Study of the decay properties of the Higgs boson into two b quarks and upgrade of the ATLAS inner tracker

Louis D'Eramo

SORBONNE UNIVERSITÉ
École doctorale des Sciences de la Terre et de l'Environnement et Physique de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE)

Étude de la désintégration du boson de Higgs en deux quarks b et amélioration du trajectographe de l'expérience ATLAS

Présenté par

Louis D'Eramo

THÈSE DE DOCTORAT DE PHYSIQUE

Dirigée par Dr Giovanni Calderini

Presentée et soutenue publiquement le 23 septembre 2019 devant le jury composé de :

MM.	Stéphane	Jezequel	Rapporteur
	Tim	Scanlon	Rapporteur
Mme	Daniela	Bortoletto	Examinatrice
MM.	Roberto	Salerno	Examinateur
	Matthew	Charles	Examinateur
	Giovanni	Calderini	Directeur de thèse
	Gautier	Hamel de Monchenault	Président

SORBONNE UNIVERSITÉ
École doctorale des Sciences de la Terre et de l'Environnement et Physique de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE)

Study of the decay properties of the Higgs boson into two b quarks and upgrade of the ATLAS inner tracker

Presented by

Louis D'Eramo

Submitted in fulfillment of the requirements for the degree of

DOCTEUR ES SCIENCES SORBONNE UNIVERSITÉ

Supervised by Dr. Giovanni Calderini

Defended on the 23^{rd} of September in front of the committee :

Mr.	Stéphane	Jezequel	Referee
	Tim	Scanlon	Referee
Ms.	Daniela	Bortoletto	Examiner
Mr.	Matthew	Charles	Examiner
	Roberto	Salerno	Examiner
	Giovanni	Calderini	Supervisor
	Gautier	Hamel de Monchenault	President

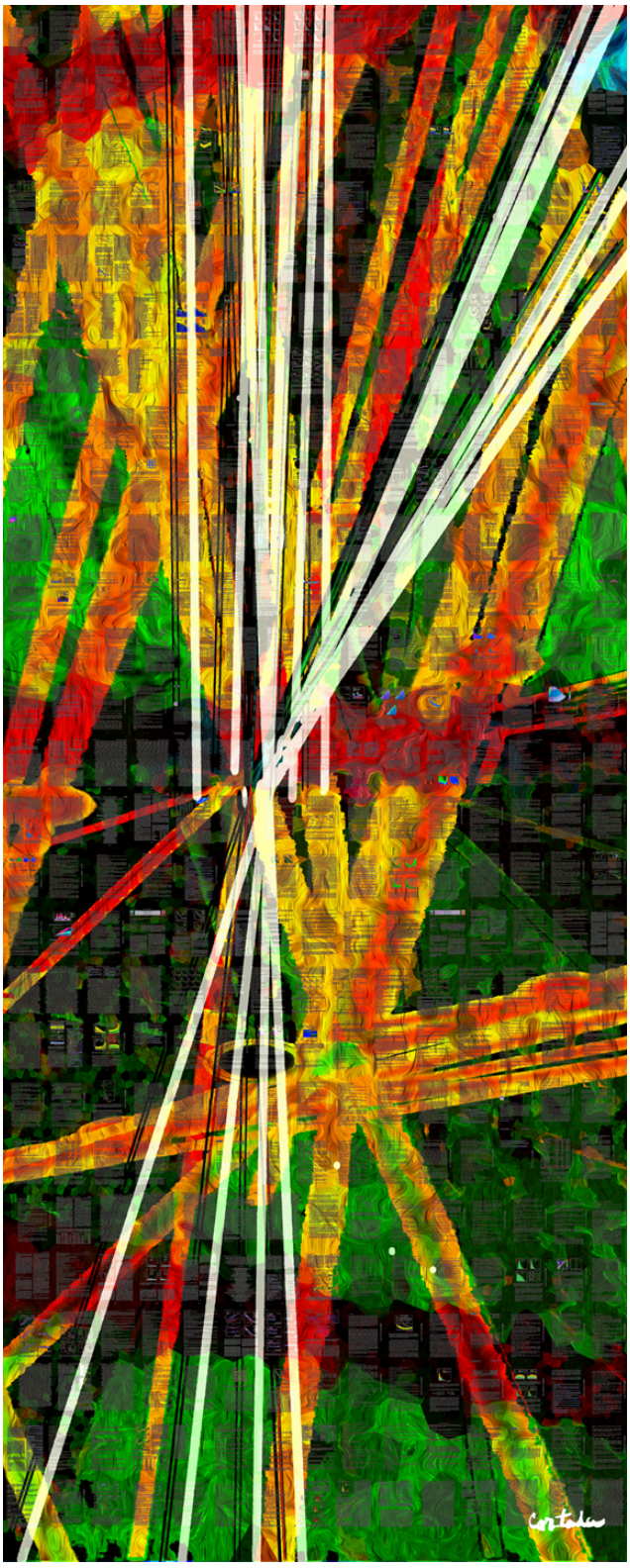


FIGURE 1 – Artistic representation of an Higgs-like particle decaying into a pair of b-quarks. The background uses published papers and represents the underground work that drives scientific discoveries [1].

Étude de la désintégration du boson de Higgs en deux quarks b et amélioration du trajectographe de l'expérience ATLAS

Résumé

Ce travail de thèse présente la recherche d'un signal correspondant au boson de Higgs dans le canal de désintégration en une paire de quarks b avec production associée d'un boson vecteur (VH). Il se base sur les données collectées par l'expérience ATLAS à une énergie de centre de masse $\sqrt{s} = 13$ TeV et correspondant à une luminosité intégrée de $\mathcal{L} = 79.8\ fb^{-1}$. La sélection des évènements, reconstruction des objets, la modélisation des bruits de fond et du signal, la discrimination des évènements, et l'analyse statistique des propriétés du signal résultante sont décrites. Un excès d'évènements compatible avec le boson de Higgs prédit par le modèle standard est mesuré avec une significance observée (attendue) de 4.9 (4.3) déviations standard et une intensité de signal $\mu^{bb}_{VH} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$. L'analyse est combinée avec le résultat obtenu au Run-1 ainsi qu'avec les analyses impliquant d'autres modes de désintégration (diphoton et $ZZ^* \to 4$ leptons, dans le mode de production VH) et d'autres modes de production (VBF et ttH, dans le mode de désintégration $\mathrm{b\bar{b}}$). L'observation statistique de la désintegration du boson de Higgs en pair de quark b est obtenue avec une significance observée (attendue) de 5.4 (5.5) déviations standard et une intensité de signal $\mu_{H\to bb} = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.15}(\text{syst.})$. Le mode de production VH est également observé avec une significance observée (attendue) de 5.3 (4.8) déviations standard et une intensité de signal $\mu_{VH} = 1.13 \pm 0.15(\text{stat.})^{+0.18}_{-0.17}(\text{syst.})$. Plusieurs améliorations sont également débattues concernant la réjection de bruit QCD et la modélisation des bruits de fond impliquant des techniques multivariées.

À partir de 2026 le LHC entrera dans une phase nouvelle, produisant des collisions à un rythme inégalé. Afin de mesurer les potentiels gains, une étude d'extrapolation est présentée concernant l'analyse $V\mathrm{h}(\mathrm{b\bar{b}})$. Deux scénarios d'hypothèses sont proposés, et les résultats sont interprétés en termes de section efficace et d'intensité de signal. Les incertitudes sur ce dernier paramètre s'échelonnent de 10 (15) % pour le mode de production WH à 43 (50) % pour le mode $\mathrm{gg}Z\mathrm{h}$ dans l'hypothèse du scénario 2 (1). Du fait de ces conditions de collision, le trajectographe de l'expérience ATLAS (Inner-Tracker, ITk) va devoir être amélioré bénéficiant des dernières technologies développées afin d'améliorer sa résolution spatiale et sa resistance aux radiations. Trois productions de capteurs planaires présentant une épaisseur réduite et une couverture géométrique étendue sont décrites. Les propriétés électriques et les efficacités de détection ont été mesurées sur faisceau après des campagnes d'irradiations.

Mots Clés

LHC,expérience ATLAS, boson de Higgs, production associée VH, quark b, ITK, FTK

Study of the decay properties of the Higgs boson into two b quarks and upgrade of the ATLAS inner tracker

Abstract

This thesis is focussing on the search for the Standard Model (SM) Higgs boson decaying to a pair of b-quarks in the vector boson associated production mode (VH) with the data collected with the ATLAS experiment at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of $\mathcal{L} = 79.8\ fb^{-1}$. The selection of the events, reconstruction of the objects, the modelling of the backgrounds and signal, the discrimination of the events, and the resulting statistical analysis of the properties of the signal are discussed for this analysis. A clear excess of events compatible with the SM Higgs boson is detected with an observed (expected) significance of 4.9 (4.3) standard deviations and a signal strength of $\mu_{VH}^{bb} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$. The analysis is combined with the results obtained in Run-1 and the analysis performed with other decay modes (di-photon and $ZZ^* \rightarrow$ 4 leptons, in the VH production mode) and other production modes (VBF and ttH, in the $\mathrm{b\overline{b}}$ decay mode). The observation of the decay to b-quarks is claimed with an observed (expected) significance of 5.4 (5.5) standard deviations with a signal strength of $\mu_{H\rightarrow bb} = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.15}(\text{syst.})$ as well as the VH production mode with an observed (expected) significance of 5.3 (4.8) standard deviations with a signal strength of $\mu_{VH} = 1.13 \pm 0.15(\text{stat.})^{+0.18}_{-0.17}(\text{syst.})$. Several improvements are also discussed concerning the QCD background rejection and background modelling using multivariate techniques. From 2026 the LHC will enter in a new phase, providing collisions at an unprecedented rate. To understand the potential gains from that period, an extrapolation study is presented. Two scenarios of extrapolation are proposed, and the results are interpreted in terms of cross-section and signal strength. The uncertainties on the signal strength are going from 10 (15) % for the WH mode to 43 (50) % for the $\mathrm{gg}Z\mathrm{h}$ one for scenario 2 (1), with a systematical limitation for the WH and $q\bar{q}Z\mathrm{h}$ processes and a predominance of the statistical uncertainties for $\mathrm{gg}Z\mathrm{h}$. Due to the new conditions, the ATLAS Inner-Tracker (ITk) will have to be upgraded to benefit from the new technologies developed to be more radiation hard and improve its resolution. Three productions of planar sensors aiming at reducing their thickness and increasing the geometrical acceptance are presented. The electrical properties and detection efficiencies have been measured on beams after irradiation campaigns.

Keywords

LHC, ATLAS experiment, Higgs boson, VH associated production, bottom-quark, ITK, FTK

Remerciements

Avant tout je souhaite remercier mes grands-parents qui, en 2008, m'ont fait découvrir le CERN. Cette visite a déclenché ce désir d'un jour travailler dans un environnement international aussi exceptionnel. Cette thèse représente l'aboutissement de ce long voyage qu'ils ont initié par leur amour et leur bienveillance. Je ne serais jamais arrivé à ce résultat sans le soutien inconditionnel de mes parents et beaux-parents. Vous m'avez donné cette curiosité et culture du savoir me permettant d'aller toujours plus loin. En plus de son impact personnel, ma soeur Juliette a questionné mon sens de la pédagogie au travers de nombreuses sessions de révisions ensemble. Ton admiration me touche et m'oblige à devoir te donner le meilleur dans la vie. Je remercie également Léola qui a partagé ma vie au début de ce parcours avec une bienveillance de chaque instant.

I wish to thank greatly my supervisor Giovanni Calderini. Since we first met in 2013, you have always guided me while allowing me to make mistakes, ask not-so-clever questions and most of all strengthening my independence. Our discussions were always of high level, helping me to take a step back on what I was doing, and even when we did not agree I had always to find good arguments to convince you. You have been there for me in crucial moments when I needed a push-up to achieve my goals. To perform the work presented in this manuscript, I was able to rely on one of the greatest physicists I have had the pleasure to meet: Paolo Francavilla. Your brilliant and state-of-the-art ideas have been a major booster for my work, exceeding the hopes I had when starting this journey. I sincerely hope that this period is the starting point of a fruitful and long-lasting collaboration. I could also count on all the members of the ATLAS team in the LPNHE, especially on Giovanni Marchiori and Kun Liu which encouraged me weeks after weeks to present my results during the local meeting, and Marco Bomben and Francesco Crescioli during all the test-beams where I bothered you with silly technical questions. At the edge of Paris, the colleagues of LAL have always been of great assistance: Nicolas Morange, Jean-françois Grivaz and Konie Al Khour. Further away, the members of the ATLAS VHbb team have created a nice working environment, especially the conveners Valerio and Tatsuya who backed all the new ideas developed while organising the work fairly amongst the members of the group. I wish also to thank Stephen Jiggings with whom I had profound and long discussions. You helped me to learn a lot, especially on points I

thought I had understood. For three years I have found also a new family composed of the past, present and future PhD students of the lab, always keen on helping, debating on our way to the canteen or at the coffee breaks: Carlo, Yee, Stefano, Daniel, Alvaro, Dilia, Changqiao, Simon, Audrey, Ilaria, Ahmed, Robert, Émilie, Anyssa, Alex, Jad, Reem, Thomas, Christelle, Julianna and many other I may have missed here. A special thank to Audrey and Reem with whom I shared special periods during shifts on the test-beam, as well as Ilaria with which we shared joy and stress and much more than a simple analysis. You are all great physicists, but also great friends with whom times flies like LLPs.

Ma vie de thésard ne s'est pas arrêtée à la vie de laboratoire. Je remercie donc également mes ami.e.s de l'ENS Cachan qui de par les voyages et les sorties diverses m'ont permis d'échanger sur nos pratiques relatives de la recherche ainsi que d'avoir supporté mes anecdotes sur la Lorraine: Marion, Pascal, Rémi, Suzanne, Antoine, Méghanne, Pierre, Claire, Édouard, Vincent, Hugues, Bertrand et Doriane; les membres de l'A[list] Marie, Étienne, Chedar, Chatan, Fred, Stephane, Rémy, ainsi que toutes les personnes que j'ai pu rencontrer lors de ces quatre années fantastiques. En arrivant à Paris j'ai pu aussi m'intégrer à la grande famille des écologistes qui m'ont fait découvrir le 10eme arrondissement, mais aussi des profils divers et variés. Un grand merci donc à: Julien, Marine, Nour, Patrick, Léa, Amar, Lara, Célia, Jean-Philippe et tant d'autres que ces quelques lignes ne suffiraient pas.

Contents

Introduction

Particle physics is probably the finest example belonging to the reductionist approach (at least from the concept of physicalism developed by Otto Neurath and Rudolf Carnap [2]). Its construction relies on the existence of fundamental particles, described by the Standard Model (SM) of particle physics, that explain the properties observed at the sub-atomic and sub-nuclear scale. The theory designed to describe the interactions between those constituents is a perfect mix, in the sense of the reductionism, of fundamental theories discovered earlier, such as special relativity/electromagnetism, quantum mechanics and group theory. The striking agreement between the experimental observations and the theoretical predictions is further strengthens this approach among the scientific community. The construction of this model also illustrates the great complementarity between the theoretical and experimental approach, the latter driving some significant developments of the former.

The circumstance that allow particles to have mass, a property that is fundamental to explain the strength of the SM weak interaction, is not intrinsic to the SM and would violate one of the fundamental gauge invariances of the theory. This property can be restored thanks to the presence of an added scalar field, that spontaneously breaks this gauge symmetry through the Brout-Englert-Higgs (BEH) mechanism [3, 4]. The presence of this field is revealed by the existence of a neutral scalar particle known as the Higgs boson. The search for this missing piece in the SM jigsaw puzzle has motivated the construction of particle accelerators, the LHC [5] being the latest. Finally, its discovery has been announced conjointly by the ATLAS [6] and CMS [7] collaborations, two multi-purpose experiments located on the LHC, on July 4, 2012 [8, 9]. The study of this recently discovered particle allows the completeness of the predictions of the SM to be probed and thus is of particular interest in this field of research.

The content of this thesis is based on studying the structure of the coupling of the Higgs boson to the heavier quark accessible in the final state decay: the b quark. Being the heaviest fermionic object accessible in the Higgs boson decay, the branching ratio of this coupling is the largest one, due to the fundamental structure of Higgs interactions. Therefore its measurement is crucial to determine the level of accuracy of the Standard

Model, while any deviation could be a hint for something beyond this theory.
However, this signature is experimentally challenging at hadron colliders due to the nature of the final state that can be mimicked by overwhelming QCD b-hadron production. To disentangle those events, the search can be restricted to a particular production mechanism of the Higgs boson. Amongst the most accessible modes at hadron colliders, the associated production with a vector boson V, is one of the most sensitive, thanks to the presence of leptons from the vector decay, which can be easily tagged. Therefore the analysis presented in this document is going to focus on this process. Searches based on the vector boson fusion (VBF) or the top-quark pair associated production ($t\bar{t}h$) can be also conducted but suffer from lower sensitivities.

Some hints of this signature have been reported in the past. The Tevatron [10] experiments D0 and CDF published a final result [11] on the $Vh(b\bar{b})$ analysis with an observed significance of 2.8 standard deviations for a Higgs mass of 125 GeV. The Run-1 data of the LHC confirmed partially the results of Tevatron. The CMS collaboration published a signal excess of 2.1 standard deviations [12], while the ATLAS collaboration observed an excess of 1.4 standard deviations [13]. A combination between the two analyses and between the Vh and $t\bar{t}h$ in the $b\bar{b}$ final state lead to an observed 2.6 standard deviations [14]. Before the beginning of the work presented in this thesis, a preliminary result was released by the ATLAS experiment [15] with an observed significance of 0.42 standard deviations. This was late superseded by an ATLAS mid Run-2 result [16] that claimed a clear evidence of the $Vh(b\bar{b})$ process with a 3.5 observed significance. The CMS collaboration also published a mid Run-2 evidence [17] with a 3.3 observed significance.
In this thesis, I describe the search for the $Vh(b\bar{b})$ process with data collected by the ATLAS detector during the partial LHC Run-2 period corresponding to an integrated luminosity of $\mathcal{L} = 79.8\ fb^{-1}$ at a centre-of-mass energy of $\sqrt{s} = 13$ TeV.

The LHC program will continue for 20 more years. Several upgrades are expected to allow the final integrated luminosity of $\mathcal{L} = 3000\ fb^{-1}$ to be reached. From 2026, collisions will enter an unprecedented regime, called the High Luminosity LHC (HL-LHC), with a mean instantaneous luminosity that will be a factor 5 larger than the present one. This collection of data will allow precise measurements of the Higgs sector by reducing both the statistical and systematic uncertainties on key parameters. Differential measurements are expected also to bring hints of possible deviations from the SM, such as in the case of the $B^0 \to K^{*0}\mu^+\mu^-$ conducted by LHCb [18] and Belle [19] that are showing up to 3.4 standard deviations effects on angular distributions. The results in the Higgs sector will therefore target differential cross-section measurements as already reported in the di-photon channel [20]. This implies that the detector has to withstand a high radiation fluence, maintaining at the same time an excellent reconstruction efficiency and readout capability. With such constraints in mind, a new ATLAS Inner TracKer (ITk) is being designed. This instrument has a crucial role in the track reconstruction and b-quark identification, and thus will be of great importance to the future of the $Vh(b\bar{b})$ analysis.
In this thesis, I also describe the perspective for the search of $Vh(b\bar{b})$ in the scenario of

$\sqrt{s} = 14$ TeV collisions and an integrated luminosity of $\mathcal{L} = 3000\ fb^{-1}$ and a contribution to the design and testing of planar silicon pixels in the scope of the ITK project.

Structure of the manuscript and personal contributions

This thesis is divided into seven chapters introducing the scope of the work and presenting my personal contributions. Chapter 1 provides an overview of the theoretical framework needed to understand the fundamental processes studied in this document. It describes how the SM is constructed and why the Higgs boson is a fundamental piece in this theory. A dedicated section introduces the phenomenology of the Higgs boson discovery at hadron colliders.
In Chapter 2, I describe the LHC chain of accelerators that are used to produce the data analysed in this document. The structure of the ATLAS detector is also presented in this chapter, focusing on the various sub-components and their physics detection principles.
Chapter 3 is divided into two subparts describing the connexions between the signal generated in the detector and the data used in the analysis. The first part concentrates on the trigger system, with an emphasis on the Fast TracKer system where I worked on the improvement of operation of the Associative Memory (AM) chip (from the electrical consumption optimisation to the design of the bank of patterns loaded). The second part introduces the reconstruction of various objects that are needed in the analysis and the description of how they are calibrated. Particular attention is given to the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction as this object is crucial for the analysis improvements detailed in the next sections.
The next three chapters describe in detail the main data analysis conducted with the improvements obtained. Chapter 4 introduces the strategy of the analysis, from the event selection and classification point of view. A significant part of the chapter is dedicated to the multi-variate tools as they are at the core of the improvements presented in this document.
In Chapter 5, I describe the strategies used to reduce and model the various backgrounds that are affecting the search. The first part is devoted to the QCD multi-jet background. I have been working on the estimation of this background yield and tried to optimise the rejection cuts in the search for a $Z\mathrm{h} \to \nu\bar{\nu}\mathrm{b}\overline{\mathrm{b}}$ signal. The second part introduces the systematics used to model the impact of the choice of a MC generator to estimate the SM backgrounds. I worked in particular to develop a new technique using multi-variate tools. A strong focus is given to the fundamental differences with respect to the technique presently used in the analysis, and the performance of the new method.
Chapter 6 introduces the statistical tools used for the search of the Higgs boson. After a digression on the definition of the statistical framework, followed by the specificities of the $V\mathrm{h}(\mathrm{b}\overline{\mathrm{b}})$ analysis, the results obtained in [21] are presented. I have personally contributed to the development of the framework and derived the cut-based result, collaborating with the other members of the team to provide the full set of results. The effect of modifying the multi-jet cuts are also presented and compared to the official results.
Last but not least, Chapter 7 describes the expectations and developments for the next stages of the LHC. First, a set of results are derived and their extrapolation is presented for a data-set corresponding to $\mathcal{L} = 3000\ fb^{-1}$. I have derived all the results presented,

in collaboration with another PhD student. Last I present the developments for the future pixel tracking detector of the ATLAS experiment by the LPNHE group. I have been involved in the measurement campaign to test the sensors' properties, from the data-taking to the analyses presented.

1 Theoretical framework

The path towards the description of most fundamental elements of the matter has lead physicists to develop an elaborate mathematical framework based on deep hypothesis, supported by strong experimental observations driven by technological breakthroughs. The latest most accurate model named the Standard Model (SM) lies on a simple zoology of observed particles described by a complex but elegant theory using the Quantum Field Theory (QFT) approach.
In this chapter I introduce the theoretical and experimental basis upon which the physics analysis presented in Chapters 4, 5 and 6 is built. I start with the description of the experimental zoology of observed particles. I then introduce the QFT model and discuss how the fundamental interactions are handled in this model, taking the example of the electroweak interaction. From this situation, I present the need for a spontaneous symmetry breaking and the Higgs mechanism for the generation of the mass. I finally emphasis the phenomenological aspects of the Higgs boson, from its production to its decay modes, relevant in the experimental situation of the Large Hadron Collider (LHC).

1.1 Zoology of observed elementary particles

From a historical point of view the description of elementary structures has driven the late 19^{th} and the 20^{th} century discoveries in physics, such as the electron discovery by Lord Thompson [22] in 1897 or composite particles like the neutron [23]. These

discoveries are deeply linked to the progress made from the theory side with the development of quantum theory, as well from the experimental side in various fields such as the nuclear physics, astroparticle observations and the development of particle accelerators. Along with the discoveries of new particles, and following the long tradition of the nuclear/chemistry field, a classification and naming process started, with a focus on the interaction. This process was made famous by the technique of Dimitri Mendeleev who used such classification on atoms, which later would be interpreted in terms of the existence of more fundamental particles that were driving the observed property differences. The fermion and boson categories were created to separate the particles according to their quantum spin, which leads to different grouping properties such as the Pauli exclusion principle for Fermions or the Bose-Einstein condensation for Bosons. This grouping is both based on experimentally observed properties as well as on the mathematical description that will be introduced later on. If most of the observed particles can be assigned to one of those two categories, it is natural to separate the electron and its heavy partners, the muon and the tau, from the proton and neutron and their partners, called hadrons (from the greek hadros, stout, thick) with the former called leptons (from the Greek leptos, light). The hadronic family can be classified as fermions, called baryons (from the Greek barys, heavy) or as bosons, called mesons (from the Greek mesos, middle). By the classification of their properties, it was proposed by Gell-Mann [24] that the hadrons were in fact composed of other elementary particles called quarks which are fermions, which were indeed discovered at SLAC by the Deep Inelastic Scattering process [25]. In total six quarks have been discovered, three with a positive electrical charge of +2/3, the remaining having a negative -1/3 electrical charge. This structure allows them to be grouped into three families. This behaviour is also observed on the leptonic side with the addition of the neutrino postulated by Pauli in 1930 [26] but discovered experimentally in 1956 by Cowan and Reines [27]. Three families of leptons are accompanied by three flavours of neutrinos, which are now proven to be different from the neutrino mass states. This led to the current interpretation of the fermions in the context of the Standard Model to be at the origin of the matter. They could all be summarised in Table 1.1.

	Leptons			Quarks		
Family	Particle	Q/e	Mass [GeV]	Particle	Q/e	Mass [GeV]
First	ν_e	0	$< 2 \times 10^{-9}$	u	+2/3	$2.16^{+0.49}_{-0.26} \times 10^{-3}$
	e^-	-1	0.511×10^{-3}	d	-1/3	$4.67^{+0.48}_{-0.17} \times 10^{-3}$
Second	ν_μ	0	$< 0.19 \times 10^{-3}$	c	+2/3	1.27 ± 0.02
	μ	-1	0.106	s	-1/3	$93^{+11}_{-5} \times 10^{-3}$
Third	ν_τ	0	$< 18.2 \times 10^{-3}$	t	2/3	172.9 ± 0.4
	τ	-1	1.777	b	-1/3	$4.18^{+0.03}_{-0.02}$

Table 1.1 – Summary of the elementary fermions properties. The charge is given per fraction of the electron charge e. Only an upper limit on the mass of the flavour eigenstates is given. The mass of the top quark is given from direct measurements on $t\bar{t}$ events [28].

Elementary particles can also have a bosonic state, such as for the photons that were postulated by Albert Einstein in [29] discovered by Robert Millikan [30] and

corroborated afterwards by the Compton effect [31]. Following this successful approach and with the theoretical development of Feynman, Tomonaga and Schwinger in the 1950s, the interaction between particles was predicted to originate from the exchange of such bosons. Several other elementary forces were imagined over the 20^{th} century to explain the cohesion of the nucleus whose charged particles are of the same sign (strong interaction), and the observed β decay of a neutron into a proton (weak interaction). The gluons binding the quarks, and thus the nucleus together were discovered in the Petra experiment in 1979 [32], while the β decay was interpreted in terms of a quark flavour changing current modelled by W/Z boson exchange. The latter were discovered at the CERN SPS experiments UA1 [33, 34] and UA2 [35, 36]. The only interaction that is not included in the standard model so far is gravity, since no good theory is able to include a description of general relativity into the formal QFT description with consistent experimental results. Table 1.2 summaries the various gauge bosons involved in the SM.

Interaction	boson	Mass [GeV]
Electromagnetic	γ	0
Weak	$W^{\pm}$	80.38
	Z	91.19
Strong	g	0

Table 1.2 – Summary of the elementary gauge bosons properties [28].

1.2 The Standard Model: a theoretical explanation

From the experimental observations, a global theory has been developed since the 1950s from the mathematical description of quantum theory based on symmetries but augmented by the deep knowledge of special relativity. The purpose of the SM is to find the biggest irreducible groups of elements that would respect those two principles, whilst the interaction would follow naturally from the symmetries imposed giving rise to conserved properties: the relevant charges of the interaction. The purpose of this Section is not to derive all the rules of the QFT, which is already spanned by abundant literature, and hence I will focus only on the main relevant points to the work described in this thesis.

1.2.1 Lagrangian description and group symmetry: the Quantum Electrodynamics (QED)

Starting from a description of quantum physics, the natural choice would be to consider the energy of the system and hence the Hamiltonian description to parametrise the system and derive the rules of interaction. However, this description, while answering directly the time evolution of the system, lacks the inherent symmetries. Thus it breaks the first principle given above, and one has to go one step earlier and actually compute the Lagrangian form of the equation. The energy conservation resulting from symmetries is then converted in a deeper principle that the time-integrated Lagrangian called action must be extremal:

$$S = \int \mathcal{L}\, dt$$

In fact, the generalisation of the action principle was given by Feynman and resulted in the path integral formulation which was shown to be equivalent to other prescription, but gives a central role to the Lagrangian definition in the computation of motion.

When considering a fermion field ψ in the free theory, the Dirac Lagrangian can be written as:

$$\mathcal{L}_{free} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi$$

where m is the mass of the fermion, γ^{μ} are the 4×4 Dirac matrices:

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & \tau^i \\ \tau^i & 0 \end{pmatrix}$$

where τ^i are the Pauli matrices:

$$\tau^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tau^2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \tau^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and $\bar{\psi} = \psi^{\dagger}\gamma^0$ is the correct hermitian transformation of the field ψ to form the Lorentz scalar $\bar{\psi}\psi$. The Einstein notation for repeating indices has been adopted here to ease the notation. From this notation, the kinetic term can be recognised in the first part, while the second denotes the mass term. The simplest complete transformation one could imagine which such a system would be to make a local phase transformation of the field (interpreted as a rotation):

$$\psi \to \psi' = e^{iq\cdot\theta(x)}\psi$$

where x is the point in the 4-dimensional phase space considered, q is a real parameter and $\theta(x)$ the local phase transform. From the hypothesis set in the introduction of the section, the Lagrangian shall be made invariant under such a local gauge U(1) symmetry forming an Abelian-Lie group. However, when applying this transformation on the Lagrangian, it

transforms to:

$$\begin{aligned}\mathcal{L}_{free} \rightarrow \mathcal{L}' &= (e^{-iq\cdot\theta(x)}\bar{\psi})(i\gamma^\mu\partial_\mu - m)e^{iq\cdot\theta(x)}\psi \\ &= \mathcal{L}_{free} - q\partial_\mu\theta(x)\bar{\psi}\gamma^\mu\psi\end{aligned}$$

Since the Lagrangian writing is defined up to an additive quantity, a corrective term can be introduced to restore the invariance. A clever choice is to consider the following covariant derivative:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i\,q\,A_\mu$$

where A_μ is an additional vector field (i.e. described as a boson) which is required to transform under the U(1) transformation as:

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\theta(x)$$

The resulting Lagrangian is therefore written as:

$$\begin{aligned}\mathcal{L}_{U(1)} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \\ &= \mathcal{L}_{free} - qJ^\mu A_\mu\end{aligned}$$

where $J^\mu = \bar{\psi}\gamma^\mu\psi$ is a term associated to the electromagnetic current, and thus the latter term can be interpreted as a contact term between the electron and the field A_μ. In order to connect the theoretical prediction to the observation, the A_μ field is interpreted as the mediator of the interaction (i.e. the photon in this case). Therefore this local gauge transformation gave rise to the existence of gauge mediating boson for an interaction between charged leptons with an electromagnetic charge q. This new massless field needs also a kinetic term for its propagation that needs to be added to the Lagrangian. Not to violate the group symmetry defined above, the solution is to introduce the field strength tensor $F_{\mu\nu}$ defined as:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

which can be obtained by computing the transformation of the commutation operator $[D_\mu, D_\nu]$ with respect to the U(1) local gauge symmetry[37]. This term is directly gauge invariant and is integrated into its quadratic form in the full QED Lagrangian:

$$\begin{aligned}\mathcal{L}_{QED} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \mathcal{L}_{free} - \underbrace{qJ^\mu A_\mu}_{\text{interaction term}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{mediator kinetic term}}\end{aligned}$$

From the block diagonal structure of the γ matrices, a natural decoupling of the 4-dimension field into a two-block structure is appearing: the Weyl representation of the

fields:

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$$

where L and R denote the left- and right-handed spinors components. A careful inspection of the free Lagrangian, shows that the terms with the γ_μ matrices (kinetic and interaction terms) are connecting the components with the same chirality while the mass term is connecting the left- to the right-handed parts:

$$\mathcal{L}_{QED} = \underbrace{\bar{\psi}_L i\gamma^\mu \partial_\mu \psi_L + \bar{\psi}_R i\gamma^\mu \partial_\mu \psi_R}_{\text{kinetic term}} - \underbrace{qA_\mu(\bar{\psi}_L \gamma^\mu \psi_L + \bar{\psi}_R \gamma^\mu \psi_R)}_{\text{interaction term}} - \underbrace{m(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R)}_{\text{mass term}} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

1.2.2 The electroweak unification

With the observation of the coupling between left-handed electrons and neutrinos [38], and the lack of observed right-handed neutrinos, a chiral representation is used to explain this effect:

$$L_L = \begin{pmatrix} \nu_L^e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_L^\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_L^\tau \\ \tau_L \end{pmatrix}, \qquad L_R = e_R, \mu_R, \tau_R$$

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \qquad U_R = u_R, c_R, t_R$$

$$D_R = d_R, s_R, b_R$$

Using a pure vectorial transformation as U(1) would not result in the desired properties. Hence an mixed vector-axial theory based on the SU(2) local gauge invariance was predicted by Bludman [39] in 1958 where the fields would transform as:

$$\psi_L \to \psi'_L = e^{i\theta(\vec{x}).\vec{\tau}} \psi_L$$
$$\psi_R \to \psi'_R = \psi_R$$

where $\vec{\tau} = (\tau_1, \tau_2, \tau_3)$ are the Pauli matrices described earlier and generators of the SU(2) group. The main difference from the QED case is the special relationships amongst the generators of the SU(2) group. The non commutative properties imply a non-Abelian symmetry which would modify the intrinsic properties of the fields. This transformation is also responsible for the generation of a massless isovector triplet $\vec{W}_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)$:

$$\partial_\mu \psi_L \to D_\mu \psi'_L = (\partial_\mu - i\, g\, \frac{\tau_i}{2} W_\mu^i) e^{i\theta(\vec{x}).\vec{\tau}} \psi_L$$
$$\partial_\mu \psi_R \to \partial_\mu \psi'_R = \partial_\mu \psi_R$$

The introduced gauge fields transform similarly to the U(1) case, with the difference that the structure of the non-Abelian group has to enter into account as suggested by Yang and Mills [40]:

$$W_\mu^i \to W_\mu^{i'} = W_\mu^i - \frac{1}{g'} \cdot \partial_\mu \theta^i(x) + f^{abc} W_\mu^b \theta^c(x)$$

where f^{abc} is defined from the generators of the subgroup algebra:

$$[\tau^a, \tau^b] = i f^{abc} \tau^c$$

resulting in the special case where f^{abc} is the Levi-Civita ϵ^{abc} symbol for SU(2). The resulting kinetic part of the Lagrangian for the leptons l can be expressed as:

$$\begin{aligned} \mathcal{L}_{\text{kin.}} &= i\bar{L_L}\gamma^\mu D_\mu L_L + i\bar{L_R}\gamma^\mu \partial_\mu L_R \\ &= \underbrace{\frac{g}{\sqrt{2}}(J^{+,\mu} W_\mu^+ + J^{-,\mu} W_\mu^-)}_{\text{charged current}} + \underbrace{g(\bar{\nu_L}\gamma^\mu \nu_L - \bar{l_L}\gamma^\mu l_L) W_\mu^3}_{\text{neutral current}} \end{aligned}$$

where the charged currents $J^{\pm,\mu}$ are defined as:

$$J^{+,\mu} = (J^{+,\mu})^\dagger = \bar{\nu_L}\gamma^\mu l_L$$

and the positive and negative vector bosons:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_1^\mu \pm W_2^\mu)$$

The parametrisation based on the $W^\pm$,W^3 parametrisation is more natural from the experimental point of view as it couples neutrinos and leptons from the same generations. The same can be obtained with the quark model replacing the charged leptons and neutrinos respectively by up- and down-type quarks. The particular similarities in the universality of couplings, the vector character and the affected particles between the electromagnetic and weak interactions made Glashow [41] imagine in 1961 a way to build a consistent theory giving rise to the two interactions with the same mechanism through

a SU(2) × U(1) group invariance. In that regard the transformation laws would act as:

$$\psi_L \to \psi'_L = e^{i(\theta(\vec{x}).\vec{\tau}+\alpha(x)\cdot Y)}\psi_L$$
$$\psi_R \to \psi'_R = e^{i\alpha(x).Y}\psi_R$$
$$\partial_\mu\psi_L \to D_\mu\psi'_L = (\partial_\mu - i\,g\,\frac{\tau_i}{2}W^i_\mu - i\,g'\,\frac{Y}{2}B_\mu)e^{i(\theta(\vec{x}).\vec{\tau}+\alpha(x)\cdot Y)}\psi_L$$
$$\partial_\mu\psi_R \to D_\mu\psi'_R = (\partial_\mu - i\,g'\,\frac{Y}{2}B_\mu)\psi_R$$
$$W^i_\mu \to W^{i'}_\mu = W^i_\mu - \frac{1}{g'}\cdot\partial_\mu\theta^i(x) + f^{abc}W^b_\mu\theta^c(x)$$
$$B_\mu \to B'_\mu = B_\mu - \partial_\mu\alpha(x)$$

where Y is the weak hypercharge defined as:

$$Y = 2\cdot Q - 2\cdot I^3_W$$

with I^3_W is the third component of the weak isospin which is $\pm 1/2$ for the left-handed doublets and 0 for the right handed singlets. The original kinetic term of the Lagrangian can thus be written as:

$$\mathcal{L}_{\text{kin.}} = \frac{g}{\sqrt{2}}(J^{+,\mu}W^+_\mu + J^{-,\mu}W^-_\mu) + g\underbrace{(\bar{\psi_L}\gamma^\mu\frac{\tau_3}{2}\psi_L)}_{\text{weak isospin current}}W^3_\mu + g'\underbrace{(-\bar{\psi_L}\gamma^\mu\frac{1}{2}\psi_L - \bar{\psi_R}\gamma^\mu\psi_R)}_{\text{weak hypercharge current}}B_\mu$$
$$= \frac{g}{\sqrt{2}}(J^{+,\mu}W^+_\mu + J^{-,\mu}W^-_\mu) + gJ^{3,\mu}W^3_\mu + g'J^{Y,\mu}B_\mu$$

where the lepton weak isospin $J^{3,\mu}$ and hypercharge $J^{Y,\mu}$ currents are defined to mimic the charged currents $J^{\pm,\mu}$. However, none of the currents can be defined as the electromagnetic one. Therefore a transformation of the two vector boson fields W^3_μ and B_μ is needed thanks to simple rotation of angle θ_W and was proposed by Weinberg to restore the A_μ field:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\cdot\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}$$

writing down the neutral part of the Lagrangian one finds:

$$\mathcal{L}_{\text{neutral}} = gJ^{3,\mu}(\cos\theta_W Z_\mu + \sin\theta_W A_\mu) + g'J^{Y,\mu}(-\sin\theta_W Z_\mu + \cos\theta_W A_\mu)$$

The electromagnetic contribution can be extracted as:

$$\mathcal{L}_{\text{em.}} = (g\sin\theta_W J^{3,\mu} + g'\cos\theta_W J^{Y,\mu})A_\mu$$

Noticing that :

$$J^{3,\mu} + J^{Y,\mu} = \bar{\psi_R}\gamma^\mu\psi_R + \bar{\psi_L}\gamma^\mu\psi_L = J^\mu$$

we can decide to fix the following condition by freezing a degree of freedom: $g\sin\theta_W = g'\cos\theta_W$, and restoring the electromagnetic part of the interaction as developed in the QED while taking into account the weak interactions:

$$\begin{aligned}\mathcal{L}_{\text{kin.}} &= \frac{g}{\sqrt{2}}(J^{+,\mu}W_\mu^+ + J^{-,\mu}W_\mu^-) + gJ^{3,\mu}W_\mu^3 + g'J^{Y,\mu}B_\mu \\ &= \frac{g}{\sqrt{2}}(J^{+,\mu}W_\mu^+ + J^{-,\mu}W_\mu^-) + \frac{g}{\cos\theta_W}(J^{3,\mu} - \sin^2\theta_W J^\mu)Z_\mu + g\cdot\sin\theta_W\, J^\mu A_\mu\end{aligned}$$

where the $W^\pm$ bosons are coupling the down part of the doublets to the upper part in a charged current, the neutral Z boson couples to the electromagnetic current and to a chiral breaking current coupling left-handed particles, and the recovered electromagnetic current coupling to the photon. When identified to the expression obtained in QED a direct link can be made between the strength of the coupling g and the angle θ_W:

$$e = g\cdot\sin\theta_W$$

As for the photon, a kinetic term can be defined for the new vector bosons, and will incorporate the non-Abelian structure part defined for the transformation of the gauge fields defined earlier:

$$\mathcal{L}_{YM} = \frac{1}{4}F^{a\mu\nu}F_{a\mu\nu} \text{ with } F^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc}W^b_\mu W^c_\nu$$

Here the field strength $F^a_{\mu\nu}$ is no longer a gauge nor a Lorentz invariant quantity, while the quadratic form $F^{a\mu\nu}F_{a\mu\nu}$ is and thus can be incorporated in the Lagrangian. One has to notice as well that the vector fields entering the kinetic terms are the original SU(2) triplets W^i and the U(1) singlet B and not their transformed partners $W^\pm$, Z^0 and A.

1.2.3 The quarks behaviour: the quantum chromodynamics (QCD)

The classification of the hadrons is defined by the fermions composing them (quarks) and their relationship that as for the QED is explained by a group symmetry which involves the creation of gauge bosons (gluons). Experimental observations are indicating that no quark can be seen isolated as they form either mesons or baryons. A recent result by the LHCb collaboration claimed to have observed pentaquark structures [42], likely formed by a coupling between a meson and a baryon. Nonetheless, a theory based on a colour number (experimentally found to be 3) using the SU(3) group theory was proposed to answer the observation, as it represents the minimalistic transformation to be designed answering the experimental observations. Thus the quark fields are represented using 3-dimensional vectors (contrary to the spinors in the EW theory) and are required

to transform as:

$$\psi_b \to \psi'_a = e^{\frac{i}{2}\theta^c(x).t^c_{ab}}\psi_b$$
$$\partial_\mu \to (D_\mu)_{ab} = \partial_\mu \delta_{ab} - i g_s A^c_\mu t_a b^c$$
$$A^c_\mu t^c_{ab} \to A^c_\mu t^c_{ab} + \frac{1}{g_s}\partial_\mu \theta^c t^c_{ab} + i[t^c_{ab}, t^d_{ab}]\theta^c A^d_\mu$$

where a, b are colour indices (from 1 to 3) and c, d the Gell-Mann matrix indices (from 1 to 8), A^c_μ being the eight massless gluons created from the symmetries. The Gell Mann matrices are defined as:

$$t^1 = \frac{1}{2}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix} \quad t^2 = \frac{1}{2}\begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix} t^3 = \frac{1}{2}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix} \quad t^4 = \frac{1}{2}\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}$$
$$t^5 = \frac{1}{2}\begin{pmatrix}0&0&-i\\0&0&0\\i&0&0\end{pmatrix} \quad t^6 = \frac{1}{2}\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix} t^7 = \frac{1}{2}\begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix} \quad t^8 = \frac{1}{2\sqrt{3}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix}$$

and couple different colours to each other. The quarks are therefore coloured objects. The QCD Lagrangian can be derived in this theory respecting the non-Abelian structure of the group for the kinetic term for the gluons:

$$\mathcal{L}_{\text{QCD}} = \bar{\psi_i}\left(i(\gamma^\mu D_\mu)_{ij} - m\,\delta_{ij}\right)\psi_j - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a$$

where $G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu$ is the kinetic term already described before, but where the structure constant f^{abc} has a more complicated form than the Levi-Civita in SU(2). When expanding the Lagrangian using the formula described above, several terms are providing the self-interactions of the gluon fields with a trilinear and quadrilinear coupling. However, additional terms known as Fadeev-Popov ghosts [43] have to be added to restore the gauge invariance and maintain the Gauge unitarity of the S-matrix.

The gluon self-interaction is also causing issues in the computation of the diagrams at a fixed order. When considering the gluon-gluon or quark-quark scattering mediated by the strong interaction, a series of NLO corrections are making the computation diverge. Those ultra-violet divergences are not only observed in the QCD case, and lead to the theory of renormalisation and regularisation:

— **Regularisation:** In order to cancel out the UV divergences a hard cut off λ is placed on the integral used in the computation. This is accompanied by a regularisation of the number of dimensions. However, this technique can not be used alone since it would lead having an un-physics parameter governing the results.

— **Renormalisation:** A new scale μ_R is introduced, such that the global result of the computation (cross-section, decay width, charge...) corresponds to the measurement at this scale. This scale is interpreted as the fraction of the computation coming from LO to the loop diagrams. It is usually taken to be the scale of the interaction (lepton mass for QED, top mass for $t\bar{t}$ measurements...), but should not appear in the final result: it is therefore absorbed in the definition of the coupling constant g.

In practice when performing the two operations in series, the regularisation scale λ is cancelled out, and only the renormalisation scale is exhibited. Its evolution is dictated by the Renormalisation Group Equations:

$$\alpha_{QED}(\mu) = \frac{\alpha_{QED}(\mu_R)}{1 - \frac{\alpha_{QED}(\mu_R)}{6\pi}\ln\left(\frac{\mu}{\mu_R}\right)}$$

$$\alpha_s(\mu) = \frac{\alpha_s(\mu_R)}{1 + \frac{\alpha_s(\mu_R)}{6\pi}(11N_C - 2N_f)\ln\left(\frac{\mu}{\mu_R}\right)}$$

where $g_{QED/s} = \sqrt{4\pi\alpha_{QED/s}}$, N_C is the number of colours of quarks and N_f is the number of generation of fermions. One important difference is that the factor in the denominator is negative for the QED structure while being positive for QCD. Therefore, the lowest the energy is, the strongest the coupling is, as shown in Figure 1.1. This results in quark confinement such that the resulting particle do not show any global colour, such as in the meson and baryon cases. Therefore no bare quark can be observed. When two quarks are moved away from each others, the potential energy from the strong interaction increases until reaching the threshold for a quark-antiquark pair production. This effect is called asymptotic freedom.

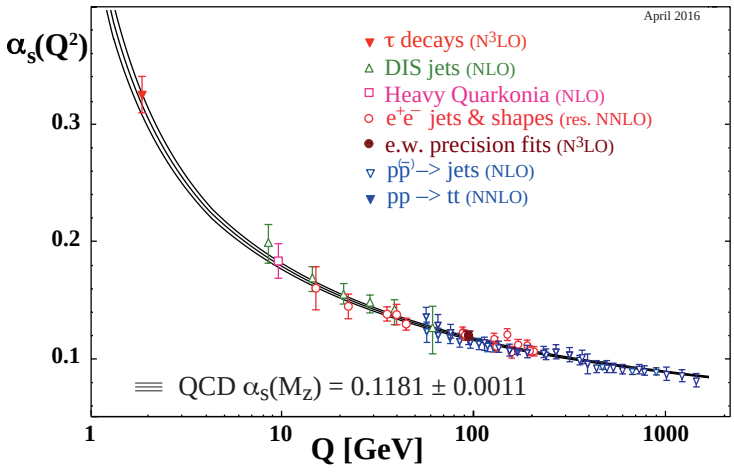


Figure 1.1 – Summary of measurements of α_S as a function of the energy scale Q [28].

1.3 Spontaneous symmetry breaking and Higgs mechanism

Although the electroweak unification gives a nice harmonisation amongst the two theories in a coherent way, it forbids any ad-hoc mass term for the gauge bosons in the Lagrangian, due to the transformation of the fields themselves. While this is adequate for the photon in the QED theory, it contradicts the observation of massive W and Z bosons. Even before the discovery of such particles, it was thought that due to the constraints of the short distance of the weak interaction, the mediator should be massive. Furthermore, it was shown that the mass term for fermions would couple the right- and left-handed part of the doublets, which would break the chiral symmetry introduced with the SU(2) transformation. Following the work derived by Peter Higgs, Robert Brout and Francois Englert, and Gerald Guralnik, C. R. Hagen, and Tom Kibble [3, 4, 44] which introduces the concept of Spontaneous Symmetry Breaking (SSB) and the link to massive gauge bosons, Salam and Weinberg proposed separately an application to the electroweak unification [45, 46] problem. It is worth noting that the original theory was developed in the context of superconducting materials where, in a second-order transition, the magnetic field would only penetrate the material on a length scale inversely proportional to the mass of the gauge Goldstone boson.

1.3.1 Illustration of a SSB in a U(1) symmetry group

All the symmetries that were introduced so far have lead to the creation of new interactions and particles to mediate those interaction. The concept introduced here lies on a different methodology, as it introduces a new field ϕ chosen to be a complex scalar field, and a corresponding potential that would act on this field. The Lagrangian can include this new term in the following way:

$$\mathcal{L}_\phi = \partial_\mu \phi^* \partial^\mu \phi - V(\phi)$$

where the general form of V is chosen to be:

$$V(\phi) = -\mu^2 |\phi|^2 + \frac{\lambda}{2} |\phi|^4$$

Depending on the sign of the μ^2, the shape of the potential could lead to a single minimum at zero ($\mu^2 < 0$) or two minima at a vacuum expected value (v.e.v) $v = \pm\left(\frac{\mu^2}{\lambda}\right)^{1/2}$ ($\mu^2 > 0$) as shown in Figure 1.2, λ being considered positive.

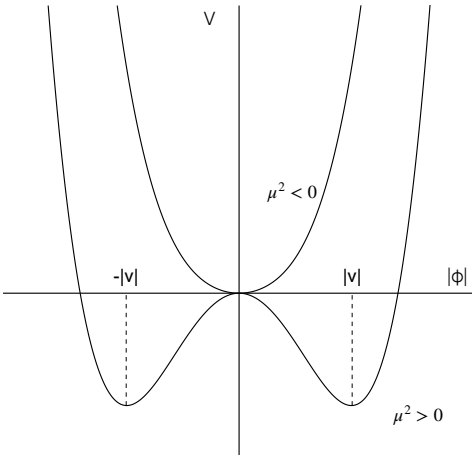


Figure 1.2 – Representation of the potential V(ϕ) only considering the norm of the complex scalar $|\phi|$.

As for the leptons in Section 1.2.1, the local gauge invariance of such field is considered and the related formulation of the covariant derivative and the boson field A_μ:

$$\phi \to \phi' = e^{iq\cdot\theta(x)}\phi$$
$$\partial_\mu \to D_\mu = \partial_\mu + i\,q\,A_\mu$$
$$A_\mu \to A'_\mu = A_\mu - \partial_\mu\theta(x)$$

Knowing that the field ϕ would take values close to the one of the v.e.v., a linear expansion around this minima is made, recasting the imaginary part of the field to ease the computation [1] :

$$\phi(x) = \frac{1}{\sqrt{2}}(v + h(x))$$

1. The case where the field is still complex has been treated extensively in the literature.

where h is a real scalar field. The Lagrangian can therefore be written as:

$$\begin{aligned}
\mathcal{L}_\phi &= (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi) \\
&= (\partial_\mu - iqA_\mu)\frac{1}{\sqrt{2}}(v+h(x))(\partial_\mu + iqA_\mu)\frac{1}{\sqrt{2}}(v+h(x)) - \mu^2(\frac{1}{\sqrt{2}}(v+h(x)))^2 + \frac{\lambda}{2}(\frac{1}{\sqrt{2}}(v+h(x)))^4 \\
&= \underbrace{\frac{1}{2}(\partial_\mu h)^2 - \frac{1}{2}2\cdot\mu^2h^2}_{\text{massive scalar h}} + \underbrace{\frac{1}{2}q^2v^2A_\mu^2}_{\text{mass term for the gauge boson}} + \underbrace{\frac{1}{2}q^2h^2A_\mu^2 + q^2vA_\mu^2h}_{\text{coupling between h and the gauge boson}} \\
&\quad - \underbrace{\frac{\lambda}{2}vh^3 - \frac{1}{8}\lambda h^4}_{\text{scalar self coupling}} + \mathcal{O}(\text{higher order term in h})
\end{aligned}$$

The first conclusion from the SSB is that a mass term for the gauge boson is appearing with $m_A = \sqrt{2}qv$, the new scalar h is a massive scalar boson with $m_h = \sqrt{2}\mu$, and furthermore interaction terms are arising between the gauge boson and the scalar boson itself through tree- and four-legs couplings. The same happens with the self-coupling of the scalar h.

1.3.2 Generalisation to the electroweak case

The strength of the SSB theory is its applicability to non-abelian group symmetries such as the SU(2) which is used to describe the electroweak interactions. As for the leptons the scalar fields must be transformed into a spinor that is complex from first principles:

$$\phi = \begin{pmatrix}\phi^\dagger \\ \phi_0\end{pmatrix} = \begin{pmatrix}\phi_1 + i\phi_2 \\ \phi_3 + i\phi_4\end{pmatrix}$$

which can be transformed using the equations obtained for the SU(2) × U(1) case. As previously, infinite ground states are available. For the sake of simplicity, the scalar field can be expended around the minimum where $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = \frac{1}{\sqrt{2}}(v+h)$. The development of each term of the Lagrangian gives:

$$\begin{aligned}
D_\mu\phi &= [\partial_\mu + i\frac{g}{2}\tau^i W_\mu^i + i\frac{g'}{2}B_\mu]\begin{pmatrix}0 \\ \frac{1}{\sqrt{2}}(v+h)\end{pmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{pmatrix}0 \\ \partial_\mu h\end{pmatrix} + i\frac{v+h}{2\sqrt{2}}\begin{pmatrix}g(W_\mu^1 - iW_m^2 u) \\ -gW_\mu^3 + g'B_\mu\end{pmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{pmatrix}0 \\ \partial_\mu h\end{pmatrix} + i\frac{v+h}{2\sqrt{2}}\begin{pmatrix}g\sqrt{2}W_\mu^- \\ -\sqrt{g^2+g'^2}Z_\mu\end{pmatrix} \\
(D^\mu\phi)^\dagger &= \frac{1}{\sqrt{2}}(0, \partial^\mu h) - i\frac{v+h}{2\sqrt{2}}(g\sqrt{2}W^{+\mu}, -\sqrt{g^2+g'^2}Z^\mu)
\end{aligned}$$

where the considered gauge fields are projected into the $W^{\pm}$, Z^0 and A basis. The first comment is that the photon field is not appearing in this expression this leads to the creation of a massless field, and results only from the choice of the ground states considered. The Lagrangian scalar term is therefore giving :

$$\begin{aligned}\mathcal{L}_{\phi} &= (D^{\mu}\phi)^{\dagger} D_{\mu}\phi - V(\phi) \\ &= \frac{1}{2}(\partial_{\mu}h)^2 - \frac{1}{2}2\cdot\mu^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 \\ &\quad + \frac{(gv)^2}{4} W^{+\mu}W^{-}_{\mu} + \frac{g^2+g'^2}{8} v^2 Z^{\mu}Z_{\mu} \\ &\quad + \frac{g^2 v}{2} h W^{+\mu}W^{-}_{\mu} + \frac{g^2+g'^2}{4} v h Z^{\mu}Z_{\mu} \\ &\quad + \frac{g^2}{4} h^2 W^{+\mu}W^{-}_{\mu} + \frac{g^2+g'^2}{8} h^2 Z^{\mu}Z_{\mu}\end{aligned}$$

The first line describe the scalar kinetic and self coupling as in the U(1) case, the second gives rise to the mass of the weak vector boson:

— the mass-degenerate W bosons: $m_{W^+} = m_{W^-} = m_W = \frac{gv}{2}$

— the massive Z boson: $m_{Z^0} = \frac{\sqrt{g^2+g'^2}}{2} v$

— the massless photon: $m_A = 0$

while the third and fourth lines are showing the linear and quadratic couplings of the higgs to the gauge bosons. The relationship between the Z and W masses also helps to get rid of the couplings g and g' with the weak angle:

$$\frac{m_W}{m_Z} = \frac{g}{\sqrt{g^2+g'^2}} = \cos\theta_W$$

These couplings are, however, linked to the Fermi constant G_F, the mass of the W boson and the electrical charge e. The Higgs v.e.v. in itself is also determined by the Fermi constant through the relation:

$$v = \sqrt{\frac{1}{\sqrt{2}G_F}} \simeq 246 GeV$$

The direct coupling of the scalar boson to the gauge fields from the third line is shown to be proportional to the square of their masses:

$$\begin{aligned}\mathcal{L}_{\text{coupling}} &= \frac{g^2 v}{2} h W^{+\mu} W^{-}_{\mu} + \frac{g^2 + g'^2}{4} v h Z^{\mu} Z_{\mu} \\ &= \frac{m_W^2}{v} h W^{+\mu} W^{-}_{\mu} + \frac{m_Z^2}{v} h Z^{\mu} Z_{\mu}\end{aligned}$$

Thus the masses of the gauge bosons are obtained from the interaction with a scalar field whose ground state is broken under the symmetry of the SU(2) × U(1) group. This field is also known as the Higgs field, and therefore the excitation of this field around the ground state h is called the Higgs boson which mass is not predicted by the SM since λ and μ are free parameters. The λ parameter giving rise to the trilinear and quartic self-coupling of the Higgs, which is still to be measured, is of particular interest for the physicists allowing to compute the full Higgs potential and thus understanding possible instabilities.

