer as a change in the
1702 cost function with respect to the weighted output of the last layer:

$$\Delta_j^L = \frac{\partial E}{\partial z_j^L}. \quad (10.11)$$

1703 At the same time the loss depends indirectly on all the preceding layers, so keeping in mind eq. 10.9
1704 we can define the error of an arbitrary node j in arbitrary layer l as the change in the cost function E

1705 with respect to the weighted input z_j^l :

$$\Delta_j^l = \frac{\partial E}{\partial z_j^l} = \frac{\partial E}{\partial a_j^l} \sigma'(z_j^l), \quad (10.12)$$

1706 where $\sigma'(z_j^l)$ is the derivative of the non-linear activation function σ with respect to its input at z_j^l . But
1707 on the other hand we can also interpret the error function Δ_j^L in terms of bias partial derivatives:

$$\Delta_j^l = \frac{\partial E}{\partial z_j^l} = \frac{\partial E}{\partial b_j^l} \frac{\partial b_j^l}{\partial z_j^l} = \frac{\partial E}{\partial b_j^l} \cdot \mathbf{1}. \quad (10.13)$$

1708 So starting from the output layer we can compute the error in any layer l , provided we know it for the
1709 subsequent layer $l + 1$:

$$\begin{aligned} \Delta_j^l &= \frac{\partial E}{\partial z_j^l} = \sum_k \frac{\partial E}{\partial z_j^{l+1}} \frac{\partial z_j^{l+1}}{\partial z_j^l} = \\ &= \sum_k \Delta_j^{l+1} \frac{\partial z_j^{l+1}}{\partial z_j^l} \left(\sum_k \Delta_j^{l+1} w_{kj}^{l+1} \right) \sigma'(z_j^l). \end{aligned} \quad (10.14)$$

1710 And finally we can get the gradient of the cost function E with respect to a weight of an arbitrary
1711 neuron:

$$\frac{\partial E}{\partial w_{jk}^l} = \frac{\partial E}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{jk}^l} = \Delta_j^l a_k^{l-1}. \quad (10.15)$$

1712 Using these four equations (10.11, 10.13, 10.14, 10.15) it is possible to "backpropagate" the error back
1713 from the output layer and once we can compute the gradient - we know how we should tune the
1714 weights and biases in order to minimize the loss function.

1715 10.1.3 Batch normalization

1716 Batch normalization is a regularization scheme that helps to improve the speed and stability of the
1717 DNN training. The main idea behind the method is to prevent an *internal covariant shift* - a change in
1718 the distribution of network activations due to the change in network parameters during training by
1719 means of normalization of the parameters transferred from layer l to layer $l + 1$ [12]. So let us consider
1720 a layer l that has d inputs $\mathbf{x} = (x^1, x^2, \dots, x^d)$, then for every x^k we perform the following transformation:

$$\hat{x}^k = \frac{x^k - E[x^k]}{\sqrt{Var[x^k]}}, \quad (10.16)$$

1721 where $E[x^k]$ and $Var[x^k]$ are the expectation and variance of the parameter x , calculated over the
1722 training dataset, respectively. Although we have to be sure that we preserve the non-linearity of the
1723 activation function output. In order to do this the two additional parameters are introduced:

$$y^k = \gamma \hat{x}^k + \beta^k, \quad (10.17)$$

1724 where the parameters γ and β are trained just like the rest of the network parameters. Practically
1725 if the training is performed within the mini-batch scheme with batch size $B = x_1, \dots, x_m$ the batch

1726 normalization layer is inserted between the DNN layers the transformations for the input x are the
1727 following:

$$\begin{aligned} \frac{1}{m} \sum_{i=1}^m x_i &\rightarrow \mu_B \\ \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 &\rightarrow \sigma_B^2 \\ \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} &\rightarrow \hat{x}_i \\ \gamma \hat{x}_i + \beta &\rightarrow y_i \equiv BN_{\gamma, \beta}(x_i), \end{aligned} \quad (10.18)$$

1728 where ϵ is a small regularization constant.