1.3.3 Concept of Yukawa lepton masses

The Higgs field can also couple to the fermions allowing to explain the mass of such particles in one global theory. Contrary to the gauge boson case, where the interaction arises from the covariant derivative term in the kinetic term of the scalar Lagrangian, the formulation here must be ad-hoc while preserving the symmetries of the total Lagrangian. The simplest form one could think for this term is:

$$\mathcal{L}_{\text{leptons}} = -Y \bar{L_L} \phi L_R + \text{h.c.}$$

which is now invariant under the SU(2) × U(1) symmetry since the chirality transformation of the L_L doublet is compensated by the presence of the Higgs field. The coupling Y also called Yukawa coupling is a parameter that is not fixed a priori and h.c. stands for the hermitian conjugate part of the expression. Expanding the scalar around its ground state gives the following result:

$$\begin{aligned}\mathcal{L}_{\text{leptons}} &= -Y(\bar{\nu_L}, \bar{l_L}) \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v+h) \end{pmatrix} l_R + \text{h.c.} \\ &= -\frac{1}{2} Y \sqrt{2} v \bar{l_L} l_R - \frac{Y}{\sqrt{2}} h \bar{l_L} l_R + \text{h.c.}\end{aligned}$$

The first term is directly a mass term for the leptons with $m_l = Y\sqrt{2}v$, showing the link between the Yukawa parameter and the mass of the lepton. The coupling with the Higgs is also noticed to be proportional to their mass.

In the case of the quark sector the term must involve a different prescription for the up-type quark to respect the U(1) charge conservation:

$$\mathcal{L}_{\text{quarks}} = -Y_d \bar{Q}_L \phi D_R - Y_u \bar{Q}_L \tilde{\phi} U_R + \text{h.c.}$$

where $\tilde{\phi} = i\tau^2\phi = \begin{pmatrix} \frac{1}{\sqrt{2}}(v+h) \\ 0 \end{pmatrix}$ has an opposite hypercharge. Deriving the expression around the ground state as for the lepton results in the same mass generation and coupling to the Higgs field. When considering the three families in the formulation, the Yukawa prescription can be extened in terms of matrices:

$$\mathcal{L}_{\text{quarks}} = -Y_d^{i,j} \bar{Q}_L{}^i \phi D_R^j - Y_u^{k,l} \bar{Q}_L{}^k \tilde{\phi} U_R^l + \text{h.c.}$$

where now $Y_d^{i,j}$ and $Y_u^{k,l}$ are 3×3 matrices generating similar mass matrices. From first principles nothing forces these matrices to be diagonal. However, the observed quarks are the eigenstates of the mass matrices. Transfer matrices can be computed such that:

$$\underbrace{u^i}_{\text{interaction states}} = U_u^{i,j} \underbrace{u'^j}_{\text{mass states}} \quad \text{and } d^k = U_d^{k,l} d'^l$$

This is particularly important when looking at the flavour changing currents such as in the electroweak theory:

$$\bar{u}_L^i \gamma^\mu d_L^i W_\mu^+ \rightarrow \bar{u}_L^i \gamma^\mu (U_u^\dagger U_d)_{i,j} d_L^j W_\mu^+$$

The $(U_u^\dagger U_d)_{i,j}$ is a non diagonal matrix also known as the Cabibbo–Kobayashi–Maskawa (CKM) matrix [47, 48] which allows terms that explain the violation of the charge and parity measurement in the electro-weak sector.

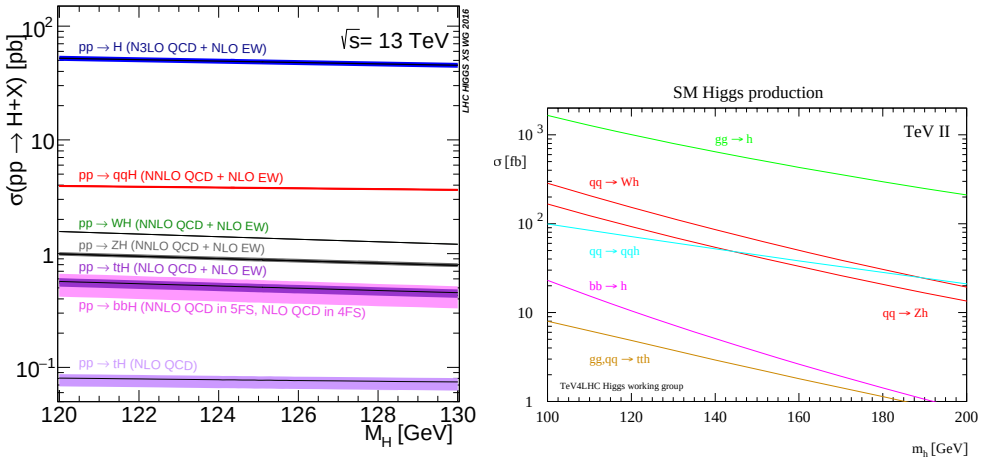
1.4 The Higgs boson at the LHC

The search of a Higgs boson that could prove that the Higgs mechanism is indeed the correct one to describe the mass of the elementary particles, was tried at the LEP and Tevatron colliders using direct searches and indirect constraints. The need for such accelerators was motivated by the lack of exact prediction for the Higgs mass, and thus a variety of energies needed to be scanned. Finally with the announcement of its discovery by ATLAS and CMS the 4^{th} of July 2012 at a mass of 126.0 ± 0.4 (stat.) ± 0.4 (syst.)

GeV ($125.3 \pm 0.4\,(\text{stat.}) \pm 0.5\,(\text{syst.})$ GeV) in the ATLAS (CMS) experiment, the SM jigsaw puzzle got completed. An age of precision measurement and verification of the predicted properties started. In this section, I describe the various production modes offered in the context of a proton-proton collider such as the LHC, and the decay modes allowing to observe such particle in this context.

1.4.1 Production modes

As the Higgs boson couples to all the massive particles, a variety of ways can be used to produce it. At the LHC, the elementary constituents of the protons (gluons, u and d valence quarks, sea quarks) are initiating the interactions. The modes are categorised by the initial states resulting in the Higgs boson, or by the final state created with the Higgs boson. The modes are ordered by their cross-section as shown in Figure 1.3a.



(a) (b)

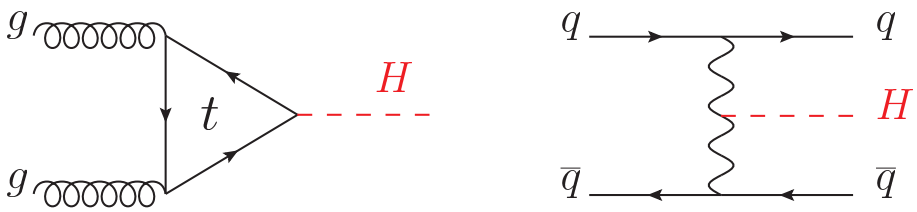
Figure 1.3 – Production cross-sections for a SM Higgs boson as a function of the Higgs mass: (a) at $\sqrt{s} = 13$ TeV at the pp LHC collider [49]; (b) at $\sqrt{s} = 1.96$ TeV at the $p\bar{p}$ Tevatron collider [50].

Gluon fusion: ggF

The dominant process at the LHC is happening through the coupling initiated by the gluons. This coupling is loop induced as shown from Figure 1.4a, as the Higgs is not coupling to the massless gluons. While this production should be rarer compared to tree-level processes, the important gluon density inside the protons compensates for this. Furthermore since the Higgs coupling to fermions is directly proportional to their masses, top- and bottom-quarks mediated creation is enhanced and can be found only thanks to loops or to the excitation of sea quarks (which have negligible densities inside the protons). This production mode is roughly 10 times larger than the second largest one. The feature of this production mechanism is also the absence of side-products in the final state.

Vector boson fusion: VBF

The VBF production mode is the second leading one and relies on the creation of off-shell vector bosons, which are coupled to the Higgs boson thanks to their high mass as presented in Figure 1.4b. Due to the structure of the collisions, the out-going quarks hold the initial longitudinal momentum of the initial ones, creating two jets located in the forward and backward regions of the detector, where the energy resolution is poorer. Therefore this clear detector signature can be used to disentangle this production from the ggF. In order to separate even more this signature from a standard QCD production, an extra radiated photon can be requested as done in [51].



(a) (b)

Figure 1.4 – Main Leading Order Feynman diagrams contributing to the Higgs production in the gluon (a), and vector boson fusion (b) production mode [28].

Associated production to a vector boson: VH

All the processes where the Higgs is radiated (Higgs-strahlung) from an initial particle which is present in the final state, are suppressed due to the low-amplitude of the generation of those initial particles. Amongst all the possible initial particles, the vector bosons are favoured due to the high coupling with the Higgs, as shown in the Leading Order (LO) diagram in Figure 1.5. This mode was particularly important for the LEP searches due to the small direct coupling of the Higgs to electrons and the absence of gluons and quarks, as well as at Tevatron where the $p-\bar{p}$ collisions were enhancing the VH production compared to others as shown in Figure 1.3b.

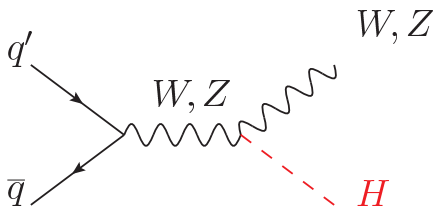


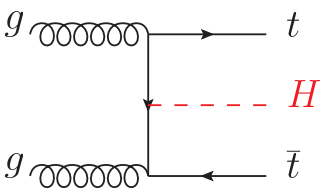
Figure 1.5 – Main Leading Order Feynman diagrams contributing to the Higgs production in the Higgs-strahlung (or associated production with a vector boson) production mode [28].

The production can be initiated with either the quarks or the gluons inside the protons. The Wh production can only be mediated with the former one, with an

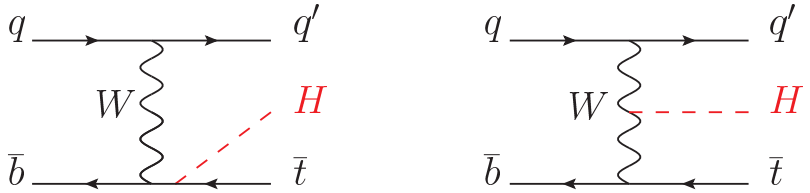
asymmetry between the W^+h and W^-h due to the different abundance of quark and anti-quark in the protons ($u\bar{d} \to W^+ / \bar{u}d \to W^-$). The Zh production is mediated at LO by both $q\bar{q}$ and gg diagrams, the former being at tree level as for the Wh ($u\bar{u}/d\bar{d} \to Z$), while the later goes through loops of quarks which is therefore quite suppressed in theory. However, as for the ggF mode, it is enhanced by the important gluon partonic distribution function PDF ($\sim 14\%$ of the total Zh production). Due to their difference in production mechanism, the ggZh production may be sensitive to the presence of new-physics particles in the loop, which would change the total cross-section. The Higgs spectrum is also impacted differently by the two modes which could be investigated when performing a momentum differential measurement. NLO and NNLO QCD corrections are applied to the system triggered by qq , $q\bar{q}$ and gg initial states with the possible radiation of extra partons in the final state, and are affecting all the production. The NLO corrections affect the cross-sections up to $\mathcal{O}(30\%)$ factor, while the NNLO corrections are at the per cent level. The NLO QCD corrections on the ggZh mode are still very challenging to compute because of the multi-scale boxes present in those corrections and are therefore obtained in an effective theory considering an infinite top-quark mass and vanishing b-quark one. Those corrections are quite important in magnitude, and therefore as often as possible the $q\bar{q}$ and gg modes are treated separated.

Production in association with heavy quarks

A small contribution to the total Higgs production cross-section comes from the direct coupling to fermions in the initial state as shown in Figure 1.6. Due to its important mass, couplings with tops are dominant, but couplings to bottom-quarks are also important since they have a higher PDF than tops.



(a)



(b)

Figure 1.6 – Main Leading Order Feynman diagrams contributing to the Higgs production in (a) associated production with a pair of top (or bottom) quarks, (b) production in association with a single top quark [28].

This mode is a unique way to understand the direct coupling to the top-quark as it only enters in loops for the other production modes. However, due to the high $t\bar{t}$ QCD production at the LHC, this channel is very challenging to tackle experimentally.

The summary of all the production mechanisms and their associated cross-section computed at the higher order available is given in Table 1.3.

Process	σ (pb)	$\frac{\sigma_i}{\sigma_{tot}}$ (%)	
ggF	48.6	87.2	
VBF	3.78	6.8	
Wh	1.37	2.5	4.0
$q\bar{q}Z$h	0.76	1.4	
gg$Z$h	0.12	0.2	
$t\bar{t}$h	0.51	0.9	
$b\bar{b}$h	0.49	0.9	
other	0.09	0.2	

Table 1.3 – Production cross-sections of the $m_h = 125$ GeV Higgs boson at the LHC for $\sqrt{s} = 13$ TeV [49].

1.4.2 Decay modes

While the production modes are essential to trigger specific topologies, the main difference for analyses is due to the final state generated from the Higgs decay. The lifetime of the Higgs boson is too short to be even measurable so that only the products of its disintegration are accessible to the experimentalist. Due to its mass, the Higgs boson can only decay to certain particles allowed by the kinematics. For instance, the $t\bar{t}$ threshold is not open at such mass, and in the case of di-boson, at least one of the two vector boson must be produced off-shell, reducing the probability for this decay. From the rules obtained in Section 1.3.2 and 1.3.3, the width of the decay of the Higgs boson in each of the channels can be derived:

$$\Gamma(\mathrm{h} \to f\bar{f}) = \frac{m_f^2}{v^2} N_C \frac{m_h}{8\pi} \left(1 - 4\frac{m_f^2}{m_h^2}\right)^{3/2}$$

$$\Gamma(\mathrm{h} \to VV^*) = \delta_V \frac{m_h^2}{v^2} N_C \frac{m_h}{32\pi} \left(1 - 4\frac{m_V^2}{m_h^2}\right)^{1/2} \left(1 - 4\frac{m_V^2}{m_h^2} - 12\left(\frac{m_V^2}{m_h^2}\right)^2\right)$$

where N_c is the colour factor ($N_c = 3$ when considering quarks, $N_c = 1$ for leptons), and δ_V a factor that normalises the number of diagrams giving the same final state ($\delta_W = 2$ while $\delta_Z = 1$). In order to interpret further on the probability of decay, a Branching Ratio

is defined as:

$$BR(\mathrm{h} \to XX) = \frac{\Gamma(\mathrm{h} \to XX)}{\sum_i \Gamma(\mathrm{h} \to ii)}$$

The value for the branching ratio for the different modes of disintegration as a function of the Higgs boson mass is given in Figure 1.7.

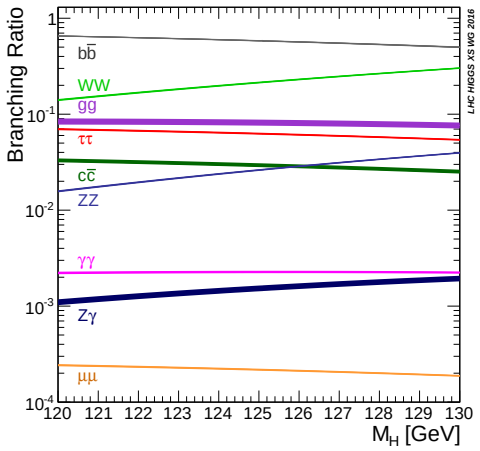


Figure 1.7 – Branching ratios for various decay modes of a SM Higgs boson, as a function of the Higgs mass m_h [49].

The $\mathrm{b\bar{b}}$ decay mode is the dominant one for the SM Higgs boson with a BR of roughly 58% and is linked in particular to the analysis presented in this thesis. Second in terms of BR is the WW decay. Depending on the decay chain of the W boson, this channel presents challenges for analysis: in the hadronic channel it is overwhelmed by the QCD background produced in a pp collisioner, in the leptonic channel neutrinos are emitted preventing to have a precise full-mass reconstruction. The coupling to leptons is accessible thanks to $\tau\tau$ channel, but it suffers from the same problems as for the WW channels since the τ decays are similar in signature to the W ones. It was first observed in 2016 thanks to an ATLAS and CMS combination [52]. The coupling to the second generation of leptons is obtained through the $\mathrm{c\,\bar{c}}$ decay, which suffers as the $\mathrm{b\bar{b}}$ from the QCD background, while not having good identification of c-quarks yet. First results have been presented by the ATLAS collaboration in the VH production channel [53] and the CMS collaboration [54] or through the direct coupling to the J/ψ mesons [55]. With a similar rate the ZZ^* decay, where both Z bosons decay into leptons, is one of the discovery channels that lead to the observation of the Higgs boson in 2012. Even if its BR is low, the small background contribution and the good lepton energy resolution of the search have made it the golden channel for the Higgs properties measurements. Finally, the coupling to the second generation of fermions could be measured also thanks to the di-muon channel. However, the reduced BR of this channel does not allow to get measurements yet.

The Higgs boson can also decay to massless particles, thanks to loops involving massive particles. As such, third in the order of BR, the final state involving gluons is quite significant at the LHC. However, similarly to the quark decays, it is overwhelmed

from the background. Therefore no searches have tried so far to look at this final state. On the contrary, the small BR $\gamma\gamma$ decay was involved as well in the discovery in 2012 with the ZZ^* decay. The search is characterised by a smooth falling background of di-photon on top of which a peak corresponding to the Higgs boson can be found. Thanks to the good energy resolution of the photons it is also used in the mass-measurement of the Higgs. With a similar BR, the $Z\gamma$ final state possessed all the good properties to give a clean signal but is penalised by the low branching ratio of the Z-boson to the leptonic decays.

2 The Large Hadron Collider and the ATLAS detector

In order to probe the elementary properties of the SM, for instance studying the coupling of the Higgs boson to quarks, high energy colliders have been developed with a variety of technical solutions. The principle of such machines is to bring stable particles as close as possible to each other, with a certain energy, such that the interactions described earlier can act on the system, resulting in the creation of new states decaying to more particles or the modification of the kinematical properties of the former particles. Colliders can be classified into two categories depending on the nature of the initial particles: leptons or hadrons. In order to probe the dynamics of the SM, as well as to reach the energy threshold of the creation of new particles, colliders need to be associated with an accelerating structure coupled to an optic system controlling the circulating beam. One of the most relevant collider parameters is the centre-of-mass energy of the system denoted $\sqrt{s}$. The beams cross each other at special locations called interaction points (IP), where detectors need to be installed in order to measure the properties of the emitted particles which will help the analysers to reconstruct the history of the event and understand the underlying processes.

The work developed in this thesis is based on the collisions recorded at the LHC collider located at the CERN facility in Geneva by the ATLAS experiment. The structure and properties of the accelerator are described in Section 2.1, while the description of the ATLAS detector is given in Section 2.2.

2.1 Large Hadron Collider LHC

The LHC is a multi-purpose collider, which uses protons and ions (Lead and Xenon) as primary particles for various sets of studies on the Standard Model as well as on the nucleus structure. Since the studies presented in this document are related to proton collisions, the ion part of the acceleration and collisions will not be described. The LHC program started in 1983 with the first studies and a first approval given by the CERN council in 1994. Due to budget reasons, the first project was conceived as a two stages accelerator with the first step at $\sqrt{s} = 10$ TeV before reaching a final $\sqrt{s} = 14$ TeV. This project was amended in 1995 with the Conceptual Design Report [56], where only the latter stage was proposed thanks to the improvement in accelerating technology, and with the failure of the Superconducting Super Collider (SSC) in the United States of America, which led more nations to contribute to the budget of the LHC. The main success of the LHC was to duild upon the existing accelerating and civil engineering structure of the Large Electron Positron (LEP) collider.

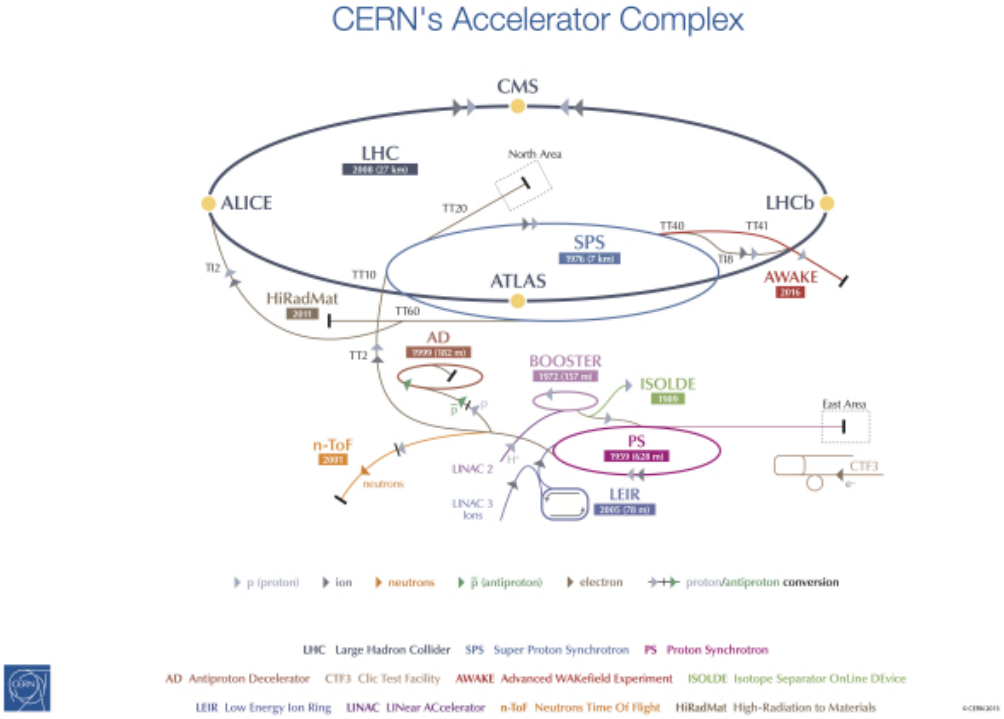


Figure 2.1 – Overview of the CERN accelerator complex [57].

The collider is composed of a cascade of different accelerators in order to reach the LHC accelerator as shown in Figure 2.1, which are also used for other purposes serving various collaborations located at CERN. The initial protons originate from bottled di-hydrogen. The hydrogen atoms are ionised using a powerful electrical field, which yields nearly 70% of protons. The outgoing protons already have a kinetic energy of 100 keV and are further accelerated up to 50 MeV. This first stage is integrated into the linear accelerator LINAC2 which is a 37 m long machine providing an excellent beam availability (higher than 98%), with the highest current reached by such technology so far (180 mA). This stage is ongoing a full replacement with a more mature technology using hydrogen anions H^- and state of the art accelerating cavities, which will allow more compact structures of proton bunches and a higher exiting beam energy of 160 MeV [58]. This new linear accelerator called Linac4 is currently under-test and is going to be operated for the next LHC runs from 2021.

The outgoing protons are then transferred to the 157 m long Proton Synchrotron Booster which is responsible for accelerating the protons to an energy of 1.4 GeV before injecting them to the 628 m long Proton Synchrotron that gives a final 25 GeV beam. Those two accelerators have circular shapes in order to repeat several times the same accelerating step to reach higher energies. The grouping of the protons into small bunches of roughly 10^{11} particles separated by a distance of 25 ns, as required by the LHC collision scheme is carried out by the PS. The beam is then sent to the Super Proton Synchrotron (SPS) which gives an extra boost to reach a final energy of 450 GeV using a 7 km ring. Accelerated protons are also sent to different experiments such as the North Area test-beam hall described in Section 7.2.2, which recieves only 0.084% of the initial protons to be sent to the LHC [59]. In order to provide the correct proton-proton collision scheme, the bunches are sent to two counter-propagating lines circulating in the 27 km long LHC tunnel.

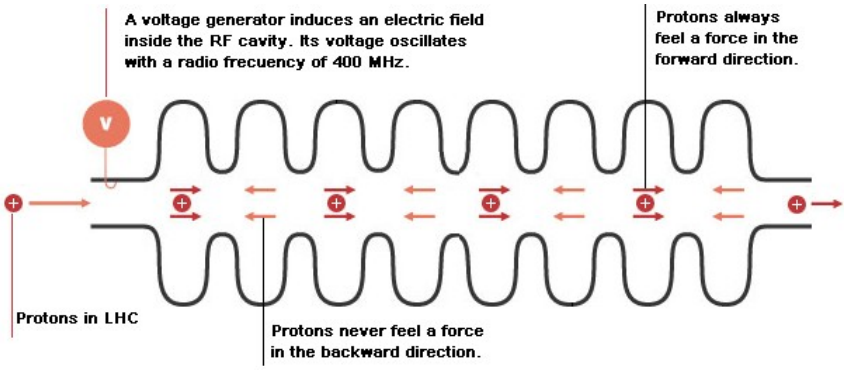


Figure 2.2 – Schema of the acceleration of a proton through an RF cavity. The frequency of electrical field is tuned such that proton bunches are always synced to an accelerating field [60].

The full acceleration scheme of the LHC is based on the two ingredients of the electromagnetism reflected in the Lorentz force: electrical fields for longitudinal accelerations and magnetic fields to act on the transversal motions. The acceleration techniques used in the LHC chain of acceleration is based on Radio-Frequency (RF) cavities whose

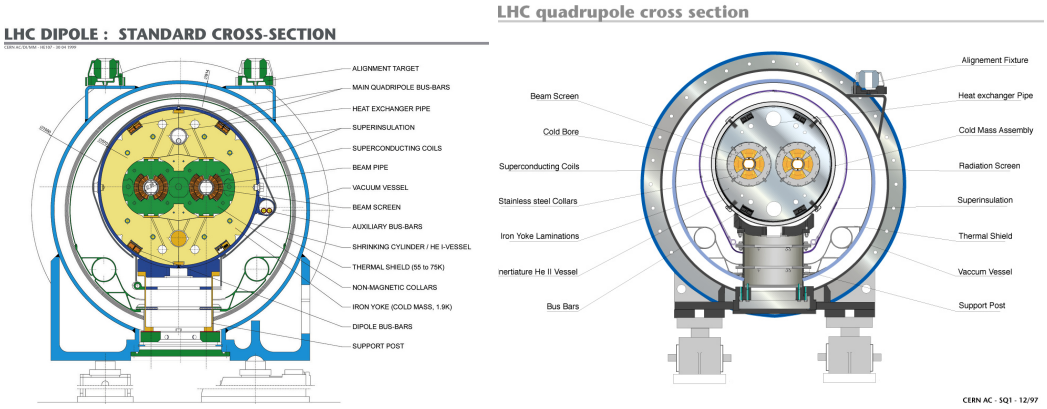
principle is presented in Figure 2.2. The transverse motion needs to be controlled in two different ways. The limitation in energy is set by the magnetic field, B, available and the radius of the circle R, the Lorentz force compensating the centrifugal one:

$$F_{Lorentz} = evB = \frac{\gamma m v^2}{R} = F_{centrifugal} \rightarrow \frac{p}{e} = BR$$

where p is the momentum of the charged particle. Because of the counter-propagating motion of the beams in a particle-particle collider, the direction of the magnetic field for the bending should be created with opposite directions. Due to the small size of the tunnel, the two pipes should be as close as possible leading to the solution presented in Figure 2.3a. A total of 64 cables are assembled around two coils shaping the magnetic field lines. In order to create a large enough field to reach energies of 7 TeV per beam, an 8.3 T field is needed which corresponds to an 11.7 kA current circulating in the circuit. In order to avoid heating from the Joule effect, each line is made of 6500 superconducting filaments of Niobium-Titanium. The whole system has thus to be cooled down to 1.9 K using superfluid Helium. Each magnet is 14.3 m long and the LHC structure is made of 1232 of them. Bending the trajectories is not enough to keep the beam circulating since bunches are made of same-charge particles, Coulomb interactions perturb the transverse shape. As for the propagation of light, magnetic lenses are used to focus the beams, consisting of magnetic quadrupoles as presented in Figure 2.3b. The principle difference with the motion of a light beam is that when a plane (horizontal or vertical) is focussed, the other one gets defocussed. The usual system is a composition of those two magnets into a FODO system (where F/D focusses vertically/horizontally the beam and O is a space deflection magnet) resulting in a 110 m long cell. The global effect of such cells is, therefore, to bend and focus the beam. Other types of multipoles are also used in the LHC chain to compensate for vertical and horizontal size correlations, as well as to correct for other effects (electron clouds effects, interactions between bunches ...). The FODO cells are grouped into 2.45 km long arcs with 545 m long straight sections of accelerating RF cavities into eight independent sectors which are powered in series. This structure was motivated by the pre-existing tunnel of LEP.

In addition to accelerating and bending the beams, the LHC is also responsible for approaching the beams as close as possible in four IPs. Before entering the IP, the beam is defocussed from a 0.2 mm transverse size to 1.6 mm. This allows the use of wide-aperture and high gradient quadrupoles called inner-triplets to focus the beam to a final size of $\sim 10\mu$m. The major figure of merit of the collision is called the instantaneous luminosity $\mathcal{L}$. It is defined as the factor linking the number of events N produced per second and the cross-section σ of the process:

$$\frac{dN}{dt} = \sigma \mathcal{L}$$



(a) (b)

Figure 2.3 – Cross section of the superconducting dipole [61] (a) and quadrupole [62] (b) magnets used in the LHC ring.

This factor is directly linked to the collision scheme provided by the accelerator machine:

$$\mathcal{L} = \frac{N_p^2 n_b f_{rev} \gamma_r}{4\pi\sigma_1\sigma_2} F$$

where N_p is the number of protons per bunches, n_b the number of bunches circulating in the collider, f_{rev} the revolution frequency of the beam, γ_r the relativistic factor associated to the particles in the bunch, $\sigma_{1/2}$ is the transverse width of the bunch 1/2 under the gaussian hypothesis, and F the geometrical factor of the collision scheme. In the accelerating community two quantities defined as the normalised transverse beam emittance ϵ_n and amplitude function β^* are widely used instead of the transverse width in order to derive the transport equation of the beam. The emittance is a measurement of the beam spread in the position-momentum space. As this quantity is proportional to the momentum of the beam, a normalised quantity can be defined when this effect is taken out. The beta function is defined as the relationship between the emittance and the actual beam size $\sigma = \sqrt{\beta\epsilon}$, and is called β^* when taking the smallest value at the IP. While the emittance is conserved when considering the effect of magnetic quadrupoles in order to respect Liouville's theorem, the modification affects the β^* quantity. The law can thus be written as:

$$\mathcal{L} = \frac{N_p^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n\beta^*} F$$

The geometrical factor F, called the Piwinski reduction factor, takes into account the fact that collisions are not happening head-on but with a small opening angle θ_c:

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sqrt{\epsilon_n \beta^*}}\right)^2\right)^{-1/2}$$

where σ_z is the longitudinal width of the bunch. Several ideas have been proposed to increase this factor such as the Crab cavities that would allow kicking the bunch to

increase the overlapping region of interaction in a head-on type collision. Usual values for the designed instantaneous luminosity were of the order of $1 \times 10^{34} cm^{-2}.s^{-1}$, whereas peak values observed in 2018 were approximately of $2.1 \times 10^{34} cm^{-2}.s^{-1}$ under normal running conditions. This improvement was obtained thanks to a lower β^* value (25 cm compared to a nominal value of 55 cm), as well as a lower emittance (1.8 μm compared to a nominal value of 3.75 μm), while not reaching the final energy of $\sqrt{s} = 14$ TeV and having less bunches than expected (2556 instead of 2808). For comparison the peak instantaneous luminosity was $\sim 1 \times 10^{32} cm^{-2}.s^{-1}$ [63] at LEP, and reached $4 \times 10^{32} cm^{-2}.s^{-1}$ with the Tevatron collider [64].

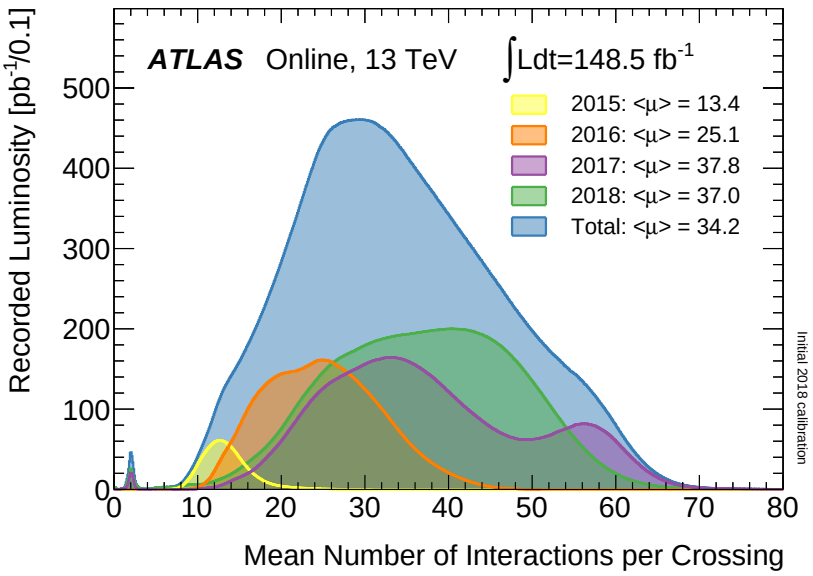


Figure 2.4 – Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the different years of operations during Run-2. The mean number of interactions per crossing is the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch [65].

The increase of the luminosity gives an increase of the number of pp interactions: at each bunch crossing, several collisions can happen increasing the probability to observe rare and interesting events. However, the surrounding soft collisions are increasing the complexity of the reconstruction of the hard-scatter objects and worsen the resolution of the quantities defined in Section 3.2. This effect is called pile-up and can be caused by particles from the same bunch-crossing (in-time), or due to remnants of the interactions of the previous bunch crossing (out-of-time). The pile-up is closely related to the mean number of interactions $\langle\mu\rangle$ represented in Figure 2.4, where the mean value for the Run-2 is approximately 34. The future evolution of the LHC, described in Chapter 7, will have to face a mean number of interactions of roughly 200 with consequent experimental challenges to reconstruct the events.

The luminosity of the LHC is decaying with time due to the loss of particles in the collisions, the interaction with the remaining gas molecules in the tube, as well as the

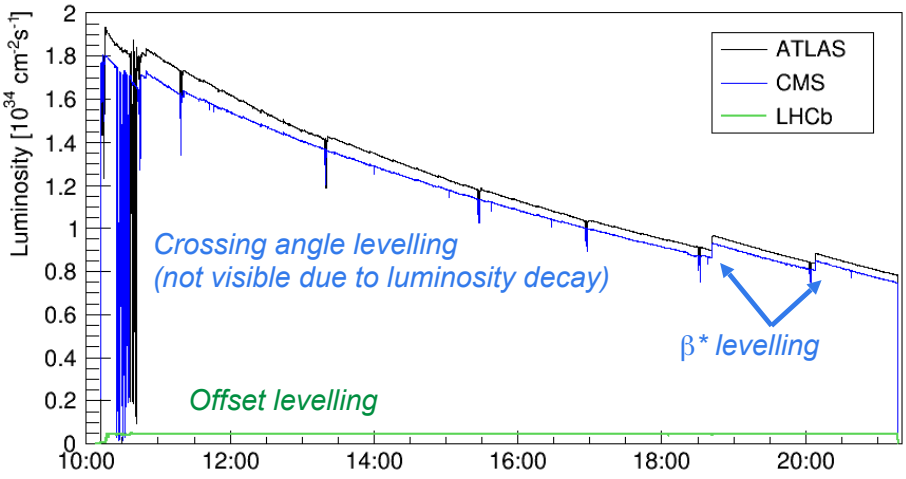
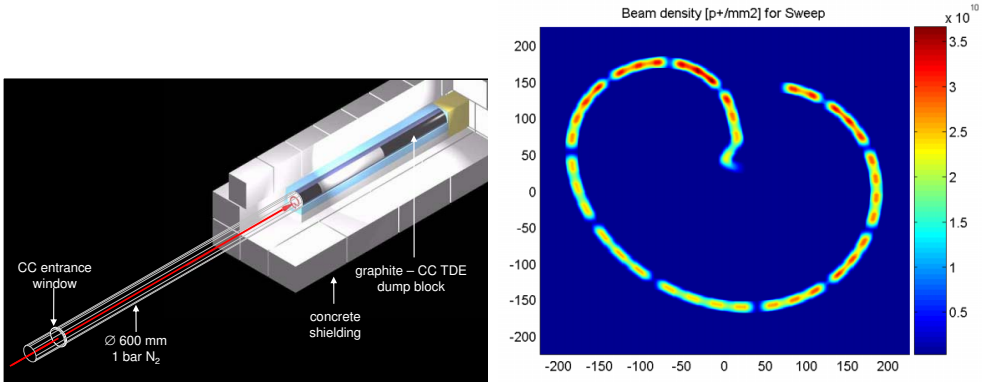


Figure 2.5 – Evolution of the luminosity during a typical LHC run, measured in three of the interaction points of the LHC [66].

degradation of the emittance in the process due to the Touschek effect [67]. A levelling of the luminosity is used throughout the fill in order for it to last longer and to hold a similar collisions status. This is done by modifying the crossing angle and β^* or with an offset levelling of the bunch collisions. The situation is presented in Figure 2.5. The typical run time of an LHC fill is 10 hours, while the filling of the machine and the ramping of the energy take approximately 20 and 25 minutes respectively. At the end of the run, when the luminosity is too low, the beam is dumped. The beam is extracted from the main ring and sent to a 6 m^3 graphite cylinder (TDE) that can heat up without melting and is kept under a small overpressure of nitrogen. A schematic view of the system is presented in Figure 2.6a. In order to dilute the power of the beam a set of horizontal and vertical dilution kickers are used to sweep the beam into an e shape as shown from Figure 2.6b.



(a) (b)

Figure 2.6 – (a) Schematic view of the dumping block structure of the TDE. (b) Nominal sweep profile density (p^+/mm^2) at the TDE entrance window [68].

Around each collision point seven experiments are located in order to record the properties of the particles emitted during the interactions:

- **A Toroidal LHC ApparatuS (ATLAS)** [6]: The experiment has an almost full coverage with a general physics purpose on which the work presented here is based. A full description is given in the next section.

- **Compact Muon Solenoid (CMS)** [7]: located at the other side of the ring compared to ATLAS, the CMS detector is also targeting a wide range of physics processes. It relies on a different conception compared to ATLAS and the analyses are conducted by different teams to avoid systematical effects that could lead to biases.

- **LHCb** [69] : This experiment is dedicated to flavour physics measurements and CP violation studies. Contrary to the two previously cited experiments, it is designed as a very forward detector.

- **A Large Ion Collider Experiment (ALICE)** [70]: This detector was designed to study the structure of QCD, focusing on the quark-gluon plasma state that happens at high temperatures in heavy-ion collisions as well as in pp collisions.

- **TOTal Elastic and diffractive cross-section Measurement (TOTEM)** [71]: as for the next experiments, the TOTEM detector is not directly located on one of the four IP of the LHC. It is used to measure proton-proton interaction and diffraction cross-sections. It relies on two sets of detectors placed on both sides of the CMS interaction point.

- **Large Hadron Collider forward (LHCf)** [72]: Similarly to the position of TOTEM, the LHCf experiment targets forward particles from the ATLAS interaction point. Its physics case focusses on the creation of ultra-energetic neutral pions whose production cross-section and energy can be used to calibrate the perturbative QCD models used for simulation of air shower experiments such as AUGER [73] or HESS [74] looking at the GZK cut-off. Compared to the UA7 [75] experiment using the SppS accelerator, the equivalent fix-target energies of pions probed at the LHC have three orders of magnitudes higher reach 10^{17} eV.

- **Monopole and Exotics Detector at the LHC (MoEDAL)** [76]: sharing the cavern of the LHCb experiment, it looks for magnetic monopoles and other highly ionising massive particles.

It must be noted that contrarily to what ATLAS and CMS are requesting, the ALICE and LHCb experiments are running with a smaller instantaneous luminosity and are located at the old positions of the L3 and Delphi experiments respectively, while ATLAS and CMS are high luminosity experiments located at nearly two opposite sites (to avoid beam resonances) in new caverns that were built for that purpose.

2.2 The ATLAS experiment

The goal of the ATLAS experiment is to look for an important variety of signatures, from SM measurements to beyond-SM discoveries in the context of the ultra-energetic collisions delivered by the LHC. While most of the searches are conducted on the pp collisions, data obtained with heavy-ions also provide useful cross-checks to the ALICE experiment. In order to tackle these very different aspects, the detector was constructed with an onion shape, with each layer providing different types of measurements as presented in the following sections.

2.2.1 The detector structure

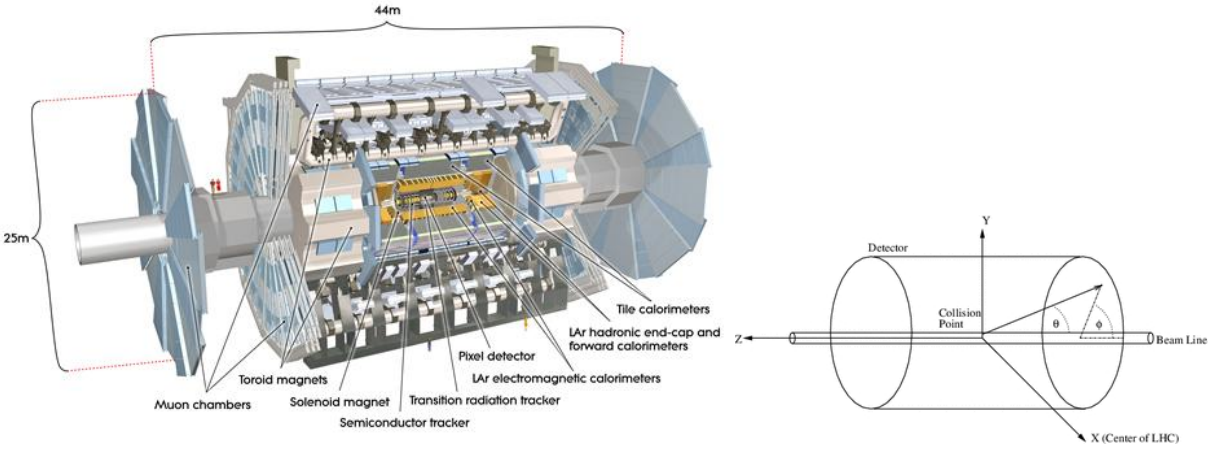
In order to increase the geometrical acceptance and the hermiticity, the detector was designed with a cylindric shape, as shown from Figure 2.7a, around the beam pipe with a diameter of 25 m and a length of 44 m, weighing approximately 7.000 tons. From its shape, a right-handed coordinate system is defined. A natural z-axis is defined as the axis of the beam pipe. The transverse plane at the collision point is defined by the x- and y-axes completing the system, the former pointing towards the centre of the LHC ring, the latter pointing upwards, defining as such the direction of the z-axis. The detector is thus divided in two parts, the A-side being defined by the positive z, whilst the C-side by the negative z. In this system of coordinates, a vector is defined with respect to two angles: the azimuthal angle ϕ from the projection in the transverse plane, while the polar angle is defined with respect to the z-axis. This representation is shown in Figure 2.7b. From the relativistic theory the rapidity can be defined as:

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$$

where the 3-momentum vector can be written $\vec{p} = (p_{\mathrm{T}}\cos(\phi), p_{\mathrm{T}}\sin(\phi), p_z)$, and the $p_{\mathrm{T}} = ||\vec{p}||\cos(\theta)$ is called the transverse momentum. In the massless limit ($||\vec{p}|| \gg m$), the pseudo-rapidity η is introduced:

$$\eta = \lim_{\frac{m}{||\vec{p}||}\to 0} y = \frac{1}{2}\ln\left(\frac{||\vec{p}||+p_z}{||\vec{p}||-p_z}\right) = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$$

The particular interesting property of the pseudo-rapidity is that any difference between two objects is Lorentz-invariant under a boost in the z-direction. From this the angular distance in the (η, ϕ) plane between two objects is defined as $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.



(a) (b)

Figure 2.7 – (a) Open view of the ATLAS detector with the relevant sub-detectors mentioned [77] (b) Representation of the ATLAS system of coordinates [78].

The apparatus is designed to host several types of sub-detectors that would interact differently with the particles crossing them, measuring their properties such as their energy and position. Figure 2.7a is showing the structure of the detector, while Figure 2.8 focusses on the interaction of the outgoing particles with the sub-detectors:

— **charged particles momentum and position:** In order to measure the momentum of the particles a large magnetic field is generated by the magnet system (see Section 2.2.5) which bends the trajectory of the charged particles due to the Lorentz force. The measurement of the curvature is critical to the charge and momentum determination. The ATLAS detector uses modern silicon and gaseous trackers located close to the interaction point in the so-called Inner Detector (ID, see Section 2.2.2), as well as in the outer part of the detector to measure the properties of the muons (see Section 2.2.4). These detectors are also used to give a precise indication of the position of the tracks allowing the primary vertex of the event to be reconstructed to reduce the effect of pile-up, as well as the possible secondary vertex from particles decaying inside the volume of the detector such as the B-hadrons. The coverage of the ID system is however limited to the central region ($|\eta| < 3.0$).

— **energy measurement:** In order to complete the measurement of the charged particles properties, as well as detecting neutral particles, calorimeters are used to measure energies. The goal is to stop the particles by nuclear or electromagnetic processes, which result in the creation of a cascade of new particles producing a characteristic signal that is used to measure the energy.

While the muons leave signals in the tracking systems, some weakly interacting particles such as neutrinos or possible BSM particles do not leave any signal in the detector. Several indirect quantities such as the $E_{\mathrm{T}}^{\mathrm{miss}}$ described in Section 3.2.3 are defined to detect the presence of such particles.

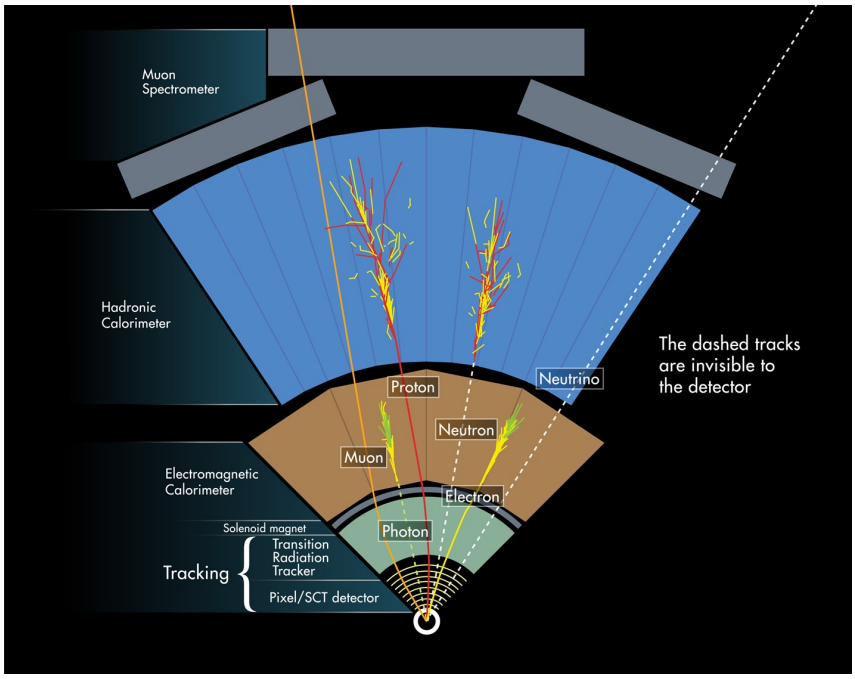


Figure 2.8 – Schematic view of the interaction of different particles with the ATLAS detector [79].

2.2.2 The Inner Detector

Covering the innermost part of the ATLAS apparatus, the Inner Detector (ID) [80], is used to measure the trajectory of charged particles. From this measurement, algorithms are designed to reconstruct the tracks of the particles as the helix parameters, and compute various properties:

— The transverse momentum: due to a uniform 2 T magnetic field generated by a solenoid (see Section 2.2.5), the p_{T} is measured from the curvature of the tracks. A lower threshold of 0.1 GeV is obtained for the tracking system.

— The sign of the charge of the particle: the direction of the curvature helps to discriminate between positively and negatively charged particles.

— Vertex identification: thanks to a good resolution for the transverse and longitudinal impact parameter of tracks, d_0 and z_0, the reconstruction can precisely determine the origin of the vertex from which the particles originate from.

— Electron/pion identification: due to the charged tracks measured in the tracker, the deposits in the calorimeters left by electrons and charged pions can be disentangled from photons and neutral pions, in the low p_{T} regime (0.2 to 150 GeV). The discrimination between the two type of particles is also obtained due to the Transition Radiation Tracker (TRT) described in Section 2.2.2.2.

The inner system itself is composed of three subsystems as shown in Figure 2.9, that covers hermetically the region in ϕ, while only covering the region up to $|\eta| < 3$. A system of solid-state detectors has been chosen to be closest to the IP, with a pixel solution for the first layers, while a strip design was chosen for the later ones. The rest of the volume (almost 75%) is occupied by a gaseous technology.

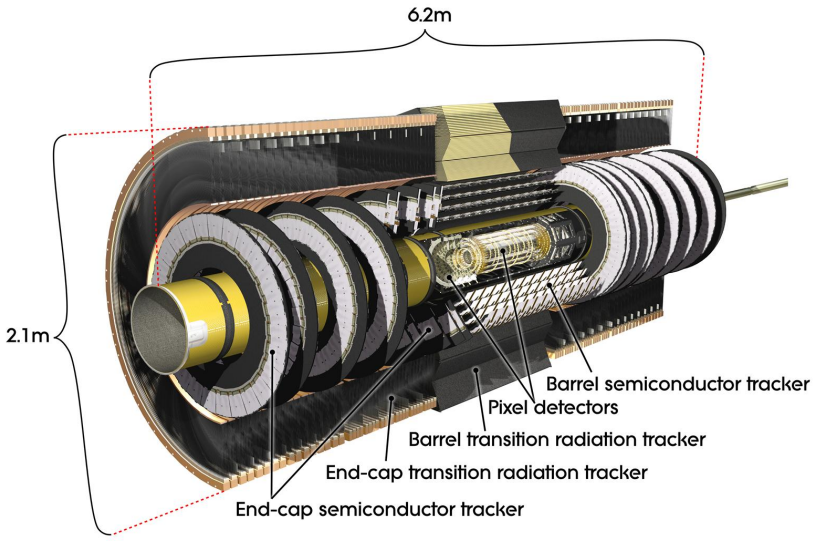


Figure 2.9 – Open view of the entire Inner Detector [81].

2.2.2.1 The solid-state detectors

The solid-state detectors installed in the innermost tracker layers are based on semiconductor technology, which is heavily used in HEP for detector design, as well as in industry for a variety of electronics applications. The generation of a readable signal is based on the creation of a PN junction. A pure silicon crystal has a gap energy of 1.12 eV, and electron-hole pairs are continuously created by thermal processes at room temperature. Their number would largely exceed any signal generated by the ionisation of the crossing of a charged particle. However impurities can be introduced in a controlled way in the bulk of the semiconductor material in order to create new energy levels in the gap. Usually, impurities are atoms with one electron more (donors) or less (acceptors) than silicon in the outermost orbital levels, creating respectively n-doped and p-doped substrates, modifying the Fermi energy level. A PN junction is obtained when n-doped and p-doped materials are juxtaposed, creating a controlled diffusion of carriers, as shown

in Figure 2.10. An equilibrium zone depleted of free charges, called the depletion zone, is reached, creating an electrical potential difference. This zone can be extended by applying an external bias voltage. When a charged particle passes through the depletion zone, it ionises the material, creating electrons and holes. Due to the electrical field in the region, the charges are drifting and diffusing. This motion creates an electrical signal in the electrodes thanks to the Ramo theorem [82], which is then transmitted by the read-out system.

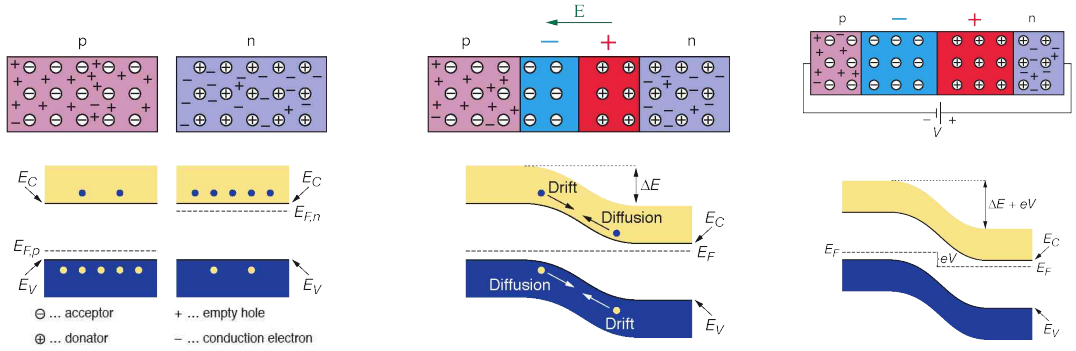


Figure 2.10 – Illustration of the PN junction with two separated n- and p-doped materials (left), the creation of the junction and appearance of the depletion zone (middle), the application of the bias voltage and the expansion of the depletion zone (right) [83].

Pixels

At the closest distance of the IP, a pixel detector is used where the flux of particles is particularly large, and thus where a radiation-hard detector is needed (see Section 7.2.1), with an extremely good spatial segmentation to avoid hits sharing tracks. The detector geometry is optimised to reconstruct tracks coming from the interaction region: the layers are parallel to the beam pipe in the barrel while disk shapes are built in the end-caps. These shapes are optimised such that the particles hit the surface at a perpendicular angle, reducing the amount of material traversed, and thus energy loss and multiple scattering. It helps also to increase the number of layers crossed at high η and thus the number of points available to extract the track parameters. The coverage of such a system in the present detector is extended up to $|\eta| < 2.5$, thanks to the disks shapes. The modules composing the system are n-on-n junctions where the pixel implants are much heavily doped with the n-type impurities with respect to the sensor bulk. The in-pixel pitch is designed to be $50 \times 400 \mu m^2$, while the thickness is 250 μm, leading to a 12 μm resolution in the r-ϕ plane and a 115 μm in the z-direction. The 80.4 million pixels, covering a total surface of 1.7 m^2, are linked to FEI3 [84] read-out chips, designed to resist the important radiation doses (100 Mrad) of such system foreseen at the different layers over the next few years.

The inner-most layer called Insertable B layer (IBL) [85] was added during the long shutdown of the LHC between the Run-1 and the Run-2. The main reason for this additional layer compared to the original plan was the increased expected luminosity

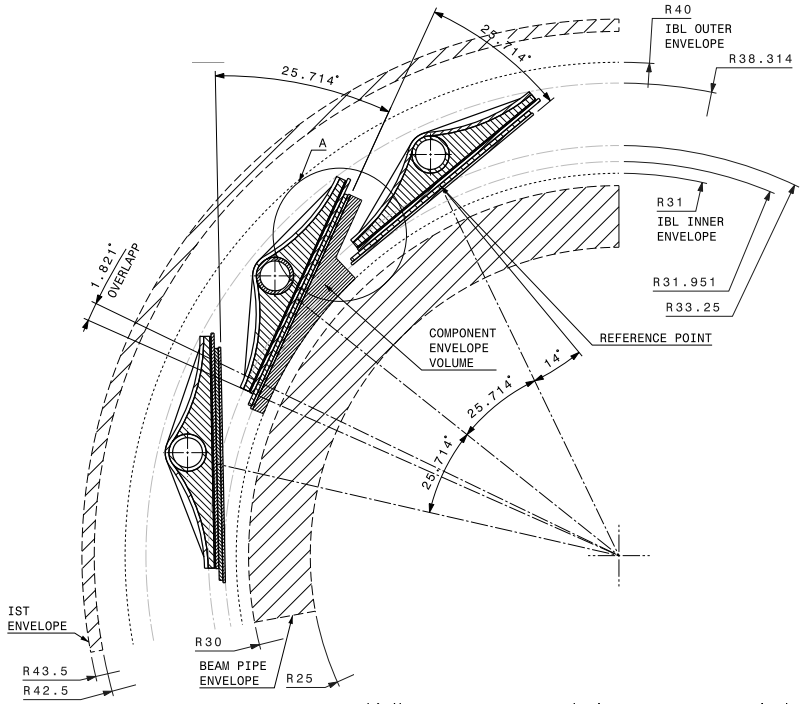


Figure 2.11 – IBL layout: rϕ view [85].

and radiation level between the two runs of the LHC, which would degrade the tracking performance of the system due to radiation damage as well as the increased pile-up, resulting in poorer b-tagging performance. Furthermore, adding a layer at a radius of 33 mm results in a higher identification capability of secondary vertices, and thus increases the b-jets identification power at low and medium jet p_{T} [86]. The motivation was also triggered by the increasing number of failures observed in the pixel system. The IBL layer consists of 14 staves tilted in the transverse plane with an angle of 14° in order to minimise the Lorentz angle [1] effect while having a $\Delta\phi = 1.8°$ overlap of the modules as shown from Figure 2.11. The extension of the IBL system allows to cover up to $|\eta| < 2.9$ of the phase space available, increasing the performance of the tracking. Due to a new generation of sensors, the staves are using 24 planar modules in the central position and 4 3-D sensors at each extremity of the staves. The 3-D sensors rely on a different R&D structure of the PN junction as explained in Section 7.2.1 with a n-on-p doping structure, while the planar sensors are based on the n-on-n technology used for the rest of the pixel layers. The thickness of the sensor are ranging between 200 and 230 μm, whilst the pixel size has been shrunk from $50 \times 400 \mu m^2$ to $50 \times 250 \mu m^2$ in order to increase the track separation power, leading to a 10 μm resolution in the r-ϕ plane and a 100 μm in the z-direction. A new readout chip called FEI4 [87] has also been designed to encompass the higher radiation dose expected at this position. Finally, the beam pipe was replaced and reduced in radius to insert the system.

1. Inside a magnetic field the charges generated in the sensors are drifting in a direction not parallel to the electrical field direction, thus leading to a small deviation of the observed position of the particle crossing. This effect is compensated a posteriori.

The overall system is cooled down to -13°C using an evaporative C_3F_8 cooling system, while the IBL system is using CO_2.

Semi Conductor Tracker (SCT)

Because of the cost of big surfaces of pixels, and complexity to maintain such systems, a stripped detector based on the same n-on-n technology was designed to instrument the above layers. The 63 m^2 surface uses 4088 double-sided modules organised in 4 layers in the barrel and 18 end-cap disks on the two sides covering up to $|\eta| < 2.5$ of the phase space. The steered angle between the strips is designed to be 40 mRad in order to benefit from a 2-D measurement to get an unambiguous position in the direction of the strips. The inter-strip distance is 80 μm resulting in a 17 μm resolution in the r-ϕ plane and a 580 μm in the z-direction.

2.2.2.2 The gaseous detector: Transition Radiation Tracker (TRT)

In the outermost part of the inner-detector, a gaseous tracker is installed. It relies on the timing of a charge drifting in the Transition Radiation Tracker (TRT) [88]. When a charged particle is passing through a gas mixture (70% Xenon, 27% CO_2, 3% O_2), electrons and ions are created. In practice, this mixture is filling straw tubes at low pressures (5-10 mb), and an electrical field is created between the border of the straw and a 31 μm wide tungsten gold plated wire located at the centre. The drift time of the electrons gives information on the position of the particle which allows spatial resolution of the order of 130 μm in the r-ϕ plane.

Furthermore, the space between the straws is filled with polypropylene fibres/foils so that the difference in reflexive index between the materials can generate transition radiations photons. Since the probability of this generation is linked to the relativistic γ factor of the particle and thus its mass, a discrimination between electrons and charged pions can be performed [88]. All the properties of the transition radiation detection are summarised in Figure 2.12.

As in the case of the previous detectors, the conception differs in the barrel from the end-caps. In the central part, almost 5000 tubes are set parallel to the beam pipe with a length of 150 cm and a radius of 2 mm, while straws are arranged perpendicular to the pipe in a fan-like arrangement in the end-caps. This organisation helps to have roughly 35 hits for the tracks improving the measurement of the p_T of the particles. However the coverage is limited to the $|\eta| < 2$ region.

2.2.3 The Calorimeters

Detecting the energy of the particles is the next fundamental step in the reconstruction of the event. It is necessary to integrate the measurement of the tracker

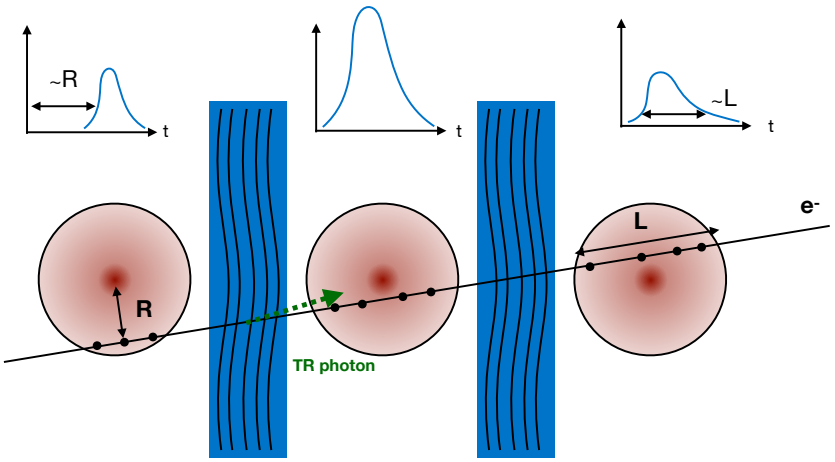


Figure 2.12 – Schema explaining the functioning of the TRT detector. The blue square represents polypropylene fibres/foils generating the transition radiation photons (in green). The exploitation of the signal shape can give information on the crossing of the charged particle.

to differentiate the neutrals from the charged-particles. This information is used to reconstruct the invariant mass in decays and to measure the $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from weekly interacting particles. The principle behind the detection is to stop particles, measuring the energy released by this process. The calorimetry is the field in HEP grouping the different techniques implementing these measurements and is not fundamentally different from the thermodynamics since the first HEP calorimeter used to measure a difference of temperature. Ionisation, nuclear reactions and atom excitations are triggered in showers generated by incoming particles, giving rise to secondary particles and degrading their initial energy. The shower development is deeply linked to the nature of the particle, as well as to the materials used. Interactions can be separated into a electromagnetic component (where bremsstrahlung and photon conversion effects dominate), and an hadronic one (where nuclear inelastic scattering dominates). Therefore separate detectors are constructed to enhance those effects, allowing a discrimination of the particles depending on whether they interact more with the former or the latter. The probability for such an interaction to happen is given by:

$$P(x) = e^{-\frac{x}{L}}$$

where x is the distance from the last interaction, and L the typical interaction length of the process denoted as X_0 for electromagnetic showers, and λ for hadronic ones. The length of the shower, therefore, depends on this parameter but also on the energy of the particle in a logarithmic scale: the 95% containment of the shower is given by the length $t_{95\%}$:

$$t_{95\%}^{EM} = \ln(\frac{E_0}{E_C}) + C_j + 0.08 \cdot Z + 9.6 \text{ in units of } X_0$$

$$t_{95\%}^{HAD} = 0.2\ln(\frac{E_0}{1\,GeV}) + 0.7 + 2 \cdot E_0^{0.13} \text{ in units of } \lambda$$

where E_C is the critical energy defining the point where losses by ionisation and by bremsstrahlung are equivalent, and C_j a factor that is ± 0.5 depending if the particle is an electron or a photon. Hence it is important to get a small radiation length to design compact detectors and measure the full deposit inside the detector volume. It must be noted that the development of hadronic showers is much more complicated than the electromagnetic ones since it includes inelastic interactions producing a variety of new particles including electrons/photons/pions that would develop an electromagnetic shower as well. A fraction of the energy of hadronic showers is also not observed due to the binding energy of released hadrons which can represent up to 40% of the total energy. Thus a good model of the reaction is necessary to make a precise measurement, nonetheless resulting in a poorer resolution for hadronic calorimeters compared to electromagnetic ones.

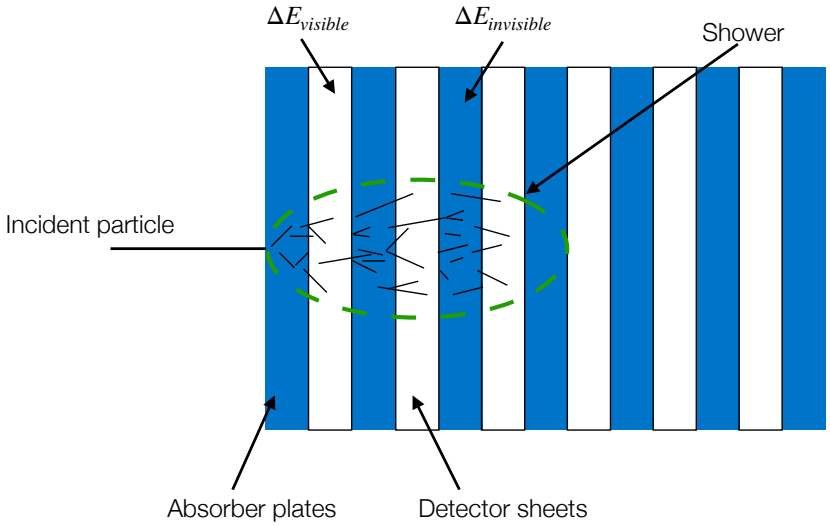


Figure 2.13 – Schema explaining principle of sampling calorimetry in HEP. The total absorbed energy is $E_{absorbed} = \sum \Delta E_{visible} + \sum \Delta E_{invisible}$. Freely interpreted from [89]

The concept of the calorimeters is based on the technology to generate the shower in absorbers and to detect the emitted particles in an active media. Two solutions are used, either using one single material for the two effects or decoupling them. The former is known as homogenous technique and is used for instance in the CMS electromagnetic calorimeter [90], which uses high-density scintillating crystals. However it is rather complicated to design compact detectors that would contain the full shower development with such materials, and they can only be used for electromagnetic showers as the radiation length is roughly one order of magnitude higher for hadronic interactions. The decoupling is performed in the ATLAS detector case by adding layers of high-density materials within the active material that is used to measure the energy as presented in Figure 2.13. It helps to reduce actively the detector size, and measure the shower development, while only allowing to measure a fraction of the total energy (around 30% for the ATLAS electromagnetic calorimeter) thus degrading the resolution in energy. Test-beam measurement and a good understanding of the detector with dedicated correction factors are required not to bias the energy measurement of such deposits. The particles created then cross the active material where the signal is detected and sent to an acquisition chain.

The calorimeter should have the largest possible coverage in order to ensure a good measurement of the missing energy which is crucial in analyses such as $Vh(b\bar{b})$. A clear separation is present between the barrel and the forward region with the possibility of cracks to pass the services, locally degrading the energy measurement. The full system is presented in Figure 2.14

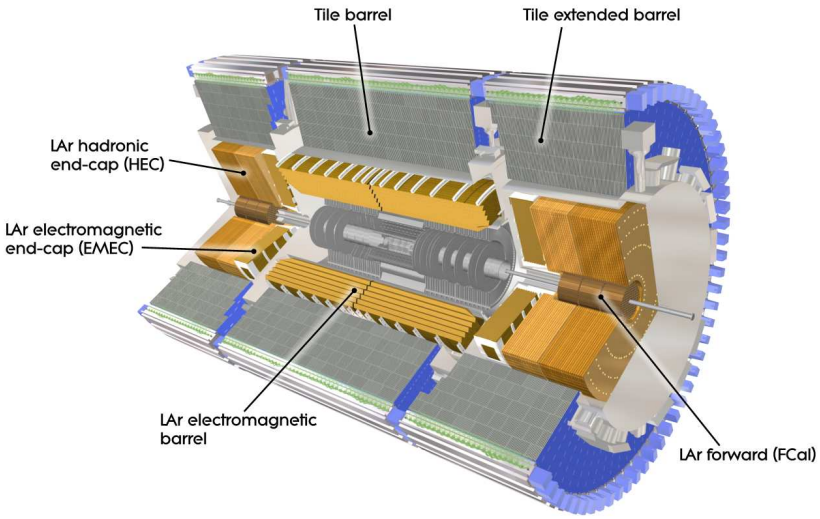


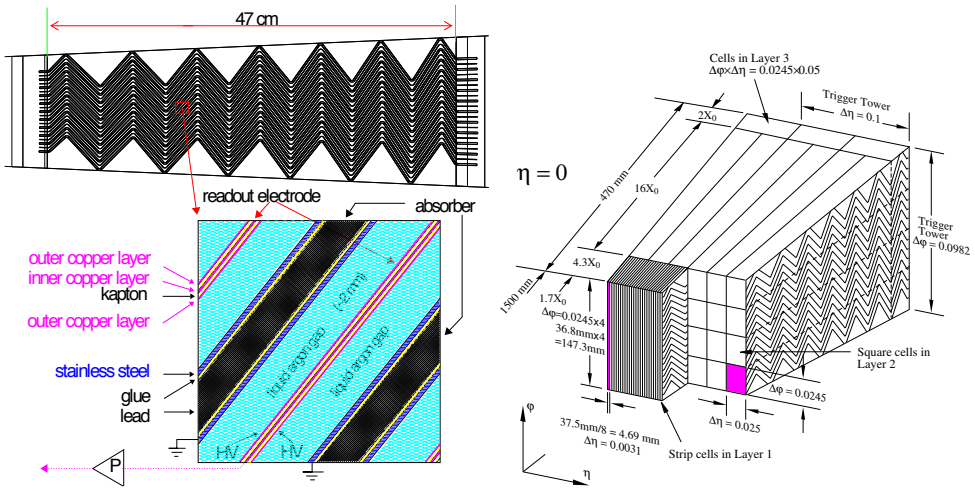
Figure 2.14 – Cut-away view of the ATLAS calorimeter system [6].

The electromagnetic calorimeter

The electromagnetic calorimeter in ATLAS is based on a Liquid Argon (LAr) active material whereas the sampling is done using lead plates reinforced with stainless steel for mechanical purposes. Once charged particles are created in the absorber, they lose energy ionising the active media (LAr) and the electrons are attracted by copper electrodes, creating a capacitive signal on an inner copper layer isolated with Kapton as shown in Figure 2.15a. The geometry has been adopted to guarantee a good ϕ coverage with no dead regions through an accordion structure as presented in Figure 2.15b. The total thickness of the system is higher than 22 X_0 in the barrel and can reach 24 X_0 in the end-caps.

The detector is further segmented in towers with 190000 cells in η and ϕ in order to sample the geometrical phase space, providing a position measurement of the deposit that can be used for object reconstruction. The longitudinal shape development is also accessible by sampling the calorimeter thickness in four sections as shown in Figure 2.15b:

— **Pre-Sampler**: this module is placed inside the solenoid volume to estimate the energy lost by the particle before entering the calorimeter. It must be noted that the tracker contribution to the final material budget varies between 0.5 and 2.5 X_0 mostly due to the TRT and the services. The detector



(a) (b)

Figure 2.15 – (a) Schema of the absorber and electrode disposition in the EM calorimeter barrel [91]. (b) Section of the barrel calorimeter showing the last three radial sections with their corresponding cell segmentation [6].

only covers the central region ($|\eta| < 1.8$) and is thin (1.1 cm) compared to the rest of the calorimeter and does not contained any absorber.

— **Layer 1**: This layer accounts for 4.3 X_0 of thickness. It has a fine granularity in order to discriminate the single-photon signal from the highly boosted decay of neutral pions $\pi^0 \to \gamma\gamma$. The granularity ranges from 0.003 to 0.1 in η, while it is kept at 0.1 in ϕ.

— **Layer 2**: This is the main layer where most of the shower developes (for photons and electrons with energy up to 50 GeV). The granularity is fixed to $(\Delta\eta, \Delta\phi) = (0.025, 0.025)$.

— **Layer 3**: The last layer accounts for an extra 2 X_0 where the highly energetic particles are measured. It has a coarser granularity with $(\Delta\eta, \Delta\phi) = (0.05, 0.025)$ and acts as a pre-sampler for the hadronic calorimeter.

The energy resolution obtained with such system is:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{\mathrm{E[GeV]}}} \oplus c_i$$

The first term represents the stochastic term linked to the pre-sampling, and the last term is due to the detector non-uniformities, alignment and electronics calibrations. While the first term is obtained from test-beam measurements [92], the second one can be measured from $Z \to$ ee [93]. Its value varies in the detector between 0.1 and 0.3 %. The noise term is found to be negligible for this LHC phase contrary to the HL-LHC runs.

The hadronic calorimeter

The hadronic calorimeter is designed with two different technologies in the barrel ($|\eta| < 1.7$) and in the end-caps ($1.5 < |\eta| < 3.9$).
The former, called the Tile calorimeter, is made out of steel absorbing plates and plastic scintillators fibres connected to photo-multipliers to readout the signal of the emitted photons. It is structured as shown in Figure 2.16 allowing a vertical and η/ϕ effective sampling. It is divided into two regions ($|\eta| < 1$ and $1 < |\eta| < 1.7$) in order to provide space for the services of the electromagnetic calorimeter. Scintillators are installed to recover some signals, but the reconstruction of jets in these regions is degraded.

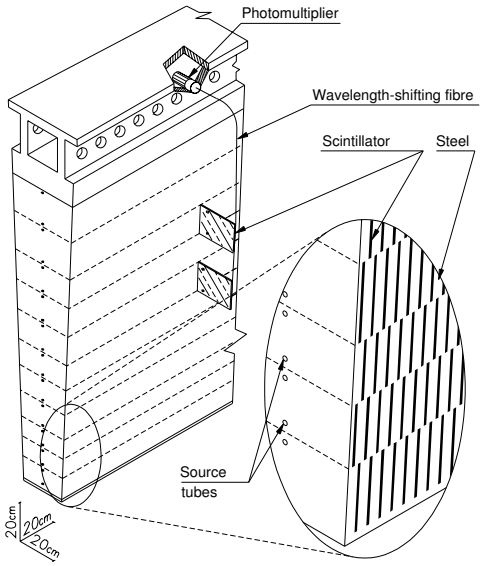


Figure 2.16 – Tile from the ATLAS tile hadronic calorimeter. [6].

The end-caps are reusing the LAr technique but use copper instead of lead as the absorber. The detectors are arranged in two independent wheels with the same diameter covering different η regions. Their modules are grouped with tower granularity ranging from $(\Delta\eta, \Delta\phi) = (0.1, 0.1)$ in the front wheels to $(\Delta\eta, \Delta\phi) = (0.2, 0.2)$ in the rear ones.

The total energy resolution obtained in test-beam [94] is shown to be:

$$\frac{\sigma_E}{E} = \frac{55\%}{\sqrt{\mathrm{E[GeV]}}} \oplus 3.4\%$$

principally due to the nature of the interaction and the sampling of the calorimeter. It is worth considering that the transparency of the scintillators will degrade with time due to the radiation, worsening the stochastic term by 10%.

The Forward calorimeter

Located 4.7 m away from the collision point, the Forward Calorimeter (FCal) is composed of an electromagnetic and hadronic part and allows measurements up to $|\eta| < 4.9$. The system must have a radiation tolerance higher than the other system due to its position. That is why the LAr technology is used. The electromagnetic part is made of copper for the absorber, while the hadronic part uses higher Z material (tungsten) to increase the number of total radiation lengths of the system. A compact design was implemented to avoid any energy leakage and consists of concentric rods for the electronics aligned with the beam pipe.

The total energy resolution obtained for electrons (pions) measured in test-beam is [95]:

$$\frac{\sigma_E}{E} = \frac{28.5(94)\%}{\sqrt{\mathrm{E[GeV]}}} \oplus 3.5(7.7)\%$$

2.2.4 The Muon Spectrometers

Due to their mass and lack of strong interaction, muons only minimally interact with the calorimeters, making it very challenging to measure their energy and momentum with such techniques. Therefore a large magnetic field is used to compute their momenta from the bending of their trajectories. However the precision of the ID tracking is not enough to get precise measurements. The muon detectors are placed as far as possible from the IP as represented in Figure 2.17, to get the best resolution

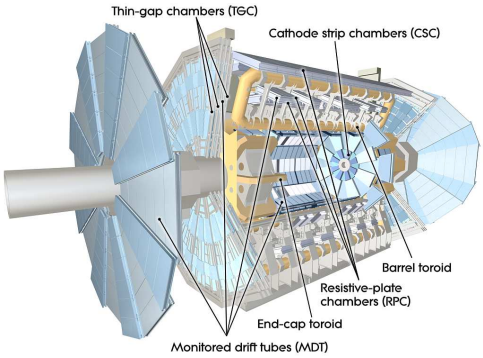


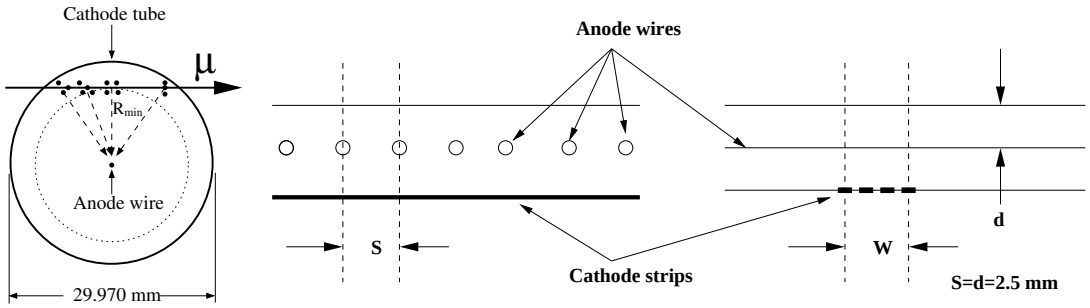
Figure 2.17 – Cut-away view of the ATLAS Muon system [6].

The first level of detectors is designed to give a high momentum precision measurement. They are placed in the proximity of the toroidal magnets (see Section 2.2.5), and consequently must respect the symmetries imposed by them. However in order to provide a hermetic coverage, some ϕ overlap is imposed. The only noticeable gap is found at $|\eta| < 0.1$ in order for the calorimeter and ID cooling and electronics services to be supplied. The barrel system is composed of three cylindrical layers at a radius of 5, 7.5, 10 m from the beam pipe, whilst a system of wheels is set for the end-caps at 7.4, 10.8,

14, 21.5 m from the IP along the z-axis. The system is thus covering tracks up to $|\eta| < 2.7$ and is composed of two different detecting systems:

— **Monitored Drift Tubes** (MDTs): are composed of drift chambers as represented in Figure 2.18a formed of 3 cm diameter aluminium tubes filled with 3 bars of an Ar-CO_2 mix which acts as an active media. The detection principle is very similar to the TRT Transition Radiation Detection system, for its position measurement part. Each chamber is made out of three to eight layers of tubes allowing a position resolution of 35μm per chamber (80μm for one tube). The detectors are covering the full η range except for the first three closest wheels to the IP which are only covering $|\eta| < 2.0$, since they cannot stand a rate higher than 150 Hz/cm^2, which is expected in this region.

— **Cathode Strip Chambers** (CSCs): are composed of multiwire proportional chambers which can stand a rate as high as 1000 Hz/cm^2. An electrical signal is induced in the wire inside an Ar-CO_2 mix as shown in Figure 2.18b, and transmitted through capacitive couplings to the strips oriented in the r-ϕ direction providing the position of the hit. The obtained resolution is 60μm which is better than the MDTs, but produces a signal that is more complicated to process.



(a) (b)

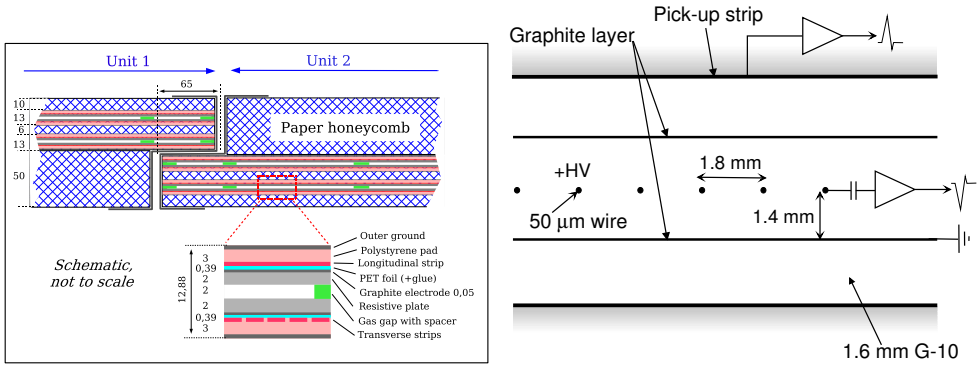
Figure 2.18 – Overview of the two muon momentum measurement devices: (a) the Monitored Drift Tubes (MDTs) (b) Cathode Strip Chambers (CSCs) [6].

In order to give a much faster signal for the trigger than what the aforementioned devices can produce, two sets of detectors complete the muon system. Their information is also used in the momentum reconstruction but their resolution is not as good. The solutions are different once again between the barrel ($|\eta| < 1.05$) and the end-cap ($1.05 < |\eta| < 2.5$):

— **Resistive Plate Chambers** (RPCs): are gaseous parallel electrode-plate detectors. The active medium is a $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 mixture which is used to trigger an avalanche thanks to a high 4.9 kV/mm electrical field. The fast signal of 5 ns width is induced via capacitive coupling to two orthogonal

sets of read-out strips as shown in Figure 2.19a. It is deployed in the barrel region.

— **Thin Gap Chambers** (TGCs): are multiwire proportional chambers as for the CSC. They show a much better timing than CSC and also provide the azimuthal information for the MDTs in the forward region. In order to get a faster signal, they rely on a highly quenching mix of CO_2 and n-pentane.



(a) (b)

Figure 2.19 – Overview of the two muon trigger system: (a) Resistive Plate Chambers (RPCs) and (b) Thin Gap Chambers (TGCs) [6].

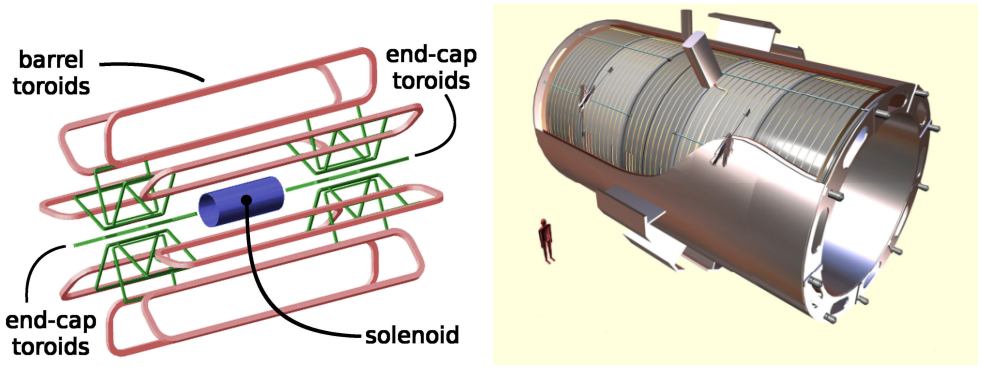
The total momentum resolution obtained with such system ranges from $\sigma_{p_T}/p_T =$ 10% for 1 TeV muons to 2-3% in the low-p_T range, where the ID information provides additional information.

2.2.5 The Magnet system

In order to measure the momentum of the charged particles, two systems of magnets are used: a solenoid [96] provides the magnetic field for the inner-detector, while a toroidal system, composed by a barrel [97] and two end-caps [98] magnets, is used for the muon spectrometer as shown in Figure 2.20a. This complexity arises from the large dimensions of the detector. The CMS detector, on the contrary, is more compact and can rely on a single 4 T solenoid as shown from Figure 2.20b.

The solenoid creates a 2 T axial field inside the inner-detector. Because it is located in front of the calorimeters, its design was made to optimise the material thickness accounting for only 0.66 X_0 of radiation length. It is made of a single superconductive coil layer which is cooled down to 4.5 K and supports a current of nearly 8 kA. Its inner and outer radius are respectively 1.23 and 1.28 m for a length of 5.8 m.

The toroidal system provides a field ranging from 0.2 to 2.5 T in the barrel and 0.2 to 3.5 T in the end-caps. It covers a much larger volume, with an inner radius of 4.7 m and an outer one of 10.05 m for the barrel part, 25.3 m long, whilst the end-caps are



(a) (b)

Figure 2.20 – (a) Illustration of the magnet system layout in the ATLAS experiment [6] (b) Perspective view of the CMS solenoid inside the vacuum tank [99].

5 m long and extend from 0.9 to 5.3 m in radius. In order to have a good radial overlap while keeping a good bending power at the interface between the two regions, a special geometry was designed. The end-cap coils are rotated by a 22.5° angle with respect to the barrel ones and distributed symmetrically around the calorimeter. The circuits are made from aluminium stabilised NbTi superconductor that could hold current up to 20.5 kA, providing a peak field of 3.9 T. While the eight barrel coils are isolated from each other, the end-caps are assembled in one block allowing an easier access to the core of the detector. The full system is cooled down to 4.7 K to ensure a good working point for the superconducting circuit. The muon chambers are located inside and around the barrel vessels, whilst upstream of the end-cap ones, in order to benefit from the highest magnetic fields.

2.2.6 The Forward ATLAS detectors

At a distance from the ATLAS interaction point, ranging from 17 to 240 m, a series of three detectors are placed. They are used to provide measurements of the luminosity for the experiment, as well as studying the very forward objects in the same context as the TOTEM experiment [71] with respect to the CMS detector. Their position and names are outlined in Figure 2.21.

Luminosity measurement Using Čerenkov Integrating Detectors (LUCID)

The nearest detector located outside the ATLAS cavern is used to provide online luminosity monitoring. It relies on a differential luminosity measurement based on the detection of inelastic p-p collision scattering. The concept is driven by the detection of the emission of Čerenkov light in an over-pressurised C_4F_{10} gas at 1.2-1.4 bar, providing a low energy threshold for the emission (2.8 GeV for pions and 10 MeV for electrons). The light is focussed with Winston cones towards PMTs. The detector is divided into twenty tubes surrounding the beam pipe and is located at $|\eta| \simeq 5.3$, therefore the system

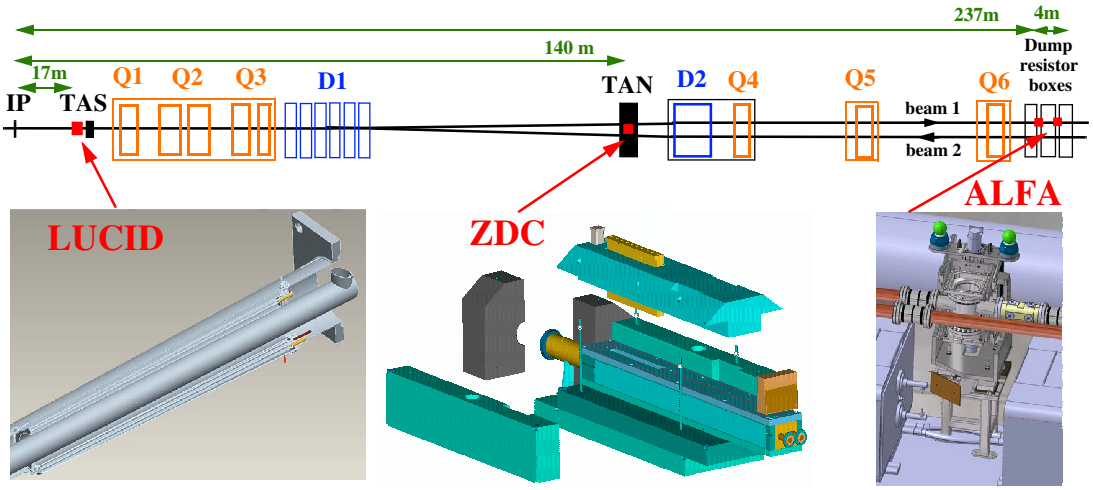


Figure 2.21 – Placement of the forward detectors along the beam-line around the ATLAS interaction point [6].

must be reliable under high radiation doses which were tested during the R&D. With the nominal prediction for the luminosity, it was also shown that the system was able to measure single particles with a probability of detecting several particles during the same measurement at the per cent level. The expected luminosity measurement resolution has been designed to be 5% to be competitive with respect to the ALFA detector.

Zero-Degree Calorimeter (ZDC)

Located at a distance of 140 m ($|\eta| \simeq 8.3$), the ZDC [100] detector provides a measurement of very-forward processes. In p-p collisions, beam-gas and beam-halo effects are responsible for up to 9% of the total inelastic collisions (from PYTHIA simulations). These events are characterised by a symmetric signature and are useful to understand the behaviour of the detector for Minimum Bias events. Hence the ZDC detector can be used as a trigger and vertex locator ($\Delta z \sim 3cm$) for such events without relying on the ID. During heavy-ion collisions, this detector can also be used as a neutron detector.

The system is a sampling calorimeter using tungsten plates perpendicular to the beam direction as absorbers and quartz rods set parallel to the beam in an 8×12 matrix to collect the emitted light. Quartz strips are also placed in the direction of the plates. Four modules are used, one acting as an electromagnetic calorimeter where all the rods are considered individual to track the shower development ($X_0 \sim 29$), while the three remaining ones are acting as a hadronic calorimeter and have a coarser resolution.

In order to be shielded against radiations, it is located inside the Target Absorber Neutral (TAN).

Absolute Luminosity For ATLAS (ALFA)

At the most forward point of the system, the ALFA detector [101] is based on the Roman pot technology to detect elastic scattering of protons that is related to the total cross-section via the optical theorem. The choice of detection relies upon special beam conditions (important β^* and low emittance), because the angle of diffusion is expected to be smaller than the usual beam divergence, and thus corresponds to a lower instantaneous luminosity, such that radiation is not playing an important role. The Roman pot technology is based on a separated vessel (pot) that is brought close to the beam without interrupting its void. Current technologies allow a distance on the order of a few mm. 64 scintillating fibres are placed in a u-v stereo configuration in double-sided planes. Each fibre has a width of 0.5 mm and ten planes are assembled and glued together. Each fibre is connected to a pixel of multi-anode phototubes.

3 Data acquisition and reconstruction in ATLAS

From the signals produced by the detectors listed in Chapter 2 to the data used in this thesis, several intermediate steps are needed. First, a selection of events is conducted at detector level before the data is written to disk. In order to account for the difference between the quantities reconstructed by the detectors and the true values, and to correct for possible imperfections in the instrumentation, data is corrected and calibrated. Physical objects can finally be reconstructed from the measurements and stored for the analysis. Having good control over this chain is thus essential to understand the effects seen in the analysis.
In this chapter, I discuss first the trigger system that is used for the global event selection decoupling the structure into the hardware and the software components. Then I introduce a recent development for a tracking trigger that is foreseen for the Run-3 of the LHC and beyond: Fast TracKing. In the second part of the chapter, the reconstruction techniques for the various objects are discussed with an emphasis on the ones mostly used in the $Vh(b\bar{b})$ analysis.

3.1 The data acquisition in ATLAS

Given that 34 collisions are happening for each bunch crossing on average, as shown in Figure 2.4, the individual event size is approximately 1.5 MB. With a bunch-crossing happening every 25 ns on average for the Run-2 of the LHC, the expected event

rate of the ATLAS detector is about 40 MHz, yielding a theoretical 60 TB/s data rate. Not even considering that no hardware device can write down or transport data with such a rate on disk or tape (the order of magnitude for a modern hard drive is 100 MB/s), the final amount of data collected could not be stored or analysed afterwards. Since most the produced events are not relevant for the analysis (soft scattered), and the ones (hard scattered) which are triggered are must interesting for further study, the trigger system [102] is designed to reduce the rate by selecting on the latter. In order to take into account the different dead time of detectors as well as the complexity of the algorithm, two stages are set for the trigger as shown in Figure 3.1: The Level 1 and High Level Trigger (HLT).

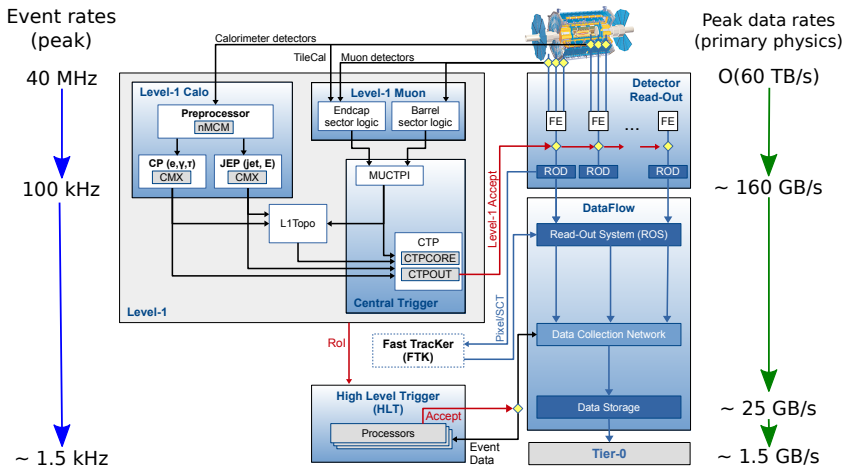


Figure 3.1 – Layout of the ATLAS Trigger and Data Acquisition system in Run-2 [102].

3.1.1 The hardware trigger system: Level 1

The first reduction step called Level 1 [103] is operated by custom-made electronics that receive information from both the calorimeter and muon chamber Front-End (FE) electronics. As the tracking information takes longer to read-out, this step only focuses on the high transverse-energy electrons, photons, muons, taus and jets as well as MET by using the analogue sum of calorimeter signals and signals from the muon chambers. The final decision to read the data is sent through the TTC (Timing, Trigger and Control) information directly to the FE (like for the LAr calorimeters) or the ReadOut Drivers (RODs) (for example in the case of the pixel detector). The main requirement for the total output rate is given by the capability of the detector read-out system that is fixed to 100 kHz. The processing time requirement of 2.5 μs is given by the length of the buffers of the various sub-detectors. However, most of the dead time is due to transferring data to L1, that is why it is placed as close as possible to the detector in the cavern itself (total travelling time of signals can be as long as 0.8 μs). To gain from the various specifics of the sub-detectors the L1 is composed of several sub-modules.

The Level 1 Calorimeter Trigger

The goal of the L1 Calo [104] is to process both electromagnetic and hadronic objects coming from the two calorimeters. It receives 7168 summed analogue signals from coarse calorimeter cells called Trigger towers (from $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ to $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$). The signal is then digitised, filtered and calibrated through the preprocessors. The outgoing data is sent in parallel to two algorithmic processors: the Cluster Processor (CP) and the Jet/Energy-sum Processor (JEP). Both processors identify candidates that pass a given trigger selection, *e.g.* the L1_EM24 item will require a group of cells from the electromagnetic calorimeter to have an E_T above 24 GeV. The CP algorithm focusses on candidate electrons, photons and taus that pass a given E_T and, if desired, isolation requirements. The JEP algorithm operates on even coarser Triggers towers ($\Delta\eta \times \Delta\phi = 0.2 \times 0.2$) and produces variables such as $\sum E_T$, $E_{\mathrm{T}}^{\mathrm{miss}}$ or the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance. The output of the L1 calo is a Region Of Interest (ROI) linked to the corresponding trigger items checked by the objects in the CP and JEP.

The Level 1 Muon Trigger

In parallel, the information from the muon RPC (for the central muons) and TPC (for the forward muons) are collected by the L1 muon trigger [105]. The goal of the algorithm is to find a coincidence of hits across several layers that are consistent with a muon originating from the interaction point. First, a hit has to be found on a pivot plane (middle station for the RPC, outer station for the TPC) as shown in Figure 3.2. Then tracks are reconstructed using coincidence windows on the other planes which are computed using MC simulation of single muons for several p_{T} thresholds. As for the L1 calo, the output is a ROI associated with a trigger item.

The Level 1 Topological Trigger

With growing pile-up and energy, and to keep a fixed L1 trigger rate, the thresholds of L1 trigger items needs to be risen. However, some more advanced variables could be used in order to keep an acceptable rate while preserving the final trigger and analysis efficiency. The L1 Topo Trigger uses the ROIs from the L1 Calo and the L1 muon to compute angular and kinematic quantities (such as $\Delta\phi/\Delta\eta$, invariant mass or even corrections to $E_{\mathrm{T}}^{\mathrm{miss}}$ using lookup tables). All of these features are possible thanks to the high performance of modern FPGAs (Field Programmable Gate Arrays). The improvement could, for instance, be very beneficial for high mass searches (such as for a BSM resonance, high p_{T}) where an invariant mass requirement can be set at the L1 step, or for the B-physics searches (low p_{T}) such as $\mathrm{J}/\psi \to \mu^{\pm}\mu^{\mp}$ where ΔR and invariant mass constraints can be set thus allowing the trigger rate to be reduced by roughly a factor 3.5 as shown in Figure 3.3.

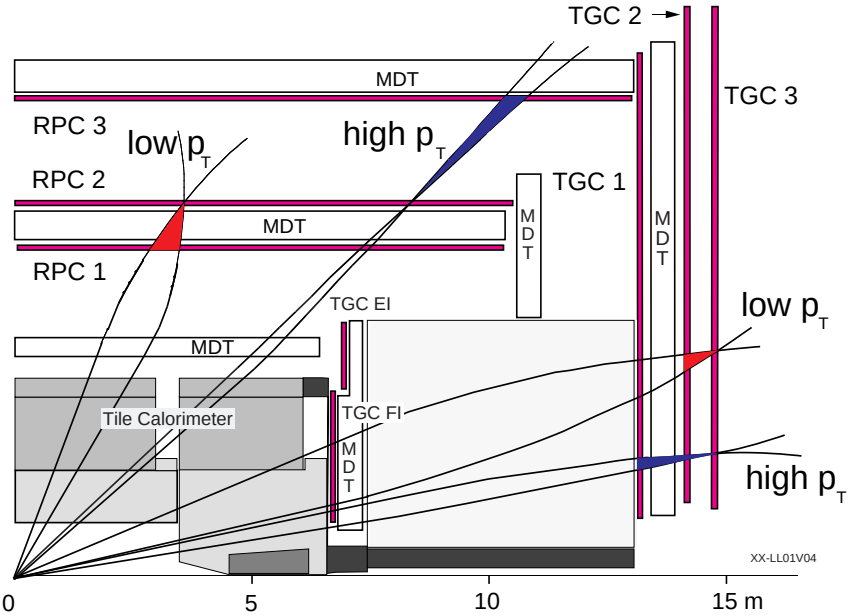


Figure 3.2 – Cross-section of the ATLAS detector, showing the locations of the L1 Muon trigger chambers. The p_{T} dependent coincidence windows are shown with the plain area [105].

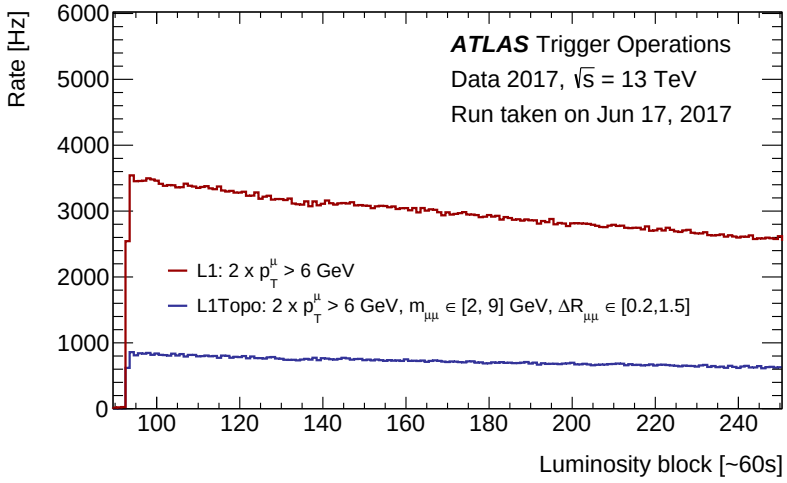


Figure 3.3 – Rate of first trigger level (L1) for $\mathrm{J}/\psi \to \mu^{\pm}\mu^{\mp}$ events in the case (red) where only two muons are selected, and the case (blue) where a more complex L1 topo trigger is selected [106].

The Level 1 Central Trigger Processor

In order to combine all the decisions from the L1 Calo, L1 muon and L1 Topo triggers, a logical card called the Central Trigger Processor (CTP) [102, 107] is used. Up to 320 different trigger keys can be received and a look-up table is then used to produce up to 512 different trigger items. Those items are composed of logical combinations of OR and AND terms of the incoming triggers.
Since the number of objects passing the thresholds can be quite high some additional

random selection is set for some trigger items. This process called prescaling, is a factor which can be tuned with the luminosity and the pile-up, as the rate changes (the lower the pile-up the less the pre-scaling is needed) as shown in Figure 3.4. This choice sacrifices some of the trigger items for which the full statistic is not essential, in favour of a lower L1 trigger rate, saving room for more important physics triggers. Some un-prescaled items are also saved and are used in the $Vh(b\bar{b})$ analysis.
In order not to request the data to be read several times from the same ROD, some extra vetoing is applied with a fix time window (the same ROD can not be called before the next N bunch crosses, called simple dead time), or in a sliding time window (the same ROD can not be called more than N times in M bunch crosses, called complex dead time). Those effects are mitigated by the addition of buffers to store the data.

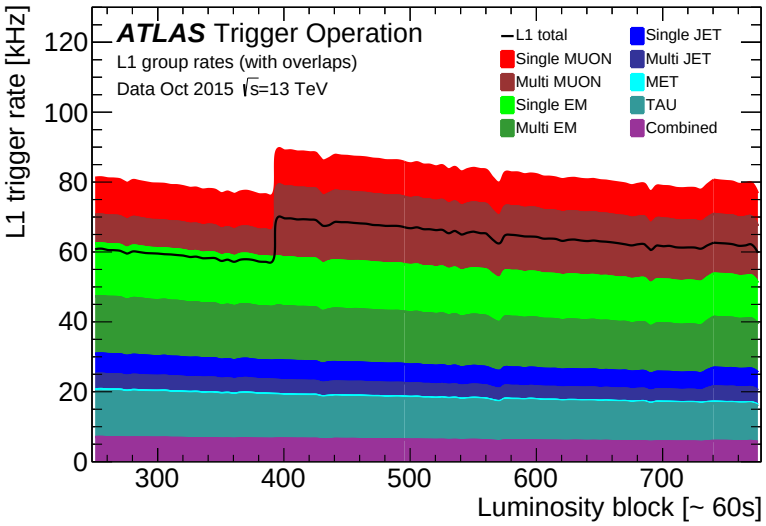


Figure 3.4 – L1 trigger rate evolution with time during one LHC fill. The increase of rates at the luminosity bloc ∼400 is explained by the removal of the pre-scaling for B-physics triggers which mainly use muon triggers. Due to overlaps the sum of the individual groups is higher than the L1 total rate, which is shown as black line [102].

3.1.2 The Fast TracKer project

Since reconstructing tracks is time-consuming, ∼O(s/event) [108], the tracking information is not included in the L1 trigger. However, in a dense environment with high in-time pile-up, it could be very useful to have some information from tracking used in the trigger decisions, to overcome the limitations of the ROI information from the L1. Simply requiring the central jets of $p_{\mathrm{T}} > 25$ GeV to come from a vertex originating at a hard interaction helps to reduce the number of candidate selected jets from 21 to 3.8 in a $< \mu >= 69$ environment [109]. Furthermore, some topologies looking for displaced vertex (as for the $Vh(b\bar{b})$ analysis) could benefit from the tracking information to have a smarter trigger. That is why a hardware-based Fast TracKer (FTK) has been proposed for integration in the ATLAS trigger system.

3.1.2.1 Description of the project

The first idea about including tracking as part of the trigger information was developed for the Collider Detector at Fermilab (CDF) experiment, with the Silicon Vertex Tracker (SVT) [110] trigger. As for the ATLAS experiment, the B-physics measurements were enhanced at CDF thanks to their displaced vertex properties, the main difference being that the instantaneous luminosity is much higher at LHC. The Fast-Tracker (FTK) project benefits from technical improvements as well as of the experience from the SVT and its architecture is presented in Figure 3.5. The hardware implementation in the two cases relies on the same principle: a simple fit conducted on hardware chips. The data from the Pixels and SCT are retrieved from the RODs and sent by the Data Formatters (DF) to one of the 64 (128 after upgrade) FTK towers depending on their $\eta - \phi$ position. In order to avoid inefficiencies, the towers are overlapping at the boundaries. Furthermore, only tracks with $p_{\mathrm{T}} > 1$ GeV are considered. Clustering is also carried out at that step, and only the cluster centroids are transmitted further on. Those centroids will be referred to as FTK hits in the following. The fitting is then performed in two steps.

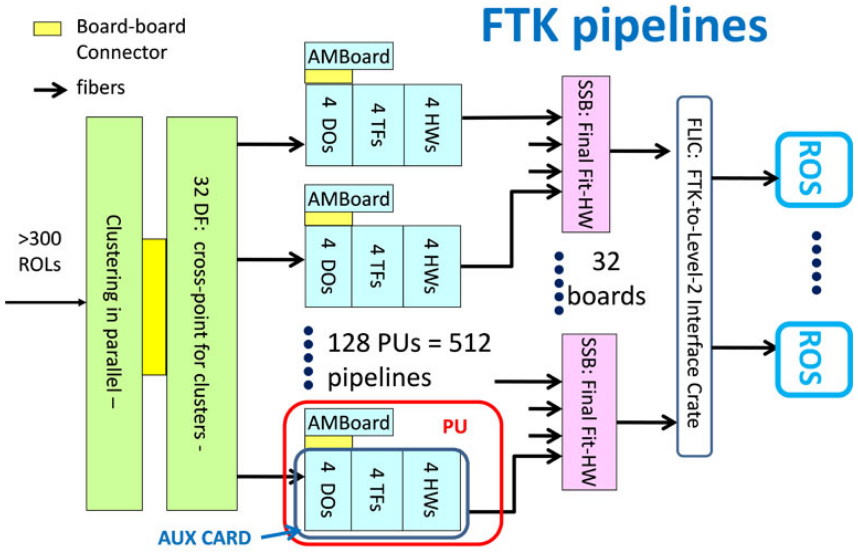


Figure 3.5 – Initial sketch of FTK system with 64 towers and two Power Units (PUs) per tower. The upgrade of FTK will double the amount of PUs. DF is the Data Formatter, DO is the Data Organizer, TF is the Track Fitter, HW is the Hit Warrior, AM is the Associative Memory [109].

First Stage Fit

The First Stage Fit (FSF) is performed simultaneously by two Processing Units (PUs) on each FTK towers and only considers three pixel layers (the B-layer and the two outer pixel layers) and five SCT layers (the four axial layers and one stereo layer) for the FTK hits.

The hits sent by the DF are received by one of the Data Organiser (DO). A copy of

the hits is converted into Super-Strips (SS) which have a coarser resolution, in order to facilitate the next steps and linked to the original FTK hit by a SS address that allows rapid retrieval afterwards.
One of the key ingredients of the system is the hits association used to get the track property. This very CPU consuming step in tracking is solved by a pattern matching hardware algorithm handled by Associative Memory (AM) chips. Those chips use pre-loaded pattern banks in order to match in a massively parallel way and in a very reduced processing time the selected pattern to the SS hits. This operation has a negligible contribution to the overall time to reconstruct the tracks. Once 6, 7 or 8 SS hits are matched to one of the patterns, the pattern road id and the SS id are sent back to the DO. The electrical performance of the chip is discussed in Section 3.1.2.2 while the performance of the pattern bank are discussed in Section 3.1.2.3.
From the road id, the full resolution FTK hits and the pattern are sent to the Track Fitter (TF) that is an FPGA based algorithm. In order to reduce the computational time, a linearised fit is performed instead of a complex helical fit, as a consequence of the small size of the $\eta - \phi$ region. The five helix parameters (d_0, z_0, p_{T}, η, ϕ) are derived from:

$$\tilde{p}_i = \sum_{l=1}^{N} C_{il} x_l + q_i \qquad \forall i \in [1,5] \tag{3.1}$$

where $\tilde{p}_i$ are the linearised helix parameters, x_l the l^{th} component of the N hit coordinates of the cluster centroids. In the FSF $N = 11$ coordinates are considered (two coordinates for each of the three pixel layers, and one coordinate for each of the five SCT layers). C_{il} and q_i are the constants associated with the sector where the fit is performed. The computation of the constants is described in Section 3.1.2.3. Evaluating the goodness of tracks is done through a χ^2 test, and only tracks that pass a fixed cut value χ^2_{cut} are passed forward. In order to recover some efficiency, if the χ^2 is under $\chi^2_{threshold}$, one layer is allowed not to have a hit. An extrapolation to get the missing hit position is then performed to minimise the χ^2, and if the final value is under χ^2_{cut} the track is kept. Since several tracks can share similar hits, a duplicate track removal function called Hit Warrior (HW) is performed after the TF. The selection of tracks is decided upon a combination of the χ^2 score and the number of layers that got a hit on the tracks.

Second Stage Fit

In order to improve the resolution of the helix parameters as well as improving the robusteness to pile-up, the information of the 4 remaining layers (1 pixel and 3 stereo SCT layers) is added back to the system in the Second Stage Board (SSB). Each SSB is fed by 4 FSF boards and gets the remaining information from the DF. First, an extrapolation function aims to identify the hits belonging to tracks obtained at the FSF. In order for the tracks to be passed onwards, they must contain at least 10 layers with real hits. Then a TF is performed as for the FSF but with constants calculated with $N = 16$ coordinates. Then a slightly more complex HW is used with respect to the FSF. Only one out of two SSBs will carry out the HW function and output the final tracks, referred to as Final

SSB (fSSB). This SSB will receive the tracks from: its TF, as well as the neighbouring SSB in the ϕ-plane ($-\phi$ and $+3\phi$ called preliminary SSB or pSSB, and the $\pm 2\phi$ fSSB). A summary of the tracks transmission is presented in Figure 3.6. The output tracks are sent through an FTK Level-2 Interface Crate (FLIC) toward the HLT ROSs were the data is then made available for the Level-2 trigger.