1729 10.2 HR regression

1730 Considering that hadronic recoil is an observable what uses many inputs from ID, EMC and HC it is
1731 reasonable to expect improvement of the result using modern MultiVariate Analysis (MVA) techniques.

1732 10.2.1 Input features and model

1733 Training, testing and validation was performed using the MC sample $W^+ \rightarrow \mu\nu$ at 13 TeV with the same
1734 selection. From the 3625136 events that have passed the selection 2 275 902 were used for training and
1735 1 349 234 for testing the performance. Below is the list of 38 input features:

- 1736 • **Hadronic recoil** is possible in a number of definitions. As it was described before, the HR may be
1737 defined using exclusively charged Particle Flow Objects (PFO), exclusively neutral PFO or both.
1738 All three definitions are included to the input features in the with two Cartesian components for
1739 each definition, making 6 input features.
- 1740 • **Transverse energy sum** $\sum E_T$ is also defined in three similar ways, adding three input features.
- 1741 • Cartesian components of the two leading jets momenta in the transverse plane. The jets were
1742 demanded to have $p_T > 20$ GeV. If one or both jets don't make the cut or there is less than two jets
1743 in the event - the corresponding features were assigned zero value.
- 1744 • Cartesian components of the five leading Neutral Particle Flow Objects (nPFOs) and five leading
1745 Charged Particle Flow Objects (cPFOs) momenta in the transverse plane.
- 1746 • Number of primary vertices in the event.
- 1747 • Pile-up value μ .
- 1748 • Total number of jets in the event.
- 1749 • Total number of nPFOs and cPFOs in the event.

1750 All input features were preprocessed using the StandardScaler module from Scikit Learn package [13].
 1751 The model contains 3 dense layers with 256 neurons each, alternated with batch normalization layers
 1752 (see Fig. 102). Using batch normalization layers has allowed to reduce the training time by the factor of 10. The model has used Adam optimizator with learning step 0.001 and batch size of 4000 data points.

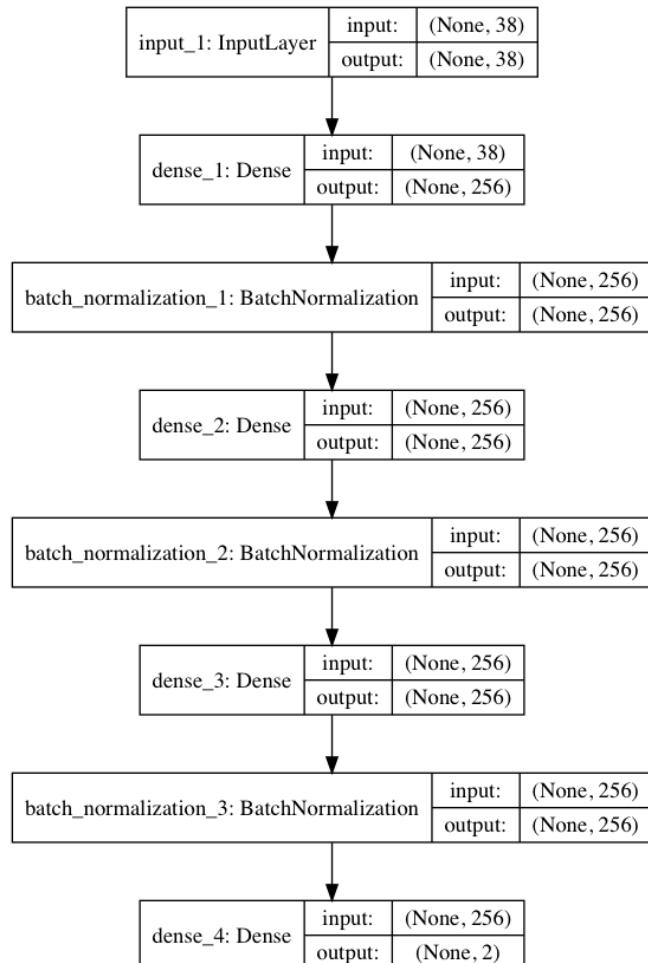


Figure 102: A model of the DNN used in the analysis.