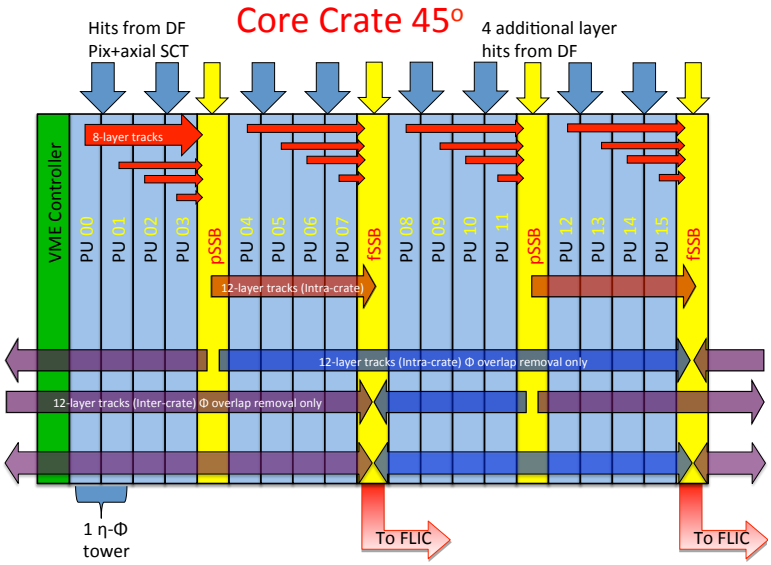


Figure 3.6 – Diagram of the transmission of data between the SSB boards to perform the HW function. fSSB boards are the final boards receiving the tracks from the ϕ neighbouring SSB board [109].

3.1.2.2 Improvement of the AM chip consumption

One of the main ingredient presented in Section 3.1.2.1 is the hardware pattern matching. This step is performed thanks to AM chips devices. Each hit is decomposed into 8 layer hits that contain 18 bits. The bits are compared to the ones contained in the pattern banks through Content-Addressable Memory (CAM) cells which are also used to register the MAC addresses of the devices connected to a router. CAM cells compare 2-bit lines and return a logical value when bits match. The cells are grouped into longer chains of bit comparison, which could be imagined as an encoded character. Those groups are further organised into a N×M matrix, where N is the number of patterns stored and M the number of characters forming the pattern. Once all the columns are returning a positive comparison for all the bits inside, the positive matching between the input and the line (*i.e.* the pattern) is made. Due to the ability to develop ternary CAM cells, a third logic state, referred to as Don't Care bit, is allowed by the system. In FTK this ternary option is used to always return a positive value for the matching. Therefore, the number of effective bits is reduced by one unit for every DC bit added. The actual AM chip used in FTK has up to 16 bits characters, *i.e.* 65536 positions and 2 DC bits per

layer. The match signal is then propagated per patterns onto a decision function that sends on the address of matched patterns and the number of matched layers. A sketch of the proof of principle of the AM chip is presented in Figure 3.7.

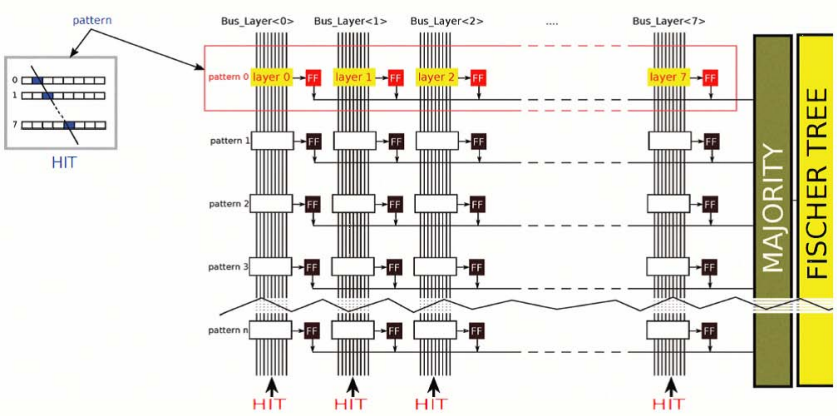


Figure 3.7 – Sketch of the AM chip principle. Each 8 coordinates hit is compared through CAM cells to pre-loaded patterns [109].

Hits are sent to the AM chip at a clock rate of 100 MHz and each chip can contain up to 128k patterns. From a purely electrical consumption point of view, the AM chip system is behaving as charging/discharging bit lines, *i.e.* like capacitors. Therefore there is an important correlation between how different the bit structure of two subsequent hits entering the AM chip is and its actual power dissipation. The basic solution set for the sequence of hits is a First In First Out (FIFO) logic implemented on a Field Programmable Gate Array (FPGA). In this scheme, there is a high probability that two subsequent hits have a very different bit structure. However, a simple sorting algorithm could be used to sort the incoming hits in order to reduce the bit flip and therefore the chip electrical consumption. The minimisation of the bit flip is often generalised as the Travelling Salesman Problem [111] which is a well known problem in graph theory to be N-P complex. Solutions can be divided into exact approaches (such as the brute force algorithms) and heuristic algorithms. For the AM chip, two algorithms are proposed in Table 3.1. In order to understand the performance of the ordering algorithms, simulations have been performed as well as measurements on the chip. Results are presented in Figure 3.8. Since the power consumption is linked to the mean number of bit flips the simulated gain is defined as:

$$\text{Gain}_{\text{sim.}} = 100 \times \frac{\langle \text{bit} - \text{flip} \rangle_{\text{non sorted}} - \langle \text{bit} - \text{flip} \rangle_{\text{sorted}}}{\langle \text{bit} - \text{flip} \rangle_{\text{non sorted}}}$$

In order to measure the performance of the ordering algorithms using the real chip itself, a simple test bench has been designed:

— measuring the consumption while sending the same hit 100% of the times $\text{P}_{\text{one hit send}}$;

— measuring the consumption while sending a list 100% of the times $\text{P}_{\text{non sorted}}$;

— sorting the list with the algorithm 2;

— measuring the consumption with the sorted list P_{sorted}.

The measurement of the $P_{\text{one hit send}}$ helps to factorise out the electrical effects that are not due to the bit-flip, such as resistive effects. This power consumption has been shown to be of the order of 10% of the total electrical consumption. The measurement has been repeated 10 times for each point in order to decrease the statistical uncertainty. The electrical gain is then defined as:

$$\text{Gain}_{\text{meas.}} = 100 \times \left(1 - \frac{P_{\text{sorted}} - P_{\text{one hit send}}}{P_{\text{non sorted}} - P_{\text{one hit send}}}\right)$$

From the performance shown in Figure 3.8, only algorithm 2 has been tested. The overall agreement between the simulated and measured electrical gain is satisfactory even though the uncertainties may have been under-estimated. For a list of 10 elements, the algorithm 2 leads to a reduction of 33% of the capacitive electrical consumption, hence helping the total electrical budget of the system and reducing the needs of cooling.

	Algorithm 1	Algorithm 2
Algorithm	• make lists of N hits; • compute the bit flip between all the hits; • order them in order to minimise the total bit flip.	• make list of N hits; • compute the bit flip between the hits and the last hit of the sorted list ; • add the closest hit to the sorted list and add new hit in the list.
Pros	similar to what would be optimal (brute force on full list)	easy to implement, fast.
Cons	difficult to generalise to long lists	depends on the length of the list

Table 3.1 – Considered algorithms to solve the hit ordering problem.

3.1.2.3 Generating the pattern bank

Generation of constants

The constants introduced in Equation 3.1 are computed from single muon Monte Carlo samples that cover all the physics regions where tracks could be found. The computation relies on the minimisation of the distance between the true parameter p_i and the linearised one $\tilde{p}_i$:

$$\left\langle (\tilde{p}_i - p_i)^2 \right\rangle = \left\langle (\sum_{l=1}^{N} C_{il} x_l + q_i - p_i)^2 \right\rangle \qquad \forall i \in [1,5] \tag{3.2}$$

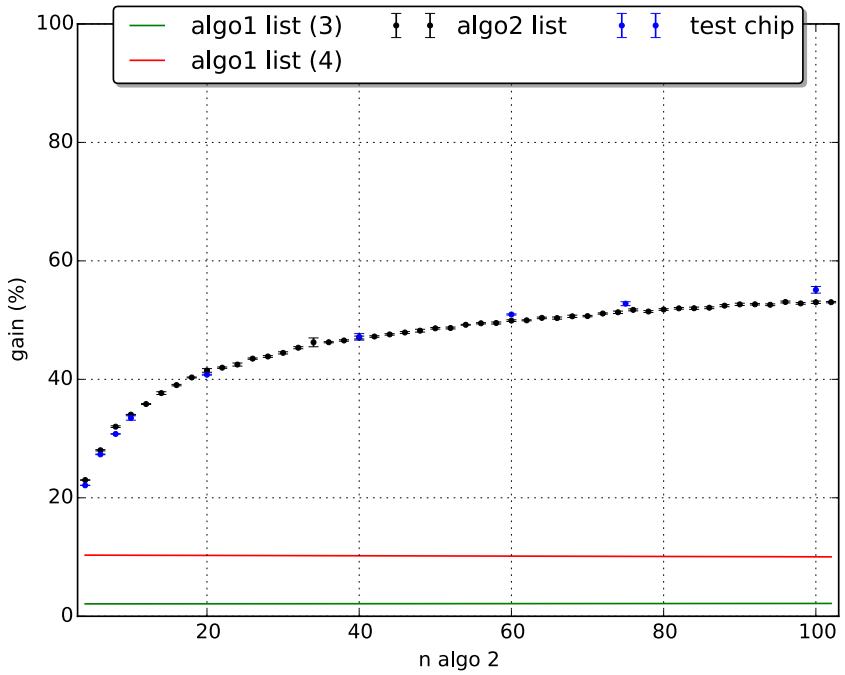


Figure 3.8 – Electrical consumption gain thanks to the ordering algorithm. Two sizes of the algorithm 1 (in red and green) have been simulated while a range of values for the algorithm 2 (in black) are presented. 9 values (in blue) for the algorithm 2 have been measured on a AMchip06.

over the parameters q_i and C_{il}, and where the average is over the training samples. The solution is given by the following system:

$$\begin{cases} \sum_{k=1}^{N} C_{ik}(\langle x_l x_k\rangle - \langle x_l\rangle\langle x_k\rangle) + \langle p_i\rangle\langle x_l\rangle - \langle p_i x_l\rangle = 0 & \forall l \in [1,N] \\ \sum_{l=1}^{N} C_{il}\langle x_l\rangle + q_i = \langle p_i\rangle \end{cases} \tag{3.3}$$

Identifying the covariance matrix

$$\{V\}_{lm} = \langle x_l x_m\rangle - \langle x_l\rangle\langle x_m\rangle$$

in Equation 3.3, the constants can be solved for:

$$\begin{cases} C_{il} = \sum_{m=1}^{N} V_{lm}^{-1}(\langle p_i x_m\rangle - \langle p_i\rangle\langle x_m\rangle) & \forall i \in [1,5], l \in [1,N] \\ q_i = \langle p_i\rangle - \sum_{l=1}^{N} C_{il}\langle x_l\rangle \end{cases} \tag{3.4}$$

The inversion of the covariance matrix can be problematic, hence the sector being rejected. From empirical observations, sectors with at least five tracks produce proper sets

of constants. The full detail of the computation can be found in Section A.1. To measure the quality of the linear approximation that is applied here, a χ^2 test can be performed from the covariance matrix defined above:

$$\chi^2 = \sum_{i,j=1}^{N} (x_i - \langle x_i \rangle) V_{ij}^{-1} (x_j - \langle x_j \rangle) \tag{3.5}$$

In order to reduce the complexity of the problem, the Principal Component Analysis (PCA) is used. This method is interpreted as finding the axis where the covariance matrix has the highest eigenvalues, and therefore where the inverse matrix will show a negligible contribution. Equation 3.5 can be rewritten:

$$\chi^2 = \sum_{i=1}^{N-5} \left(\sum_{j=1}^{N} A_{ij} x_j + k_i \right)^2 \tag{3.6}$$

where A_{ij} and k_i are constants defined in A.1 with the detailed computation.

Production of patterns

In order to have a sufficient coverage of the system, the production of patterns did not rely on MC samples since a very basic estimation would require 400 billion full simulation tracks to get sensitive results. Instead pre-computed constants can be used and inverted to generate coordinates (x_j) from 5 uniformly distributed helix parameters $(\hat{p_i})$ as well as N-5 gaussianly distributed constrains for the χ^2 values $(\hat{\chi_l})$:

$$\begin{cases} \sum_{j=1}^{N} C_{ij} x_j + q_i = \hat{p_i} & \forall i \in [1,5] \\ \sum_{j=1}^{N} A_{lj} x_j + k_i = \hat{\chi_l} & \forall l \in [1, N-5] \end{cases} \tag{3.7}$$

This allows as many patterns as wanted to be generated (not considering the CPU time needed for the generation). One pattern can therefore be generated multiple times (with multiple tracks giving the same list of SS), and its multiplicity is called coverage. This coverage can go as high as 10^2 as represented in Figure 3.9. A large coverage helps to get a higher pattern efficiency since more tracks could be associated with the same pattern. When generated, the patterns are ordered in decreasing order by coverage in so called thin pattern (TSP) banks.

However, there are limitations given by the hardware of the system. First, the actual number of patterns which is possible to store in an AM chip is limited. Second, due to the finite size of the SS, fake roads that correspond to shared hits by different particles are generated. The number of fake roads increases with the number of generated patterns. This impacts the number of fits to be performed and thus is limited by hardware

performance. The easiest solution would be to keep only the first patterns from the TSP, but low coverage patterns can also depict some particular topology that cannot be covered by the other patterns, and not considering them could decrease the sensitivity of the system, for instance with respect to some new physics events.

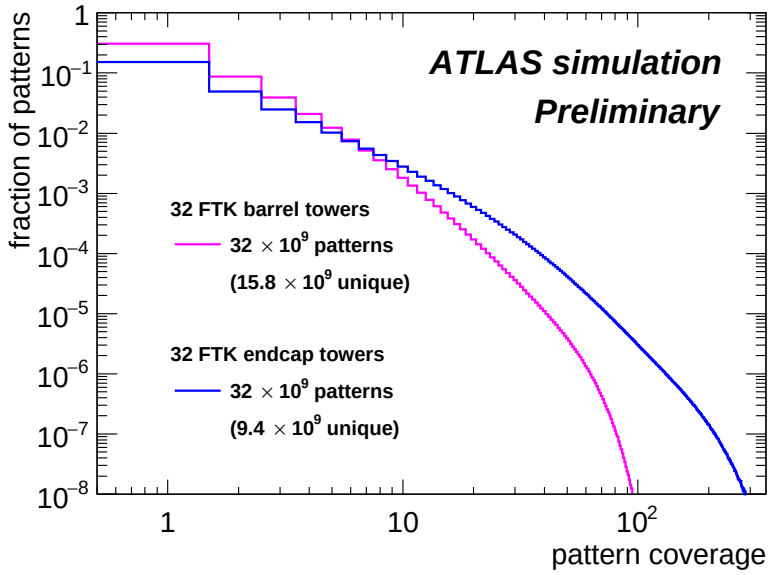


Figure 3.9 – Coverage distribution of patterns. Each FTK tower has 1 billion patterns generated. Barrel towers ($|\eta| < 1.6$) are represented in blue while endcap towers ($|\eta| > 1.6$) are represented in magenta [112].

In order to decrease the number of stored patterns while keeping good efficiency, some SS can be merged per layer thanks to the Don't Care (DC) bit of the AM chip. The DC bit acts as a ternary option beyond the usual binary bit and therefore can make a variable width pattern as shown in Figure 3.10. The number of fake roads is then decreased while keeping a relatively good efficiency. The banks generated with the DC bit option are called AM banks. To assess the quality of the AM banks, simulations with single muons and $t\bar{t}$ in a pile-up environment are performed. The FTK efficiency is then measured as the efficiency of matching a truth track with the full FTK algorithm, thus containing the clustering, the matching and the fitting of the FSF. The single muon efficiency is shown as the red line of Figure 3.11a. A big crack corresponding to the transition region can be observed in the efficiency. A second step in the generation takes this η distribution to adapt the number of patterns per tower and thus equalising the efficiency:

$$\mathrm{N_{pattern}} = \mathrm{N_{noPart}}\ \max\left(\frac{1}{\ln(1-\epsilon/\epsilon_{\max})}, \mathrm{w_{min}}\right)$$

with $N_{pattern}$ the number of newly saved AM patterns per section, N_{noPart} the number of previously saved AM patterns per section, ϵ the mean efficiency of the sector, ϵ_{max} the targeted efficiency, and w_{min} a global parameter to avoid having a sector with too few patterns. The procedure, called partitioning, can be repeated several times and it

was found that the optimal situation was obtained after 2 iterations. The effect can be observed on the blue and green curves of Figure 3.11a.

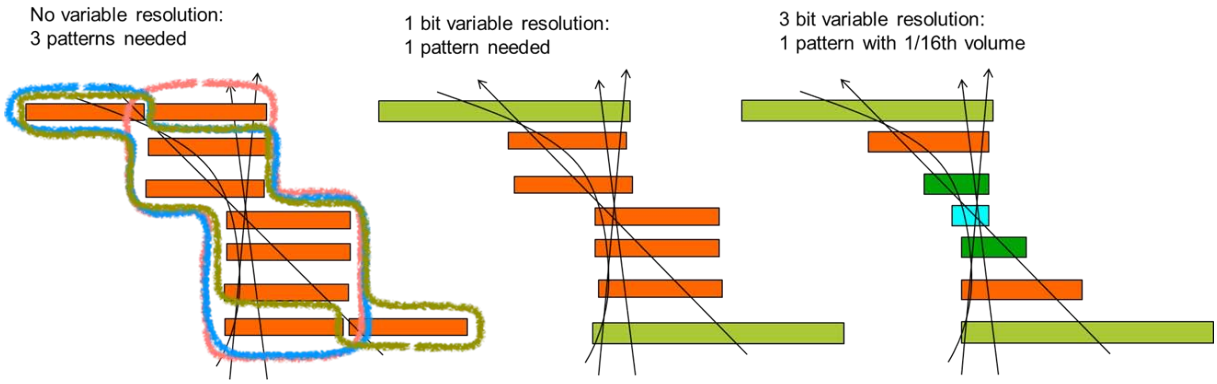
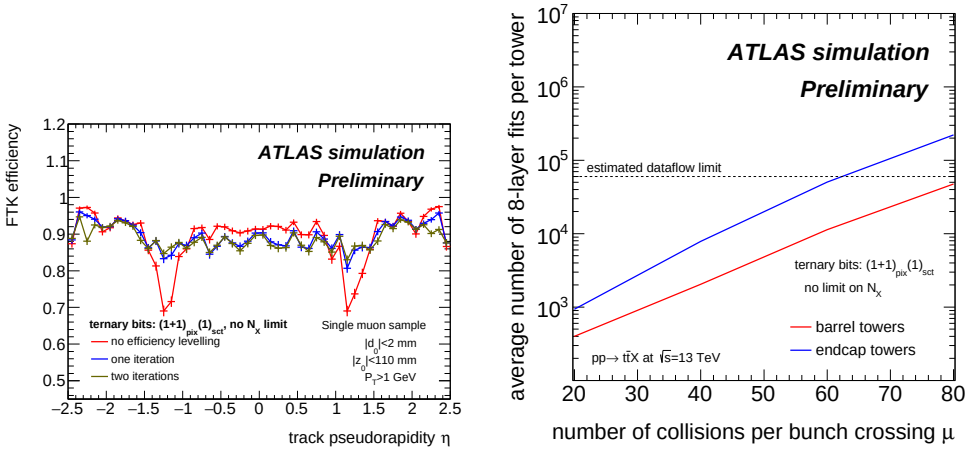


Figure 3.10 – Sketch showing the impact of the DC bit on the size the SS for the patterns [109].

The optimisation on the number of DC bits as well as their situation has been conducted. The efficiency after one iteration of partitioning is a good indicator of performance as shown in Table 3.2. As a first conclusion, the major impact has been obtained when adding DC bits on the outermost layers of the system and when adding as many DC bits as possible. One important drawback of the DC bit is the final number of fits obtained in a busy environment as it is shown in Figure 3.11b. In order to preserve the system from having dead-time, the number fits per tower is limited and having too many DC bits could make the system reach the limit for Run-3 pile-up conditions. Therefore a clever limitation of DC bits was implemented: the overall number of DC bits is limited (with different limits between the barrel and the end-cap), but the DC bits configuration per layer can vary. So the maximum number of DC bits per layer can be increased while fixing the sum of the DC bits. Figure 3.12a presents the effect on the single muon FTK efficiency. The overall efficiency is flatter and the crack of the transition region is less affecting the result. This can be explained thanks to the initial better coverage of the patterns. The final efficiency in the barrel is also increased thanks to the higher number of DC bits on the outermost layers as shown in 3.2. Figure 3.12b presents the effect on the number of fits. The reduction obtained thanks to a smaller volume of the patterns, helps to increase the pile-up end-point by 15 %.

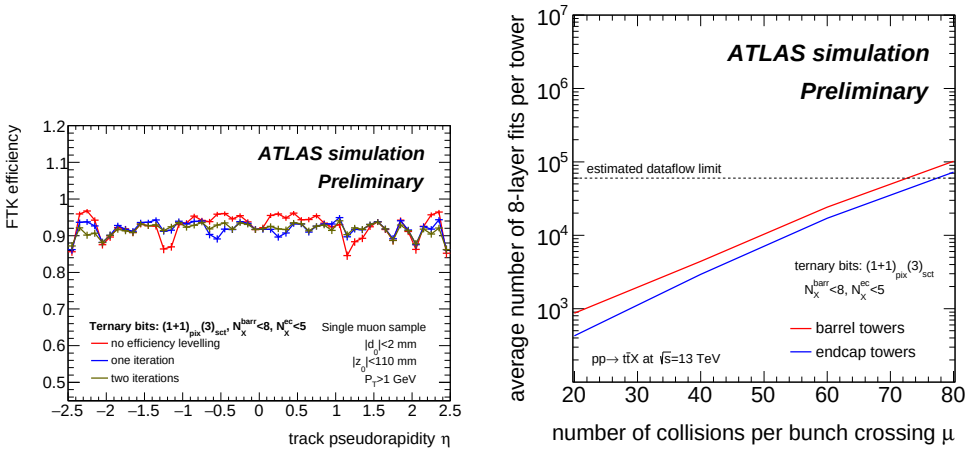
Layers	6	7	34	35	36	37	45
Eff. (%)	86.5	85.5	86.3	86.1	88.8	88.1	86.1
Layers	46	47	56	57	347	357	367
Eff. (%)	88.6	87.8	88.9	88.5	88.1	90.9	92.3
Layers	456	457	467	3456	3457	3567	4567
Eff. (%)	91.2	91.0	90.4	93.4	93.2	94.3	94.3

Table 3.2 – Single muon FTK efficiency after one iteration of partitioning. The line layers indicates on which layer DC bits have been added from the initial configuration [222 11111].



(a) (b)

Figure 3.11 – Offline performance of the FTK AM banks with a DC bits configuration of respectively 2 and 1 ternary bits in the pixel and SCT layers. (a) tracking efficiency with respect to truth of single muons, (b) number of fits as a function of the pile-up obtained with $t\bar{t}$ MC samples at 4 pile-up conditions (20,40,60,80) [112].



(a) (b)

Figure 3.12 – Offline performance of the FTK AM banks with a DC bits configuration of respectively 2 and 3 ternary bits in the pixel and SCT layers, but with a limitation of a total of 8 DC bits in the barrel and 5 in the end-cap. (a) tracking efficiency with respect to truth of single muons, (b) number of fits as a function of the pile-up obtained with $t\bar{t}$ MC samples at 4 pile-up conditions (20,40,60,80) [112].

3.1.3 The software trigger system: High Level Trigger

Following the Level 1 and FTK systems, the High Level Trigger [113] is implemented using a software chain. The main goal is to reduce the L1 rate from 100 kHz to a final output rate of roughly 1 kHz as shown in Figure 3.13. The system uses a combination

of the L1 CTP ROI (muon, electron/photon, τ and jet) and the ROS information. More complex computation can be performed at this stage such as $E_{\mathrm{T}}^{\mathrm{miss}}$ calculations, but also can take into account detector reconstruction effects. The HLT itself is composed of two subsystems.

Level 2

The Level 2 trigger receives the information from the L1 system as ROIs. It uses both a hardware and a software implementation. The ROIs from L1 are combined in a single block using the RoI Builder (RoIB). Depending on the ROI type, different information will be retrieved from the ROS. For instance, if an electron ROI is sent, the information from the electromagnetic calorimeter corresponding to the ROI, as well as the inner-detector information will be retrieved. Doing so helps to reduce the size of the data requested, and thus speeds up the process (the typical fraction of events requested is 5%). The information is then passed along to a multi-threaded CPU farm that exploits various algorithms set by the user to process the event on a time-scale of 10 ms per event. Contrarily to Level 1 more complicated information can be used, for instance, the HLT farms can set B-physics streams by using the track reconstruction. A more refined definition of the objects is also provided through the ROS compared to the Level-1, helping for instance to strengthen the isolation criteria. The flexibility of the code allows having almost in-time pre-scaling and correction, especially for the track-based trigger menus.

The Event Filter

Once the L2 decision is known, the accepted event is passed to the Event Builder. All data fragments from the ROS are then collected and gathered to be passed along thanks to the Event Builder (EB). The fragments are also deleted from the ROB during that step. A dedicated partial reconstruction and calibration can be also performed at this point for special calibration streams, expressed physics streams used for data monitoring, and a debug stream used to monitor the performance of the TDAQ system. The Event Filter (EF) itself is constructed with the same structure as the Level 2 and uses a multi-threaded CPU farm with similar algorithms. Compared to Run-1, Level 2 and EF algorithms have been merged in the same farm in order to exploit the full flexibility of the system. The final information is then transmitted to the Tier-0 facility at CERN's computing centre for offline reconstruction.

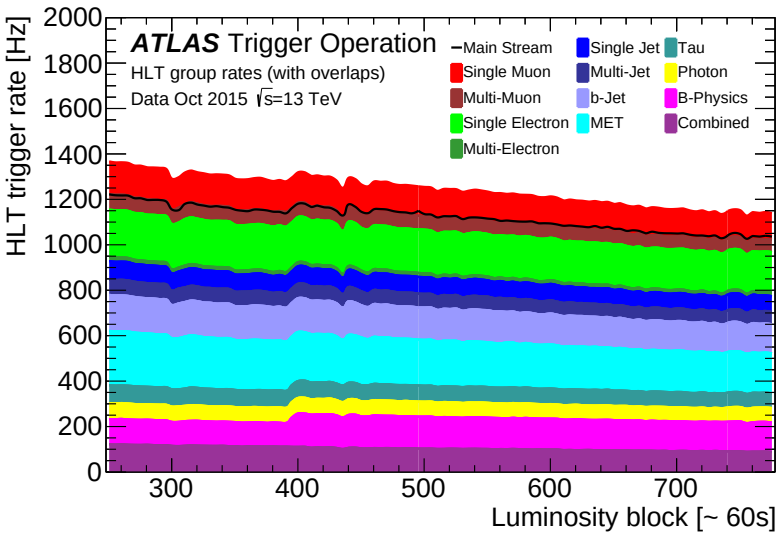


Figure 3.13 – HLT trigger rates evolution with time during the same LHC fill as Figure 3.4 [102].

3.2 Object reconstruction

The events passing the trigger selections described above are composed of physics signals from the detector. In order to conduct the final $Vh(b\bar{b})$ analysis, it is crucial to group those signals into objects used in particle physics such as single elementary particles or more complex objects. Their properties must also be reconstructed and calibrated to take into account the detector resolution and possible biases introduced. The objects considered here are electromagnetic objects (section 3.2.1), hadronic objects with a special emphasis on flavour tagging (section 3.2.2) and objects based on the global $E_{\mathrm{T}}^{\mathrm{miss}}$ object (section 3.2.3) and the $E_{\mathrm{T}}^{\mathrm{miss}}$sig.

3.2.1 Electromagnetic objects

Outside the tracker, the first ATLAS sub-detector interacting with charged and neutral objects generated in the event is the electromagnetic calorimeter. Special selections are optimised to identify photons and leptons. The former are not relevant for the $Vh(b\bar{b})$ analysis and are therefore not discussed here. Charged leptons are used instead to classify the final states studied in this work and their selection is reported here.

3.2.1.1 Electrons

Electrons are reconstructed from the electromagnetic calorimeter clusters using a sliding window algorithm [114]. The clusters are then matched to tracks refitted to take into account bremsstrahlung energy losses [115]. In order to reject converted photons, only tracks with vertices compatible with the primary interaction vertex of the hard collision are kept. Electron candidates are then classified using a likelihood-based method exploiting both information from the inner-detector and the electromagnetic calorimeter such as shower shapes, track-cluster matching quantities, information from the TRT, other track properties and bremsstrahlung-related variables. The list of variables is reported in [116]. Since the calorimeter response to electrons can be parametrised as a function of the E_T and the pseudo-rapidity η, the likelihood pdfs are derived in this 2-D map. The results are obtained from MC simulations, corrected to reproduce data efficiencies, so data-MC SF are extracted [116]. Efficiencies of identification for three likelihood purities are shown in Figure 3.14. In order to increase the purity of identification, an isolation criteria can be requested. It is defined both for the tracks and for the calorimeter clusters as the sum of the momentum of tracks/clusters around the electron. Several cuts define working points and are tuned from Z to two leptons MC samples. The loose (tight) criteria is used to keep 99% (95%) of real electrons. The $Vh(b\bar{b})$ analysis defines three identification working points based on those variables to classify the electrons:

- Loose: electrons are required to have a $E_T > 7$ GeV and $|\eta| < 2.47$, to pass a loose cut on the likelihood ratio, and to satisfy a loose isolation requirement.
- Medium: electrons are required to pass the loose criteria but with a tighter $E_T > 25$ GeV criteria.
- Tight: electrons must pass the medium VH criteria aling with a medium likelihood cut, and a tight isolation criteria.

A summary of the selection criteria is presented in Table 3.3.

The energies from the electron candidates are then corrected to take into account mismeasurement from the detector and data-MC differences [93]. First a multi-variate analysis trained on calorimeter variables such as the energy deposited per layers is performed to have a better estimate of the real energy of the electron starting from the measured one. Adjustments and corrections are then added to take into account non-uniformities, relative energy scales of the different layers, and data-MC relative scales. These corrections can be based on data or MC studies: events such as Z boson decays to an electron–positron pair are used for calibrating high energy electrons, while for instance $J/\psi \to ee$ are useful for low-energy calibration. Energy calibration and resolution are parametrised through:

$$E^{data} = E^{MC}(1+\alpha_i) \qquad \left(\frac{\sigma_E}{E}\right)^{data} = \left(\frac{\sigma_E}{E}\right)^{MC} \oplus c_i$$

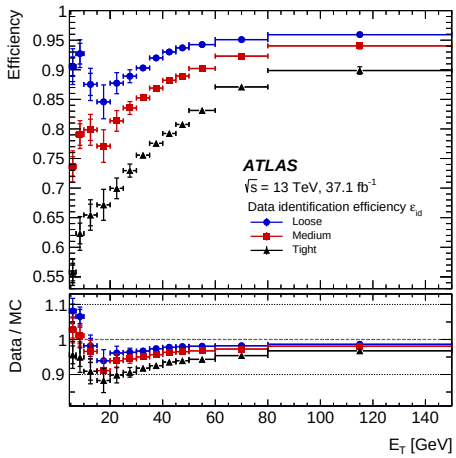
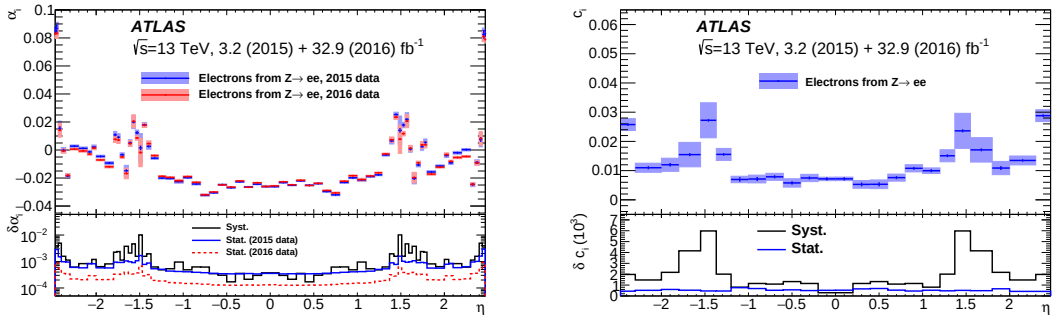


Figure 3.14 – Measured electron-identification efficiencies in Z→ ee events for the Loose (blue circle), Medium (red square), and Tight (black triangle) operating points as a function of η. The data efficiencies are obtained by applying data-to-simulation efficiency ratios that are measured in J/$\psi \rightarrow$ ee and Z→ ee events to the Z→ ee simulation [117].

where $\oplus$ denotes the sum in quadrature and where α_i and c_i are the energy scale factor and resolution for the electrons divided into η regions denoted by the index i. Their values are shown respectively in Figure 3.15a and Figure 3.15b.



(a) (b)

Figure 3.15 – Performance of the electron energy resolution using $\mathcal{L} = 36\ fb^{-1}$ of 2015 and 2016 ATLAS data [93].

3.2.1.2 Muons

Contrary to electrons, muons are feebly interacting in the calorimeters while having quite a long lifetime. Different reconstruction algorithms are applied depending on the η of the muon [118].

— Combined muon: In the region $|\eta| < 2.5$ most of the muons are reconstructed from the information of the inner detector and the muon chambers fitted separately first and then combined with an outside-in fit;

— Segment-tagged muon: If one or more chamber(s) do not have tracks but the muon candidate is consistent with a track in the inner detector, the muon is saved to recover efficiency for low p_{T} muons;

— Calorimeter-tagged muon: The central region ($|\eta| < 0.1$) is partially instrumented in muon chambers, therefore to recover some efficiency muon candidates are matched to minimum ionising compatible deposit in the calorimeters and tracks in the inner detector

— Extrapolated muons: On the forward region ($2.5 < |\eta| < 2.7$), the inner-detector is not instrumented. A standalone script is then developed to recover the muon p_{T} from the tracks in the muon chambers.

Furthermore, extra criteria are applied on the Combined and Extrapolated muons, a minimal number of hits in the chambers is set and, to avoid muons from in-flight decay of kaons or other hadronic particles, a charge over momentum significance $(\mathrm{q/p})$ [1] cut is used. When matched to an inner-detector track, only muons with vertices compatible with the primary interaction vertex of the hard collision are kept to reject muons from pile-up and cosmic rays. The identification efficiencies are presented in Figure 3.16a. The selected working point corresponds to medium muons except in the $|\eta| < 0.1$ where loose muons are used. Since a real muon can be matched to several reconstruction categories, an overlap removal technique is applied in the central region to avoid duplicates. When muons share a same inner-detector track a preference is given to combined muons over segmented-tagged muons and finally calorimeter-tagged muons.

On top of that three identification levels are set:

— Loose: muons are required to have a $\mathrm{E_T} > 7$ GeV, to pass a Loose muon selection [118], and to have an isolation requirement similar to that of electrons designed to achieve a 99 % muon efficiency.

— Medium: muons are required to pass the loose criteria but with a tighter $\mathrm{E_T} > 27$ GeV and $|\eta| < 2.5$ criteria.

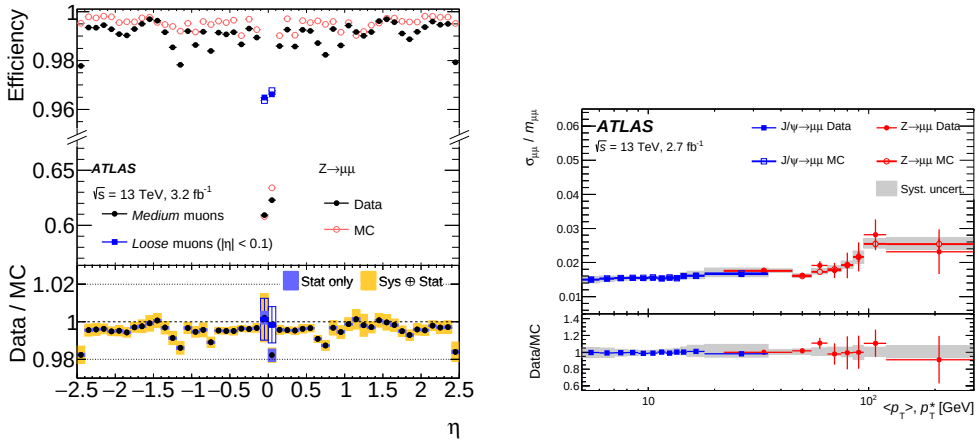
— Tight: similar to the medium criteria, muons must pass the loose VH criteria but with a tighter $\mathrm{E_T} > 25$ GeV, $|\eta| < 2.5$ and a tighter isolation criteria with a 95% efficiency for real muons.

A summary of the selection criteria is presented in Table 3.3.

With the use of the di-muon final state for the Z and J/ψ decay, the muon momentum in MC is then calibrated. Although the reconstruction step of the simulation is quite accurate, it cannot reach the performances obtained with the detector (scale and resolution at the permile and percent level). Once corrected data/MC agreement is

1. defined as the absolute value divided by the sum in quadrature of all the uncertainties.

consistent within the various uncertainties, and the mass resolution is extracted for the two decays with a resolution at the per cent level as shown in Figure 3.16b.



(a) (b)

Figure 3.16 – Performance of the muon reconstruction (a) and momentum resolution (b) using $3.2 fb^{-1}$ of 2015 ATLAS data [118].

lepton	Selection	p_{T}	η	ID	d^0_{sig}	$\lvert\Delta z_0 \sin\theta\rvert$	Isolation
e	Loose	>7 GeV	$\lvert\eta\rvert < 2.47$	loose	< 5	< 0.5 mm	loose
	Medium	>27 GeV	$\lvert\eta\rvert < 2.47$	loose	< 5	< 0.5 mm	loose
	Tight	>27 GeV	$\lvert\eta\rvert < 2.47$	tight	< 5	< 0.5 mm	tight
μ	Loose	>7 GeV	$\lvert\eta\rvert < 2.7$	loose	< 3	< 0.5 mm	loose
	Medium	>27 GeV	$\lvert\eta\rvert < 2.5$	loose	< 3	< 0.5 mm	loose
	Tight	>25 GeV	$\lvert\eta\rvert < 2.5$	medium	< 3	< 0.5 mm	tight

Table 3.3 – Summary of the charged lepton identification criteria.

3.2.1.3 Taus

Being the heaviest lepton discovered so far, the taus have a very short lifetime ($c\tau = 87.03\mu m$ [28]) which even in boosted topologies results in them typically decaying before reaching the inner-detector. Almost all the taus decay to either one charged particle (1-prong, 84.71%) or three charged particles (3-prongs, 15.20%), with the remaining decays at the per-mille level. Furthermore, the final state is either leptonic (mainly $\tau_{lep} \to l\bar{\nu}_l\nu_\tau$, $\tau_{lep} \to l_1\bar{l}_1 l_2\bar{\nu}_{l_2}\nu_\tau$ being strongly suppress) in 35% of the cases, or hadronic ($\tau_{had} \to hadron(s)\nu_\tau$) in 65% of the cases, with charged pions representing 94 % of the hadrons created. In the first case leptonic taus enter the analysis through the electron and muon selection described previously. Only hadronic taus are explicitly reconstructed. The hadronic taus are reconstructed in the hadronic calorimeter using the anti-Kt algorithm defined in Section 3.2.2.1 with a radius parameter of 0.4. Each tau cluster is required to contain either one or three small R=0.2 sub-clusters corresponding to the decay products. The reconstructed taus must pass some kinematic cuts such as $p_{\mathrm{T}} >$

20 GeV and $|\eta| < 2.5$ excluding the transition region between the barrel and forward calorimeters ($1.37 < |\eta| < 1.52$). Tau candidates are then required to pass a classification BDT to further reject background from standard hadronic jets [119]. This BDT is trained on information from calorimeters and tracking detectors. Three BDT working points are defined for different signal efficiencies. In the $Vh(b\bar{b})$ analysis, the medium working point is used and corresponds to an efficiency of identification and reconstruction for hadronic taus of about 55% in the 1-prong case and 40% in the 3-prong case. The performance of the BDT is shown in Figure 3.17, the solid lines are derived from constant BDT threshold requirement while the points are derived from a p_{T} flat efficiency, hence explaining the differences between the two.

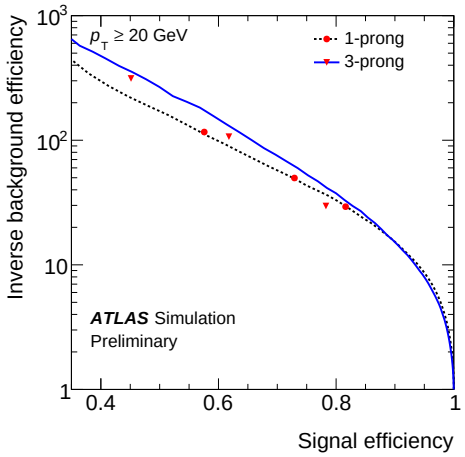
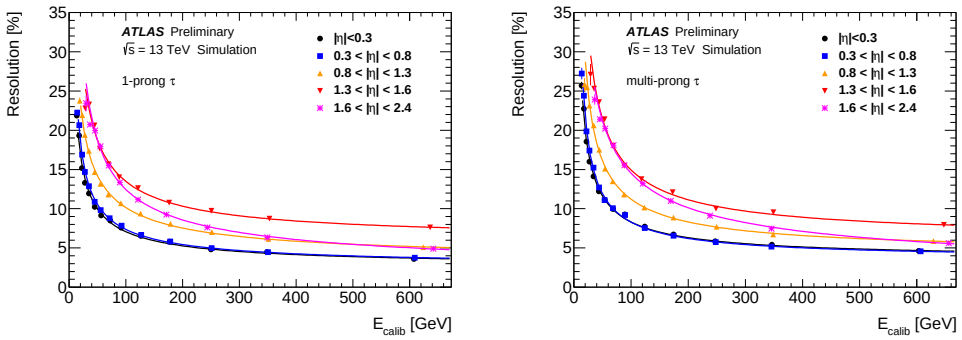


Figure 3.17 – Performance of the tau reconstruction using simulated samples aimed at reproducing early Run-2 conditions [119]. The solid lines are derived using a constant BDT threshold requirement while the points are using a p_{T} flat efficiency criteria.

Once identified, the hadronic tau jets are calibrated through the Local Cluster Weighting (LCW) technique [120]. However, to compensate for pile-up effects and to take into account the specific characteristics of the tau jets (such as the hadronic composition or the cone size of the sub-jets), a dedicated set of scale factors for the $\mathrm{E}^{\tau}_{\mathrm{LC}}/\mathrm{E}^{\mathrm{vis}}_{\mathrm{true}}$ as well as the pseudo-rapidity of the candidate η and the number of vertices n_{vtx} corrections are computed [119, 121]. The energy resolution is then computed from MC samples thanks to the width of the tau candidates at the Local Cluster scale to their true energy. The $\mathrm{E}^{\tau}_{\mathrm{LC}}/\mathrm{E}^{\mathrm{vis}}_{\mathrm{true}}$ resolution is shown in respectively Figure 3.18a and Figure 3.18b for the 1- and multi-prong decays. An extra offline correction is set on the absolute tau energy scale (TES, $\mathrm{E}_{\mathrm{calib}} - \mathrm{E}^{\mathrm{vis}}_{\mathrm{true}}$) and the data-MC modelling on the visible mass of the tau candidates. This allows as well to assess systematics on the final reconstruction of the tau energy that are propagated through the analysis.

3.2.2 Hadronic objects

In the proton-proton collisions, the elementary processes are governed by the interactions of quarks and gluons. Due to the structure of the strong force, these objects



(a) (b)

Figure 3.18 – Performance of the tau energy resolution for the 1-prong decay (a) and the multi-prong decay (b) using simulated samples aiming at reproducing early Run-2 conditions [119].

cannot be observed isolated in the detector, but rather hadronise around the initial parton creating a group of hadronic particles called jets. From the experimental point of view, the resulting objects will be defined as a collection of hadronic sub-constituents with a signature in the hadronic calorimeter and the possibility of adding the track information. The definition of the collection is developed in Section 3.2.2.1, the calibration of the jet energy and its resolution is summarised in Section 3.2.2.2 and the identification of the initial parton is presented in Section 3.2.2.4.

3.2.2.1 Jet reconstruction

In order to get a consistent definition over the full analysis and between data and MC, the jets need to be constructed with an algorithm which is safe with respect to experimental issues such as mismeasurement or miscalibration effects, and from the theoretical point of view (infra-red or ultra-violet corrections).

Identification of the sub-constituents

Depending on the nature of the events (data or truth MC), sub-constituents of the jets can be defined, originating from different aspects. In the case of data or in reconstructed MC, clusters of cells are identified from both the hadronic and electromagnetic calorimeters. In order not to consider signals originating from noises (electronic and from pile-up), the cluster formation is dictated by the cell signal significance defined as:

$$\zeta = \frac{\mathrm{E_{cell}}}{\sigma_{\mathrm{cell}}}$$

where $\mathrm{E_{cell}}$ represents the calorimeter cell ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$) energy measurement, and σ_{cell} the average expected cell's noise measured at the electromagnetic energy scale [122]. A seeding algorithm is used to form topological clusters from $|\zeta| > 4$ cells. Adjacent cells with $|\zeta| > 2$ cells are added to the topo-clusters. The cluster is growing until no more cells

with $|\zeta| > 2$ are found. The final topo-cluster can show several maxima with $E_{cell} > 500$ MeV. In such situation, if the local maxima are surrounded by four cells in the cluster, a splitting algorithm is performed to disentangle the local topo-clusters in order to improve the identification of close objects such as in boosted regime.

In the case of truth MC samples the truth-jets are using the truth stable particles which have a lifetime exceeding $c\tau > 10$ mm, but rejecting truth muons and truth neutrinos.

Clustering the particles into jets

From the sub-components identified above, the final jets are constructed using the anti-k_t algorithm [123]. This sequential algorithm uses the distance between the sub-components i and j $d_{i,j}$ and the distance to the beam $d_{i,B}$ defined as:

$$d_{i,j} = \min\left(\frac{1}{p_{T_i}^2}, \frac{1}{p_{T_j}^2}\right)\frac{\Delta_{i,j}^2}{R^2}$$

$$d_{i,B} = \frac{1}{p_{T_i}^2}$$

where $\Delta_{i,j}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ is the angular distance between sub-components i and j, and R is a free parameter related to the jet opening. The $d_{i,j}$ and $d_{i,B}$ distances are computed for the i-th object with respect to all the j-th objects. The smallest quantity is taken: if it is one of the $d_{i,j}$, the two objects are merged into a new object that is kept in the list; otherwise i is called a jet and removed. The procedure is repeated until no elements are found in the list. From the definition of the distance $d_{i,j}$, the jets tends to structure into circular shapes in the $\eta - \phi$ plane with a typical radius governed by the parameter R. In most of the ATLAS analysis, such as the $Vh(b\bar{b})$ analysis, the value 0.4 is set for the R parameter.

Other sequential algorithms are available for jet definition such as the k_t [124] and the Cambridge/Aachen [125] algorithm. From the theoretical point of view, the anti-k_t algorithm is robust against soft particle radiation as the soft particles are added to the jet without modifying the final shape. From the experimental point of view the anti-k_t algorithm is particularly robust against pile-up, but as well in term of jet reconstruction efficiency and purity [126].

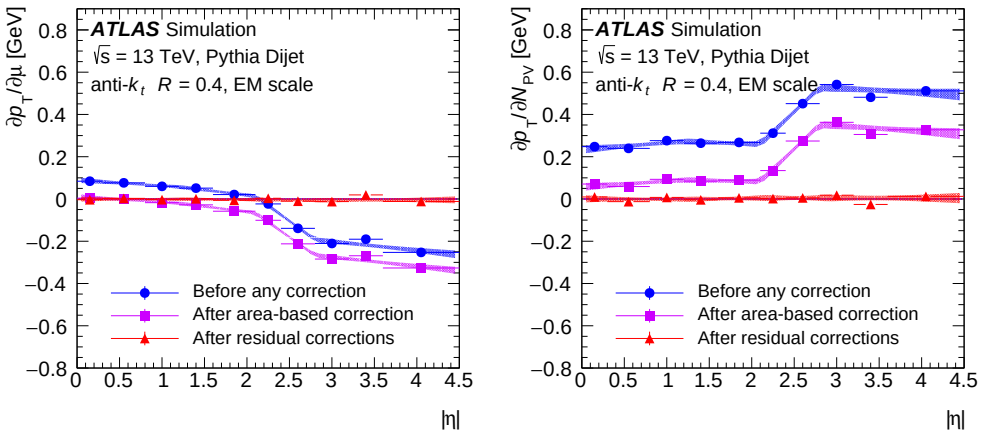
3.2.2.2 Jet energy calibration

The topo-clusters introduced in Section 3.2.2.1 use the electromagnetic scale for their energies. To correct for the differences from the cluster shape and calorimeter response of hadronic objects, a Jet Energy Scale correction [127] is performed on the jets. Several steps are performed sequentially:

— Origin correction: aims at correcting the jet position while keeping the energy constant. This results in having a better η resolution.

— Pile-up correction: aims at subtracting the excess energy due to in-time and out-of-time pile-up. Two corrections are derived. A component linked to the jet-area [128] is derived from simulated MC in the central region ($|\eta| < 2$). To compensate for the missed sensitivity in the forward region and higher occupancy of high-p_{T} jets, a residual correction binned in $|\eta|$ is calculated from MC. The variation of the corrected p_{T} versus μ and the number of primary vertices $\mathrm{N_{PV}}$, is shown respectively in Figure 3.19a and in Figure 3.19b.

— MC-based calibration: aims at adjusting the energy scale from the EM to the particle level scale as well as correcting for η biases caused by the calorimeter geometry. The JES corrections are derived from MC di-jet samples and parametrised in terms of η and $\mathrm{E^{reco}}$ of the jets.



(a) (b)

Figure 3.19 – Performance of the pile-up correction algorithms. The pile-up corrected p_{T} variation over μ (a) and $\mathrm{N_{PV}}$ (b) versus η illustrate the impact of the corrections [127].

— Global sequential calibration: aims at correcting effects from the jet composition and energy distribution. A sequential calibration based on five calorimeter and associated tracks based quantities [127] is obtained through MC di-jet simulation. The global energy scale is not affected by these corrections.

— Residual in-situ calibration [129]: aims at recovering the data-MC disagreements coming from imperfect detector simulation, collision mismodelling (such as pile-up or hard scattering) as well as shower development in the calorimeter. Corrections are extracted from the in-situ ratio of $\mathcal{R}_{in-situ} = \langle \frac{P_T^{jet}}{P_T^{ref}} \rangle$, where the ref is the known object used to calibrate the jet. This ratio

is calculated in data and MC samples to extract the MC corrections. The p_{T} of the jets in the central region are calibrated using a photon, a Z boson or multi-jet recoiling against the probe jet. Each method targets a specific p_{T} range and are combined using a single weight as shown in Figure 3.20. An η inter-calibration is performed in order to correct the jet p_{T} at high-η, using well-calibrated reference jets recoiling against a probe jet in the forward region.

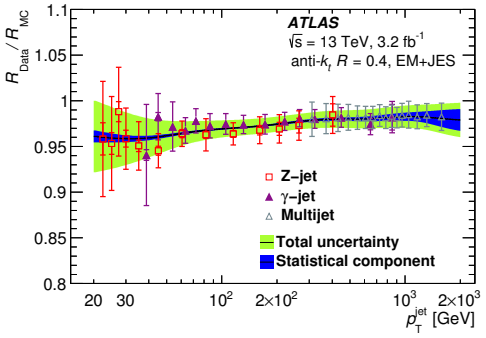


Figure 3.20 – Performance of the residual in-situ calibration algorithm, which represents the scale factors used to correct for the data-MC disagreement [127]

3.2.2.3 Jet selection

Once the calibration is applied, a few requirements on the quality of jets are required. First, any fake-jets not originating from the preliminary collisions must be discarded. They are likely to come from cosmic ray showers, proton losses upstream of the interaction point or from large scale coherent calorimeters cell noise [130]. The discrimination carried out using some cuts on key variables linked to the quality of the signal pulse shape in the LAr calorimeter, the fraction of energy deposited in the calorimeters with respect to the total energy of the jet, and the fraction of p_{T} coming from the tracks associated to the primary vertex. Two working points are available (BadLoose and BadTight), where the tight requirement is designed for searches at high jet p_{T} and large $E_{\mathrm{T}}^{\mathrm{miss}}$. The loose working point is used in the $Vh(b\bar{b})$ analysis.

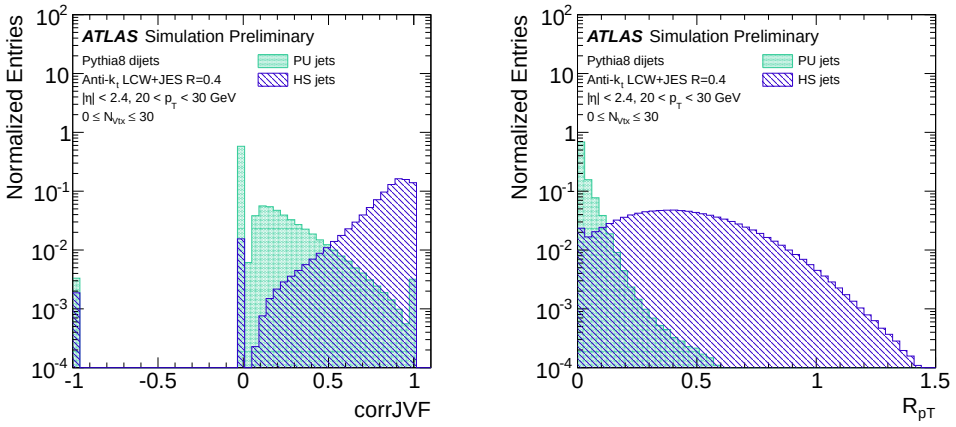
Then, the contribution from in-time and out of time pile-up is suppressed. These contributions can arise from hard QCD collisions as well as coming from local fluctuations which result from the superposition of random particles from pile-up jets. Therefore the tracking information can be used to reject such jets (only for jets of $|\eta| < 2.4$). From the Run-1 prescription to the Run-2 usage a corrected Jet Vertex Fraction (corrJVF) has been designed as [131, 132]:

$$\mathrm{corrJVF} = \frac{\sum_k p_T^{trk_k}(PV_0)}{\sum_l p_T^{trk_l}(PV_0)} + \frac{\sum_{n\geq 1}\sum_l p_T^{trk_l}(PV_n)}{k.N_{vtx}}$$

where $\sum_k p_T^{trk_k}(PV_0)$ is the scalar p_{T} sum of the tracks originating from the hard scattered vertex and asociated to a jet, the $\sum_{n\geq 1}\sum_l p_T^{trk_l}(PV_n)$ denotes the scalar p_{T} sum of the tracks originating from the pile-up interactions that is corrected by $k.N_{vtx}$ where $k = 0.01$ is used to correct the behaviour from linear increase with the number of tracks. The distribution of the variable for pile-up jets and hard scattered jets is shown in Figure 3.21a. Another discriminating variable was set as the ratio [131, 132]:

$$R_{P_T} = \frac{\sum_k p_T^{trk_k}(PV_0)}{p_T^{jet}}$$

which represents the fraction of the p_{T} coming from the hard scattered vertex and is shown in Figure 3.21b.



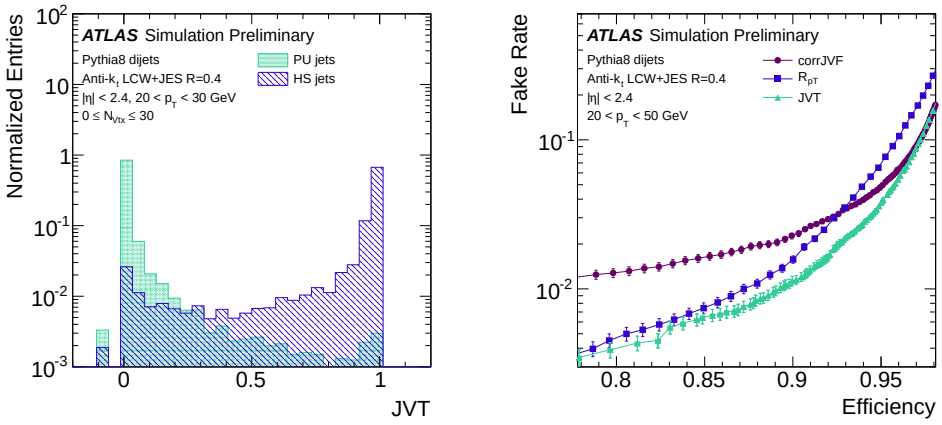
(a) (b)

Figure 3.21 – Distribution of the two discriminating variable corrJVF (a) and R_{P_T} (b) obtained through MC dijet simulation for both Pile-Up jets (PU jets) and Hard Scattered jets (HS jets) [131].

The combination of the two variables inside a 2- dimensional likelihood, based on a k-nearest neighbour (kNN) algorithm [133], results in jet-vertex-tagger (JVT) discriminant [131, 132] and is shown in Figure 3.22a. The final discrimination allows a good rejection of fake pile-up jets while retaining a good jet selection. In the $Vh(b\bar{b})$ analysis, the medium working point chosen corresponds to a final hard scattered jets efficiency of 92% and can be compared to the pure JVF cut that was applied in Run-1 from Figure 3.22b.

3.2.2.4 b-jet tagging

By construction, jets typically originate from quarks and gluons. Some specific algorithms are developed to understand the initial nature of such jets such as the



(a) (b)

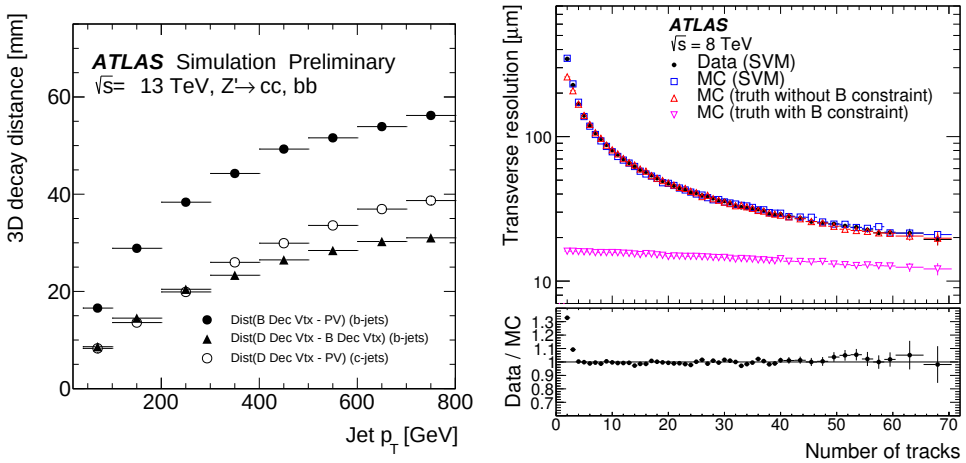
Figure 3.22 – Performance of the JVT algorithm. The distribution of the JVT score is shown in (a), while the final rejection *vs.* efficiency is shown in (b) [131].

quark/gluon taggers [134]. The importance of identifying the jets created by the fragmentation of b-quarks is at the root of the b-quark final state analysis such as the $Vh(b\bar{b})$. It relies on the relative long lifetime of b-hadrons ($c\tau(B^{\pm}) = 491.1\mu m$ [28]), which makes it possible to identify a secondary displaced vertex. For example, given the mass of the $B^{\pm}$ $m_{B^{\pm}} = 5.279$ GeV [28] with a momentum of 50 GeV, the decay length will be ~ 4.7 mm which is an order of magnitude larger than the usual primary vertex transversal position resolution ($\sim 100\mu$m) [135] as shown in Figure 3.23b. Thanks to the Lorentzian boost, the distance can even increase up to a few dozen of mm as shown in Figure 3.23a.

The b-tagging algorithms are designed to discriminate between b-jets and light-, τ-, and c-jets, the latter has similar properties to b-jet and are also separated from light- and τ-jets in the performance. Algorithm use the tracking and some calorimeter properties. They can be classified into three classes:

— Impact Parameter Tagging Algorithms [137]: use the point of closest approach to the primary vertex: d_0, the transverse impact parameter defined in the $r-\phi$ plane and $z_0 sin(\theta)$ the impact parameter in the longitudinal direction. A significance is extracted from the two impact parameter components from the signed ratio to the respective track resolution. The sign helps to distinguish if the secondary vertex is in front or behind the primary vertex with respect to the jet direction. A global log likelihood ratio is determined from the distribution of the signed impact parameter significances:

$$\log(\mathrm{P_i}/\mathrm{P_j}) = \log\left(\frac{\prod_{\mathrm{m=1}}^{\mathrm{N}} \mathrm{PDF_i}(\mathrm{d_0}/\sigma_{\mathrm{d_0}}, \mathrm{z_0}\,\sin(\theta)/\sigma_{\mathrm{z_0}\,\sin(\theta)})}{\prod_{\mathrm{m=1}}^{\mathrm{N}} \mathrm{PDF_j}(\mathrm{d_0}/\sigma_{\mathrm{d_0}}, \mathrm{z_0}\,\sin(\theta)/\sigma_{\mathrm{z_0}\,\sin(\theta)})}\right)$$



(a) (b)

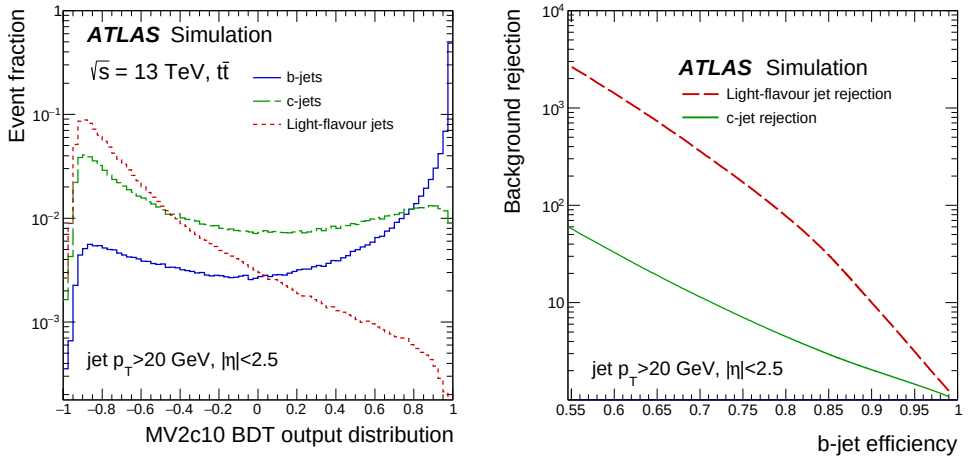
Figure 3.23 – (a) Average decay distance of the B- and D-hadrons with respect to the primary interaction vertex, and the relative distance between the B- and D-hadron decay vertices as a function of the jet p_{T} ([136]). (b) Resolution of the transverse primary vertex position as a function of the number of fitted tracks estimated using the split-vertex method (SVM) for minimum bias events and MC simulation. A comparison with the Beam spot constraint method is highlighted ([135]).

where the jet has N associated tracks that could be assigned to the flavour i or j. This algorithm, called IP3D tagger, uses several categories of tracks quality that are combined to scan various b-, c- and light-jet topologies [137].

— Inclusive Secondary Vertex Tagging Algorithms [138]: uses tracks as well as energy information to reconstruct the secondary decay point. All tracks that can be originated from interaction with the material (hadronic interactions in detector material, photon conversions...) or from prompt decays such as Λ or $\mathrm{K}^0_{\mathrm{S}}$ are rejected. Tracks are grouped by pairs and only the most relevant ones (passing a strong χ^2 cut) are kept and merged into one inclusive vertex. Some kinematical distributions of the secondary vertex are constructed and passed as the output of the algorithm.

— Decay Chain Reconstruction Algorithms [136]: uses track as well as energy information to compute the full decay channel of the hadrons in order to identify the b $\rightarrow$ c transition as most of the B-hadrons decay to a D-meson. The main assumption of the algorithm is to consider that the D-hadrons decay vertex is aligned with the B-hadron flight direction. After finding the vertex topologies (track-to-vertex association), a Kalman Filter fit is performed to extract the decay positions and associated uncertainties. The simulated reconstructed secondary vertex resolution obtained is slightly worse than the measured primary vertex one shown in Figure 3.23b (from ~ 0.4 mm to ~ 1.1 mm). Some kinematical distributions of the secondary and ternary vertex are constructed and passed as the output of the algorithm.

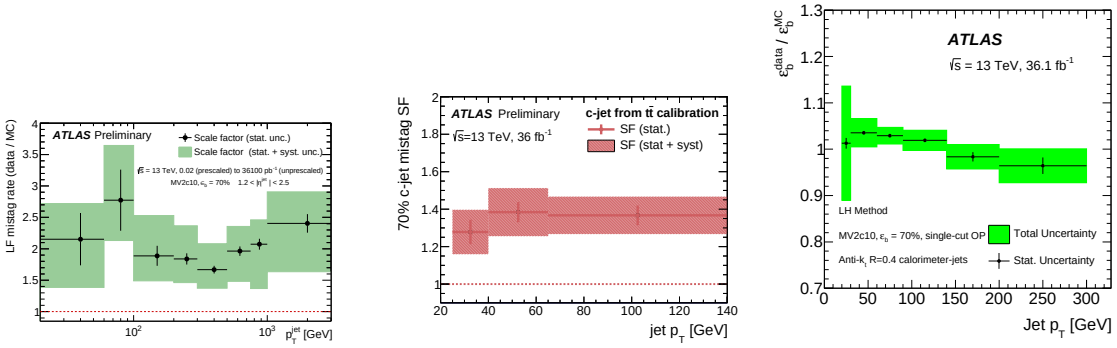
In order to perform a better rejection of c- and light-jets, the outputs of the individual algorithms are combined inside a single discriminating variable provided by a multivariate class algorithm called MV2c. This algorithm is trained on $t\bar{t}$ events where events are classified according to their flavour: b-jets as signal and c- and light-jets as background. The fraction of c- to light-jets can be tuned from the MC sample and determines the performance of the two rejection curves. The fraction chosen in the discriminant used by the $Vh(b\bar{b})$ analysis is 7% and corresponds to the MV2c10 algorithm [139]. The final BDT score is shown in logarithmic scale in Figure 3.24a. Several working points are defined based on the b-jet efficiency. In the $Vh(b\bar{b})$ analysis, the 70% efficiency working point is chosen and corresponds to a c-jet rejection of 12 and a light-jet rejection of 381. The rejection curve is shown in Figure 3.24b.



(a) (b)

Figure 3.24 – (a) Output score of the MV2c10 BDT for the three jet flavour coming from a $t\bar{t}$ MC sample. (b) b-jet efficiency *vs.* the c- and light-jet rejection coming from a cut on the MV2c10 BDT score [140].

The b-tagging algorithm is then corrected for data/MC scale factors. Those are derived for each jet-flavour, using a negative tag method for light-jets [141], semi-leptonic $t\bar{t}$ decays for the c-jets [142], and either a tag-&-probe or a combinatorial likelihood method for the b-jets [140]. Extracted SFs are shown in Figure 3.25. Specific uncertainties are also extracted by the same method and propagated through the analysis. Dedicated 2-D maps (p_{T} and η) are generated at the analysis level from the prescription of the ATLAS flavour tagging group to compute the MC b-, c-, light-, τ-tagging efficiencies.



(a) (b) (c)

Figure 3.25 – Data-MC scale factors for the light- (a) [141], c- (b) [142] and b-jets (c) [140] for the 70% working point of the MV2c10 algorithm.

3.2.3 Missing transverse energy

3.2.3.1 Definition

The kinematics of the proton bunches circulating in the LHC accelerator is characterised by a distribution of small transverse velocities. During the inelastic collisions happening at the LHC, hard scattered particles can bring a large fraction of the longitudinal momenta. However, momentum conservation results in the sum of the transverse momentum of all the particles in the final state to be small.

$$\sum \overrightarrow{p_T^{initial}} = \sum \overrightarrow{p_T^{final}} \approx \overrightarrow{0}$$

However, due to experimental reasons the balance can be imperfect:

— due to the nature of the interaction with the detector it may be impossible to reconstruct the momentum of some particle. This is the case for the SM neutrinos, involved in the $Vh(b\bar{b})$ analysis whose interaction cross-section is too small to get any signal. The same signature would be observed for dark matter candidates or for supersymmetric particles such as the neutralinos.

— In the same way, particles which are undetected or badly measured because they end up in any inefficient region of the detector, can contribute to the imbalance of energy in the transverse plane. Furthermore, the effect of pile-up events that are not fully rejected can lead also to imbalances.

The measurement of the imbalance is called Missing transverse momentum (MET) and is defined as:

$$\sum_{i \in visible} \overrightarrow{p_T^i} + \overrightarrow{E_T^{miss}} = 0 \Rightarrow \overrightarrow{E_T^{miss}} = - \sum_{i \in visible} \overrightarrow{p_T^i}$$

where $\overrightarrow{p_T^i}$ represent each of the measured momentum in the transverse plane. The sum of these momenta can be decomposed by each object category that has been defined previously:

$$\overrightarrow{E_T^{miss}} = -\sum_{i\in \mathrm{e}} \overrightarrow{p_T^i} - \sum_{i\in\mu} \overrightarrow{p_T^i} - \sum_{i\in hadronic\,\tau} \overrightarrow{p_T^i} - \sum_{i\in\gamma} \overrightarrow{p_T^i} - \sum_{i\in jets} \overrightarrow{p_T^i} - \sum_{i\in Soft\,Terms} \overrightarrow{p_T^i}$$

where the Soft Terms are defined as all the objects that are not identified in the other categories. Their contributions can be evaluated through the calorimeter topo-clusters (Calorimeter-based Soft Term) or through the ID tracks (Track-based Soft Terms, TST). In the first case, all particles creating a deposit in the calorimeter are taken into account, however due to the poor position resolution the contribution of pile-up is quite high. The second case has a good position resolution and can select tracks only coming from the hard vertices thus reducing the pile-up contribution, but cannot take into account the neutral particles. A second limitation is the reduced coverage of the ID system ($|\eta| < 3.0$) compared to the calorimeter ($|\eta| < 4.0$). The latter definition is used for the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction in Run-2 while the first was used during Run-1.

More variables can be defined in relation to the $\overrightarrow{E_T^{miss}}$ definition:

— the magnitude $E_{\mathrm{T}}^{\mathrm{miss}} = |\overrightarrow{E_T^{miss}}|$ that is the variable used in the $Vh(b\bar{b})$ analysis.

— the scalar sum of the momenta:

$$\sum E_T = \sum_{i\in \mathrm{e}} p_T^i + \sum_{i\in\mu} p_T^i + \sum_{i\in hadronic\,\tau} p_T^i + \sum_{i\in\gamma} p_T^i + \sum_{i\in jets} p_T^i + \sum_{i\in Soft\,Terms} p_T^i$$

that provides an assessment of the hardness of the collision which is linked to the resolution on the $E_{\mathrm{T}}^{\mathrm{miss}}$.

— the hard object scalar sum of the momenta:

$$H_T = \sum_{i\in \mathrm{e}} p_T^i + \sum_{i\in\mu} p_T^i + \sum_{i\in hadronic\,\tau} p_T^i + \sum_{i\in\gamma} p_T^i + \sum_{i\in jets} p_T^i$$

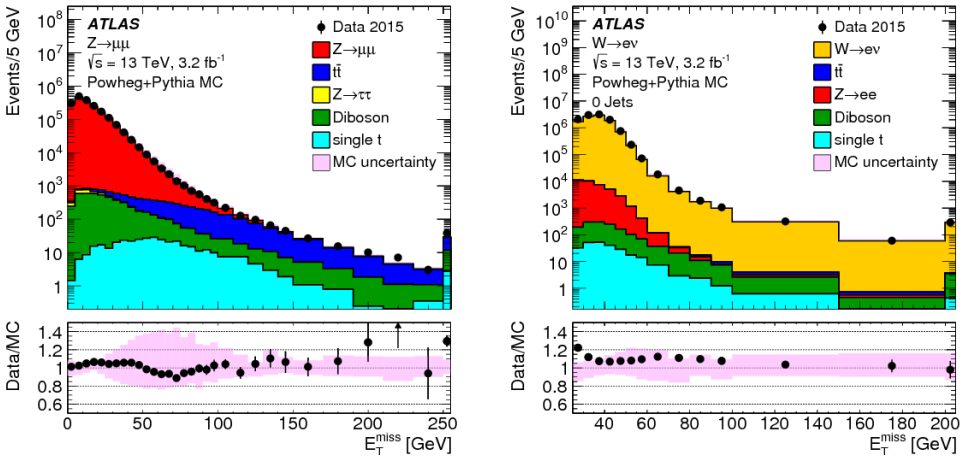
that provides similar criteria as the $\sum E_T$ while avoiding the possible mismodelling of the soft term.

— the track missing transverse momentum p_T^{miss} which is defined as the magnitude of the vector $\overrightarrow{p_T^{miss}} = -\sum_{i\in tracks} \overrightarrow{p_T^i}$ where only the momenta of tracks are used. This produces a more robust evaluation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in terms of pile-up but is limited by the coverage of the inner-detector and by the non

reconstruction of neutral particles. In the $Vh(b\bar{b})$ analysis this quantity is used to reject the QCD background contribution.

3.2.3.2 Performance and calibration

In order to assess the performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable and the various related quantities, two final state cases have been studied. The $\mathrm{Z} \to \mu\mu$ final state helps to assess the effect of detector resolution in the $E_{\mathrm{T}}^{\mathrm{miss}}$ since it does not present any true $E_{\mathrm{T}}^{\mathrm{miss}}$. The data-MC performance is shown in Figure 3.26a. The number of associated jets is used to understand the effect of soft terms (0 additional jets) or the effect of jets and pile-up. To assess the effect of genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ from neutrinos, the $\mathrm{W} \to \mathrm{e}\nu_{\mathrm{e}}$ final state is also studied. The data-MC performance is shown in Figure 3.26b.

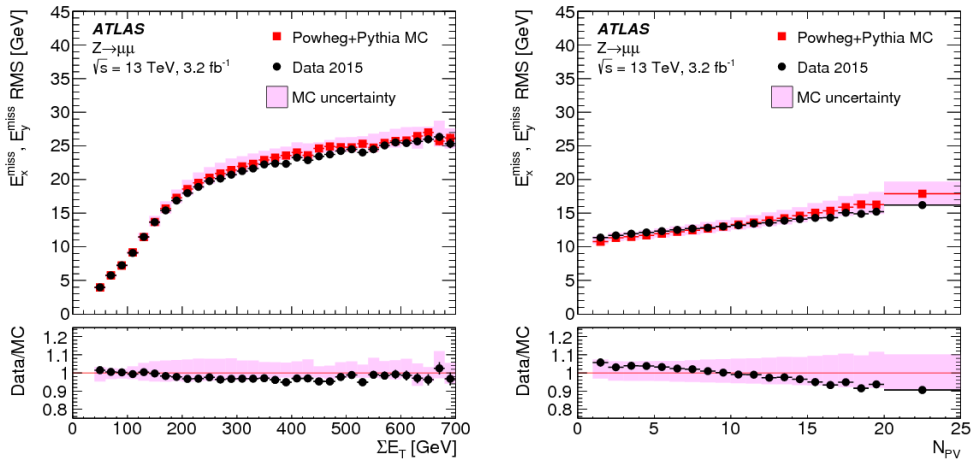


(a) (b)

Figure 3.26 – Distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable for two final state topologies, with no genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ for $\mathrm{Z} \to \mu\mu$ (a) or with genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ for $\mathrm{W} \to \mathrm{e}\nu_{\mathrm{e}}$ (b). Data-MC disagreements are consistent with statistical fluctuations except for low-p_{T} $\mathrm{W} \to \mathrm{e}\nu_{\mathrm{e}}$ where missing di-jet MC samples could explain the deficit [143].

Due to the long tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, the resolution is determined by the Root Mean Square value of the projection along with the x and y transverse coordinates of the $E_{\mathrm{T}}^{\mathrm{miss}}$. From Figure 3.27a, three regimes can be identified. At first, the resolution is dominated by the muon p_{T} resolution ($\sum E_T < 70$ GeV), then the possibility of adding one additional jet increases the resolution up to $\sum E_T \simeq 180$ GeV where the two jet events dominate. At high $\sum E_T$ the evolution is dictated by a $RMS \sim \sqrt{\sum E_T}$ law. The overall data-MC is quite good. The effect of pile-up is shown in Figure 3.27b. The slope of the energy resolution in data as a function of the number of primary vertices is flatter than in MC and is mainly coming from events containing additional jets.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ being a composed object, the only element that does not involve calibration and uncertainties is the soft term. The soft term $E_{\mathrm{T}}^{\mathrm{miss}}$ contribution uncer-



(a) (b)

Figure 3.27 – Distribution of the Root Mean Square value of the transverse projection of the $E_{\mathrm{T}}^{\mathrm{miss}}$, in bins of $\sum E_T$ (a) and in bins of the number of primary vertices (b) in an inclusive sample of $\mathrm{Z} \to \mu\mu$ events. The resolution term can be approximated by a square root behaviour for high enough $\sum E_T$ [143].

tainty is extracted from the soft term p_{T} projection onto the hard object direction and parametrised as a function of the hard object p_{T}. The analysis is conducted on a $\mathrm{Z} \to \mu\mu$ final state. The results are shown in Figure 3.28a. Similarly, the soft term resolution is derived in both the longitudinal and transversal direction with respect to the hard object p_{T} and the resulting uncertainty is extracted from Figure 3.28b and 3.28c.

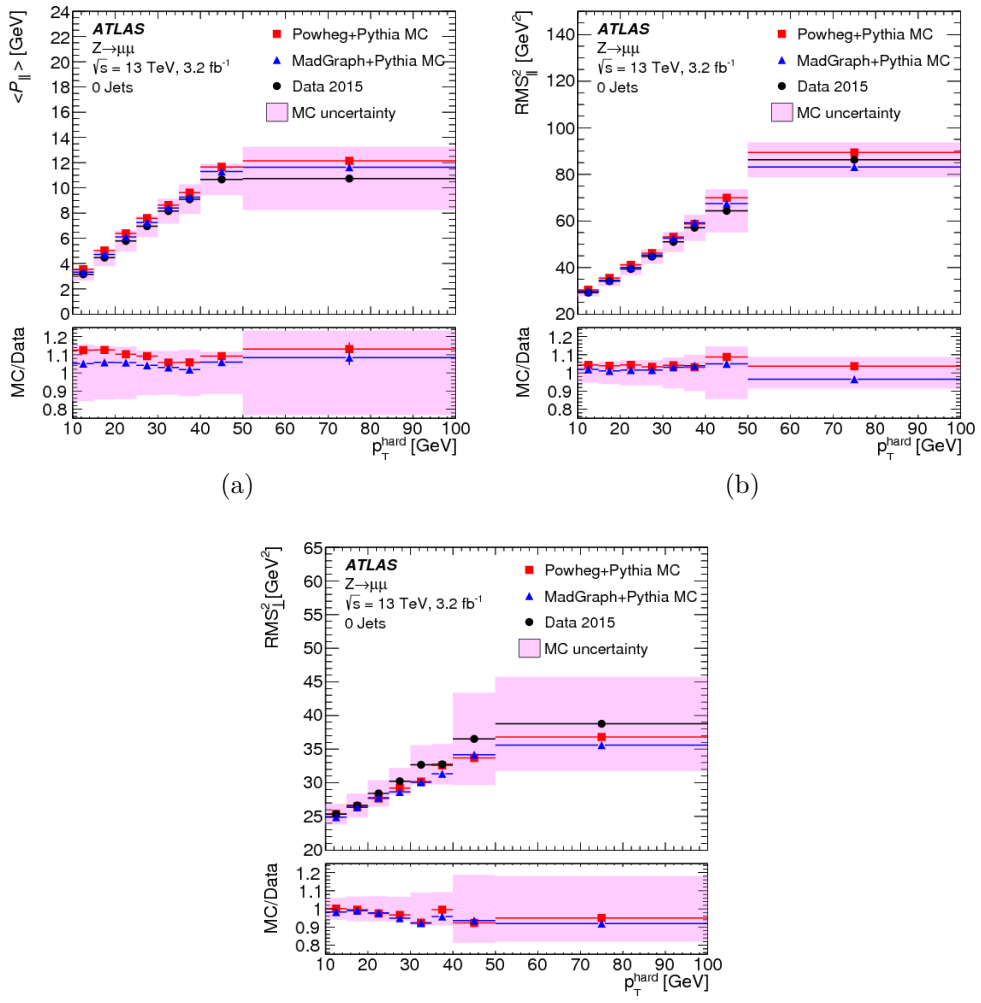
3.2.3.3 Missing transverse momentum significance

From the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of Figure 3.26, it is clear that both genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ resulting from mismeasurements can have large values. In order to separate the events that contain real $E_{\mathrm{T}}^{\mathrm{miss}}$ (in order to select events with neutrinos or dark matter particles) from fake $E_{\mathrm{T}}^{\mathrm{miss}}$, a new variable based on the resolution of the objects entering the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is created, called MET significance. The goal is to penalise events with objects with a bad energy resolution that could cause high $E_{\mathrm{T}}^{\mathrm{miss}}$ values. From Figure 3.27a, one choice for the approximation of the resolution of all the objects entering the $E_{\mathrm{T}}^{\mathrm{miss}}$ would be $\sqrt{\sum E_T}$, or in order not take into account the soft terms the relaxed $\sqrt{H_T}$ definition:

$$\mathcal{S} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{\sum E_T}} \text{ or } \mathcal{S} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{H_T}}$$

This simple event-based quantity is commonly used in the ATLAS collaboration as in the $Vh(b\bar{b})$ analysis where the resolution is approximated by using the p_{T} of the leptons in the 2-leptons category.

However, the true $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is only approximated by the $\sqrt{\sum E_T}$ for high $\sum E_T$, as shown in Figure 3.27a, this being more true for the TST definition than for the



(a) (b)

(c)

Figure 3.28 – Distribution of the soft term longitudinal p_{T} w.r.t. the hard object p_{T} (a). The MC simulation results are lying well inside the statistical bands. The soft term p_{T} resolution is shown both in the longitudinal (b) and transversal (c) direction [143].

CST one [144]. In order to improve the approximation, an object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is developed by taking the individual resolutions of all the objects entering the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation:

$$\mathcal{S}^2 = \left(\overrightarrow{E_T^{miss}}\right)^{\mathrm{T}} \left(\sum_{\mathrm{i}} \mathrm{V^i}\right)^{-1} \left(\overrightarrow{E_T^{miss}}\right)$$

where the V^i are the covariance matrices of the momentum of each object i that is entering the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. For each object, these covariance matrices can be written easily in the basis of the p_{T} of the object:

$$V^i = \begin{pmatrix} \sigma^2_{p_T^i} & 0 \\ 0 & p_T^{i\,2}\sigma^2_{\phi^i} \end{pmatrix}$$

where ϕ^i is the azimuthal angle of the object p_{T} in the regular (x,y) basis. All the resolutions of the objects considered are summarised in Table 3.29. The soft term resolution has been re-derived from $\mathrm{Z} \to \mu\mu$ MC samples with a jet veto and was found to be 8.9 GeV [145].

Object	Kinematic	Relative resolution	
		Parallel	Perpendicular
Electrons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.7%	0.4%
Photons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.9%	0.4%
Hadronic τ	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	5.5% – 6.7%	1%
Jets	$p_{\mathrm{T}} = 20$ GeV, $\eta = 0$	22%	4.6%–7.1%
	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	7%	1.1%–1.6%
Muons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	2%	0.1%
Track Soft Term		8.9 GeV	8.9 GeV

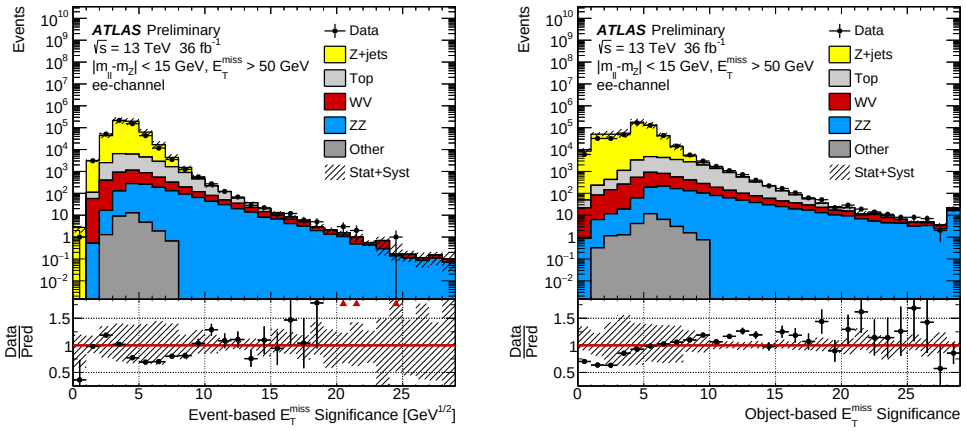
Figure 3.29 – Resolution for the objects entering the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for a representative p_{T} for each object coming for each individual calibration [145].

Once computed, all the covariance matrices are rotated to the natural basis for the $E_{\mathrm{T}}^{\mathrm{miss}}$ where one of the axes is aligned with the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. The final $E_{\mathrm{T}}^{\mathrm{miss}}$sig. can, therefore, be rewritten:

$$\begin{aligned}\mathcal{S}^2 &= \begin{pmatrix} E_T^{miss}, & 0 \end{pmatrix} \begin{pmatrix} \sigma_L^2 & \rho_{LT}\sigma_L\sigma_T \\ \rho_{LT}\sigma_L\sigma_T & \sigma_T^2 \end{pmatrix}^{-1} \begin{pmatrix} E_T^{miss} \\ 0 \end{pmatrix} \\ &= \frac{|E_T^{miss}|^2}{\sigma_L^2(1-\rho_{LT}^2)}\end{aligned}$$

where ρ_{LT} is the correlation factor of the longitudinal L and transverse T measurements, $\sigma_{L,T}$ being the variances in the two directions. Those quantities are calculated event-by-event with all the objects available. The modelling of this variable has been studied with the $\mathrm{Z} \to \mathrm{ee}$ final state selection on both data and MC [145]. The distribution of the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. is shown in Figure 3.30a while the object-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. is shown in Figure 3.30b. A $E_{\mathrm{T}}^{\mathrm{miss}}$ cut at 50 GeV has been set on the samples and data to extract high-met region. Some mismodelling appears at the low object-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. values, which is thought to come from a bad modelling of the pile-up in the MC generators. This mismodelling is spread over a larger range of values for the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig., which is a good sign of the localisation of the pile-up contribution.

The object-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. shows a better separation between the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ from the $\mathrm{Z} \to \mathrm{ee}$ while the true $E_{\mathrm{T}}^{\mathrm{miss}}$ $\mathrm{ZZ} \to \mathrm{ee}\nu\nu$ samples are shifted toward higher values. This separation power can be investigated by looking at the selection of the true $E_{\mathrm{T}}^{\mathrm{miss}}$ sample $\mathrm{ZZ} \to \mathrm{ee}\nu\nu$ over the rejection of the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ $\mathrm{Z} \to \mathrm{ee}$. The results are shown in Figure 3.31. For an 80% signal efficiency, the object-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. is able to have a 90% rejection power while the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. only shows a 70% rejection of fake $E_{\mathrm{T}}^{\mathrm{miss}}$.



(a) (b)

Figure 3.30 – Distributions in data compared to MC predictions including all relevant backgrounds for events in the $Z \to ee$ selection and $E_T^{miss} > 50$ GeV for: (a) event-based E_T^{miss}sig. (b) object-based E_T^{miss}sig. [143].

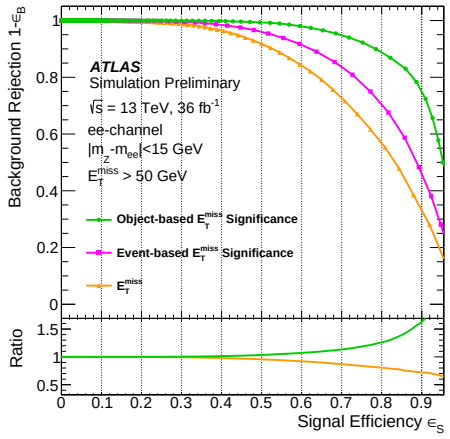


Figure 3.31 – Background rejection versus signal efficiency in simulated $Z \to ee$ and $ZZ \to ee\nu\nu$ samples with a $Z \to ee$ selection. [143].

4 Event selection and categorisation in the Vhbb analysis

As outlined in the introduction and Chapter 1 of this thesis the fundamental search for a Higgs boson decaying into a pair of b quarks is both theoretically exciting and experimentally challenging. Even with its large branching ratio of 0.581 ± 0.007 at a Higgs mass of $m_H = 125.09$ GeV [49], the $\mathrm{h} \to \mathrm{b\bar{b}}$ decay mode is overwhelmed by a huge $\mathrm{b\bar{b}}$ background. Run-1 measurements of the total fiducial cross-section of $\mathrm{b\bar{b}}$ pair production is on the order of magnitude of 10 nb [146, 147], while the total Higgs production cross-section is six orders of magnitude lower [49]. So as to increase the signal-to-background ratio, the approach of using the associate production with a vector boson is considered in this analysis. The leptonic decay of this boson provides an efficient way to trigger on such events. Thus the analysis is divided into three search channels depending on the number of observed charged leptons in the final state, excluding de-facto the hadronic τs: 0-lepton targeting the $\mathrm{Z} \to \nu_\ell \bar{\nu}_\ell$ decay (with a $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger signature due to the neutrinos), 1-lepton targeting the $\mathrm{W} \to \ell \nu_\ell$ decay and 2-leptons targeting the $\mathrm{Z} \to \ell^{\mp}\ell^{\pm}$ decay. While the global strategy is kept the same across the different channels, the background composition, the signal characteristic and the systematics differ leading to small but important analysis optimisations.

The key ingredient in the observation of the $V\mathrm{h}(\mathrm{b\bar{b}})$ process is a good separation between background and signal events. A Multi-Variate Analysis (MVA) using Boosted Decision Trees (BDTs) has thus been developed to serve as a final discriminant on which the statistical analysis described in Chapter 6 is conducted. The selection of the variables used for the definition of the BDTs as well as their performance is discussed in Section 4.4. A separate cross-check, aiming to study $\mathrm{VZ}(\mathrm{b\bar{b}})$ events using the same MVA approach, is also described here.
The understanding of the backgrounds is a crucial point in the analysis as it will determine our confidence in the description of the discriminant, and thus in the signal observation.

The background simulation is described in Section 4.1, along with the data selection. Analysis dependent reconstruction effects are described in Section 4.3: since b-jets are at the core of the analysis, their identification is primordial. In order to improve the statistical power of the MC samples used, truth- and direct-tagging methods are discussed (Section 4.3.2). The b-jets energies are then corrected in order to improve the di-jet mass resolution (Section 4.3.3). Finally, the overlap techniques to avoid double-counting of objects are discussed in Section 4.3.1.

This approach has shown in previous results to be more efficient than a simple Cut-Based Analysis (CBA) using the di-jet mass discriminant. Some of the selections have been kept to strengthen the performance of the MVA approach and are described in Section 4.2. A separate CBA analysis is still kept as a cross-check and is discussed inside this section.

4.1 Data and simulated samples

Data

This analysis uses datasets including pp collisions recorded with the ATLAS detector in stable beam conditions during Run-2 in 2015, 2016 and 2017 corresponding to an integrated luminosity of $\mathcal{L} = 79.8\ fb^{-1}$ collected at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. In order to ensure a good quality of the collected data, all the sub-detectors are required to be operational with good efficiency. The good quality of the analysed data is taken care of by the ATLAS Data Quality group. On top of this requirement, events corrupted from bad TileCal and LAr noise bursts are also removed.

Data events are also required to pass a trigger selection as mentioned in the Section 3.1. This selection is dependent on the analysis channel considered and summarised in Table 4.1.
In the 0-lepton channel, the decision is based on the magnitude of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. For each data taking period, the lowest unprescaled trigger is chosen. The HLT threshold was raised from 70 GeV in 2015 to 90 GeV in the first part of 2016 and finally to 110 GeV in 2017. With the latter, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is shown to be fully efficient for an offline reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ of 180 GeV.
In the 1-lepton channel, different triggers are used depending on whether the lepton is a muon or an electron. In the first case, the same $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger as in the 0-lepton channel is used. No important gain has been observed by requiring a single muon trigger for those events, and the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is fully efficient for high-p_{T} muons. As for the electron

selection, events are required to pass one of the lowest unprescaled electron triggers with p_{T} ranging from 24 to 120 GeV in 2015 and from 26 to 300 GeV in 2017. Low p_{T} electrons must pass an isolation criteria at the HLT step.
In the 2-leptons channel, the leptons are required to pass either the lowest unprescaled electron triggers defined in the 1-lepton case or the lowest unprescaled muon triggers with p_{T} ranging from 20 to 50 GeV in 2015 and from 26 to 60 GeV in 2017. As for the electrons in the 1-lepton channel, low p_{T} muons must pass an isolation criteria at the HLT.

Object	L1 thr. (GeV)	HLT thr. (GeV)	Extra criteria	Period
$E_{\mathrm{T}}^{\mathrm{miss}}$	50	70		2015
	50	90		2016 (A-D3)
	50	110		2016 ($\geq$ D4)
	55	110		2017
e	20	24		2015
		26	loose ID, loose isolation	2016 & 2017
	20	60		2015
		60	medium ID	2016 & 2017
	20	120		2015
		140	loose ID	2016 & 2017
		300		2017
μ	15	20	loose isolation	2015
		24	medium isolation	2016 (A-D3)
		26		2016 ($\geq$ D4) & 2017
		40		2015 & 2016 (A)
		50		2015 & 2016 & 2017
		60	MS only	2017

Table 4.1 – Lowest unprescaled triggers used during the 2015, 2016 and 2017 data collection periods. 'MS only' stands for the case where only muon spectrometer reconstruction occurs (no inner-detector information is combined). This was used to avoid discrepancies at high energies.

MC samples

In order to get a correct description of the data, MC samples are generated for the signal and most of the backgrounds (the special case of QCD jets is described in 5.1.2). Depending on the sample, different generators are used to properly model the PDF, Matrix Element, Parton Shower and hadronisation, or the Underlying Event part of the process [148]. All the generators are then processed with a GEANT4-based simulation of the ATLAS detector and the standard ATLAS reconstruction software to allow comparison with the data. The sample weights are then scaled to the best-known cross-section for the process, and corrected to get the generated number of interactions to match the distribution of the dataset considered thanks to the official Pile-Up Reweighting tool. The effect of triggers is also propagated to the MC samples through reconstruction of the trigger decision and specific corrections [149].

The signal samples include three modes: $\mathrm{Zh} \to \nu\nu\mathrm{b\bar{b}}$, $\mathrm{Wh} \to \ell\nu\mathrm{b\bar{b}}$, $\mathrm{Zh} \to \ell\ell\mathrm{b\bar{b}}$. While the Wh production mode is only a quark induced process, the Zh modes can

be quark or gluon induced. Therefore five samples are generated using the POWHEG generator for the Matrix Element (ME), with the addition of the MiNLO (Multiscale Improved NLO) procedure for quark induced processes [150], interfaced to the PYTHIA 8 [151] generator for Parton Shower (PS) and Underlying Events (UE) applying the AZNLO tune [152] and using the NNPDF3.0 PDF set [153]. In those simulations, a Higgs mass of 125 GeV is assumed. The total cross-section used to normalise the samples are calculated at NNLO in QCD and NLO in EW for the quark induced processes [49, 154, 155] and NLO+NLL(QCD) for the gluon induced ones [49, 156, 157]. However, some EW NLO corrections are supposed to have sizeable effects on the p_{T} spectrum of the vector boson. Hence a dedicated correction is extracted to correct for shape effects and applied for quark induced processes.

The V+jets samples are crucial for the analysis as they are a dominant background in all the lepton channels. All the processes are generated using the SHERPA 2.2.1 generator [158] using the NNPDFs 3.0 set for both the ME and PS/UE steps. b- and c-quarks are considered to be massless at the ME step, while massive quarks are produced at the PS step. The precision of the generator depends on the number of extra partons: for up to 2 extra partons the ME is computed at NLO in EW while for 3 or 4 extra partons the precision is only LO. The merging of the two regimes is provided through a scheme based on the Catani, Krauss, Kuhn, Webber and Lonnblad (CKKW-L) technique [159, 160]. Extra partons are generated at the PS level. Since the p_{T} spectrum of the vector boson is rapidly falling, the generation is divided into slices to increase the statistics in the tails. The variable used to define the slices is the maximum between the p_{T} of the vector boson and the sum of the momentum of all the hard objects called $\mathrm{H_T}$. The following slices are used:

$$[0-70, 70-140, 140-280, 280-500, 500-1000, >1000]\ \mathrm{GeV}$$

Furthermore, to enhance the number of events generated in the regions to which the analysis is most sensitive, dedicated flavour filters have been created at the generation level:

- b-filter: At least 1 b-hadron present, with $|\eta| < 4$.
- c-filter, b-veto: At least 1 c-hadron present, with $p_{\mathrm{T}} > 4$ GeV and $|\eta| < 3$, and a veto on events that would pass the b-filter.
- c-veto, b-veto: All the events rejected by the two previous selections.

Those filters are not applied to the highest $\max(p_{\mathrm{T}}^{\mathrm{V}}, \mathrm{H_T})$ events.

The resonant diboson production ZW, WW, ZZ is a sub-dominant process. While being well known and modelled, the hadronic decay of the Z boson into a $\mathrm{b\bar{b}}$ pair can mimic the signal using a resonant dijet mass close to the region of interest in the Higgs search. Events with a W can also mimic the signal when one of the two quarks

from the W decay is misidentified as a b. Those processes are simulated with the same generator as for V+jets. The only difference is the separation of the two partonic regimes that happens with one less extra parton. Extended VZ $\rightarrow b\bar{b}$ samples are also available and are used in addition to the nominal samples to reduce the statistical uncertainties for those processes.

The $t\bar{t}$ pair production is a dominant background in the 0- and 1-lepton channels. The processes can be classified depending on the decay mode of the W boson coming from the decay of the top quark: in non-all-had samples, at least one W boson decays leptonically, while both must in the dilepton samples. Both processes are generated using POWHEG as the ME generator using the NNPDF 3.0 PDF set (NLO) that is interfaced to PYTHIA 8 for the PS/UE generator applying the A14 set and using the NNPDF 2.3 PDF set (LO). The total cross-section used for normalisation is calculated at NNLO+NNLL [161].
top quarks can also be produced in association with another particle. Those subdominant processes can be classified in three categories: s- (tb final state), t- (tq final state) and Wt-channel (tW final state) as represented in Figure 4.1d. The same generators are used for the single-top generation. All the W bosons generated are required to decay leptonically. The cross-section used for the normalisation is calculated at NLO [162, 163]. The only peculiarity is the t-channel single top in the 2-leptons channel is simulated using a fast simulation of the detector (AF2) instead of the regular full simulation. The mass of the top quark is set to 172.5 GeV in all simulations.

In order to simulate the effect of the pile-up in all the MC production, samples are overlaid with minimum bias events that are generated using PYTHIA 8 applying the A3 tune and using the NNPDF 2.3 PDF set. Their simulation is set accordingly to the luminosity profile of the data period. As this profile differed between years, MC samples are generated independently for two periods: "a" for the luminosity profile of 2015 and 2016 and "d" for the luminosity profile of 2017. A higher number of events was requested for the second period to take into account the higher data statistics of this period.

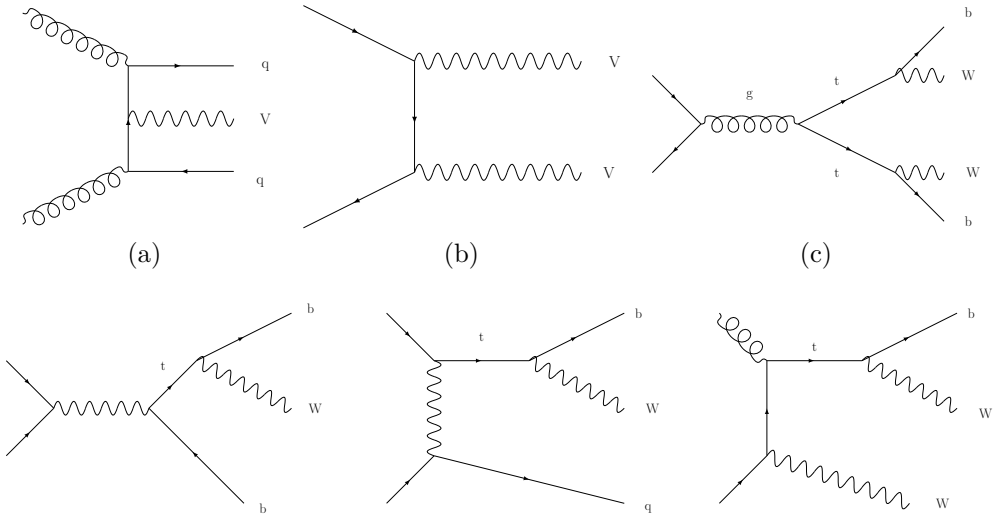
All the available MC are listed in Table 4.2 and all the Feynman diagrams of the backgrounds are shown in Figure 4.1.

Process	PDF	ME	PS+UE	tune
$q\bar{q} \rightarrow$ Vh	PDF4LHC15NLO	POWHEG +GoSaM+MiNLO	Pythia 8	AZNLO
gg $\rightarrow$ Zh	PDF4LHC15NLO	POWHEG	PYTHIA 8	AZNLO
V+jets/VV	NNPDF 3.0 NNLO	SHERPA	SHERPA	SHERPA
$t\bar{t}$/single-top	NNPDF 3.0 NLO	POWHEG	PYTHIA 8	AZNLO

Table 4.2 – Monte Carlo generators and their parameters in the $Vh(b\bar{b})$ analyses.

The case of QCD background

Due to the jet enriched final state of the search, the main backgrounds should have been processes leading to the massive production of di-jet. However, as explained



(a) (b) (c) (d)

Figure 4.1 – Feynman diagrams for the production of SM processes that contribute as background events to the $Vh(b\bar{b})$ analysis: a) weak vector boson production with two additional quarks, b) production of a pair of vector boson, c) production of a $t\bar{t}$ pair, d) production of single top quarks in the s-, t- and Wt-channel (from left to right).

in the introduction, the requirement of leptons, as well as the b-tagging, significantly reduces the number of such events. Indeed if any leptons should appear, they would be collimated with the jets and therefore would not pass the isolation criteria. Furthermore, the b-tagging requirement makes it very challenging to model this background with MC samples, since the number of events would be very low. Therefore a dedicated procedure based on data-driven estimation is used as described in Section 5.1.

4.2 Event selection

In order to increase the signal to background ratio, a set of requirements is set on the events. The analysis is based on the generation of two objects (the vector boson V and the Higgs boson) whose decays impose useful criteria for the selection.

Starting with the Higgs boson, two b-tagged jets are required in the final state. The jets entering the selection must pass some basic kinematic criteria. Their p_{T} must be greater than 20 GeV, and pass a JVT (defined in Section 3.2.2.3) as well as an $|\eta| < 2.4$ cut if their p_{T} is below 60 GeV, or simply $|\eta| < 2.5$ above. This so-called "central-jet" topology is completed by a "forward-jet" category defined by the jet in the $2.5 < |\eta| < 4.5$ region

and passing a 30 GeV p_{T} cut. The selection and classification criteria are summarised in Table 4.3. The central-jets have their b-tagging score evaluated from the MV2c10 classifier introduced in Section 3.2.2.4 with a working point selected that corresponds to roughly 70% of the signal efficiency. Only events with two b-tagged jets are kept. Thanks to the high mass of the Higgs boson and in order to reject low-p_{T} multi-jet and V+jets events, more stringent momentum cuts are set. The b-tagged labelled jets are ordered by p_{T} and the leading one is required to pass a $p_{\mathrm{T}} > 45$ GeV cut. Those two jets are called signal-jets and are used in the Higgs boson reconstruction. An extra jet can be accepted from the central- or forward-jet category. Categories are therefore defined depending on the final number of jets kept. Events with no-extra-jets will populate the 2tag2jets Signal Region (SR). In order to reject high multiplicity events coming mainly from $\mathrm{t\bar{t}}$ only events with one additional jet are kept in the 0- and 1- lepton channels to populate the 2tag3jets SR. In the 2-leptons channel, where the $\mathrm{t\bar{t}}$ contamination is quite low, all events with more than 3 jets are accumulated in a 2tag-3+jets SR. In order to be consistent in the naming across the different regions, this SR will be denoted by 2tag3jets.

Jet Category	Selection Requirements		
	$p_{\mathrm{T}} > 30\,\mathrm{GeV}$		
Forward Jets	$2.5 \leq \lvert\eta\rvert < 4.5$		
	$20\,\mathrm{GeV} \leq p_{\mathrm{T}} < 60\,\mathrm{GeV}$		$p_{\mathrm{T}} \geq 60\,\mathrm{GeV}$
	$\lvert\eta\rvert < 2.4$	or	$\lvert\eta\rvert < 2.5$
Signal Jets	JVT cut		

Table 4.3 – Requirements for the jet selection and classification.

In the vector boson final state, a certain number of electrons or muons are expected as explained in the introduction. Leptons are classified into 3 categories: VH-loose, ZH-signal and WH-signal. The distinction between the classes is summarised in Table 4.4. The 0-lepton channel is defined as events containing 0 VH-loose leptons. The 1-lepton channel must contain only 1 WH-signal lepton. The 2-leptons channel must contain 2 VH-loose same-flavour leptons, one of which at least must be a ZH-signal lepton. The two leptons are required to have opposite charges in the di-muon channel. This is not requested in the di-electron channel due to the high rate of charge misclassification (up to a few per cent [117]) due to electron-positron pair creation after a bremsstrahlung effect. In order to reject the QCD background and enhance the S/B ratio, the p_{T} of the reconstructed vector boson ($E_{\mathrm{T}}^{\mathrm{miss}}$,$E_{\mathrm{T}}^{\mathrm{miss}} + p_{\mathrm{T}}^{lepton}$, or $\sum p_{\mathrm{T}}^{lepton}$) is required to be at least 150 GeV. In the 0-lepton channel, this requirement selects events which are not on the trigger efficiency plateau (obtained at a reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ of 180 GeV). Therefore dedicated SF are derived to take into account the data/MC mismodelling in this region, with dedicated systematics. In the 2-leptons channel, the lepton triggers allow a low-p_{T} region ([75,150] GeV) to be created which improves slightly the sensitivity of the channel.

Further cuts are imposed in order to better reject the poorly modelled QCD multi-jet background or to define control regions to better model and control the main backgrounds. Those cuts are depending on the physics phase space, and are therefore defined in each of the lepton-channels.

Lepton Selection	p_{T}	η	ID	d_0^{sig} w.r.t. BL	$\lvert\Delta z_0 \sin\theta\rvert$	Isolation
			electrons			
VH-Loose	>7 GeV	$\lvert\eta\rvert < 2.47$	LH Loose	< 5	< 0.5 mm	LooseTrackOnly
ZH-Signal	>27 GeV			Same as VH-Loose		
WH-Signal	Same as VH-Loose		LH Tight	Same as VH-Loose		FixedCutHighPtCaloOnly
			muons			
VH-Loose	>7 GeV	$\lvert\eta\rvert < 2.7$	Loose quality	< 3	< 0.5 mm	LooseTrackOnly
ZH-Signal	>27 GeV	$\lvert\eta\rvert < 2.5$		Same as VH-Loose		
WH-Signal	>25 GeV	$\lvert\eta\rvert < 2.5$	Medium quality	< 3	< 0.5 mm	FixedCutHighPtTrackOnly

Table 4.4 – Requirements for the classification into the different lepton categories.

— **0-lepton channel:** in order to suppress events in a region where the trigger efficiency dependence on jet activity is poorly modelled, a cut on the scalar sum of the p_{T} of the jets ($\mathrm{S_T}$) is set to be greater than 120 GeV in the 2 jets SR and 150 GeV in the 3 jets SR. In order to reduce the multijet backgrounds a set of kinematic and geometrical cuts are applied. Those cuts involve quantities such as the signal-jets p_{T}, the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_T^{miss} (described in Section 3.2.3). The optimisation that provides the cut values is described in Section 5.1.1.

— **1-lepton channel:** in order to reject QCD multi-jet events in the electron channel a cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$ is set at 30 GeV. A $W + hf$ control region (CR) is designed to better constrain this important background. Events passing the two cuts $m_{bb} < 75$ GeV and $m_{top} > 225$ GeV enter the CR while the rest populates the SR. The m_{top} is defined in Section 4.4.

— **2-leptons channel:** since all the products of the vector boson decay can be reconstructed, a window cut is placed on the di-lepton mass: $81 < m_{ll} < 101$ GeV. It helps to reduce the non-resonant backgrounds such as QCD multi-jets, which then become negligible, and $\mathrm{t\bar{t}}$. For the latter, a dedicated CR is set by requiring opposite flavour leptons and called Top-$e\mu$ CR. It uses the same kinematic cuts as the SR.

These selections define eight signal regions (two for each number of jet bin times one or two $p_{\mathrm{T}}^{\mathrm{V}}$ regions times three lepton channels). All the cuts are summarised in Table 4.5.