1753
 1754 Twenty percent of events were used for validation. The two target values were Cartesian components
 1755 of the truth HR vector.

1756 10.2.2 Results

1757 10.3 Technical details

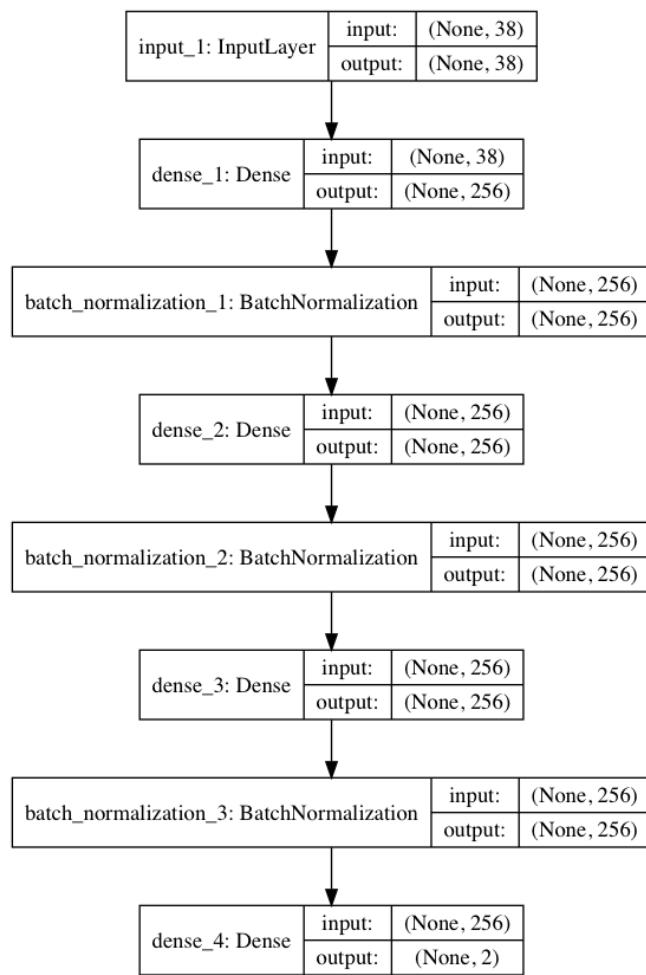


Figure 103: learning curve.

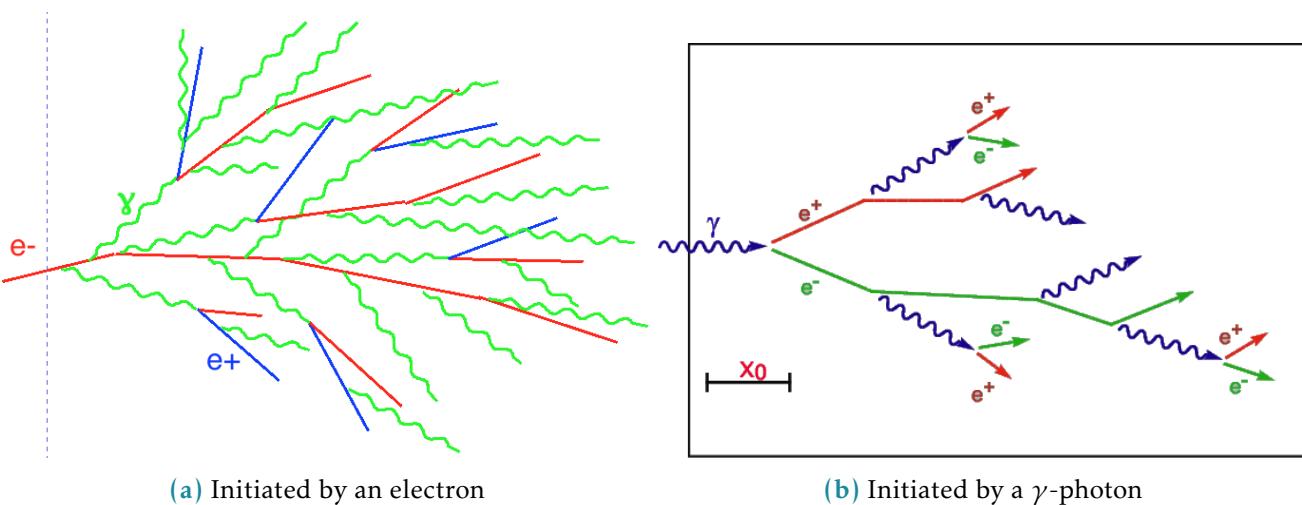


Figure 104: The schematic portrayal of EM shower development

1758 Bibliography

- 1759 [1] Yann Lecun et al. “Gradient-based learning applied to document recognition”. In: *Proceedings of*
1760 *the IEEE*. 1998, pp. 2278–2324.
- 1761 [2] Léon Bottou. “Stochastic Gradient Descent Tricks”. In: vol. 7700. Jan. 2012, pp. 421–436. doi:
1762 10.1007/978-3-642-35289-8_25.
- 1763 [3] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by
1764 back-propagating errors”. In: *Nature* 323 (1986), pp. 533–536.
- 1765 [4] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. cite arxiv:1412.6980Com-
1766 ment: Published as a conference paper at the 3rd International Conference for Learning Repre-
1767 sentations, San Diego, 2015. 2014. URL: <http://arxiv.org/abs/1412.6980>.
- 1768 [5] Yu Nesterov. “A method of solving a convex programming problem with convergence rate
1769 $O(1/k^2)$ ”. In: vol. 27. Jan. 1983, pp. 372–376.
- 1770 [6] Boris Polyak. “Some methods of speeding up the convergence of iteration methods”. In: *Ussr*
1771 *Computational Mathematics and Mathematical Physics* 4 (Dec. 1964), pp. 1–17. doi: 10.1016/0041-
1772 5553(64)90137-5.
- 1773 [7] Yann LeCun et al. “Efficient BackProp”. In: *Neural Networks: Tricks of the Trade*. Ed. by Genevieve
1774 B. Orr and Klaus-Robert Müller. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 9–50.
1775 ISBN: 978-3-540-49430-0. doi: 10.1007/3-540-49430-8_2. URL: https://doi.org/10.1007/3-540-49430-8_2.
- 1777 [8] Kurt and Hornik. “Approximation capabilities of multilayer feedforward networks”. In: *Neural*
1778 *Networks* 4.2 (1991), pp. 251–257. doi: 10.1016/0893-6080(91)90009-T. URL: <http://www.sciencedirect.com/science/article/pii/089360809190009T>.
- 1780 [9] G. Cybenko. “Approximation by superpositions of a sigmoidal function”. In: *Mathematics of*
1781 *Control, Signals and Systems* 2.4 (Dec. 1989), pp. 303–314. ISSN: 1435-568X. doi: 10.1007/BF02551274. URL: <https://doi.org/10.1007/BF02551274>.
- 1783 [10] D. E. Rumelhart and D. Zipser. “Feature Discovery by Competitive Learning”. In: *Parallel Dis-*
1784 *tributed Processing*. MIT Press, 1986, pp. 151–193.
- 1785 [11] Pankaj Mehta et al. “A high-bias, low-variance introduction to Machine Learning for physicists”.
1786 In: *Phys. Rept.* 810 (2019), pp. 1–124. doi: 10.1016/j.physrep.2019.03.001. arXiv: 1803.08823
1787 [*physics.comp-ph*].
- 1788 [12] Sergey Ioffe and Christian Szegedy. “Batch Normalization: Accelerating Deep Network Training
1789 by Reducing Internal Covariate Shift”. In: *Proceedings of the 32nd International Conference on*
1790 *International Conference on Machine Learning - Volume 37*. ICML15. Lille, France: JMLR.org, 2015,
1791 pp. 448–456.

- 1792 [13] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning*
1793 *Research* 12 (2011), pp. 2825–2830.