4.2.1 The Cut Based Analysis approach

In order to provide a complementary result in the light of the search that was conducted in Run-1 [13], a resonant di-jet mass analysis is performed with a lower sensitivity than the MVA analysis. More than just a cross-check, this analysis also provides the theoreticians with an analysis including a clearer method for the signal selection. Because of the important background remaining after the cuts described previously, a second set of cuts based on the ones from Run-1 are applied to the events. In the simple $1 \to 2$

Selection	0-lepton	1-lepton		2-lepton
		e sub-channel	μ sub-channel	
Trigger	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single lepton	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single lepton
Leptons	0 *VH-loose* leptons	1 *WH-Signal* electron	1 *WH-Signal* muon	2 *VH-loose* leptons ≥ 1 *ZH-Signal* lepton
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 150 GeV	> 30 GeV	–	–
m_{ll}	–	–		81 GeV $< m_{ll} <$ 101 GeV
Jets	Passing selection cuts in Table 4.3			
Number of jets	Exactly 2 / Exactly 3 jets			Exactly 2 / ≥ 3 jets
b-jets	Exactly 2 b-tagged jets			
Leading b-tagged jet p_{T}	> 45 GeV			
H_{T}	> 120 (2 jets), >150 GeV (3 jets)	–		–
$\min[\Delta\phi(\overrightarrow{E_T^{miss}},\vec{\text{jets}})]$	$> 20^\circ$ (2 jets), $> 30^\circ$ (3 jets)	–		–
$\Delta\phi(\vec{E}_T^{miss},\vec{b\bar{b}})$	$> 120^\circ$	–		–
$\Delta\phi(\vec{b}_1,\vec{b}_2)$	$< 140^\circ$	–		–
$\Delta\phi(\vec{E}_T^{miss},\vec{p}_{\mathrm{T}}^{\mathrm{miss}})$	$< 90^\circ$	–		–
p_{T}^V regions	> 150 GeV			75 GeV $< p_{\mathrm{T}}^V <$ 150 GeV, > 150 GeV
Signal regions	All events	$m_{bb} \geq 75$ GeV or $m_{\mathrm{top}} \leq 225$ GeV		Same-flavour leptons Opposite-sign charges ($\mu\mu$ sub-channel)
Control regions	–	$m_{bb} < 75$ GeV and $m_{\mathrm{top}} > 225$ GeV		Different-flavour leptons Opposite-sign charges

Table 4.5 – Summary of the selections applied in the different lepton-channels in the $Vh(b\bar{b})$ analysis.

body decay, a relation can be established between the angular separation of the 2 bodies $\Delta R(b_1, b_2)$ and the initial p_{T} and mass of the object:

$$\Delta R(b_1, b_2) \propto \frac{2m}{p_T}$$

Since the Higgs is recoiling against the vector boson in the reference frame of the detector, the Higgs boson momentum $p_{\mathrm{T}}^{\mathrm{H}}$ is equal to the vector boson momentum $p_{\mathrm{T}}^{\mathrm{V}}$. Therefore at high $p_{\mathrm{T}}^{\mathrm{V}}$, the angular separation of the two b-quarks is reduced, which is not the case for the other backgrounds, where the V+jets distribution is quite flat, and the $t\bar{t}$ is peaking at high values as seen from the post-fit Figure 4.2 obtained in the MVA analysis (therefore without the CBA specific cuts). Hence a $p_{\mathrm{T}}^{\mathrm{V}}$ dependent $\Delta R(b_1, b_2)$ is designed to increase the signal-to-background ratio as presented in the Table 4.6. The $p_{\mathrm{T}}^{\mathrm{V}}$ binning is modified with respect to the MVA analysis by splitting the > 150 GeV bin into two bins [150,200] GeV and > 200 GeV to benefit from more flexibility in the $\Delta R(b_1, b_2)$ cut and to take advantage of enhanced S/B at high $p_{\mathrm{T}}^{\mathrm{V}}$.

In order to reduce even more the backgrounds in the 1- and 2-leptons channels, some lepton-channel dependent cuts are designed:

— **1-lepton channel:** the transverse mass of the reconstructed W boson, defined as the transverse component of the sum of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ 4-vectors, is selected to be lower than 120 GeV;

— **2-leptons channel:** the event based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance defined as the ratio of the $E_{\mathrm{T}}^{\mathrm{miss}}$ over the sum of the leptons p_{T} called S_{T} is selected to be lower than $< 3.5\sqrt{\mathrm{GeV}}$.

Since the m_{bb} distribution is sufficient to constrain the W+hf background in the 1-lepton channel, no specific control region is needed. On the contrary, the 2-leptons Top-$e\mu$ CR is kept to constrain the $t\bar{t}$ contribution to the shape.

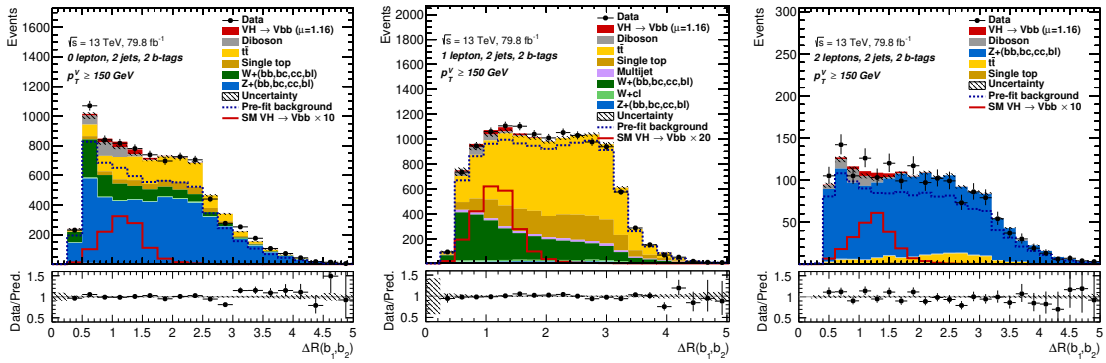


Figure 4.2 – MVA post-fit distribution of the $\Delta R(b_1, b_2)$ in the 0- (left), 1- (centre) and 2-leptons (right) channels obtained with $\mathcal{L} = 79.8\ fb^{-1}$ of data.

p_T^V region			
	[75,150] GeV (2-leptons only)	[150,200] GeV	$>$ 200 GeV
$\Delta R(b_1, b_2)$	< 3.0	< 1.8	< 1.2
Lepton channel			
	0-lepton	1-lepton	2-leptons
m_T^W	–	< 120 GeV	–
$E_T^{miss}/\sqrt{S_T}$	–	–	$< 3.5\ \sqrt{\text{GeV}}$

Table 4.6 – Summary of the selections applied for the CBA specific cuts

4.3 Event reconstruction

From the reconstruction of the objects selected in the $Vh(b\bar{b})$ analysis, a few corrections are applied to the events both in the MC and in the data. First, an Overlap Removal procedure is applied to avoid the double-counting of objects in the analysis (Section 4.3.1). As the analysis is targeting a b-jet final state, considerations on the b-tagging are crucial to improve the final result. In order to increase the statistical power of the produced MC facing a b-tagging requirement the truth-tagging method is employed (Section 4.3.2). Then the energy of the b-tagged jets is corrected to take into account physics effect such as Final State Radiation (FSR) or energy resolution degradations (Section 4.3.3).

4.3.1 Overlap Removal procedure

During the object reconstruction, some information can be shared between objects and some ambiguity can appear in their properties. An iterative procedure, called Overlap Removal (OR), is used to remove the duplicated objects from the initial containers. The algorithm is designed in the following way:

1. τ - lepton: if $\Delta R(\tau, lep) < 0.2$, unless if the lepton is a muon that is not a combined muon and has $\mathrm{p_T}^{\tau} > 50$ GeV, the τ is removed to account for τ misreconstruction.
2. electron - muon: if the two objects are sharing an ID track, the muon is removed if it is combined, otherwise the electron is removed, to account for the misconstruction of muons.
3. electron - jet: if $\Delta R(e, jet) < 0.2$ the jet is removed, else if $\Delta R(e, jet) < min(0.4, 0.04 + 10\mathrm{GeV}/\mathrm{p_T^e})$ the electron is removed, as it is likely to originate from the soft emission.
4. muon - jet: if $\Delta R(muon, jet) < 0.2$ or the muon and the jet are sharing the same ID track, and one of the two conditions are met: ($N^{tracks}_{p_T<0.5GeV} < 3$ or $p^{\mu}_T/p^{jet}_T > 0.5$ and $p^{\mu}_T/\sum p_T^{tracks\, p_T>0.5GeV} > 0.7$) the jet is removed, since it is likely to come from the muon showering in the calorimeter.
   Else if $\Delta R(\mu, jet) < min(0.4, 0.04 + 10\mathrm{GeV}/\mathrm{p_T^{\mu}})$ the muon is removed.
5. τ - jet: if $\Delta R(\tau, jet) < 0.2$ the jet is removed, as the signatures are similar while the τ reconstruction holds more information.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ uses all the objects as it is evaluated before the OR procedure. However, to avoid a double-counting, objects that have a topo-clusters already used in the computation, are rejected.

4.3.2 b-jet tagging strategies

The strategy presented in Section 3.2.2.4 for the selection of jets coming from a B-hadron, and referred to as direct-tagging, strongly rejects events containing any other quarks in the MC sample. Considering using those samples in the training of an MVA method to better classify the events, having an adequate number of events is crucial to get a good classification. Therefore a different technique based on weighting the events instead of rejecting them is applied and referred to as truth-tagging. In this method, all the events will be kept in the analysis but will get a weight that will reflect the probability that they satisfy the b-jet tagging criteria. This weight is based on b-tagging efficiency maps provided by the Flavour Tagging Combined Performance group as described in Section 3.2.2.4 that are divided in b-,c- and light-flavour regions. An individual jet weight ϵ_j is extracted from its reconstructed p_{T} and η and by choosing the correct map thanks to its truth content. If more jets are required than the number of b-tagged jets, the weight assigned to those extra-jets is the complementary to individual weight $1 - \epsilon_j$. Then a combination of the weights is calculated for all the jets selected in the event to get the final truth-tagging event weight w_{tot}:

— if 2 jets are selected and 2 are required to be b-tagged : $w_{tot} = \epsilon_{j1} \cdot \epsilon_{j2}$

— if 3 jets are selected and 2 are required to be b-tagged :

$$\begin{aligned} w_{tot} = \; & (\epsilon_{j1} \cdot \epsilon_{j2}(1-\epsilon_{j3})) + \\ & (\epsilon_{j1} \cdot \epsilon_{j3}(1-\epsilon_{j2})) + \\ & (\epsilon_{j2} \cdot \epsilon_{j3}(1-\epsilon_{j1})) \end{aligned}$$

An example of the impact on the final yield in the W+jets samples is given in Table 4.7. The statistical errors are greatly reduced in all the events not containing a b-quark, however, some important discrepancies can be observed in event yields for the other categories. By considering those two effects, it was decided to apply truth tagging only on the V+jets samples depending on the true flavour of the jets: V+cc, V+cl, V+l, and the di-boson WW sample.

4.3.3 b-jet corrections for mass resolution

Among all the variables reconstructed in the analysis, the invariant mass of the two b-jet is of great importance. It either acts as the final discriminant (CBA) or is used as a discrimination tool to separate non-resonant backgrounds such as $t\bar{t}$, single-top or V+jets in the MVA analysis. The di-boson resonant background can also be distinguished from the Higgs boson signal thanks to the difference in invariant mass. Some procedures are therefore used to improve the b-jet invariant mass resolution, acting on jets or events that passed the basic selection criteria defined in Section 4.2. As the analyses are using the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable it is important that those corrections are not propagated in the calculation of this variable.

Muon-in-jet

The nature of the b-hadron decays involves two W bosons which can decay leptonically. While the energy of electrons is measured through deposits in the calorimeters, the muons' energy deposit in calorimeters is negligible because they behave as minimum ionising particles. 12% of the b-jets passing the selection are shown to contain at least one reconstructed muon. Those muons are then ordered by the distance with respect to the jet axis, and only the closest one is kept. The muon 4-vector is then simply added to the jet 4-vector after suppressing its deposited energy in the calorimeter to avoid double-counting.

The effect of those two procedures is to shift the di-jet mass peak toward the predicted $m_h = 125$ GeV value as well as to reduce the width of the distribution, as shown in Figure 4.3.

Process	Direct Tag Yield	Truth Tag Yield
Wbb	1586.52 ± 15.82	2097.21 ± 13.19
Wbc	247.86 ± 7.28	306.35 ± 2.06
Wbl	128.02 ± 5.86	140.63 ± 0.41
Wcc	117.89 ± 13.24	173.45 ± 1.55
Wcl	146.70 ± 16.94	125.78 ± 0.30
Wl	9.65 ± 11.33	29.62 ± 0.04

Table 4.7 – W+jets direct tag and truth tag yields in the 2tag2jets region in the 1-lepton channel

p_{T}-reco

After correcting for possible muons, a p_{T} correction is applied to the jets. This aims to correct for various effects such as the low-p_{T} out-of-cone energy deposit, resolution effects, or neutrinos' energy. This correction is derived from 2-leptons signal $\mathrm{Zh} \to \ell^+\ell^-\mathrm{b\bar{b}}$ events as a p_{T} dependent scale factor comparing the distributions of the truth jets and the reconstructed ones using the same b-tagging algorithm. The SFs are determined separately for jets that contain an electron or a muon inside the jets (Semi-leptonic events) and in the fully-hadronic case. These corrections are applied to the 0- and 1-lepton channels.

Kinematic fitter

In the case of the 2-leptons analysis, the full event can be reconstructed thanks to the lack of neutrinos in the final state. Since the lepton resolution is of the order of 1% while the jet resolution is of the order of 10%, the balancing of the Z system to the Higgs can be used to constrain the b-jet resolution. The procedure varies the b-jets and leptons momentum in order to minimise a likelihood that takes into account the following terms:

- a Breit-Wigner term for the mass of the di-lepton to be consistent with the mass of the Z-boson.
- a Gaussian term for the component of the ZH system to be centred at 0 with a width of 9 GeV (determined from ZH MC truth information).
- a Gaussian term for the b-jet and lepton p_{T} to be consistent with the measured value within its resolution.
- an additional prior to take into account the true jet p_{T} in the same way as done in the p_{T}-reco corrections.

Since the lepton resolution is better than the jet ones, the main effect is observed in the di-jet mass. The addition of a third jet smears the balancing effect, and thus corrections

are weaker in this case. In the 4+ jet category, the effect is so small that the p_{T}-reco is used.

The overall effect of the corrections is presented in Figure 4.3, and a final 20 to 40 % improvement can be observed on the mass width.

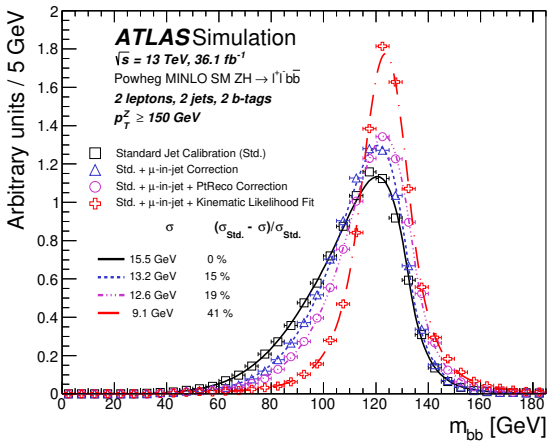


Figure 4.3 – Comparison of m_{bb} for the Nominal, muon-in-jet, p_{T}-reco and Kinematic Fit for the 2-leptons channel high p_{T} (V) channel.

4.4 Event categorisation: the Boosted Decision Trees

As presented in Section 4.2.1, in the search for a $Vh(b\bar{b})$ signal, the m_{bb} variable is of particular interest to separate signal from backgrounds. However, to increase this sensitivity additional cuts are needed on variables such as $p_{\mathrm{T}}^{\mathrm{V}}$ or $\Delta\mathrm{R}(\mathrm{b}_1, \mathrm{b}_2)$. The design of the cuts is optimised by doing 1- up to 2-dimensional cuts, very often not properly using the correlation scheme between the variables. In the light of the recent development of smart algorithms used in classification and predictions areas of industry as well as academy, Multivariate Algorithms are constructed to make full usage of all the information available in order to extract weak correlations between variables. The special Boosted Decision Trees algorithm that is used in the $Vh(b\bar{b})$ analysis is described in Section 4.4.1, while its application for the VH result is presented in Section 4.4.2.

4.4.1 Multivariate Algorithms

Behind the idea of Multivariate Algorithms is that the exploitation of a reduced set of variables is always sub-optimal to the case where the full phase space is taken into account in the description. Calling a physics outcome from this phase space y from an n-dimensional phase space $(x_1,\ldots,x_n)$, the main job is to relate the estimated value $\hat{y}$ to the observed values $(x_1,\ldots,x_n)$ from the estimator $\hat{y} = \hat{f}(x_1,\ldots,x_n)$. The point of Multivariate Algorithms is to find this estimator. Depending on the type of solution which is expected from the algorithm, two categories can be defined. If the question aims at predicting a quantitative result, as "given this event what is the most likely value for y", then the algorithm is called regression. Alternatively, if the question aims at categorising the event into 2 or more classes ($y = \{\text{ClassA, ClassB} \ldots\}$), then the algorithm is called classification. The second one is of particular interest for the $Vh(b\bar{b})$ analysis as it allows a discriminator variable to be created that separates the backgrounds from the signal.

The Multivariate Algorithms are also based on a two-phase approach. In order to construct the predictive model it needs to be trained by events which contain partially or the full phase space variables $(x_1,\ldots,x_n)$ and the truth output y. The goal is to automatise its construction in order to learn from this n-dimensional space the features requested by the user, depending on whether it is a regression or classification algorithm. Once trained the algorithm can be tested by how well it reproduces the distribution of outcomes y. This application phase is used in either the efficiency tests on the training or the final evaluation on the data. One of the main issues is not to overtrain the system, where it learns individual features from the set of events given, that is finite by construction. In most cases, the initial sample is divided into orthogonal subsets which are used to make cross-validations for each training by evaluating the outcomes obtained between the training and testing sets. The higher the number of folds, the higher number of events in the training sample is, and therefore the more resilient the training is to overtraining. However, increasing the number of folds comes at the cost of complexity in the analysis and speed of the training algorithm. In the case of the $Vh(b\bar{b})$ analysis, the choice was made to have a 2-fold training.

From all the libraries providing MVA tools, the choice of the $Vh(b\bar{b})$ analysis has been to use the TMVA package [164] thanks to its good interface with the Root [165] framework used.

Decision Tree

Amongst the classification algorithms, decision trees are the simplest predictive models and have the benefit to be similar to the cut-based analyses. The Decision Tree (DT) is an algorithm based on nodes regrouping events from a specific region of the phase-space. They are created recursively and related to each other by a series of cuts on one of the variables available. A typical representation of a DT and its nodes is given in Figure 4.4. The choice of the value of the cut and the variable is determined thanks to a

predefined criterion that provides the largest separation between Signal and Background. In the case of the $Vh(b\bar{b})$ analysis, the Gini index is evaluated:

$$Gini = \left(\sum_{i}^{N_{events}} W_i \right) p(1-p)$$

where $p = \frac{n_s}{n_s+n_b}$ is the signal purity in the considered node and $\sum_i^{N_{events}} W_i$ is the weighted sum of events in the node. The sum of weights is normalised for the Signal and Background separately before the training not to give the background events a bigger influence. This index is evaluated for both the parent and children nodes and by varying the cut on a binned version of the variables (typically using 100 bins) considered:

$$\widetilde{G} = Gini_{parent} - Gini_{child,1} - Gini_{child,2}$$

The optimal cut is chosen to maximise this $\widetilde{G}$ value, therefore the same variable can be reused several times in the training as shown in Figure 4.4.
The growth of the node is ended when specific conditions are met:

- Minimum Node Size: if the number of events falling into the children node is below a threshold (could be 5% of training events or 200 events),
- Max Depth: if the distance from the root node is equal to a threshold (typically 3 or 4), as shown in Figure 4.4

The node is set to be a leaf and is no longer sub-divided. It is therefore classified as a Signal or Background node and is given a score that is either the signal purity defined above or, in the case of the $Vh(b\bar{b})$ analysis, a +1 score if the leaf contains a majority of Signal and -1 for Background. Very small differences are reported by the TMVA collaboration on the use of either option [164].

To avoid overtraining, the trees can be pruned, removing leaves and branches with a small number of events per node and thus sensitive to statistical fluctuations. This reduces the performance of the classification but improves the metrics used to check for overtraining. However, since no clear sign of overtraining is observed in the $Vh(b\bar{b})$ analysis, no pruning methods are applied here.

Boosted Decision Trees

As the Decision Trees are sensitive to the input number of events, their structure might fluctuate leading to overtraining. Furthermore, the effects of one DT are not that different from the cut-based approach, if not for the automatisation of the variable and cut value choice. A so-called forest of trees can be constructed to overcome those problems. The idea is to construct weaker classifiers that will learn from the mistakes of the previous

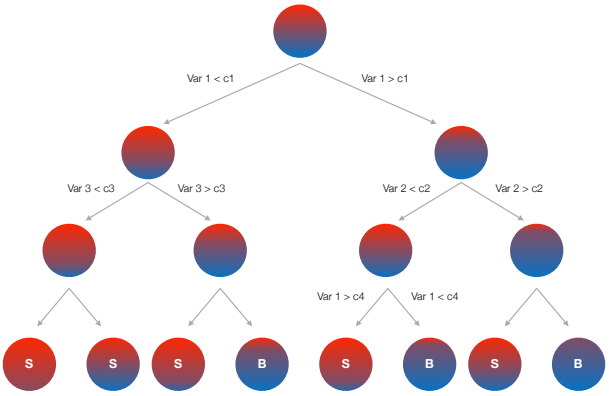


Figure 4.4 – Schematic figure of a decision tree. Each node is coloured proportionally to its content in Signal (Red) and Background (Blue). Each node is divided into sub-nodes by the evaluation of a simple cut as represented by $var_i > / < cut_i$. The final node output S or B is decided on the majority of events composing the node. Here the maximal depth of the decision tree is fixed to 4.

trees and to combine them in a smart way to build a stronger classifier. This method is called boosting and can improve drastically the performance of the analysis, making it at the same time more complex to understand. A large number of trees can also lead to overtraining. The misclassified events by a DT (signal events falling into a Background leaf and vice versa) are reweighted before entering a new DT training so that this new training tries to improve the classification of those events. This newly produced tree will enter in the final combination of trees and the procedure is repeated until some criteria, such as the total number of trees, is met.

The boosting procedure used in the $Vh(b\bar{b})$ analysis is called AdaBoost [166] and is one of the most popular algorithms available. It uses the following definition for the misclassification:

$$err_m = \frac{\sum_i W_i \cdot (1 - \delta(y_i - \hat{y}_i^m))}{\sum_i W_i}$$

where δ is the Dirac distribution, y_i the true classification of the event i and $\hat{y}_i$ the result of the classification of the tree m. The reweighting factor to be applied to the misclassified events is then defined as:

$$b_m = e^{\alpha_m} \text{ with } \alpha_m = \beta \log \frac{1 - err_m}{err_m}$$

where the β factor is the learning rate ranging between 0 and 1 and chosen by the user. In the $Vh(b\bar{b})$ analysis, this factor is chosen to be 0.15. The higher the learning rate, the more likely it is to over-shoot, while the lower it is, the more trees it takes to converge on a proper solution. The final weighted sum of trees is defined as:

$$T = \sum_{m=1}^{N_{trees}} \alpha_m T_m$$

where T_m is the n$^{\text{th}}$ trained DT. Since the α_m boosting weights are supposed to decrease with the number of trees, as the misclassification is smaller and smaller, all the added trees can be perceived as corrections to the initial DT targeting specific zones of the phase space. Since the output is a weighted sum of the response of each tree, the final BDT output is a pseudo-continuous distribution of scores between [-1,1].

4.4.2 Application to the Vh resonance

The introduced BDTs are used in the $Vh(b\bar{b})$ analysis taking advantage of the initial event selection described previously. They use the MC samples introduced in Section 4.1, combining all the Vh samples in a unique Signal sample while the rest is combined in the Background one. Each background process is normalised to its relative contribution in the final analysis phase space in order not to bias the discriminator for differences in the number of events.

Only a few variables are used in the training, as any addition of poorly modelled ones could disturb the good construction of the BDTs. In principle the addition of less discriminating variables in the BDTs would not decrease the performance, however, it increases the computational time of the training and the evaluation, which is a criterion in the choice of the number of variables used. The procedure applied to the selection is a recursion. It starts by evaluating the performance of the BDT with the best discriminating variable: the di-jet mass m_{bb}. Then the iteration proceeds by adding one by one kinematic and geometrical variables that could enhance the good separation while being correctly modelled. The process is stopped when variables do not bring any improvement in the statistical binned sensitivity in the process:

$$S = \sqrt{\sum_i (2 \cdot ((s_i + b_i) \cdot \log 1 + s_i/b_i - s_i))}$$

The list of variables is shown in Table 4.8 and is mainly inherited from the Run-1 analysis which is shown to be still optimal. Furthermore in order not to be affected by the tail of distributions which would waste BDT degrees of freedom for a small number of events a hard cut on the 99% of the distribution is set. All the overflow events are thus summed into one bin. The good data/MC modelling of the variables used in the training is shown in the 2tag2jets for events with $p_{\text{T}}^{\text{V}} > 150$ GeV the 0-lepton (Figure 4.5), 1-lepton (Figure 4.6), 2-leptons (Figure 4.7) channel.

The events are then separated into the 8 SRs defined previously, where one BDT is trained using a 2-fold scheme per region. The two MC data periods "a" and "d" are summed together to increase the statistics. At the evaluation step, both BDT scores will be stored in the same histograms since they represent the same physics sample. The options used for the BDT training are presented in Table 4.9. They are the result of some hard choices (like for the boosting procedure or separation type), or due to some dedicated studies based on optimisations. For instance, the number of trees is chosen

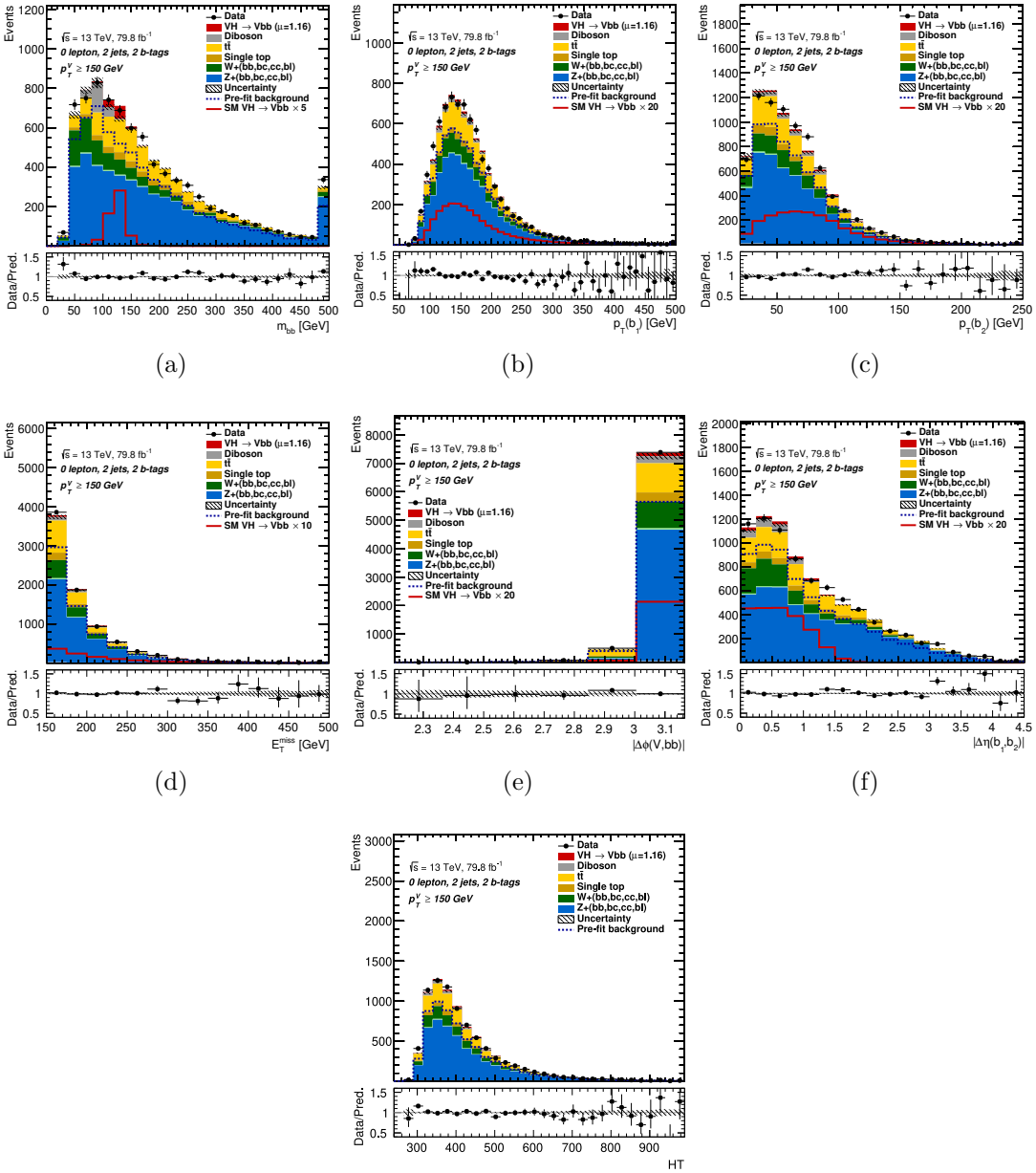


Figure 4.5 – Post-fit distributions of the variables used in the BDT training in the 0-lepton channel using only the 2tag2jets events. A data-MC comparison is shown in the bottom pad of each plots, and no important mismodelling is observed. Plotted variables are a) the dijet invariant mass m_{bb}, b) the leading jet p_{T} c) the second leading jet p_{T} d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ that is also the vector boson $p_{\mathrm{T}}^{\mathrm{V}}$ e) the ϕ angle between the vector boson and the reconstructed Higgs boson f) the difference in pseudo-rapidity between the two b-tagged jets g) the scalar sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T} of all the signal jets called M_{eff} or H_T.

such that adding more trees doesn't bring any sensitivity improvement. An overtraining is performed by looking at the test and training sample shapes as presented in Figure 4.8. To assess the discriminating power of the technique the Receiver Operating Characteris-

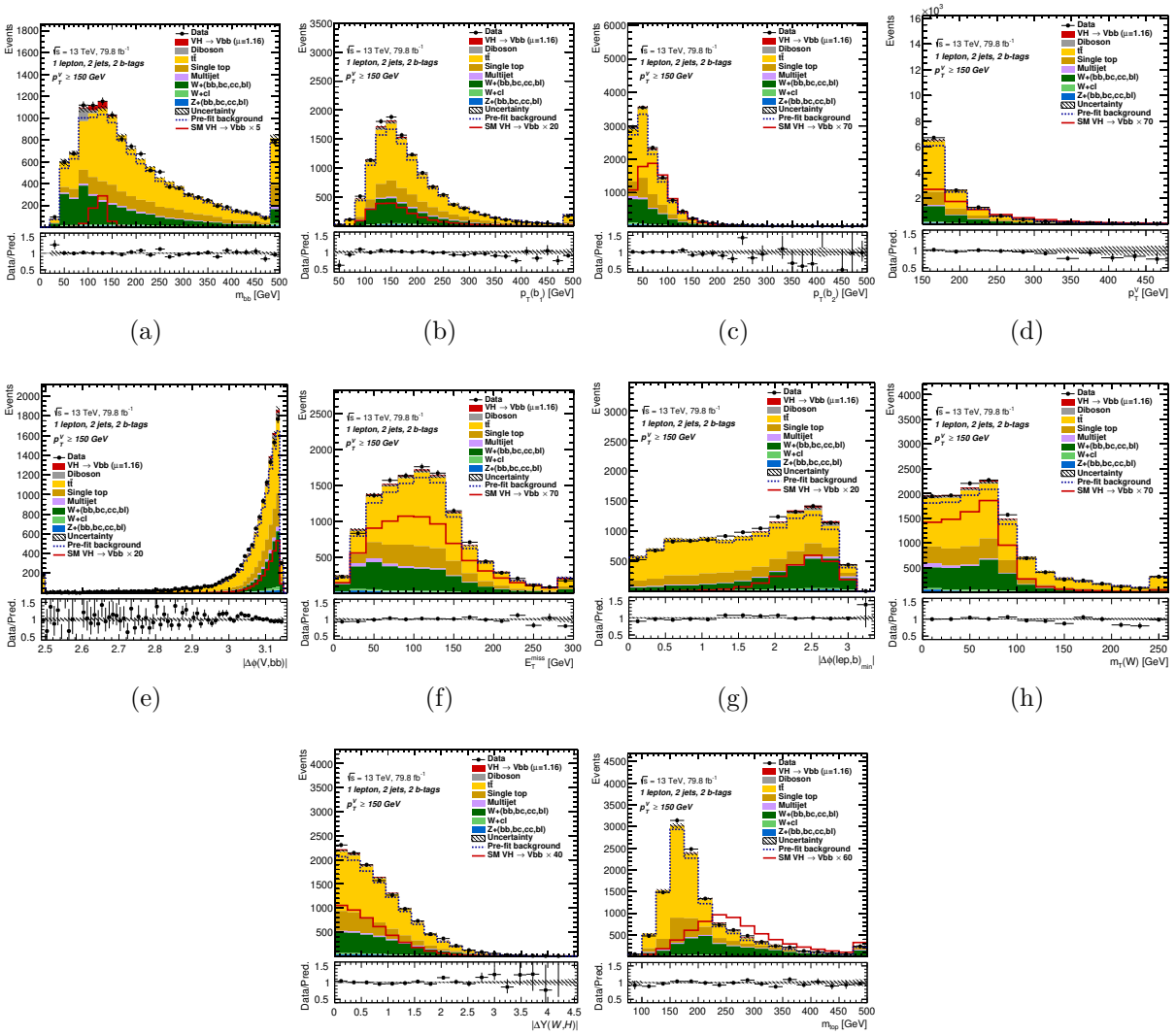


Figure 4.6 – Post-fit distributions of the variables used in the BDT training in the 1-lepton channel using only the 2tag2jets events. Plotted variables are a) the dijet invariant mass m_{bb}, b) the leading jet p_{T} c) the second leading jet p_{T} d) the reconstructed vector boson $p_{\mathrm{T}}^{\mathrm{V}}$ e) the ϕ angle between the vector boson and the reconstructed Higgs boson f) the $E_{\mathrm{T}}^{\mathrm{miss}}$ g) the minimal ϕ angle between the lepton and the b-tagged jets h) the transverse mass of the reconstructed W boson i) the difference in rapidity between the reconstructed Higgs and vector bosons j) the reconstructed top mass

tic (ROC) curve which shows the background rejection power with respect to the signal efficiency when cutting on the BDT score is computed as shown in Figure 4.9. A good rejection would maximise the area, while a random classifier would have a diagonal response.

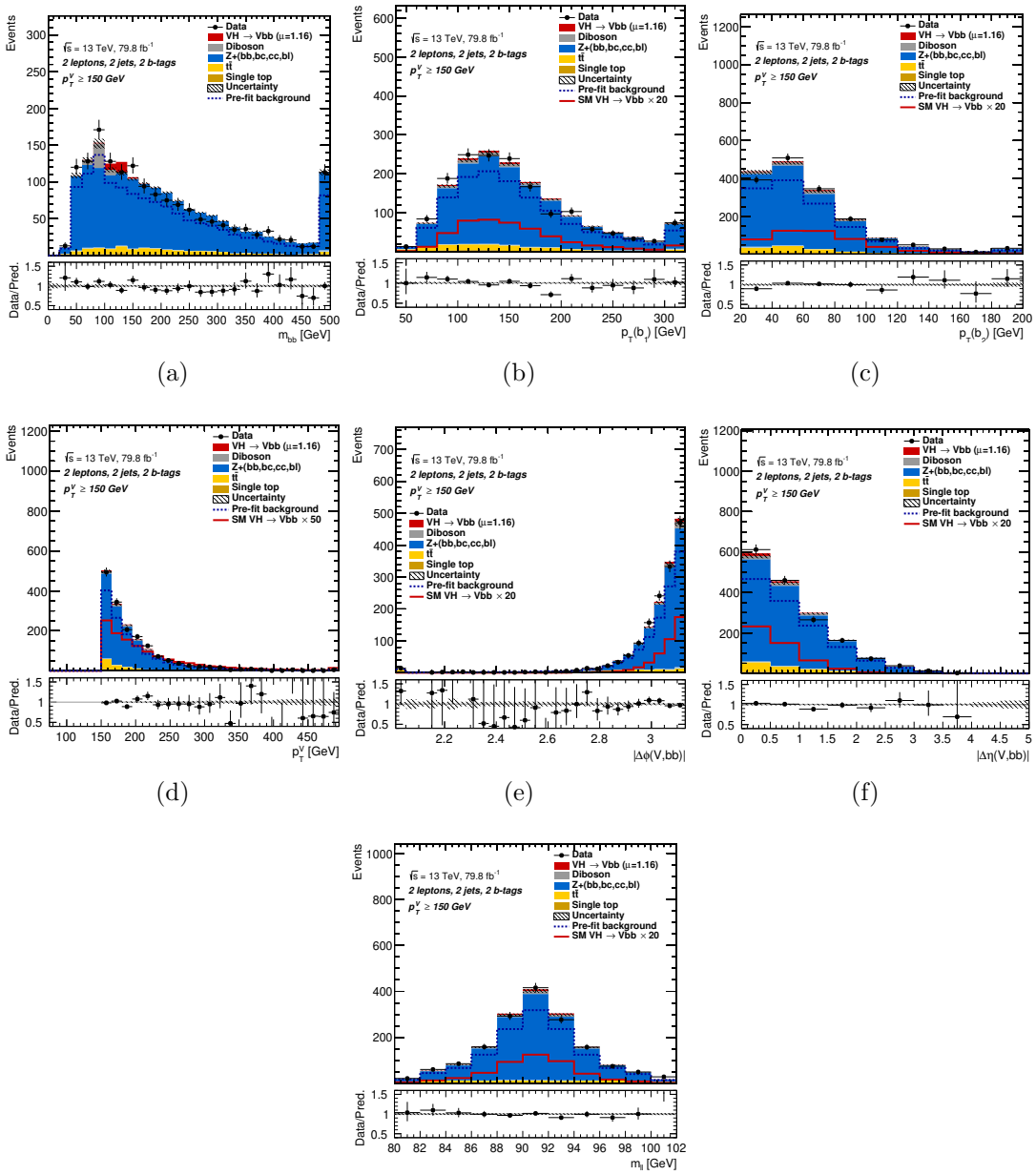


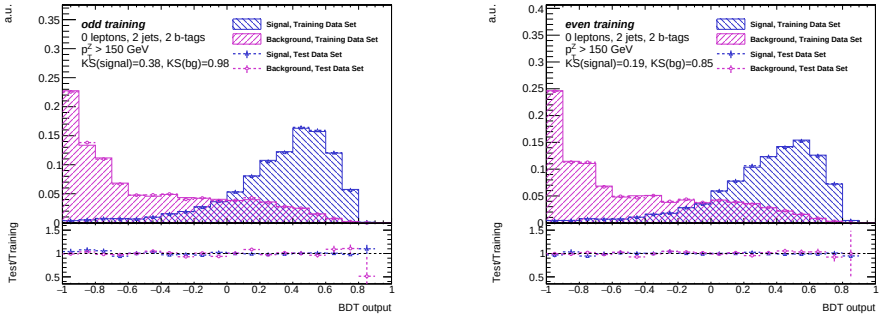
Figure 4.7 – Post-fit distributions of the variables used in the BDT training in the 2-leptons channel using only the 2tag2jets events in the high-$p_{\mathrm{T}}^{\mathrm{V}}$ bin. Plotted variables are a) the dijet invariant mass m_{bb}, b) the leading jet p_{T} c) the second leading jet p_{T} d) the reconstructed vector boson $p_{\mathrm{T}}^{\mathrm{V}}$ e) the ϕ angle between the vector boson and the reconstructed Higgs boson f) the η angle between the vector boson and the reconstructed Higgs boson g) the invariant mass of the two leptons m_{ll}

Variable	Name	0-lepton	1-lepton	2-lepton
m_{jj}	mBB	✓	✓	✓
$\Delta R(jet_1, jet_2)$	dRBB	✓	✓	✓
p_T^{jet1}	pTB1	✓	✓	✓
p_T^{jet2}	pTB2	✓	✓	✓
p_T^V	pTV	✓	✓	✓
$\Delta\phi(V, H)$	dPhiVBB	✓	✓	✓
$\lvert\Delta\eta(jet_1, jet_2)\rvert$	dEtaBB	✓		
$M_{eff}(M_{eff3})$	HT	✓		
$E_{\text{T}}^{\text{miss}}$	MET	$\equiv p_T^V$	✓	
$\min(\Delta\phi(\ell, jet))$	dPhiLBmin		✓	
m_T^W	mTW		✓	
$\Delta Y(W, H)$	dYWH		✓	
m_{top}	mTop		✓	
$E_{\text{T}}^{\text{miss}}$ significance	METSig			✓
$\Delta\eta(V, H)$	dEtaVBB			✓
$m_{\ell\ell}$	mLL			✓
	Only in 3 Jet Events			
$p_T^{\text{jet}_3}$	pTJ3	✓	✓	✓
m_{jjj}	mBBJ	✓	✓	✓

Table 4.8 – Variables used to train the Vh BDT.

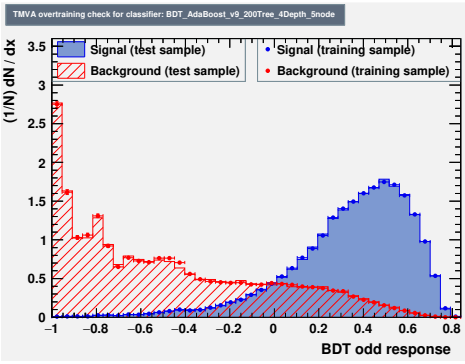
TMVA Setting	Value	Definition
BoostType	AdaBoost	Boost procedure
AdaBoostBeta	0.15	Learning rate
SeparationType	GiniIndex	Node separation gain
PruneMethod	NoPruning	Pruning method
NTrees	200	Number of trees
MaxDepth	4	Maximum tree depth
nCuts	100	Number of equally spaced cuts tested per variable per node
nEventsMin	5%	Minimum number of events in a node

Table 4.9 – Configuration parameters used for the BDT training.

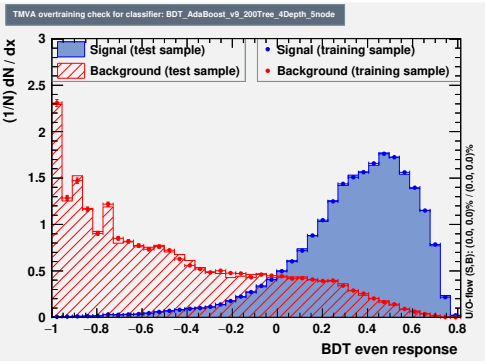


(a)

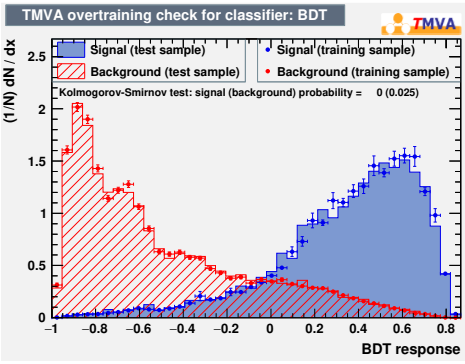
(b)



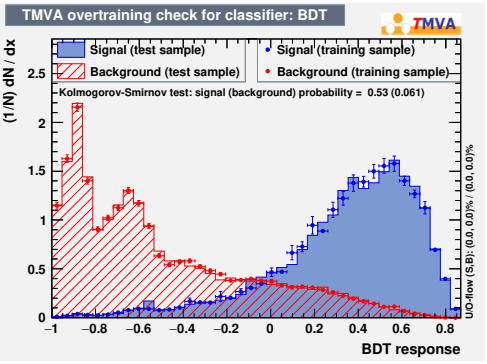
(c)



(d)

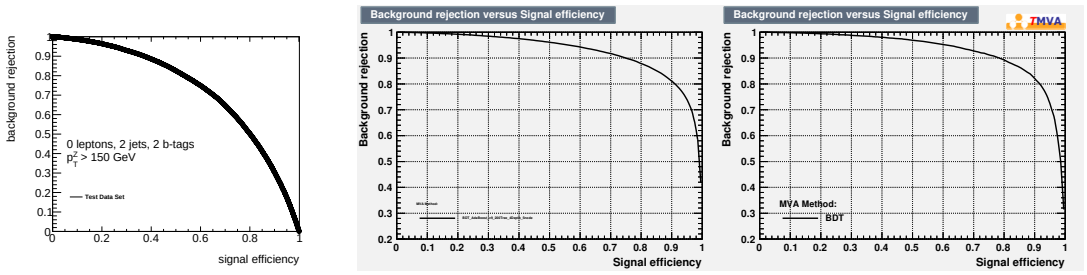


(e)



(f)

Figure 4.8 – BDT distributions of the signal (blue) and the sum of all the backgrounds (red) in the 2tag2jets regions for the 0- (top), 1- (middle), and 2-leptons (bottom) channels. In each channel, the training and testing samples for the odd (right) and even (left) folds are plotted. Since the distributions tend to agree, no sign of overtraining is observed.



(a) (b) (c)

Figure 4.9 – ROC curves of the BDT classifier in the a) 0-lepton channel b) 1-lepton channel c) 2-leptons channel. Similar good background rejections can be observed

5 Background estimation and modelling

The determination of the behaviour of the background is of particular importance in the observation of a Higgs signal, from the fit to the MVA discriminant or the distribution of the low-level variable m_{bb}. Section 4.1 has already introduced the main backgrounds as well as their determination based on MC samples, except for the QCD multi-jet one that is based on data-driven techniques which will be described in Section 5.1. The estimation extracted from MC simulations raises as many questions as a data-driven method, as undertaken for the Higgs discovery in the $\gamma\gamma$ channel. The choice of a MC generator implies several assumptions on the models governing the simulations. In order to get as close as possible the real shapes for the real processes, variations are considered in both the normalisations and the shapes. These processes are discussed in Section 5.2.

5.1 The rejection of multi-jet events

With the requirements set on Section 4.2, the residual contributions from the multi-jet backgrounds are very hard to model by MC generators, due to the b-tagging requirement as well as because of the significant cut on $p_{\mathrm{T}}^{\mathrm{V}}$. The remaining events passing the lepton identification cuts are due to photons from jet decays reconstructed as electrons, or semi-leptonic decays of heavy-flavour jets in the 1-lepton channel, the 2-leptons channel being safer due to the reduced probability of having such processes twice. In the latter

channel, a dedicated same-sign lepton analysis is conducted to assess the amount of QCD events. Being a negligible contribution (respectively 0.3% and 1.4% in the electron and muon channel), those events are not considered in the analysis. The poor reconstruction of the objects can also lead to an important shift in the measured parameters, thus creating fake $E_{\mathrm{T}}^{\mathrm{miss}}$, which would leak multi-jets events in the 0- and 1-lepton channel. Such events are discarded thanks to specific cuts leaving a negligible contribution in the 0-lepton channel as described in Section 5.1.1, while a data-driven estimate based on template methods is used in the 1-lepton channel as explained in Section 5.1.2.

5.1.1 The 0-lepton channel

Badly reconstructed jets in QCD events are a source of spurious $E_{\mathrm{T}}^{\mathrm{miss}}$, which tends to be aligned in the jet direction. In order to reduce this source of background, selections inherited from the Run-1 analysis have been re-optimised using the Run-2 data and MC samples. The following variables are used for this selection: $|\Delta\Phi(\overrightarrow{E_T^{miss}},\overrightarrow{p_T^{miss}})|$, $|\Delta\Phi(b_1,b_2)|$, $|\Delta\Phi(\overrightarrow{E_T^{miss}},\vec{\mathrm{h}})|$, $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$. Here Φ is the azimuthal angle and b_1 and b_2 are the two selected b-jets forming the Higgs candidate h. $\overrightarrow{p_T^{miss}}$ is defined as the negative sum of the transverse momenta of the tracks associated to the primary vertex of the event and passing a set of quality cuts as described in Section 3.2.3. In the $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$ calculation, only the three leading signal jets are considered, but when no third signal jet passes the selection requirements, the leading p_{T} forward jet is considered instead.

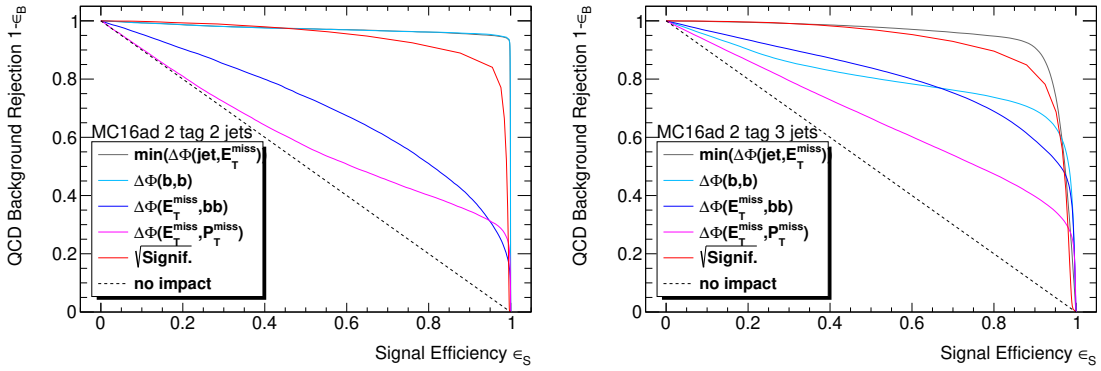
From the MC QCD di-jet samples, it is not possible to obtain a reliable estimate of the multi-jet contamination in the signal categories due to limited statistics following the full application of the 0-lepton event selection. Therefore, a data-driven estimation is used instead.

The nominal 0-lepton event selection uses four anti-QCD cuts in order to reduce the multi-jet contamination within the 0-lepton signal regions:

- $|\Delta\Phi(\overrightarrow{E_T^{miss}},\overrightarrow{p_T^{miss}})| < 90^\circ$, to remove contributions from $E_{\mathrm{T}}^{\mathrm{miss}}$ mismeasurements since the two quantities should be aligned with each others.
- $|\Delta\Phi(b_1,b_2)| < 140^\circ$, due to the large Higgs momentum, the two signal jets should be close in Φ contrary to the QCD events which should be spread over different values as shown in Figure 5.8.
- $|\Delta\Phi(\overrightarrow{E_T^{miss}},\vec{\mathrm{h}})| > 120^\circ$, since the $E_{\mathrm{T}}^{\mathrm{miss}}$ arising from the Z boson in the signal should be back-to-back with the Higgs boson.

— $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|] > 20^{\circ}(30^{\circ})$) for the 2(3)-jet category, since the $E_{\mathrm{T}}^{\mathrm{miss}}$ should not be aligned with one the jets, one of the signs that the $E_{\mathrm{T}}^{\mathrm{miss}}$ arises from mismeasurements.

An overview of the performance of the previous cuts is shown in Figure 5.1, by representing the QCD rejection obtained from an EW background-subtracted data distribution in each of the variables with respect to the MC signal efficiency. Despite the role of correlations between the variables, the most powerful discrimination is obtained with the $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$ variable. Therefore, in order to estimate the remaining multi-jet contribution in the signal region, the anti-QCD cuts are loosened by removing the $\min[|\Delta\Phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|] > 20^{\circ}(30^{\circ})$) cut for the 2(3)-jet categories.



(a) (b)

Figure 5.1 – QCD background rejection with respect to the signal efficiency for the various variables used in the anti-QCD rejection. The plot is presented for the 2tag2jets case (a) and the 2tag3jets case (b). The background shape is obtained from the EW background-subtracted data. MC16a and MC16d MC samples as well as 2015–2017 data have been used to produce this plot.

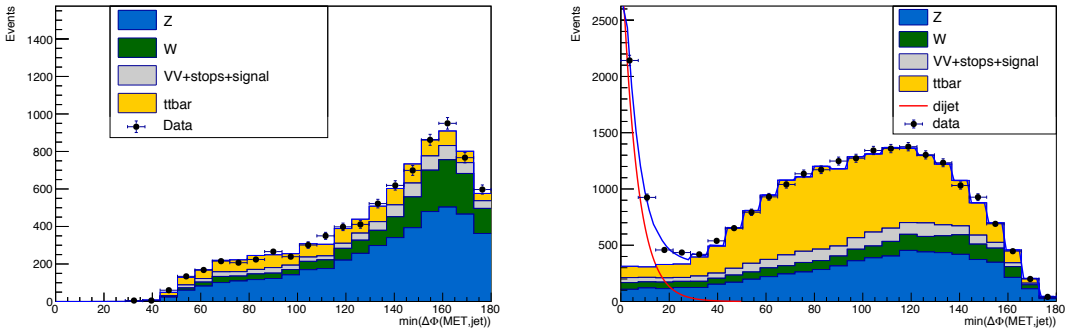
The $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$ distribution can be used to evaluate the remaining multi-jet contribution. To account for normalisation differences between the EW MC background and data in this specific phase space region, scale factors for the Z+jets, W+jets and $t\bar{t}$ backgrounds are derived separately for the 2-jets and 3-jets categories by performing a fit to data in the $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|] > 40^{\circ}$ regime. In this fit, the Z+jets, W+jets and $t\bar{t}$ normalisations are allowed to float, whilst the diboson and single-top normalisations are fixed to their MC prediction. The corresponding scale factors for the Z+jets, W+jets and $t\bar{t}$ backgrounds are summarised in Table 5.1. In order for the fit to be stable the scale factor can be fix to 1 in order not to bias the rest of the distribution.

For the 2-jets category, no multi-jet is observed (see Figure 5.2a) due to the effect of the remaining event selection applied in the 0-lepton channel. For safety, a cut at 20° is set for the 2-jets category, chosen because it is shown not to impact the yield in the simulation.

background	scale factor (2jets)	scale factor (3jets)
Z+jets	1.28 ± 0.10	1.14 ± 0.06
W+jets	1.73 ± 0.63	1. (fixed)
$t\bar{t}$	1. (fixed)	1.13 ± 0.04

Table 5.1 – Scale factors for the backgrounds from the fit for the multi-jet in $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$ distributions from MC16ad MC and 2015–2017 data. Z+jets and W+jets SF are correlated and could compensate each others. SF can be fixed for fit stability reason.

From Figure 5.2b the multi-jet contribution in the 3-jets category is found to populate the low $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$ region. A decreasing exponential distribution is used to model the shape of the multi-jet background in the low region of $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$. The yield of multi-jet events is extracted by fitting data in the region of $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|] < 50^{\circ}$ with such a decreasing exponential distribution and scaled EW background templates.



(a) (b)

Figure 5.2 – Post-fit $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$ distributions in the 2tag2jets (a) and 2tag3jets (b) category when removing the selection on this variable. The black points are representing the data used to scale the MC distributions. The multi-jet in the 2tag3jets is modelled using an exponential shape $A.e^{-c.x}$, the values of the parameters can be found in Table 5.2.

From this fit, the multi-jet yield in the signal region is estimated as a function of the cut applied to $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|]$, as presented in Table 5.3. By requiring less than 10% multi-jet contamination relative to the expected VH signal yield, a cut of $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|] < 30^{\circ}$ is retained for the 3-jets category.

Furthermore, the m_{bb} shape of the multi-jet background is studied by selecting the events within the $\min[|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|] < 20^{\circ}$ region, and by subtracting the expected electroweak and top backgrounds from the data. In this way, a m_{bb} multi-jet template is built. The shape of the m_{bb} spectrum for the combined Z+jets, W+jets, and $t\bar{t}$ background is then compared to this multi-jet template for the 2015+2016 and 2017 data collection periods, see Figure 5.3. Due to the similar shapes found between the multi-

MC period	A	c ($^\circ$)
MC16a	1614 ± 91	6.26 ± 0.33
MC16d	1652 ± 93	6.26 ± 0.34
MC16ad	3264 ± 130	6.27 ± 0.24

Table 5.2 – Parameters of the exponential fit of the multi-jet distribution in the 2tag3jets category.

Cut	VH	Multi-jet	Multi-jet/VH	Multi-jet/VH 80 GeV< m_{bb} < 160 GeV
15 deg	115.9	259.2 ± 25.6	2.23 ± 0.30	1.02 ± 0.15
20 deg	113.2	116.7 ± 15.0	1.03 ± 0.16	0.47 ± 0.08
25 deg	110.6	52.5 ± 8.3	0.48 ± 0.09	0.22 ± 0.04
30 deg	108.1	23.7 ± 4.5	0.22 ± 0.05	0.10 ± 0.02
35 deg	105.7	10.7 ± 2.3	0.10 ± 0.02	0.05 ± 0.01
40 deg	103.4	4.8 ± 1.2	0.05 ± 0.01	0.02 ± 0.01

Table 5.3 – Comparison of the multi-jet and VH yields for the 0 lepton 3jet category determined from the fit to the min[|$\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets)|] distribution. The second and third columns show the yields for the given selection, the fourth column gives the ratio, and the last column shows the ratio after scaling each yield by the fraction determined from the m_{bb} distributions in the range 80 to 160 GeV. A statistical error is propagated from the fit.

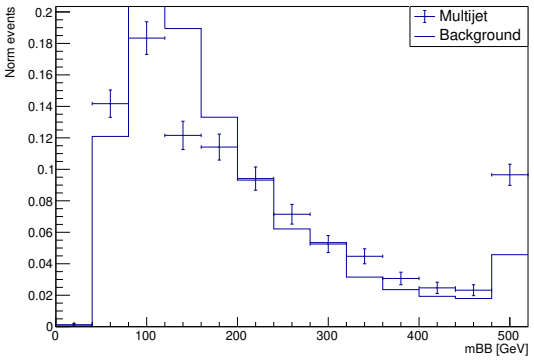
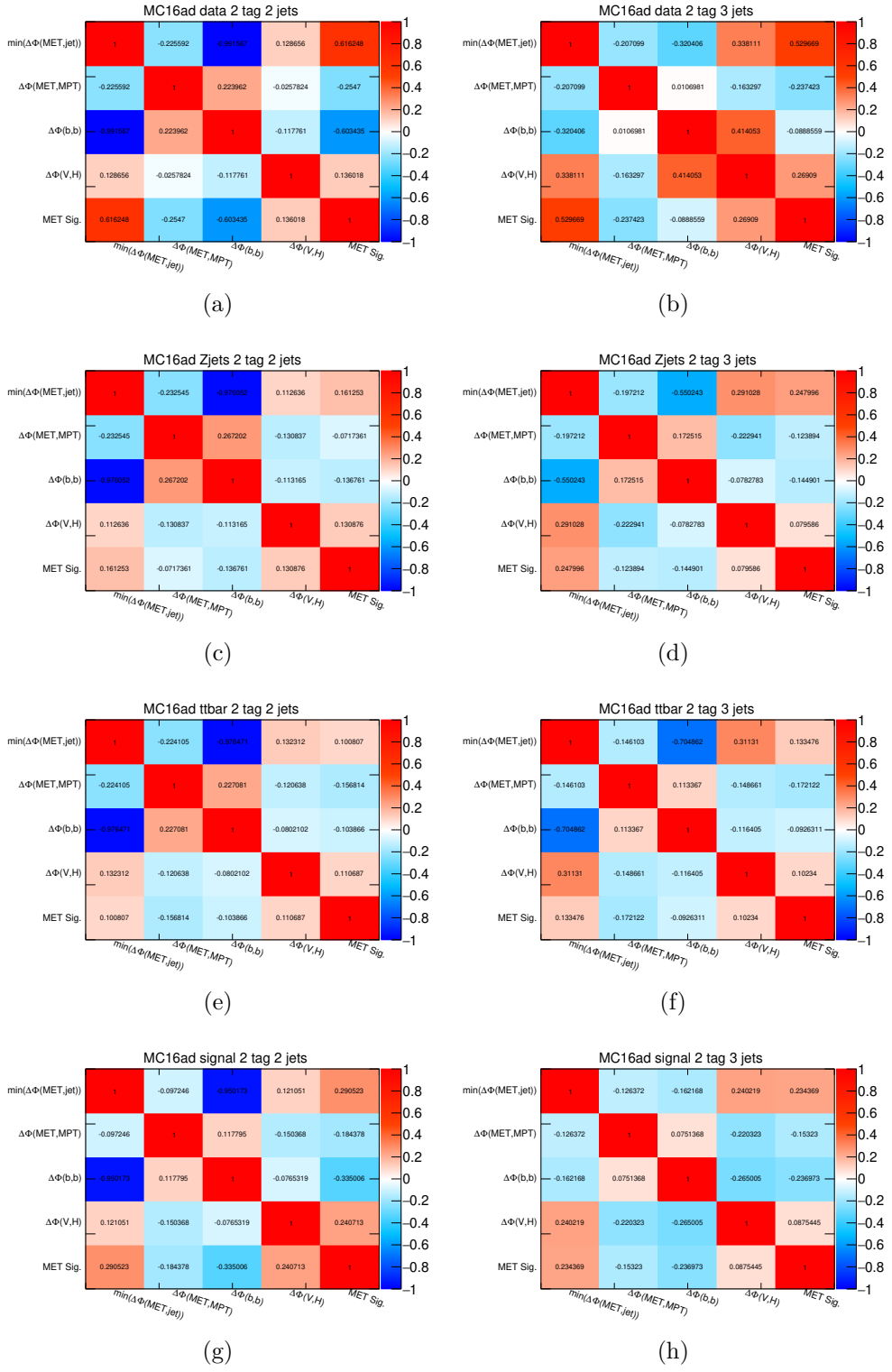


Figure 5.3 – Comparisons of the shapes of m_{bb} distributions for multi-jet and the combined W, Z, and $t\bar{t}$ backgrounds for the MC16ad period in the 2tag3jets region. The multi-jet template is determined by selecting data below 20° in min[|$\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets)|] and subtracting the backgrounds obtained from simulation. It is compared to the combined MC backgrounds m_{bb} distribution after normalising to unity.

jet and EW backgrounds, it is assumed that the remaining multi-jet contamination is absorbed by these backgrounds in the nominal fit.



(a) (b) (c) (d) (e) (f) (g) (h)

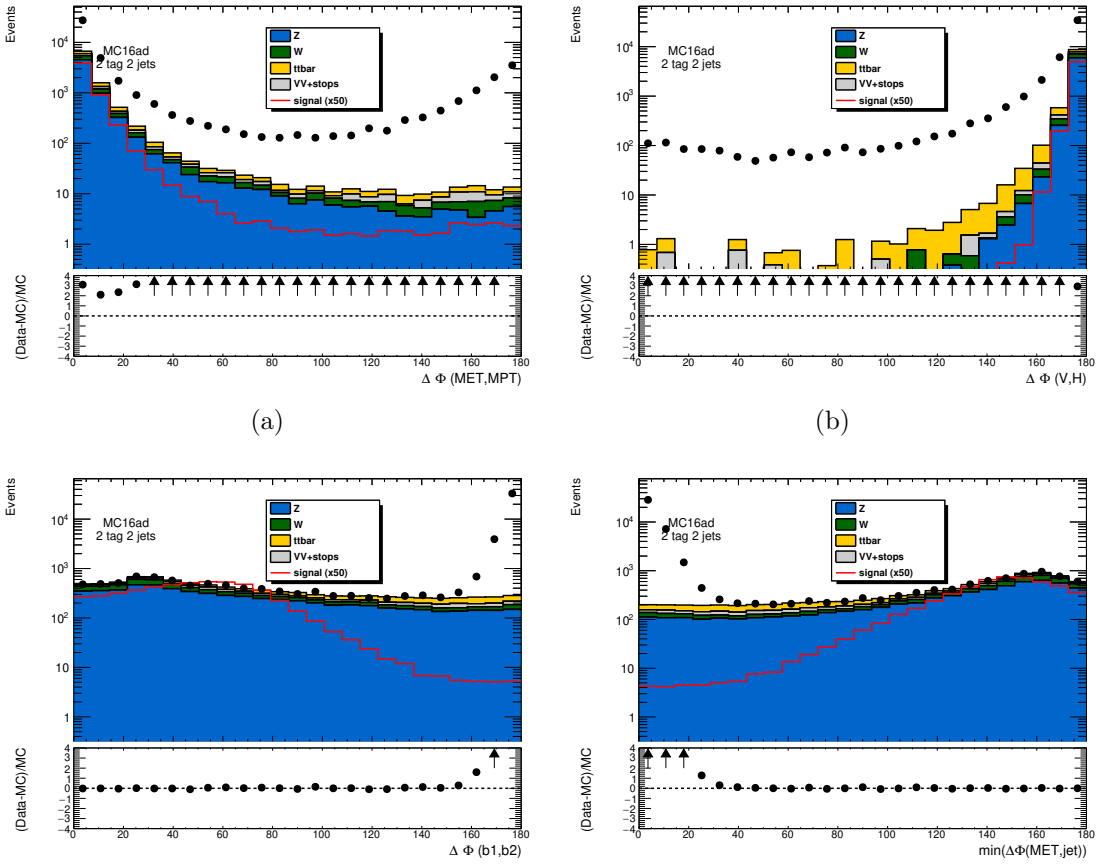
Figure 5.4 – Correlations between the variables used in the ROC curves of Figure 5.1 for data (a and b), Z+jets (c and d), $t\bar{t}$ (e and f), signal (g and h) in the 2tag2jets (left) and 2tag3jets (right). The MC samples were obtained combining period a an d, while data from 2015–2017 is considered.

Redesigning the anti-QCD selection in the 0-lepton channel

As seen from Figure 5.1, the variables used in the anti-QCD selection have diverse performance in reducing the QCD contribution. The min[$|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|$] variable is the best performing variable in general, except at high signal efficiency in the 2tag3jets category, where the other variables perform better. However when looking at the correlations in Figure 5.4, the variables are shown not to be decorrelated. For instance, $|\Delta\Phi(b_1,b_2)|$ is anti-correlated to the min[$|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|$] in the 2tag2jets category, and thus presents the same good separation power. Therefore the effect of this cut in this jet region is not adding additional discrimination power. Overall the correlations between the precut distributions of the four variables used in the cuts are found to be similar, in all the regions (data, Z+jets, $\mathrm{t\bar{t}}$, signal), which indicates that the variation of one cut parameter would lead to the same QCD rejection power as varying a different cut. The only noticeable difference is found in the correlation between min[$|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|$] and $|\Delta\Phi(b_1,b_2)|$ in the 2tag3jets regions with a higher value for the $\mathrm{t\bar{t}}$. In order to understand the relationship between the variables and their power to cut the QCD events, the distributions are produced before applying the cuts (pre-cut level) and shown in Figures 5.5 and 5.6. The MC samples are scaled according to the SF determined in Table 5.1 and the difference between the data and the MC components is solely coming from QCD processes. The data/MC ratio is plotted at the bottom of each plot and allows to spot region dominated by QCD events from a region dominated by EW backgrounds.

The number of QCD events is smaller in the 2-jets category than in the 3-jets from a first observation and seems easier to separate by the $|\Delta\Phi(b_1,b_2)|$ and min[$|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|$] distribution, as already observed from the ROC curve in Figure 5.1 and this is what is used in the implemented cuts. In the 3-jets category, the situation is indeed more complex to handle and no variable except the min[$|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|$] can isolate most of the QCD. The cross-section and kinematic properties of $\mathrm{b\bar{b}}$ production have also been studied in [167] which defines four creation schemes as shown in Figure 5.7. Figure 5.8 shows the distribution for $|\Delta\Phi(b_1,b_2)|$, depicting regions where the $\mathrm{b\bar{b}}$ pair creation proceeds through flavour creation (populating the 2-jets category in the high $\Delta\Phi$ region), or through flavour excitation of gluon splitting (populating the 3-jets category in the low $\Delta\Phi$ region). This corresponds to what is observed in the pre-cut Figures 5.5 and 5.6, where the distribution is peaking at high values in the 2-jets category, and peaking both at high and low values in the 3-jets category (the high values could come from the flavour excitation as well as from other non-considered processes as c or light misidentification).

The set of cuts could also be redesigned with the scope of simplifying the list, but also to avoid using the $\mathrm{p}_{\mathrm{T}}^{\mathrm{miss}}$ that is not well maintained by the calibration group and only used at this step in the analysis. This would be also the opportunity to increase the signal yield and to scrutinise the effect on the EW backgrounds. Furthermore, the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. defined in the Section 3.2.3, is a good variable to discriminate between QCD and signal as shown on the Figure 5.1 in both jet categories, and confirmed by looking at the pre-cut distribution in Figure 5.9. However, its correlation scheme with the main used variable min[$|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{jets})|$] from Figure 5.4 may lower the performance. Furthermore,

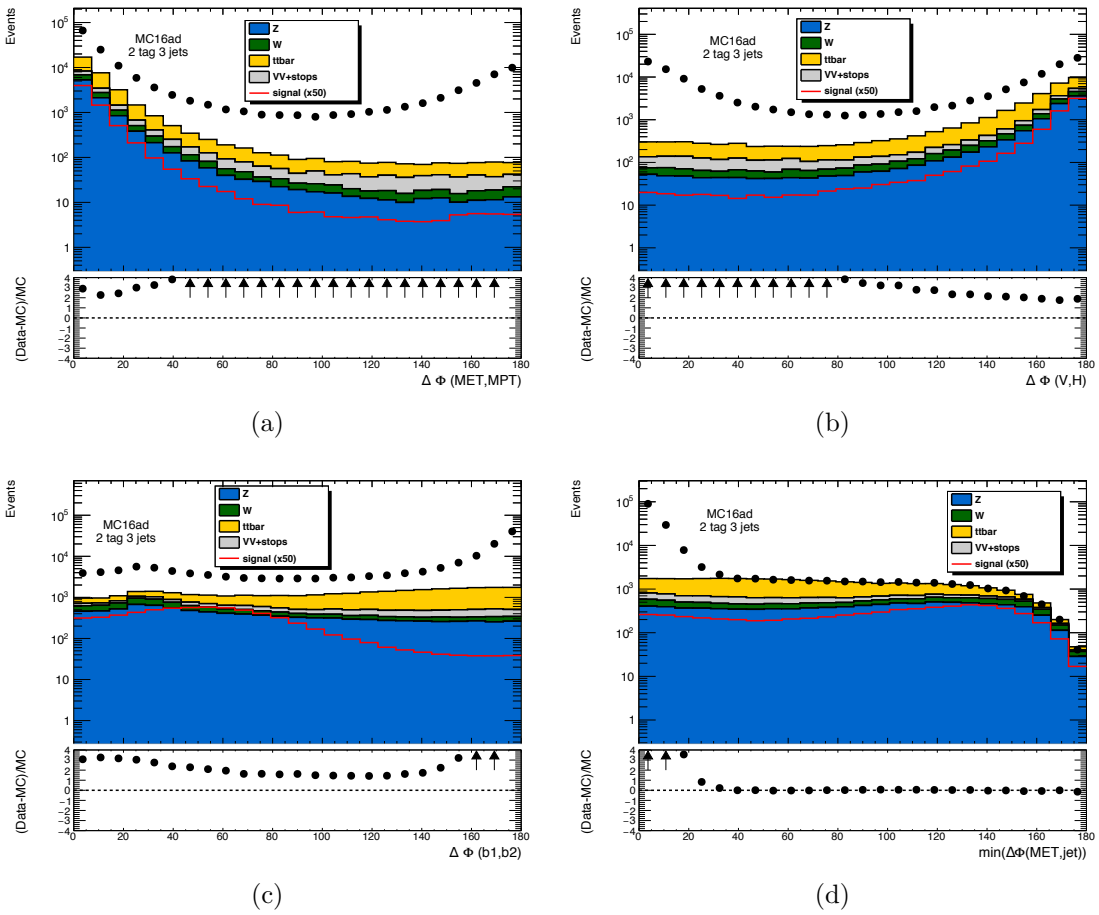


(a) (b) (c) (d)

Figure 5.5 – Distributions of the four variables used to reject the QCD backgrounds before the cuts being applied for the 2tag2jets category. Data is represented with the black dots. SM EW processes are scaled with the SF detailed in Table 5.1. MC periods a and d are considered as well as the data period 2015–2017. Arrows are drawn when the data/MC ratio absolute value is above 4.

when looking at the ROC curve of the inclusive EW backgrounds (V+jets, $t\bar{t}$, single-top and di-boson) with respect to the signal in Figure 5.10, the $|\Delta\Phi(b_1, b_2)|$ variable is shown to have some discrimination power for those backgrounds, and is far less correlated with the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. as shown in Figure 5.4. Therefore a combination between $E_{\mathrm{T}}^{\mathrm{miss}}$sig. and $|\Delta\Phi(b_1, b_2)|$ is proposed.

As for the QCD backgrounds evaluation presented before, the chosen strategy is to keep a ratio QCD/Signal in the m_{bb} signal window lower than 10%. The number of QCD background events is now evaluated using an exponential fit to the residual $|\Delta\Phi(b_1, b_2)|$ distribution redefining the scale-factors for the EW backgrounds for each defined $E_{\mathrm{T}}^{\mathrm{miss}}$sig. cut. A 2-D grid of cuts on the two variables is defined for the optimisations with 60 bins per variable, varying the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. between 0 and 15 $\sqrt{GeV}$, and the $|\Delta\Phi(b_1, b_2)|$ between 0 and 180 °. Having the two degrees of freedom the choice was made to keep the same signal yield (with a possible increase limited at 10%), but to cut harder on the number of $t\bar{t}$ events, still keeping a similar level of QCD events. An alternative was proposed to

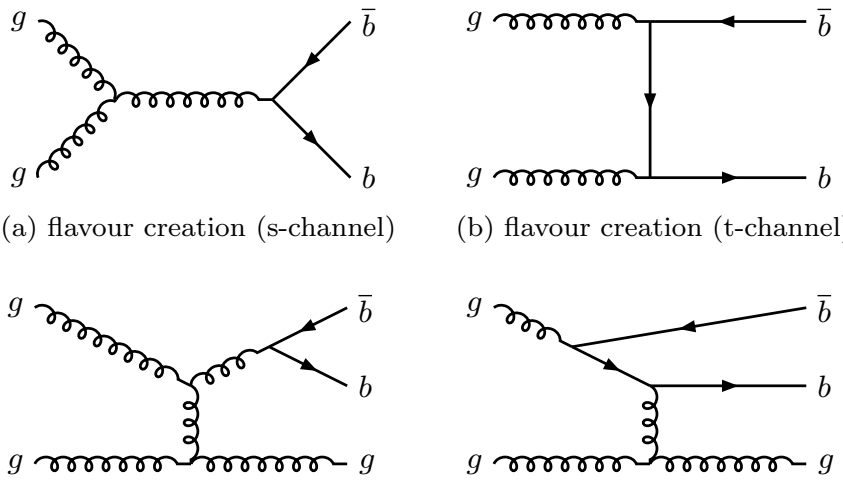


(a) (b) (c) (d)

Figure 5.6 – Distributions of the four variables used to reject the QCD backgrounds before the cuts being applied for the 2tag3jets category. Data is represented with the black dots. SM EW processes are scaled with the SF detailed in Table 5.1. MC periods a and d are considered as well as the data period 2015–2017. Arrows are drawn when the data/MC ratio absolute value is above 4.

minimise the number of Z+jets events, leading to similar cuts and similar performance. The variation of the four quantities defined in Table 5.3 is shown in Figure 5.11.

In the 2-jets region, for small enough $E_{\mathrm{T}}^{\mathrm{miss}}$sig. cut (<10), the ratio QCD/signal in the m_{bb} window SR is shown to be independent from the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. while the signal yield is slightly increasing with the cut value. The 2-D shape of the ratio QCD/signal is different in the full physics phase-space with respect to the m_{bb} SR as the Figure 5.11 is showing, due to the different correlations between the m_{bb} and the two cut variables $E_{\mathrm{T}}^{\mathrm{miss}}$sig. and $|\Delta\Phi(b_1,b_2)|$. Therefore, the optimisation the EW background suppression can be done without a strong constraint from the QCD rejection. The optimal cut is found to be $E_{\mathrm{T}}^{\mathrm{miss}}$sig. > 1.5 and $|\Delta\Phi(b_1,b_2)| < 126^{\circ}$.
In the 3-jets region, the level of QCD events is higher as already noticed in the previous evaluation. In this case, the relationship between the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. and $|\Delta\Phi(b_1,b_2)|$ cut to a 10% contamination in the SR is less straightforward. For this reason, keeping the same $|\Delta\Phi(b_1,b_2)|$ cut would lead to a signal yield reduction, with the same QCD-contamination



(a) flavour creation (s-channel) (b) flavour creation (t-channel)

(c) gluon splitting (d) flavour excitation

Figure 5.7 – Feynman diagrams of the main schemes of $b\bar{b}$ creation, populating the 2tag2jets region (a and b), and the 2tag3jets (c and d) [167].

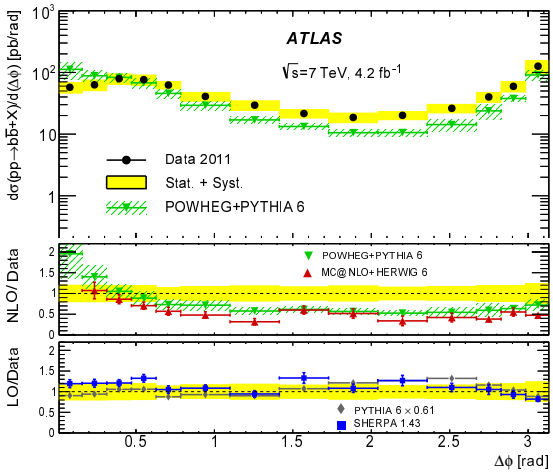
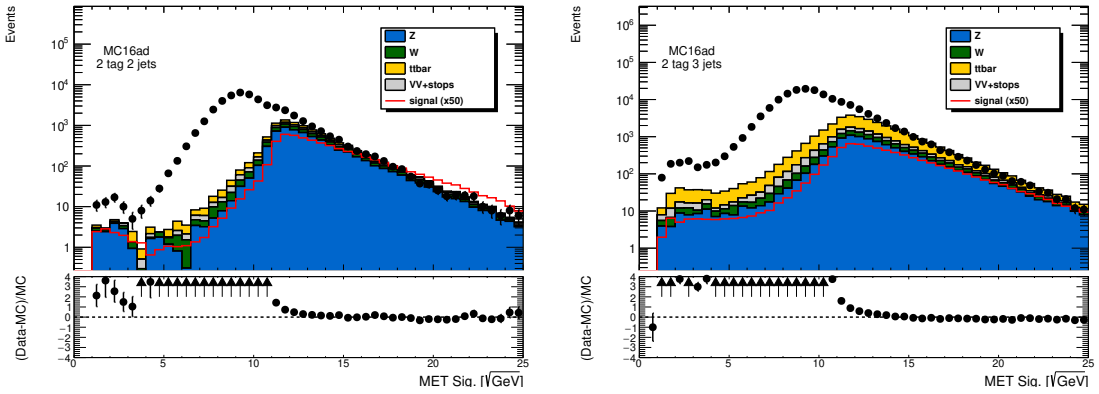


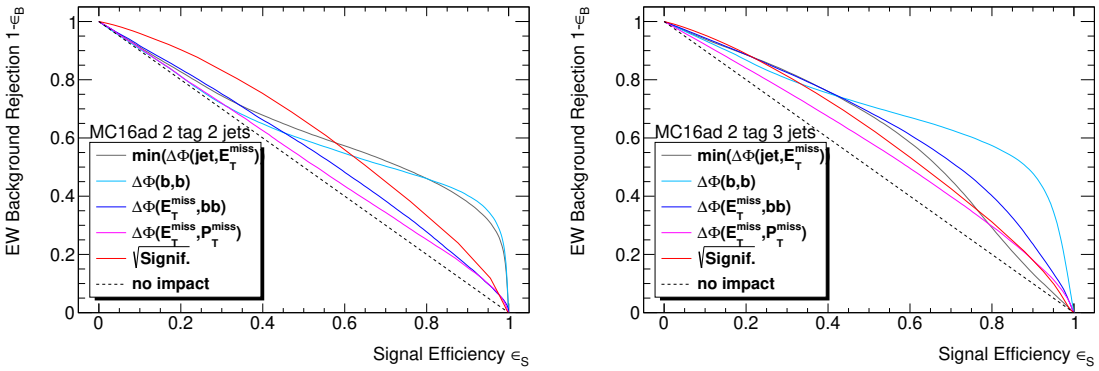
Figure 5.8 – Φ angle between the two b-tagged jets. The flavour creation scheme defined in Figure 5.7 is likely to cause the peak observed at high $\Delta\Phi$ while the broad peak at low $\Delta\Phi$ is more likely to be caused by the gluon splitting and flavour excitation schemes [167].

criteria (black star in Figure 5.11), or to a higher QCD contamination in the SR, with the signal yield (blue star). Contrary to the 2-jets case, the 2-D shape of the QCD/signal ratio is found to be much more similar between the m_{bb} window SR and the full physics phase space. The phase space available for the EW backgrounds reduction study is much more limited than in the 2tag2jets case, but a working point is found to be $E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} > 10.5$ and $|\Delta\Phi(b_1, b_2)| < 90^{\circ}$. The ratio of QCD/signal in the m_{bb} SR with that cut is found to be 8.55%, so lower than the amount reported in Table 5.3. The amount of EW backgrounds can be found in Table 5.4. Except for the W+jets background, the main improvement is obtained in the 2tag3jets region due to an important reduction of top associated backgrounds. This is thought to come from the separation obtained thanks to the $E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.}$ between the $t\bar{t}$ and V+jets backgrounds. This is probably due to the



(a) (b)

Figure 5.9 – Distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. before the anti-QCD cuts being applied. SM EW processes are scaled with the SF detailed in Table 5.1. MC periods a and d are considered as well as the data period 2015–2017 in the 2tag2jets (a) and 2tag3jets (b) region. Arrows are drawn when the data/MC ratio absolute value is above 4.



(a) (b)

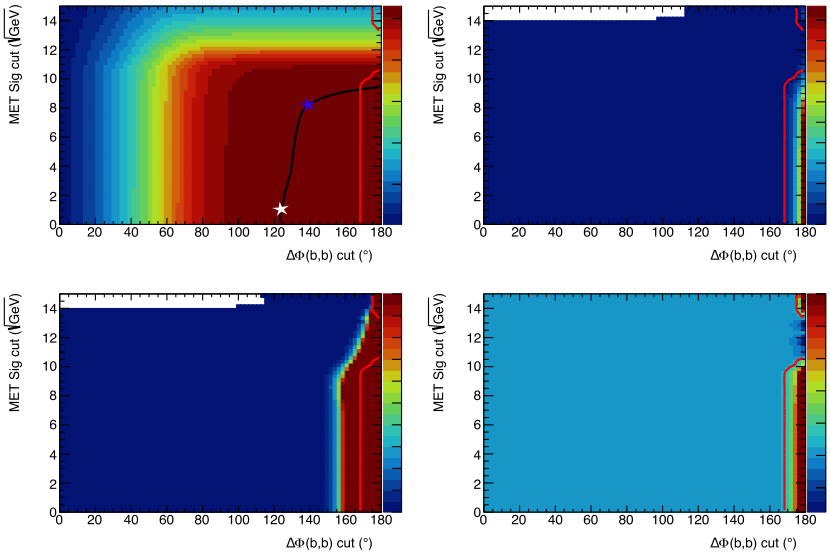
Figure 5.10 – Inclusive EW background rejection with respect to the signal efficiency for the various variables used in the anti-QCD rejection. Plots are presented for the 2tag2jets case (a) and the 2tag3jets case (b). MC16a and MC16d MC samples have been used to produce these plots.

presence of extra-objects, whose resolutions appear in the denominator of the $E_{\mathrm{T}}^{\mathrm{miss}}$sig., are shifting the distributions apart for the two processes.

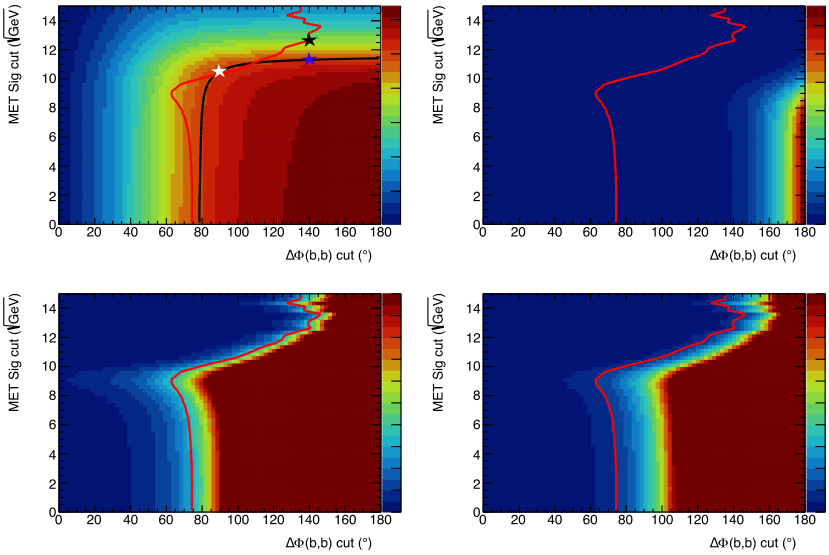
The impact of this new cut scheme in the final fit is presented in Section 6.4.

5.1.2 The 1-lepton channel

In the 1-lepton channel, the contribution of the multi-jet background is largely reduced thanks to the high $p_{\mathrm{T}}^{\mathrm{V}}$ requirement. However, it still contributes to a significant fraction of the background events and a robust procedure is necessary to estimate the



(a)



(b)

Figure 5.11 – Variation of the quantities defined in Table 5.3: Signal yield (top left), QCD Yield (top right), QCD/Signal ratio (bottom right), QCD/Signal ratio in the m_{bb} window (bottom left). The distributions are separated in the 2tag2jets (a) and 2tag3jets (b). The red line represents the cuts that would lead to a 10% QCD contamination in the m_{bb} window signal region. The black line represents the region where the cuts would lead to the same signal yield as in the old cut scheme. The blue star represents the working point that keeps the same signal yield and the same $\Delta\Phi(b,b)$ cut. In the 2tag3jets case, a black star is drawn for the working point that keeps the same QCD/signal ratio in the m_{bb} SR and the same $\Delta\Phi(b,b)$ cut. The white star is the proposed working point coming from the yield optimisation. The analysis is conducted using MC period a and data from the years 2015–2016.

Sample	2-jets			3-jets		
	nominal cuts	new cuts	dif. (%)	nominal cuts	new cuts	dif. (%)
Signal	102.5	102.5	0	102.5	102.8	0.2
$t\bar{t}$	881	743	- 16	8699	4474	- 49
Z+jets	5363	5204	- 3	6746	5722	- 15
W+jets	1096	987	- 10	1925	1814	- 6
VV+stops	521	477	- 8	1488	1112	- 25

Table 5.4 – Number of events passing the selection cuts in the 0-lepton channel. The anti-QCD cuts considered are either the nominal ones described in Section 4.2, or the redesigned ones quoted above. The numbers are extracted from the MC16 period "a" and "d", and the EW backgrounds numbers are weighted by the fitted scale factors as described in Table 5.1. The modification of SF between the two selections is found negligible compared to the impact of the cut to explain the difference in yields.

contribution of this background both in the electron and muon decay modes. In both modes, a template method is used to reproduce the shape of the QCD events.

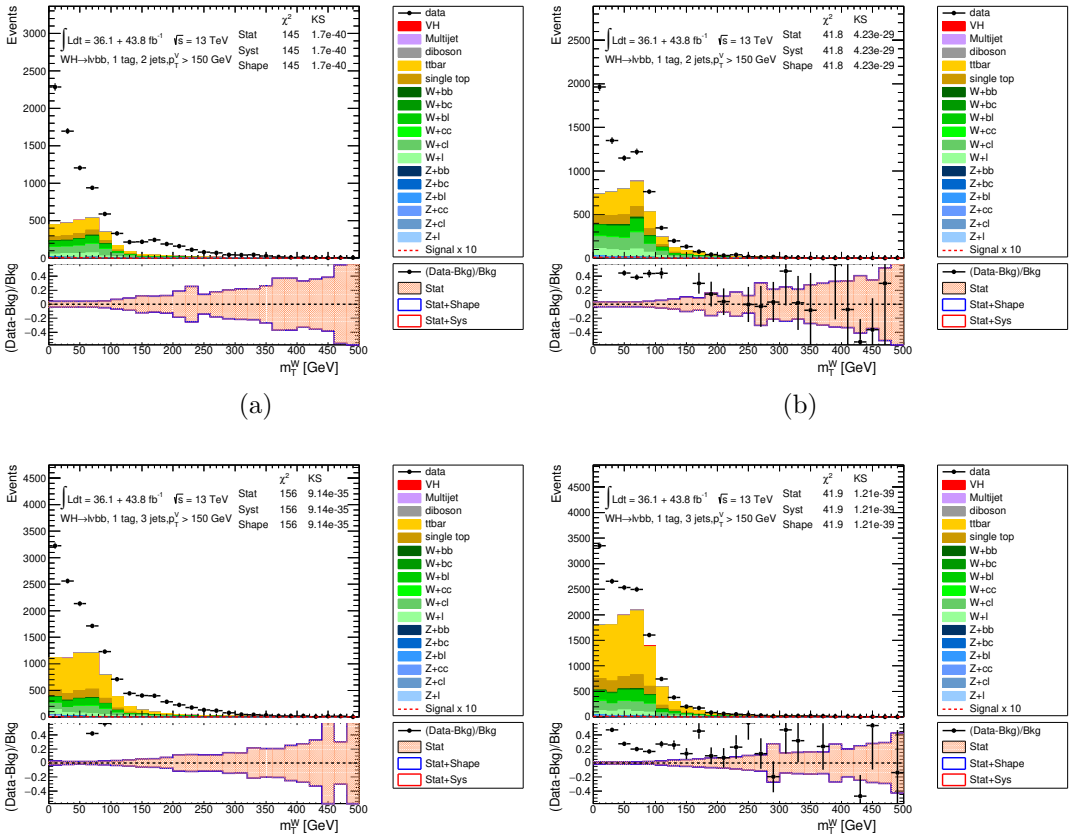
The main success from the QCD rejection in the 1-lepton channel is due to the lepton selection presented in Section 4.2, and more precisely in Table 4.4. This mainly rejects fake leptons from photons and light jets and non-prompt heavy flavour decays (in the μ channel). However, the real multi-jet contamination in the signal region cannot be extracted using simulations, both because the simulation has a limited number of events and because the simulation is not expected to reproduce fakes accurately.

A template method is therefore used to estimate the multi-jet contribution in the signal region, using data in a multi-jet enriched control region defined using inverted lepton isolation cuts. Table 5.5 summarises both the isolation cuts applied in the signal region and the inverted selection used for the multi-jet enhanced control region. The number of events in this second region is limited: it is expected to be around 9 (2) times the signal region for the electron (muon) channel. To reduce the impact of statistical fluctuations when deriving the template, only one b-tag jet instead of two is required in the control region. To get the most precise estimation of the QCD yield, the transverse mass of the reconstructed W boson is used in this study. This object is defined as the sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from the neutrino and the p_{T} of the lepton detected. Events that do not include any W boson are likely to have a low value for the transverse mass, hence allowing a good rejection of QCD events. The plots in Figure 5.12 show the distributions of this variable for the data and electroweak processes in this control region.

For each of the variables used in the analysis (BDT score, m_{bb}, ...), the EW backgrounds obtained from the MC predictions are subtracted from the data to obtain the multi-jet template in the CR. These templates are then transferred into the SR by a simple scale-factor method. This scale-factor is obtained from a fit to the m_T^W that provides discrimination mainly between processes without and with a W boson and is distributed differently for the $t\bar{t}$ and W+jets processes, where the former presents a long tail due to di-leptonic decays. This is crucial due to the importance of $t\bar{t}$ and W+jets

	Isolated Region	Inverted Isolation Region
Electron isolation	TopoEtCone20 < 3.5 GeV	TopoEtCone20 > 3.5 GeV
Muon isolation	PtCone20 < 1.25 GeV	PtCone20 > 1.25 GeV

Table 5.5 – Summary of differences in lepton isolation between the isolated and inverted isolation regions used for the template method. TopoEtCone20 variable is defined as the calorimeter transverse energy deposit in a cone of ΔR radius 0.2 around of the axis of the electron, while PtCone20 is defined as the p_{T} of the tracks in a cone of ΔR radius 0.2 around of the axis of the muon.



(a) (b) (c) (d)

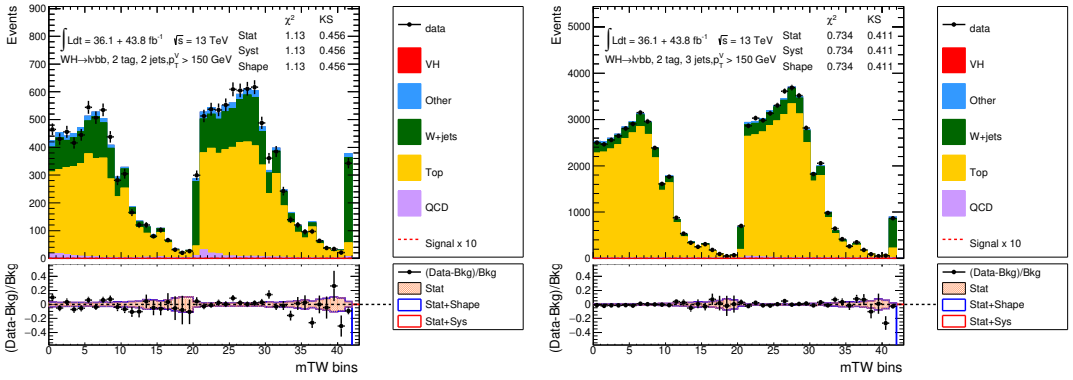
Figure 5.12 – The m_T^W distribution in the inverted isolation 1-lepton, requiring exactly 1 b-tag with 2 signal jets in the e channel (a) and in the μ channel (b), 3 signal jets in e channel (c) and in the μ channel (d) [168].

processes amongst the various EW backgrounds in the signal region. However, their overall normalisation is not correctly modelled and thus normalisation factors that can have a significant impact on the multi-jet estimate are extracted simultaneously to the multi-jet template itself. Therefore in order to avoid a bias in the multi-jet estimate, separate normalisation factors are extracted for the top ($t\bar{t}$ +single top) and W+jet contributions, using the W+hf enhanced control region to better discriminate between the two processes.

In practice, the different regions (W+hf CR, e/μ regions) are summed up in a cumulative distribution and binned for statistical reasons: the multijet-CR region is binned to get a constant MC statistical error while the W+hf CR is limited to one bin to take into account only the yield modification. The left-most bins are representing the electron region, while the right bins are representing the μ region. Separate templates are used for the electron multi-jet, muon multi-jet, top and W+jets components; the normalisation factors are extracted for each contribution is presented in Table 5.6. Post-fit plots of the m_T^W in the SR are shown in Figure 5.13. The total fake electron (muon) contamination in the 2-jets channel is estimated to be 1.9% (2.8%) and in the 3-jets channel 0.2% (0.4%). Dedicated systematics are derived on these shapes and presented in Section 6.2.3.

Region	top ($t\bar{t}$ + single top)	W+jets
2tag2jets	1.02	1.27
2tag3jets	0.99	1.13

Table 5.6 – Summary of normalisation scale factors for top ($t\bar{t}$ + single top) and W+jets derived in the isolated lepton region.



(a) (b)

Figure 5.13 – The m_T^W distribution in the isolated 1-lepton $p_{\mathrm{T}}^{\mathrm{V}} > 150$ GeV channel in the 2tag2jets (a), 2tag3jets (b) region after applying top ($t\bar{t}$ + single top) and W+jets normalisation factors. Bins 1-21 correspond to the e only channel, bins 22 to 42 correspond to the μ only channel, and bins 21 and 42 represent the W+hf control region [168].

5.2 Modelling uncertainties of the MC-based backgrounds

As presented in Section 6.1, the final profile likelihood fit uses the nominal background predictions as well as the measured data samples to get the final statistical result on the signal observation. Therefore a good modelling and dedicated uncertainties for the background shapes are one of the core aspects of the analysis. An introduction on the link between profiled likelihood fits and the background modelling is presented in Section 5.2.1, then the normalisation aspects are treated in Section 5.2.2, while shaping is divided into 1-D reweighting in Section 5.2.3 and N-D reweighting in Section 5.2.4.

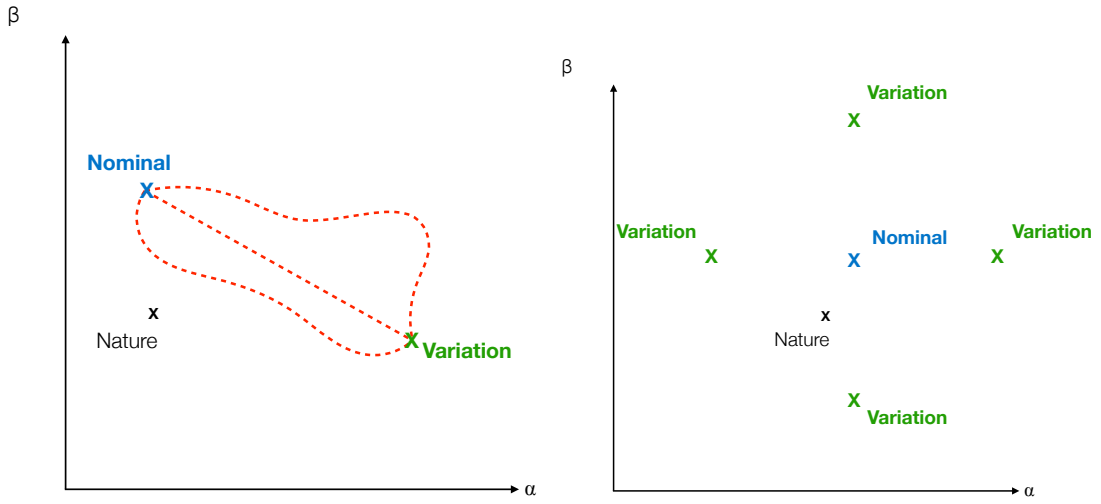
5.2.1 Introduction to the background modelling

The $Vh(b\bar{b})$ analysis is based on the description of various physics processes using MC samples whose shapes should ideally describe data in a precise way. The whole construction relies on the accuracy of the inputs used for this description. In reality, MC simulations not only suffer from the limited number of events generated, but they are also relying on non-trivial assumptions on the physics behind the model. In the same way, the detector is not measuring the true energy or position of particles, MC simulations are not describing with perfect precision the shapes of the variables related to the process. The differences can be parametrised in two different ways:

1. **internal weight variations**: some parameters are allowed to vary inside one MC prescription, such as the factorisation and renormalisation scales (μ_F and μ_R) used as arbitrary cut-off parameters in the underlying QCD processes.
2. **comparison with external generators**: some more conceptual variations such as the clustering algorithm in the PS or the ME computations can be evaluated thanks to the comparison between two generators.

The former variation can be either used with an envelope approximation (catching the maximal variation in each point of the phase space) or as a quadratic sum of all the variations. Both methods present some caveats since the global meaning of the systematics in the envelope is not always defined, and the quadratic sum implies no internal correlations between the effects. The latter variation is more complex to interpret as it takes into account several modifications. Though, it is the only evaluation possible to measure the impact of a MC design on the simulation of the process. This evaluation is often called a 2-point systematics as it relies on the mapping of the phase-space between only 2-points

as represented in Figure 5.14a. These differences are used as biases to correct for the difference between the prediction or the measurement and the reality. They are called systematic uncertainties and transform the measurement of the Parameter Of Interest (POI, signal strength, mass of the discovered particle...) into a joint measurement using the Nuisance Parameter (NP) approach.



(a) (b)

Figure 5.14 – Representation of the evaluation of a 2- (a), and N-points (b) systematics. α and β are two physics parameters prone to be modified by the choice of the MC generator.

The residual point is to understand the changes that determine a difference in the distributions of the MC events. Modifications can be classified into normalisation, acceptance, and shape variations. The first two only affects the global number of events in the distribution and leave the shapes untouched. Normalisations are used to quantify the overall uncertainty coming from effects that would affect all the events in the same way as, for instance, the uncertainty on the global cross-section of the process. The basic form of this effect can be written as:

$$\theta_{norm} = 1 - \frac{N_{var.}}{N_{nom.}}$$

where $N_{nom.}$ is the global number of events obtained with the nominal MC sample, while $N_{var.}$ is the same quantity with the variation. Acceptance uncertainties are quite similar in their effects, as they change the global number of events. However, their interpretation is linked to the categorisation of the events based on the low-level variables such as the number of jets and the $p_{\mathrm{T}}^{\mathrm{V}}$ of the events. When a similar normalisation of the backgrounds is applied between different fit regions (such as the Z+hf normalisation between the 0- and 2-leptons SR, or the W+hf between the 1-lepton SR and Whf CR), a dedicated acceptance systematic is derived:

$$\theta_{acc} = 1 - \left(\frac{N[Category_A(nom.)]}{N[Category_B(nom.)]} \Big/ \frac{N[Category_A(var.)]}{N[Category_B(var.)]} \right)$$

Finally, the shape variations are treating all the effects that cannot be covered by normalisation. Contrary to the previous uncertainties, the shape effect is variable dependent and a binned effect. Specific internal weights can be computed for each considered effect, modifying in a bin-by-bin way the value of the distribution. In the fit, the NPs are defined from those weights by either summing the contributions of the effects in quadrature (a), or taking the leading effect in each bin (b). The Physics Modelling Group (PMG) in ATLAS recommends to use the solution (b) whenever possible, but in a profile likelihood fit the meaning of the individual NPs is questionable. The solution (a) instead would lead to an important number of parameters with fit robustness issues and whose correlations are not trivial to handle. The solution could be to replace internal weights by a grid of N-point systematics as shown in the Figure 5.14b. For each sample, only the biggest effect is kept amongst all the variations as:

$$\theta^i_{shape} = 1 - \frac{h^i_{var.}}{h^i_{nom.}}$$

where $h^i_{var./nom.}$ is the value of the variation/nominal histogram at the bin i.

Since we are interested in the measurement of the POI, the NPs are "profiled-out" by using subsidiary measurements:

$$\mathcal{L} = f_0(\vec{x}|\mu, \overrightarrow{\theta}) \cdot f_{subs}(\overrightarrow{\tilde{\theta}} \,|\, \overrightarrow{\theta})$$

where f_0 is the parametrised probability of measuring the POI μ and the systematics $\overrightarrow{\theta}$ knowing the measurement $\vec{x}$, and f_{subs} the subsidiary measurement on the $\overrightarrow{\theta}$s. A naive approach would have been to reproduce the POI measurement with both $\pm 1\sigma$ variation and let the total impact be the quadratic sum of the variations. However, this rather simple technique is both CPU consuming and is mistreating the correlation of the systematics in both the subsidiary and nominal measurement. A more reasonable approach is to go for the continuous approach of the NP, which is working well for most of the systematics. However in the case of modelling systematics, the profiling function is not apriori known by the user, and its interpretation is challenging for the analyser. A few examples are shown in Figure 5.15.

In the case of a general Gaussian prior, where the alternative variation lies on the $\pm 1\sigma$ point, there is a 37% probability for the NP to be outside the variation range around the nominal. This causes issues of interpretation of pulls of the NP outside the window, especially in case 2 described previously. Furthermore, there is no reason to consider the nominal MC to be more correct than its variation. Therefore a flat prior, as represented in the left part of Figure 5.15, can be designed.

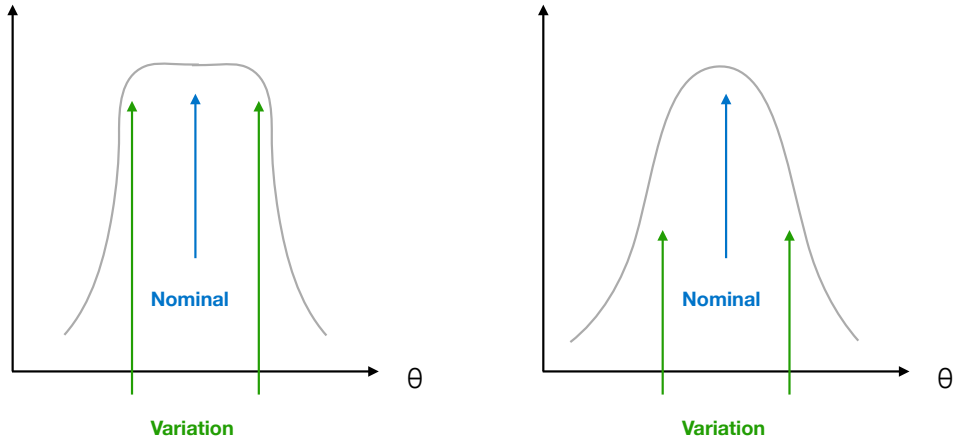


Figure 5.15 – Representation of the subsidiary measurement modelisation on the one dimensional nuisance parameter θ. A flat prior is set on a) while a standard Gaussian prior is set on b).

5.2.2 Acceptance and normalisations

In the context of the $Vh(b\bar{b})$ analysis, the acceptance and normalisation uncertainties are process specific. To cut down the CPU time of the reconstruction step for the alternative samples, they are derived from truth-based studies. The systematics used to be derived from the RIVET software [169] which is a well-known tool in the ATLAS collaboration. A more modular approach, called the TruthFramework, was developed in the $Vh(b\bar{b})$ analysis accounting for the specifics of b-tagging. All the modelling systematics, extracted from those two frameworks, are parametrised with a Gaussian profile. The prior on the width of the distributions are described in the following paragraphs.

Higgs signal

The theory provides global systematics on the cross-section and branching ratio following the LHC Higgs working group prescriptions [170] [171]. QCD scales (μ_R and μ_F) are varied independently by a factor 1/3 and 3 from their original values, and the PDF+α_S values are varied using the internal PDF4LHC15 weights (taking the 68% band of the variation). Since the Vh processes contain several sub-processes with different prescriptions, the quark induced processes ($qq \to Zh$) have been separated from the gluon induced ones. Whenever possible the quark induced processes uncertainties in the Zh channel are taken as the Wh ones from [170], leaving the gluon induced Zh being the difference in quadrature between the total prescription for Zh in [170] in the calculated gluon process. When this technique is not possible, the prescription described in [171] is used.
An analysis dependent normalisation and acceptance study is conducted to evaluate the effect of the UE/PS, scales and PDF+α_S variations:

— **UE/PS**: two ways of assessing the systematics are used as shown in the lists in Section 5.2.1: the MadGraph 5_aMC@NLO + Pythia 8 generator is compared with the A14 tune variations, and the nominal Powheg + Pythia 8 is compared to the alternative Powheg + Herwig 7. In the first case, all the variations are added in quadrature following the recommendations and the result is compared to the uncertainty of the 2-point comparison. The maximal variation is kept as the normalisation uncertainty. The same method is used to derive the acceptance systematic between the 2- and 3-jets SR. The same values are used for the quark and gluon induced processes and are correlated in the fit.

— **QCD scales**: the nominal Powheg + Pythia 8 is compared to its internal weight variation of scales, following the envelope method using the Stewart-Tackmann-method [172]. In this way, the uncertainties are derived in each jet-bin, and thus no 2- to 3-jets acceptance is required. However, since a 4-jets veto is applied, a dedicated acceptance systematic is computed. The same values are used for the quark and gluon induced processes and are decorrelated in the fit, as from the Run-1 experience where the values were different between the two processes.

— **PDF+α_S**: the quadrature sum of the separate effects due to the PDF and α_S variations from the PDF4LHC15 internal weights as for the cross-section uncertainty are computed. The same uncertainty is derived in the 2- and 3-jets without acceptance systematics since both regions yield the same systematical values.

The ME variation effects, observed in the comparison between full Pythia 8 and MadGraph 5_aMC@NLO + Pythia 8 samples, are negligible compared to all the effects observed so far

All the systematics with their values are summarised in Table 5.7.

Di-boson backgrounds

The di-boson backgrounds are composed of three distinct processes, WW, WZ and ZZ, the former having the smallest contribution due to the very stringent selection. Global normalisation factors inherited from early Run-2 papers are set on the three processes from the quadrature sum of the scale variations (factorisation, normalisation and resummation), the PS (Pythia 8 v.s. Herwig ++) and the ME/PS (Sherpa v.s. Powheg + Pythia 8) variation.

Then depending on the lepton selection only a few final states are considered for the acceptance due to the 2 b-tagged requirement on the jets j: in 0-lepton $ZZ \to \nu\bar{\nu}$jj, in 1-lepton $WZ \to \ell\nu$jj, in 2-leptons $ZZ \to \ell^+\ell^-$jj. These systematics are then applied to the sub-dominant backgrounds $WZ \to \ell^+\ell^-$jj and $WZ \to \nu\bar{\nu}$jj. All the uncertainties are

systematic	source	0L: $Zh \to \nu\bar{\nu}b\bar{b}$ 2j	0L: $Zh \to \nu\bar{\nu}b\bar{b}$ 3j	1L: $Wh \to \ell\nu b\bar{b}$ 2j	1L: $Wh \to \ell\nu b\bar{b}$ 3j	2L: $Zh \to \ell^+\ell^- b\bar{b}$ 2j	2L: $Zh \to \ell^+\ell^- b\bar{b}$ ≥3j
$\theta_{BR(H\to b\bar{b})}$	HO effects, m_b, α_S	1.7 %					
$\theta_{\sigma(q\bar{q}\to VH)}$	QCD scale	0.7 %					
$\theta_{\sigma(gg\to ZH)}$	QCD scale	27 %		–		27 %	
$\theta_{\sigma(q\bar{q}\to VH)}$	PDF+α_s	1.6 %		1.9 %		1.6 %	
$\theta_{\sigma(gg\to ZH)}$	PDF+α_s	5 %		–		5 %	
θ_{norm}	PS/UE	4.1%	4.1%	6.2%	6.2%	2.9%	2.9%
θ_{acc}(2/3 jets)	PS/UE	–	2.2%	–	1.8%	–	11.2%
θ_{norm} (2-jets)	QCD scales	6.9%	–	8.8%	–	3.3%	–
θ_{norm} (3-jets)	QCD scales	-7%	+5%	-8.6%	+6.8%	-3.2%	+3.9%
θ_{acc}(jet veto)	QCD scales	–	-2.5%	–	3.8%	–	–
θ_{norm}	PDF+α_S	1.1%	1.1%	1.3%	1.3%	0.5%	0.5%

Table 5.7 – Summary of all systematic uncertainties on the VH acceptance and normalisation originating from changing the PS/UE, QCD scales and PDF+α_S uncertainties.

considered uncorrelated between the lepton regions bins. Dedicated systematics are set for WZ as the 1-to-0 lepton acceptance due to the constraint from the 1-lepton channel, and for the ZZ as the 2-to-0 lepton acceptance in a similar way.
Similarly to the signal samples, the acceptances are derived for each variation:

— **PS/UE**: a 2-point comparison using Pythia and Herwig is added in quadrature to the internal shower variation as recommended by the Physics Modelling Group. The derived uncertainties in the 2-jets region are applied in the full region and dedicated 2-to-3 jets acceptances are derived.

— **QCD scales**: the nominal Sherpa 2.2.1 is compared to its internal weight variation of scales (doubling and halving the μ_R and μ_F), following the envelope method using the Stewart-Tackmann-method [172]. Within this approach, three systematics are produced on: a global normalisation, a 2-to-3 jets acceptance and a $\geq$ 4jets veto. The latter one is only in the 0- and 1-lepton channel, since the 2-leptons is inclusive in the number of jets.

No ME systematic is evaluated due to the very different strategy in extra-jet modelling of the Powheg and Sherpa generators. The PDF and α_S uncertainties have been proven to be negligible at this step and thus ignored in the fit. Acceptances to correlate the effects across the lepton channels are evaluated adding in quadrature all the previously cited effects. All the priors are summed up in Table 5.8

W+jets backgrounds

For all the V+jets backgrounds the systematics have been differentiated between the true final flavours of the decay: V+hf (includes bb, bc, cc and b, light final states), V+cl and V+lights. This flavour is set by the heaviest flavoured true hadron in a $\Delta R < 0.3$ cone from the reconstructed jet. The different flavoured categories are used simultane-

		0L: $ZZ \to \nu\bar{\nu}b\bar{b}$		1L: $WZ \to \ell\nu b\bar{b}$		2L: $ZZ \to \ell^+\ell^- b\bar{b}$	
	2j	3j	2j	3j	2j	≥3j	
θ_{acc}	UE/PS	5.6%	5.6%	3.9%	3.9%	5.8%	5.8%
θ_{acc} (2/3-jets)	UE/PS	–	7.3%	–	10.8%	–	3.1%
θ_{acc} (2-jets)	QCD scales	10.3%	–	12.7%	–	11.9%	–
θ_{acc} (3-jets)	QCD scales	-15.2%	+17.4%	-17.7%	+21.2%	-16.4%	+10.1%
θ_{acc} (≥ 4-jets veto)	QCD scales	–	+18.2%	–	+19.0%	–	–
		ZZ		WZ		WW	
θ_{norm}	Scales, PS, ME	20 %		26 %		20 %	

Table 5.8 – Summary of the normalisation and acceptance uncertainties on the di-boson predictions in the various jet categories and for the different final states.

ously in the fit. Some V+jets mismodelling being observed in the fit, the normalisation of the W+hf is allowed to float separately in the 2- and 3-jets in the fit with a constraint coming from the 1-lepton W+hf CR. Since the total fraction of V+cl and V+lights is below 1%, they are only constrained with gaussian priors as shown in Table 5.9.
Dedicated acceptance systematics have been derived to account for differences of normalisation:

— **CR/SR**: since the main constraint on the W+hf scale factor is coming from the 1-lepton W+hf CR, a comparison is made with the 1-lepton SR.

— **0-/1-lepton**: the constraint is then propagated to the 0-lepton SR thanks to a common 2-/3-jets acceptance uncertainty

A flavour uncertainty is also set on the flavour composition of the V+hf category by comparing the fraction of Wcc, Wbl, Wbc and Wbb in the 0- and 1-lepton SR. Due to the high fraction of events, the Wbb category has been chosen as a common reference. Since very similar numbers are observed between 2- and 3-jets regions and between SR and CR, the highest variation has been chosen in all the categories to be a used uncertainty.

The variations used to compute the various acceptance uncertainties are obtained through QCD scales (μ_R, μ_F, CKKW), PDF+α_S and ME/PS:

— **QCD scales**: the renormalisation and factorisation scales (μ_R, μ_F) are varied by a factor 1/2 and 2 in the Sherpa 2.2.1 sample. The CKKW and resummation scales of the parton merging scheme are varied internally with the help of Sherpa 2.1 samples since the information was not yet available in Sherpa 2.2.1 samples.

— **PDF+α_S**: the comparison is made between the internal NNPDFs (100 NNPDF and CT14 NNLO) variations of the nominal Sherpa 2.2.1 sample. The maximal value is kept to get a conservative estimation.

— **ME/PS**: a 2-point comparison is made with MadGraph 5_aMC@NLO + Pythia 8 samples to modify both the ME and PS generator.

All the effects are summed in quadrature to assess each systematics more conservatively. All the priors can be found in Table 5.9.

	0-Lepton		1-Lepton			
Systematic	2-Jets	3-Jets	SR 2-Jets	SR 3-Jets	CR 2-Jets	CR 3-Jets
θ_{norm} W+l	32%					
θ_{norm} W+cl	37%					
θ_{norm} W+hf	Floating normalisation (2- and 3-jets separated)					
θ_{acc}(CR/SR) W+hf	–		10%		–	
θ_{acc}(0-/1-lepton) W+hf	5%		–			
θ_{acc} W+(bc/bb)	15%		30%			
θ_{acc} W+(bl/bb)	26%		23%			
θ_{acc} W+(cc/bb)	10%		30%			

Table 5.9 – Summary of the normalisation and acceptance uncertainties on the W+jets predictions in the various categories of the 0 and 1-lepton channels.

Z+jets backgrounds

The Z+jets case is very similar to the W+jets one. The main difference comes from the regions considered for the background. If W+jets were dominant in the 0- and 1-lepton channels, the Z+jets are dominant in the 0- and 2-leptons channels. As for the W+jets the Z+hf are free to float in the 2- and 3-jets separately, while the Z+cl and Z+l get a prior from the measured variations. No explicit Z+hf control region is designed in the 0- or 2-leptons channels, since there is an important data constraint from the m_{bb} sidebands. Therefore only a 0-to-2 leptons acceptance is derived from the high-$p_{\mathrm{T}}^{\mathrm{V}}$ bin of the 2-leptons channel. A similar source of flavour uncertainties are derived as for the W+jets case and are summarised in Table 5.10

	0-Lepton	2-Leptons	
Systematic	2-/3-jets	2-Jets (low-/high-$p_{\mathrm{T}}^{\mathrm{V}}$)	$\geq$ 3-Jets (low-/high-$p_{\mathrm{T}}^{\mathrm{V}}$)
θ_{norm} Z+l	18%		
θ_{norm} Z+cl	23%		
θ_{norm} Z+hf	Floating normalisation (2- and 3-jets separated)		
θ_{acc}(0-/2-lepton) Z+hf	7%	–	
θ_{acc} Z+(bc/bb)	40%	40%	30 %
θ_{acc} Z+(bl/bb)	25%	28%	20 %
θ_{acc} Z+(cc/bb)	15%	16%	13 %

Table 5.10 – Summary of the normalisation and acceptance uncertainties on the Z+jets predictions in the various categories of the 0 and 1-lepton channels.

Top pair backgrounds

Due to the very different behaviour between the 0- and 1-lepton channels, where $t\bar{t}$ events are only partially reconstructed contrary to the 2-leptons channel where the full event is reconstructed, the uncertainties are derived separately between the channels and kept uncorrelated in the fit. The $t\bar{t}$ backgrounds are mainly constrained from the W+hf CR in the 1-lepton channel, the 2tag3jets region in the 0- and 1-lepton channels and the Top-$e\mu$ CR in the 2-leptons case. A floating normalisation is set for the 0- and 1-lepton regions, while for the 2-leptons a floating normalisation per jet category is computed. Then dedicated acceptance factors are set in the 0- and 1-lepton case:

— **CR/SR** (1-lepton): since the main constraint on the $t\bar{t}$ scale factor in the 0- and 1-lepton is coming from the 1-lepton SR, a comparison is made with the 1-lepton W+hf CR to control the normalisation in that region. A bad modelling of the $t\bar{t}$ could impact the W+jets constraint obtained. The values extracted between the 2- and 3-jets are quite different and the most conservative value is kept across the two regions and is fully correlated.

— **1-/0-lepton**: the scale factor is then propagated as in the W+jets case to the 0-lepton channel thanks to a dedicated acceptance systematic.

— **3-/2-jets** (0- and 1-lepton): the most significant contribution is found in the 3-jets region, a dedicated acceptance is set for the 3-to-2 jets constraint. Since a similar value is found in the 0- and 1-lepton case the same systematic is set and fully correlated between the regions.

In the 2-leptons case, a CR/SR acceptance have been tested, however, the distribution in the Top-$e\mu$ CR and SR are found statistically compatible, hence no systematic is assigned.

The variations used to compute the various acceptance uncertainties are obtained through scale variation (QCD/PS), ME and PS/UE:

— **QCD/PS scales**: two samples are generated with low and high QCD emission parameters. The low radiation uses a doubling of QCD scales (μ_{F} and μ_{R}) while taking the Down variation of the A14 variations. The high radiation uses the halving of the QCD scales while taking the Up variation of the A14 variations. The damping factor is also raised by a factor two compared to the nominal. Only the half variation between High and Low is taken into account.

— **PS/UE**: a 2-point comparison is made with Powheg + Herwig 7 sample to modify the PS model.

— **ME**: a 2-point comparison is made with MadGraph 5_aMC@NLO + Pythia 8 sample to modify the hard scattering generation process.

All the effects are summed in quadrature, as recommended by the PMG group. All the priors can be found in Table 5.11.

	0-Lepton		1-Lepton				2-Leptons	
Systematic	2-jets	3-jets	SR 2-jets	SR 3-jets	WCR 2-jets	WCR 3-jets	2-Jets (low-/high-$p_{\mathrm{T}}^{\mathrm{V}}$)	$\geq$ 3-Jets (low-/high-$p_{\mathrm{T}}^{\mathrm{V}}$)
θ_{norm}	floating normalisation						floating normalisation	floating normalisation
θ_{acc}(SR/CR)	–		–		25 %		–	
θ_{acc}(1-/0-lepton)	8 %		–				–	
θ_{acc}(3-/2-jets)	9%	–	9%	–	9%	–	–	

Table 5.11 – Summary of the normalisation and acceptance uncertainties on the $t\bar{t}$ predictions in the various categories of all the lepton channels.

Single-top backgrounds

As described in the Section 4.1, the single-top background is divided into three sub-components: Wt-, t-, and s-channel. A first overall normalisation inherited from the LHC top modelling working group is applied on each channel independently of the number of final leptons. These uncertainties include in quadrature the effect of varying the renormalisation and factorisation scales, as well as the PDF+α_{S} variations.

As for the analysis specific systematics, the procedure is very similar to the one used for the $t\bar{t}$ sample, to the extent that this background is not dominant in any lepton channel. The only category with a significant contribution is the 1-lepton, mainly for the t- and Wt- channel. The normalisation in the s-channel obtained from recommendations is only considered, and no acceptance systematics is derived in this channel. In the Wt-channel, it has been observed that the shape and the values of systematics are varying by a non-negligible factor when considering different final states. The main source of systematics is very different in the $b\bar{b}$ final state compared to the events where one or none b-tagged jets are true b-jets (others). In the latter case, the b-pair is likely to come from the top decay while in the former the b-jets are coming from two different decays, hence making the topologies and the systematics impact very different. In both the Wt- and t-channels, the acceptances have been separated between the 2- and 3-jets categories. However, since the two effects are going in the same direction, they are considered correlated.

The variations used to compute the various acceptance uncertainties are obtained through scale variation (QCD/PS), ME and PS/UE:

— **QCD/PS scales**: two samples are generated with low and high QCD emission parameters. The low radiation uses a doubling of QCD scales (μ_{F} and μ_{R}), taking the Down variation of the PERUGIA2012 variations. The high radiation uses the halving of the QCD scales, taking the Up variation of the

PERUGIA2012 variations. Only the half variation between High and Low is taken into account.

— **PS/UE**: a 2-point comparison is made with POWHEG + HERWIG ++ sample to modify the PS model.

— **ME**: a 2-point comparison is made with MADGRAPH 5_aMC@NLO + HERWIG ++ compared with the POWHEG + HERWIG sample to modify the hard scattering generation process. For the Wt- samples a POWHEG + PYTHIA 6 sample applying the Diagram Subtraction (DS) scheme is used to test the ME hypothesis of the Diagram Removal (DR) applied in the nominal sample.

All the effects are summed in quadrature, as recommended by the PMG group. All the priors can be found in Table 5.12.

	0-, 1- and 2-Leptons	
Systematic	2-jets	($\geq$) 3-jets
θ_{norm} (s-channel)	4.6 %	
θ_{norm} (t-channel)	4.6 %	
θ_{acc} (t-channel)	17%	20%
θ_{norm} (Wt-channel)	6.2 %	
θ_{norm} (Wt$\rightarrow$b$\bar{\text{b}}$)	55%	51%
θ_{norm} (Wt$\rightarrow$ other)	24%	21%

Table 5.12 – Summary of the normalisation and acceptance uncertainties on the single-top predictions in the various jet categories and for the different final states.

5.2.3 Shapes as 1-D reweighting

The evaluate the MC systematics, the shape of the distributions is used as a second degree of freedom. The evaluation of shape uncertainties is decoupled from the normalisation effect by requesting the weights to be centred at one, for instance by normalising the distribution when doing a 2-point systematic. This method allows having an easier understanding of the main effect that would cause the modification of the MC distribution, either from a global scale factor effect or from a shape modification. However, the main question that needs to be answered is the variable(s) that should be used to compute those effects. Indeed the $Vh(b\bar{b})$ analysis uses a final discriminant composed by the properties of a dozen of low-level variables. However, the modelling at truth level using the RIVET approach can only consider the low-level quantities. Even with the new TruthFramework tool, where such BDT variables can be computed, it is rather difficult to interpret the observed mismodelling in terms of the variations considered in the previous section. Thus the analysis has been restricted to the case of two variables: m_{bb} and $p_{\mathrm{T}}^{\mathrm{V}}$.

The final weight is evaluated using the formula in Section 5.2.1, extracted per event from the two values and two pre-computed histograms and applied to the BDT score. A few arguments lead to this choice. First, they are highly ranked in the list of used variables in the discrimination process of the MVA, therefore will have a non-negligible impact on the final BDT shape. Second, they show very little correlation in the analysis phase space which helps to avoid double counting of similar effects, while being correlated with the other variables used in the analysis. The only cross-correlated variable is the $\Delta R(b,b)$. Third, their coverage in terms of BDT effects is targeting two different regions: $p_{\mathrm{T}}^{\mathrm{V}}$ variation are mostly impacting the high MVA scores while the m_{bb} variation is impacting the low scores showing a complementary behaviour. Fourth and last, the two distributions are impacted by different sources of systematics as shown in the next paragraphs. The $p_{\mathrm{T}}^{\mathrm{V}}$ variation is also used in the 2-leptons case to assess an acceptance uncertainty between the low- and high-$p_{\mathrm{T}}^{\mathrm{V}}$ region, and thus shape and acceptance systematics are used coherently. The variations considered in the following paragraphs are similar to the ones introduced in Section 5.2.2, and unless stated otherwise no other sources are considered.

Higgs signal

For the UE/PS variation, the sources are from a tune variation of the MadGraph 5_aMC@NLO sample or the comparison between the Powheg + Pythia 8 and the Powheg + Herwig 7. It was found that the m_{bb} shape is dominated by the 2-point comparison where the shape is modelled by a second-order polynomial function, while for the $p_{\mathrm{T}}^{\mathrm{V}}$ shape, the tune variation is dominating and the highest deviation amongst all the tunes is kept. In the former case all the shapes, across lepton channels and jet bins, are found similar and a single shape is computed, while in the second the 2- and 3-jets bins are combined and derived in the various lepton channels.

For the MC scales, six variations are considered (μ_{F} =μ_{R} =0.5,2.0, μ_{F} =1. and μ_{R} =0.5,2.0 and μ_{R} =1. and μ_{F} =0.5,2.0). As for the UE/PS variation, the m_{bb} is fitted with a second-order polynomial function while a linear function is used for the $p_{\mathrm{T}}^{\mathrm{V}}$ shape. Since the distributions are similar between lepton channels, only two shapes are derived for the two jet-bins, taking the largest deviation amongst all the variations and channels.

For the PDF+α_{S}, the 30 PDF and 2 α_{S} variations from the PDF4LHC15_30 PDF set are compared to the nominal to extract the shape systematic. The same fitting scheme as for scale variation is used. The m_{bb} shape systematic is found to be negligible compared to the other shape uncertainties and therefore not included. For the $p_{\mathrm{T}}^{\mathrm{V}}$ distribution, the same method as for the scale variation is used.

For all the variations the same shape uncertainty is assigned between the g g $\rightarrow$ Zh and $q\ \bar{q} \rightarrow$ Zh, with a possible correlation between the two. Even if the distributions can be derived per regions, all the lepton and jet regions are correlated in the evaluation of the systematics. The summarised list of the systematic variations is found in Table 5.13.

systematic	0L: $Zh \to \nu\bar{\nu}b\bar{b}$ 2-jets	0L: $Zh \to \nu\bar{\nu}b\bar{b}$ 3-jets	1L: $Wh \to \ell\nu b\bar{b}$ 2-jets	1L: $Wh \to \ell\nu b\bar{b}$ 3-jets	2L: $Zh \to \ell^+\ell^- b\bar{b}$ 2-jets	2L: $Zh \to \ell^+\ell^- b\bar{b}$ $\geq$ 3-jets
$\theta_{m_{bb}}$ (UE/PS)	shape		shape		shape	
$\theta_{p_{\mathrm{T}}^{\mathrm{V}}}$ (UE/PS)	shape		shape		shape + norm	
$\theta_{m_{bb}}$ (QCD scales)	shape		shape		shape	
$\theta_{p_{\mathrm{T}}^{\mathrm{V}}}$ (QCD scales)	shape		shape		shape + norm	
$\theta_{p_{\mathrm{T}}^{\mathrm{V}}}$ (PDF+α_{S})	shape		shape		shape + norm	

Table 5.13 – Summary of all systematic uncertainties on the VH m_{bb} and $p_{\mathrm{T}}^{\mathrm{V}}$ shape modification originating from changing the UE/PS, QCD scales and PDF+α_{S} uncertainties.

Di-boson backgrounds

Due to the small contribution of the WW, and following the method applied for the normalisation systematics, no shape systematics are derived for that process.

For the UE/PS variation, the sources are from a tune variation of the Sherpa 2.2.1 sample or the comparison between the Powheg + Pythia 8 and the Powheg + Herwig 7. It was found that the $p_{\mathrm{T}}^{\mathrm{V}}$ shape deviations are well covered by the MC statistical uncertainties, hence the statistical errors are used to compute the $p_{\mathrm{T}}^{\mathrm{V}}$ shape systematics. As for the signal m_{bb} shape systematic, it was found that the 2-point comparison was dominating in the case of the di-boson process, with a more complex distribution to fit. Therefore it was decided to use directly the ratio histogram after smoothing.

A ME algorithm comparison is made by comparing the nominal Sherpa 2.2.1 with the alternative sample Powheg + Pythia 8 essentially for the final state radiation treatment that would modify the two considered distributions. The $p_{\mathrm{T}}^{\mathrm{V}}$ shapes are fitted by either a linear or third-degree polynomial function capped at 500 GeV. The m_{bb} shape is fitted with a hyperbolic tangent function in the 2-jets region to encompass a global shift of the mass peak and by a third-degree polynomial function capped at 250 GeV (range of the study at truth level) to avoid any divergences. Shapes are thus separated by di-boson process and jet regions but treated inclusively in lepton channel.

Scale variations are also considered by being doubled and halved as presented for the signal. However the deviations are well covered by the ME comparison, therefore no dedicated systematics is computed.

Even if the shapes can be derived per regions, all the lepton and jet regions are correlated in the evaluation of the systematics. The summarised list of the systematic variations is found in Table 5.14.

W+jets backgrounds

From the list of variations presented in the previous section, it was noticed that the dominant was the ME/PS variation coming from the 2-point systematic. The fitting

	0L: $ZZ \to \nu\bar{\nu}\mathrm{b\bar{b}}$		1L: $WZ \to \ell\nu\mathrm{b\bar{b}}$		2L: $ZZ \to \ell^+\ell^-\mathrm{b\bar{b}}$	
systematic	2j	3j	2j	3j	2j	≥3j
$\theta_{m_{bb}}$ (UE/PS)	shape					
$\theta_{p_{\mathrm{T}}^{\mathrm{V}}}$ (UE/PS)	shape				shape + norm	
$\theta_{m_{bb}}$ (ME)	shape					
$\theta_{p_{\mathrm{T}}^{\mathrm{V}}}$ (ME)	shape				shape + norm	

Table 5.14 – Summary of all systematic uncertainties on the di-boson m_{bb} and $p_{\mathrm{T}}^{\mathrm{V}}$ shape modification originating from changing the UE/PS and ME uncertainties.

function used in both the m_{bb} and $p_{\mathrm{T}}^{\mathrm{V}}$ case is a straight-line that is capped at 300 GeV for m_{bb}. The comparison between the shapes in the various lepton regions and jet bins showed that the various fits have the same slopes. For a more conservative approach, it was decided to use the highest variation across all the regions as a single shape. Dedicated studies have been conducted to understand residual effects, which were found to have a negligible impact on the truth closure.

Z+jets backgrounds

The 2-leptons channel has a dedicated CR for the modelling of the Z+jets, using an event-based $E_{\mathrm{T}}^{\mathrm{miss}}$sig. cut to reduce the $\mathrm{t\bar{t}}$ fraction ($E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} < 3.5\ \sqrt{GeV}$). Instead of using a second generator which could present some data-MC mismodelling and to use this difference as a systematic it was decided to use directly the nominal MC vs. data comparison. The various known MC contributions are thus subtracted from the data, and the Z+jets is scaled to the data. Three tagged regions (0-,1- and 2-tagged) are considered at the reconstructed level to assess the fit, where a blinded m_{bb} window (110 GeV $< m_{bb} <$ 140 GeV) is used in the case of the 2-tagged region. For the m_{bb} shape, the low- and high-$p_{\mathrm{T}}^{\mathrm{V}}$ regions are separated in the evaluation. A linear function is used in the m_{bb} case, whilst a logarithmic function is applied to the $p_{\mathrm{T}}^{\mathrm{V}}$ one. The systematic is applied as well in the 0-lepton channel, as the $p_{\mathrm{T}}^{\mathrm{V}}$ effects seem negligible.

A 2-point comparison with the MadGraph 5_aMC@NLO generator has also been conducted showing a variation more significant in the comparison to data. It is therefore not considered a reliable source of 2-point systematic in this analysis.

Top pair backgrounds

From all the variations presented of the $\mathrm{t\bar{t}}$ process, the 2-points systematic comparing the nominal generator to its ME variation MadGraph 5_aMC@NLO is the dominating effect observed across lepton channels as well as jet regions. A simple linear fit is conducted in both the $p_{\mathrm{T}}^{\mathrm{V}}$ and m_{bb} systematics as it is enough to parametrise the differences.

The available statistics in the 0-lepton channel is not enough to compute any values. Since the phase space and the shapes are similar to the 1-lepton case, it was decided to use a single shape for the two channels. The 2-leptons is evaluated separately and is indeed showing different behaviour. In this channel, a dedicated CR is isolating the $t\bar{t}$ processes. However, no clear sign of differences between the two regions is observed and the regions are merged. The two channels are treated as decorrelated in the fit as the correlation was not bringing any differences in the fit results.

In both the 0-/1-lepton and 2-leptons channels, the differences between njet-regions are small, and the highest variation amongst the two is kept.

Single-top backgrounds

In the single-top, the shape effects are separated by processes by distinguishing the t- and Wt- channels. The Wt channel is further split in $b\bar{b}$ or other final states. A linear fit is done for both final states, capping the m_{bb} distribution at 275 GeV. In both cases, the 2- and 3-jets regions are fitted simultaneously. The largest contribution is obtained either from the PS variation or from the Diagram Substracted and Diagram Removal ME comparison. In the fit, the $b\bar{b}$ and other final states are entering into the same NP. In the t-channel, the same procedure is considered, with a linear shape for the two distributions. The PS variation is dominating for the $p_{\mathrm{T}}^{\mathrm{V}}$ shape while the scale variation is more important for the m_{bb} one.

When considering the effect of the ME/PS variation in all the variables used in the BDT analysis for the W+jets and single-top Wt samples, a non-closure in the m_{top} distribution was found, especially at low values. This non-closure could not be covered by the $p_{\mathrm{T}}^{\mathrm{V}}$ and m_{bb} derived systematics. Therefore a dedicated systematic should have been derived to solve this problem. This clearly shows the limitation of the two 1-D shapes reweighting approach as more variables could present non-closure. Introducing new systematics to cover these effects could break some of the requirements listed in the introduction, such as the decorrelations between the variables.

5.2.4 Shapes as N-dimensional reweighting

A possible solution to the shape systematics problem is to consider a global method that could make use of all the variables entering in the final BDT discriminant and derive a weight based on this N-dimensional phase space. This idea of N-dimensional reweighting is not novel in high energy physics. One recent example can be found the discovery of CP violation in charms [173] observed with the LHCb detector. The measurement relies on the comparison of yield asymmetries of $\mathrm{D}^0 \to \mathrm{K}^- \mathrm{K}^+$ or $\pi^- \pi^+$ in the prompt tag channel $\mathrm{D}_{\mathrm{s}}^+ \to \mathrm{D}^0 \pi^+{}_s$ and in the semi-leptonic case $\overline{\mathrm{B}} \to \mathrm{D}^0 \mu^-$. The goal is to measure the asymmetry coming from the CP effects. However, other sources of asymmetries are present, such as the detection asymmetry of the by-product of the D^0 creation and the production asymmetry of the initial particle. To avoid those additional

terms, an event-based correction factor is derived based on the background subtracted distribution of three variables: the transverse momentum, the pseudo-rapidity η and the azimuthal angle Φ. The results are shown in Figure 5.16.

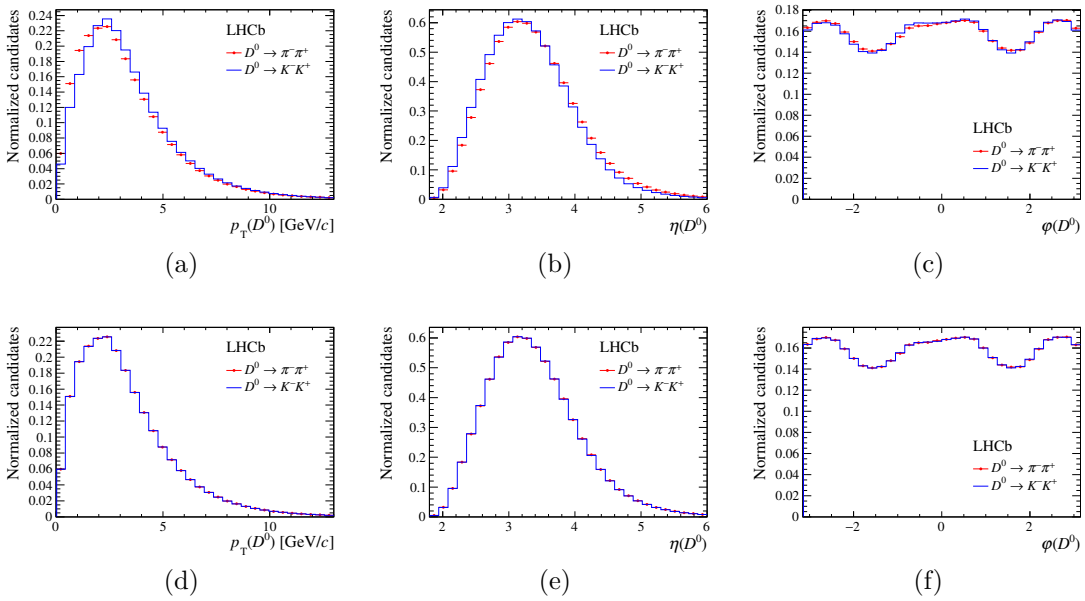


Figure 5.16 – Background subtracted data plot from the $D^0 \to \pi^- \pi^+$ (red) and $D^0 \to K^- K^+$ (blue) data selection. The three reweighting kinematic variables p_T (a, d), η (b, e) and Φ (c,f) are represented before the reweighting (top) and after (bottom) [173]

The aim of the N-dimensional reweighting is to compute the density ratio of probabilities assigned to each event from generator A and B for all the N variables x, denoted $f_B(x)/f_A(x)$, to be used as a weight in the same way as for the 1-dimensional reweighting technique. However manipulating this high-dimensional phase space is not straightforward, starting from the binning definition, and the mathematical and computational handling of such space. For instance, to get a reliable reweighting, the amount of statistics needed is exponential with respect to the number of bins per variable and the number of variables considered (the so-called "curse of dimensionality" effect in analysis theory of high-dimensional spaces). To reduce this dependency, MVA algorithms are good and efficient tools to tackle the challenge. Out of all the solutions proposed, only decision trees are considered here since they are widely used in HEP, and easier to understand. The concept behind the technique is to find a way to bin the phase space and to transform the N-dimensional problem into a 1-dimensional one. Several techniques can benefit from this approach with specific implementations described in the following paragraphs.

Classifier technique

The first usage of BDTs introduced in the Section 4.4, is to discriminate between signal and background processes. This idea can be kept but using the same physics process in signal and background, so the trees learn from the differences between the generators.

A first example of how machine learning programs can use BDTs to reweight is presented in [174]. It uses the BDT output as the purity p:

$$p = \frac{MC_A}{MC_A + MC_B}$$

where $MC_{A/B}$ is the number of events belonging to the generator A/B in the final leaf where the event falls in. From this quantity, the systematic weight can be constructed by

$$w = \frac{1-p}{p} \simeq \frac{f_B}{f_A}$$

The problem comes when boosting the decision trees and the non permutable operation between the average and the weight definition:

$$< w >= \left\langle \frac{1-p}{p} \right\rangle \neq \frac{1- < p >}{< p >}$$

In the framework of TMVA [164], this could be overcome by some rewriting of the code, or by using a second method based on ratios.

The output of the BDT would still be considered as a way to classify the events but the weight extracted is using the ratio of the BDT outputs of the two generators:

1. a BDT is trained to obtain the differences in the two MCs that will enter in the 2-point systematic evaluation. The same training parameters and variables, as in the BDT training of the $Vh(b\bar{b})$ signal/background separation found in Table 4.8 and Table 4.9, are used.
2. Using a cross-folding method, the events not used in the training are evaluated, and a ratio is computed out of the score of MC_B over MC_A.
3. the events used in the evaluation of MC_A get their systematics weight from the ratio histogram evaluated at the value of their BDT score, making a perfect closure by construction on this reweighting BDT.

This approach is further referred to a BDT-ratio or BDTr method and is illustrated in Figure 5.17.

The main interest in this technique is the perfect closure by construction when applying the reweighting on the trained BDT. The implementation of this technique is also straightforward since the BDTs are already widely used in the HEP community and are derived from standard tools compatible with the analysis framework. Even if this argument should not be the main criteria in science to decide whether or not an idea should be implemented, it eases the choice when focusing only on the performance of the method.

Since the method relies completely on a ratio, due to the limited statistics of the samples, the boundaries of the distribution of BDT scores is a problem. For the ratio to

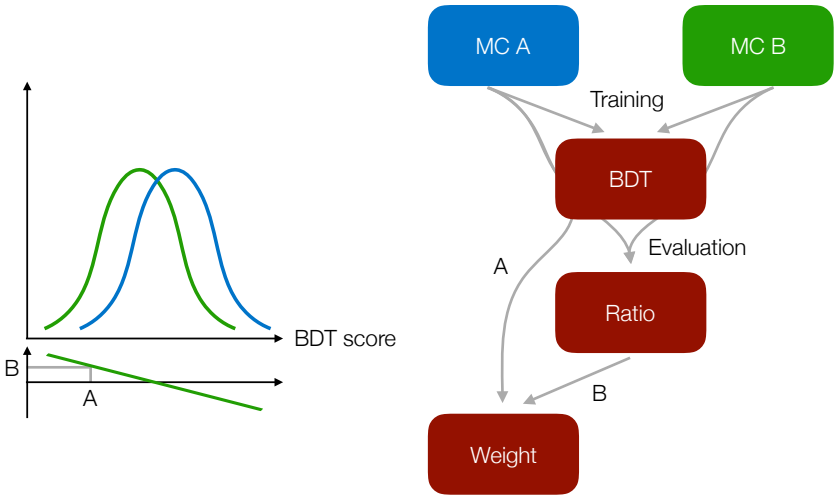


Figure 5.17 – Sketch of the procedure used to evaluate the 2-point systematic from the BDT ratio approach. After a training phase on part of the events of the two MC generators, the remaining events are used to evaluate the BDT score and compute a ratio that reflects the weight to be assigned. A represents the evaluated score of MC_A events while B the evaluated ratio at the point A.

be defined, the BDT score of MC_A must be different from zero, hence limiting the range of values for the event to be assigned a weight, since the evaluated samples may have some outliers beyond the main distribution. This method also suffers from inaccurate predictions at high weights. Those regions correspond to where the two generators disagree the most, and thus where it is easy to classify the events. Therefore those events are not going to be considered by adding more trees since the algorithm would concentrate on misclassified events.

Gradient boosted reweighting technique

To answer some of the points raised with the previous method, a novel technique developed by Alex Rogozhnikov [175] attempted to combine several ideas into a single BDT reweighter. Inspired by the bin-by-bin reweighting, the method uses a BDT to select the bins, with a boosting procedure that mixes the gradient boosting and the AdaBoost [166] methods:

1. A decision tree is built as described in Section 4.4, but with a special metric described later on.
2. In each leaf a prediction weight is computed: $pred. = \log \frac{w_{leaf,MC_B}}{w_{leaf,MC_A}}$. This corresponds to the ideal weight to be applied for the reweighting if only one tree were used.
3. Then the distributions are reweighted during the process using the predefined weight:

$$w = w \cdot \begin{cases} 1 & \text{if event comes from } MC_B \\ e^{pred.} & \text{if event comes from } MC_A \end{cases}$$

where *pred.* is the prediction weight of the leaf where the events falls in. The reweighted distribution are further on used for the training of the next tree.

The decision trees are using a dedicated splitting criteria called symmetrised χ^2:

$$\chi^2 = \sum_{leaf} \frac{(w_{leaf,MC_A} - w_{leaf,MC_B})^2}{w_{leaf,MC_A} + w_{leaf,MC_B}}$$

that will serve as the metric of the gradient boosted tree. This estimator is greedily maximised by the algorithm to find the regions where the differences are most important.

In this method, the reweighting procedure is applied at each iteration making the MC_A distributions more similar to the MC_B ones as the number of trees increases, in the same way as the misclassified events get stronger weights in the AdaBoost technique. The final sum of weights for an event is given by:

$$w_{final} = e^{pred.^{total}} \text{ where } pred.^{total} = \sum_{leaf} pred.^{leaf}$$

where the $pred.^{leaf}$ are the predictions computed at point 2 of all the leaves where the event falls in during the evaluation.

Compared to the previous method the big improvement concerns events in the outlier which are correctly treated here since there is no need for a secondary distribution to compute the weight: every event will get a score. Furthermore compared with other techniques, the BDT reweighter seems to perform more smoothly and with the need for fewer statistics [175].

The main caveat is the code implementation. This method being not widely shared at present in HEP, the technique relies on well-developed python libraries such as scikit-learn [176] and thus could benefit from the large ML communities developing those packages. However, this project is also a single-person development [177], and is, therefore, lacking in maintenance. It would be crucial for the sake of this method to be recognised by the community to implement it inside the scikit-learn framework or eventually in TMVA so it is compatible with the tools commonly used.

Example of performance with the top pair process

The goal of the two described methods is to get a weight covering the differences of the two generators and to assess the good quality of this weight the technique has to reweight the nominal MC samples and to compare it to the variation. Several metrics are then available to measure the difference between the reweighted nominal distributions and the variations.

The Pearson's χ^2 test [178] is one the oldest and straightforward estimators of the quality of the fit. This method can be reemployed in the comparison of two weighted

or unweighted histograms [179], where the null hypothesis is that the two distributions are obtained from a bin-by-bin Poissonian fluctuation of the same inputted histogram. A modified X^2 function is used as test-statistics containing the full information of the weights:

$$X^2 = \sum_{i=1}^{r} \frac{(W_1 w_{2,i} - W_2 w_{1,i})^2}{W_1^2 s_{2,i}^2 + W_2^2 s_{1,i}^2}$$

where $w_{j,i}$ is the weight of the i^{th} bin of the j^{th} distribution containing r bins which summed up give W_j having an estimated variance of $s_{j,i}$. This test-statistics is thought to have a $\chi^2_{(r-1)}$ distribution. However, it is limited by the number of bins, and more specifically by the small-weighted bins, since the $\chi^2_{(r-1)}$ distribution approximation is no longer true and the p-values are therefore less conservative.

To solve this low-statistics problem, the Kolmogorov-Smirnov test based on the cumulative distribution function instead of the probability density function in the Pearson's χ^2 test can be used. The main point from this method is that it is not based on the underlying probability distribution, making it more resilient to the global change of shapes and also takes into account cross-bin effects. The test statistic is based on the distance D:

$$D = \max_{x \in [x_{min}, x_{max}]} (|F_1(x) - F_2(x)|)$$

where $F_j(x)$ is the cumulative distribution (number of events with $x_i < x$ normalised to the total number of events) of the j^{th} MC distribution obtained in the range $[x_{min}, x_{max}]$. Contrary to the Pearson's χ^2 test which is approximated, the KS test is exact when using un-binned data. However, due to the choice of using histograms, which are more practical to store the information, it is not the case here. The test result can be influenced by the binning, biasing the outcome in case of a wrong bin choice. The required binning to avoid such effect should be compared to the scale of the physics effects, such as the object resolutions, which is believed to be the case in this study. Furthermore, since we are interested in comparing probabilities associated to the same binning, the effect is expected to cancel out.

The two methods described are still incomplete since they are only able to measure the effects on 1-dimensional distributions but don't provide combinatorial methods to assess the N-dimensional closure. Two methods are conceived to overcome that issue. The Earth Moving Distance (EMD) [180], also known in mathematics as Wasserstein metric, is a typical example of a cross-bin measurement that can be used in an N-dimensional space. This metrics can be seen as minimisation of work problem, where the work is computed from moving events of a distribution A to a distribution B. This total work is based on a flow function f_{ij} that is constrained by four conditions described in [180] and is solved by transportation algorithms studied for centuries (the first one to formalize this problem was the French mathematician Gaspard Monge in 1781[181]). It allows taking

into account the cross-bin effects. The final value is obtained by:

$$EMD = \frac{\sum_{i,j=1}^{r} d_{ij} f_{ij}}{\sum_{i,j=1}^{r} f_{ij}}$$

where d_{ij} is the distance (to be defined) between the N-dimensional bins i and j from the distribution A and the distribution B. Although this method is much more complex than the previous ones it has a natural extension toward more than 1-dimensional evaluation. The main limitation comes from the needed computational time and resources since the length of the object used in the computation is N^r. A few tests on the local cluster of the IN2P3 have shown that a maximum 3 or 4 bins for each of the 8 variables can be reached, thus preventing the test to have significant employment.

To overcome this computational limitation, one could look at the MVA side to reduce the N-dimensional problem and have a comprehensive metric. Following the idea of the BDT ratio approach, a new BDT could be trained after having reweighted the MC_A to learn any remaining differences. This idea was inspired by adversarial techniques, a growing trend in MVA, that uses a second MVA to unlearn certain characteristics from the first one. The point here is not to retrain the first MVA method, but to have a look at its properties. From this retraining, a ROC curve can be plotted showing the ability to separate the MC_A events from the MC_B ones, and a good metric for this test is the Area Under Curve (AUC). Since the ROC curve is delimited by the 0,0 and 1,1 points, the maximal value for the AUC would be 1 corresponding to a perfect ability to discriminate between the two MC samples with this BDT. On the contrary, in cases in which a BDT is not able to perform a separation, it will do a randomised choice leading to a ROC curve similar to a diagonal line (Background rejection = Signal efficiency), and thus to an AUC of 0.5. Smaller values than 0.5 can be obtained due to overfitting as illustrated in Figure 5.18.

The main limitation of this metric is the unique test procedure, i.e. for each reweighting a new metric is computed (BDT training) which raises questions of fair comparisons between two techniques. For instance, the retrained BDT could be overtrained, biassing the AUC value. Second, the AUC suffers from the lack of a unique definition of its uncertainty. A list of some possible definitions can be found in [182].

In the following paragraphs, the emphasis is put on the $t\bar{t}$ process that is evaluated in the 0-lepton channel. The 1-dimensional shape systematic uncertainty was initially dominated by the ME comparison with a similar contribution made by the PS variation. However, due to low statistics, it was decided to use the 1-lepton results. New samples were generated for this study at truth level with the following statistics:

In this study, only the BDT ratio method has been evaluated and is therefore presented since the main effort has been put on this technique. The training has been conducted separating the truth flavour of the two b-tagged jets in the same way as it is done in the single-top case because important differences have been noticed between the channels, the new statistics of the ntuples bringing some degrees of freedom. The

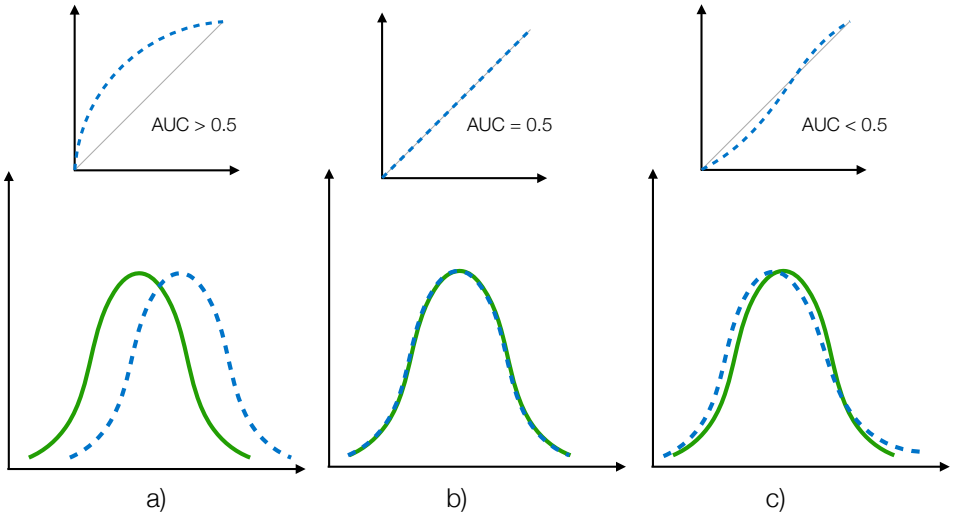


Figure 5.18 – Three toy examples to illustrate the relationship between the BDT output distribution of MC_A (shaded blue) and MC_B (plain green) in the bottom plots and the ROC curve in the upper plots (shaded blue). The random classification is shown as the grey diagonal. The three situation corresponds to a good classification (a) with AUC > 0.5, no classification (b) with AUC $= 0.5$, and an example of a bad classification (c) where the AUC is bellow 0.5 even though part of the phase space allows some good discrimination.

regions	Nominal	ME variation	PS variation
2tag2jets	190 665	94 938	122 060
2tag3jets	636 538	322 742	364 234

Table 5.15 – Statistic available (non weighted) for the $t\bar{t}$ nominal (Powheg + Pythia 8), ME (aMC@NLO + Pythia 8) and PS variation (Powheg + Herwig 7) in the 2tag2jets and 2tag3jets regions

three categories are bb, bc and others (bl, cc, cl, ll). The weighted statistics are given in Table 5.16 obtained per flavour, where the weights take into account the luminosity, cross-section and b-tagging effects.

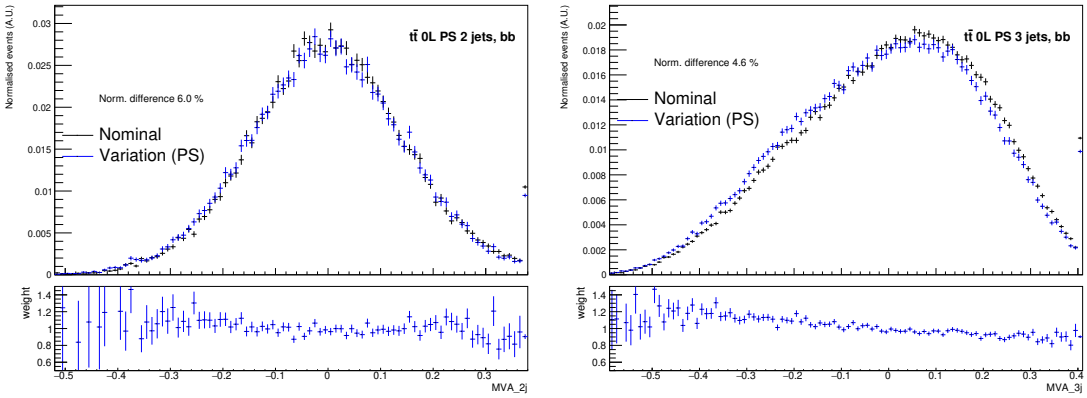
The difference of flavour fraction in the PS variation is likely to be due to differences in the b-tagging truth efficiencies. Flavour acceptance effects can be incorporated by having a normalisation effect as it is done with the $p_{\mathrm{T}}^{\mathrm{V}}$ shape systematic in the 2-leptons channel, and advocate for a separate BDT training.

The obtained BDT outputs are shown in Figure 5.19, while more plots can be found in Appendix B. While the ratio of shapes is consistent with 1 in the 2-jets case leading to small reweighting effects, a trend can be seen in the 3-jets case showing a better separation power of the technique.

Figures 5.20 and 5.21 present the truth closure for the two variables used in the 1-dimensional reweighting procedure in the 2tag3jets region for the three flavour categories.

flavour	Nominal	ME variation	PS variation
2tag2jets			
bb	21 312	17 929	20 100
bc	2 801	2 396	2 603
other	269	232	263
2tag3jets			
bb	110 590	96 241	105 689
bc	16 651	14 696	14 947
other	1 184	1 102	1 145

Table 5.16 – Statistic available (weighted) for the $t\bar{t}$ nominal (Powheg + Pythia 8), ME (aMC@NLO + Pythia 8) and PS variation (Powheg + Herwig 7) derived per flavour in the two regions



(a) (b)

Figure 5.19 – Distribution of the BDT scores for the bb flavoured events. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots. Only the PS variation is presented. On each plot the ratio between the variation and the nominal is shown on the bottom pad.

All the relevant plots can be found in the Appendix B for all the channels, flavours and jet categories.

In general, for all the distributions, the reweighting improves the χ^2 and KS scores while not reaching a perfect closure. However, all the important mismodelling are caught by the algorithm, such as m_{bb} effects for the PS variation in the flavour category labelled as other. The tables 5.17 and 5.18 gather the values for those two scores for the m_{bb} and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in all the regions. The values for all the variables can be found in Appendix B.

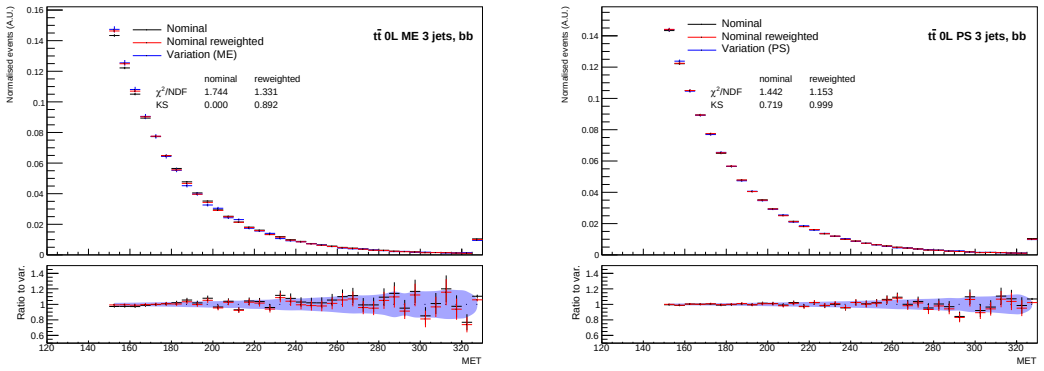
The first lesson from these tables is the evidence of complementary of the measurements of the two scores. Relatively good χ^2/NDF score such as in the ME variation 2tag3jets bb case (0.88) can yield to bad KS score (0.071), probably due to cross-bin effects. If not all the distributions in all the regions are getting good scores, a global improvement can be observed still. This is even more spectacular in the case of the in-

m_{bb}	ME variation				PS variation			
	2-jets		3-jets		2-jets		3-jets	
flavour	bfr.	after	bfr.	after	bfr.	after	bfr.	after
χ^2/NDF								
bb	1.63	1.54	0.88	0.78	1.08	0.98	1.41	1.0
bc	0.95	0.88	1.18	1.05	1.44	1.06	1.58	0.95
other	1.4	1.26	1.64	1.28	1.9	1.11	2.77	1.27
KS score								
bb	0.001	0.002	0.071	1.0	0.002	0.011	0.0	0.42
bc	0.778	0.955	0.262	0.525	0.0	0.001	0.0	0.703
other	0.001	0.036	0.0	0.033	0.0	0.845	0.0	0.671

Table 5.17 – Summary of the χ^2/NDF and KS score obtained comparing the nominal and variation m_{bb} distributions before and after the reweighting

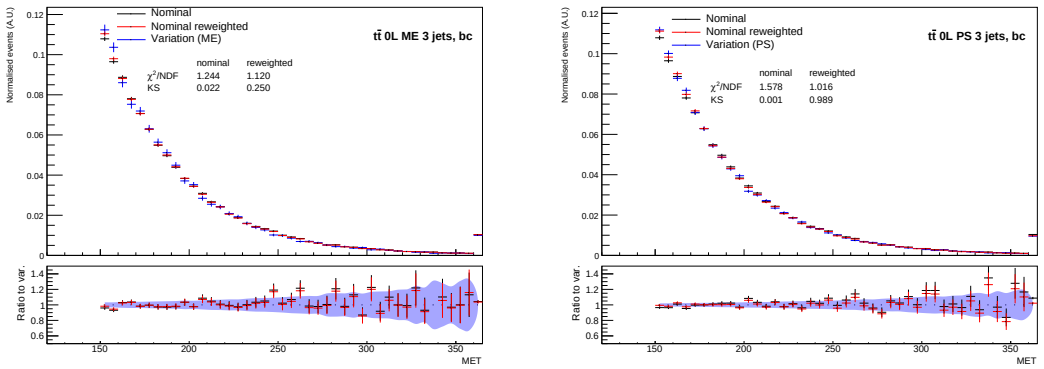
$E_{\mathrm{T}}^{\mathrm{miss}}$	ME variation				PS variation			
	2-jets		3-jets		2-jets		3-jets	
flavour	bfr.	after	bfr.	after	bfr.	after	bfr.	after
χ^2/NDF								
bb	2.1	2.05	1.74	1.33	1.58	1.52	1.44	1.15
bc	1.44	1.24	1.24	1.12	1.54	1.23	1.58	1.02
other	1.39	1.46	1.29	1.13	1.67	1.11	1.2	1.06
KS score								
bb	0.235	0.335	1.74	1.33	0.874	0.911	0.719	0.999
bc	0.001	0.031	0.022	0.25	0.085	0.518	0.001	0.989
other	0.0	0.075	0.0	0.523	0.003	0.523	0.553	1.0

Table 5.18 – Summary of the χ^2/NDF and KS score obtained comparing the nominal and variation $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution before and after the reweighting



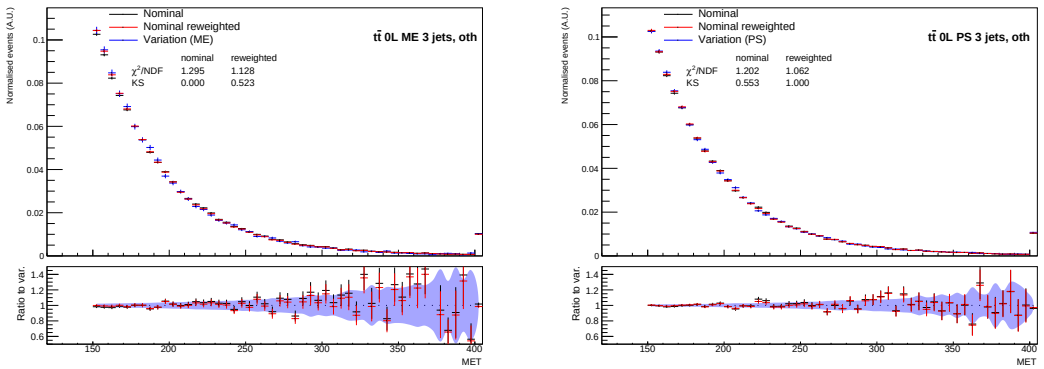
(a)

(b)



(c)

(d)



(e)

(f)

Figure 5.20 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level for the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions ($p_{\mathrm{T}}^{\mathrm{V}}$ in 0-lepton channel). The bottom part of the plots shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The distributions are plotted in the bb (top), bc (middle) and other (bottom) flavours of the two b-tagged jets in the 2tag3jets region. Considered variations are the ME (left), and the PS (right).

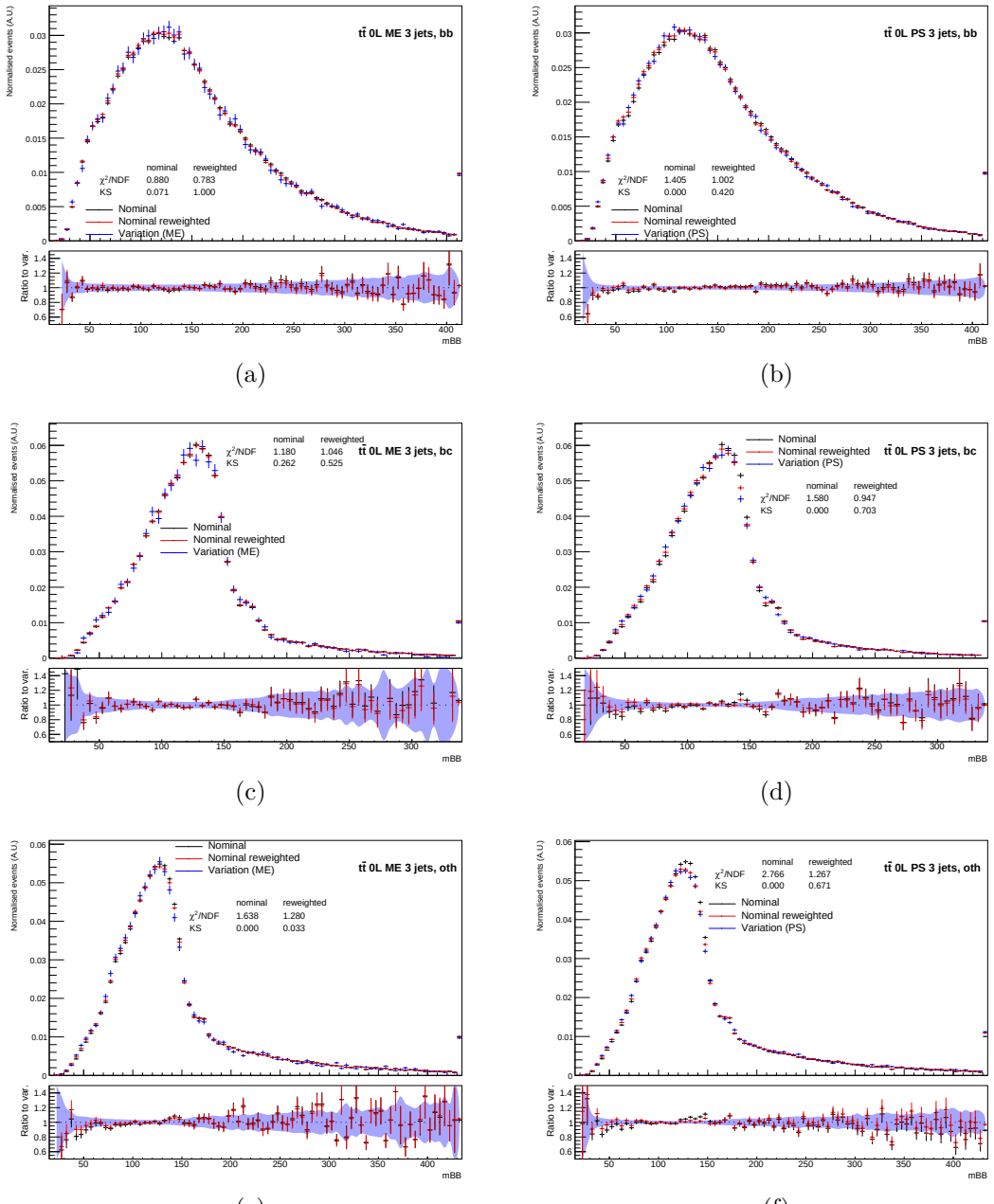
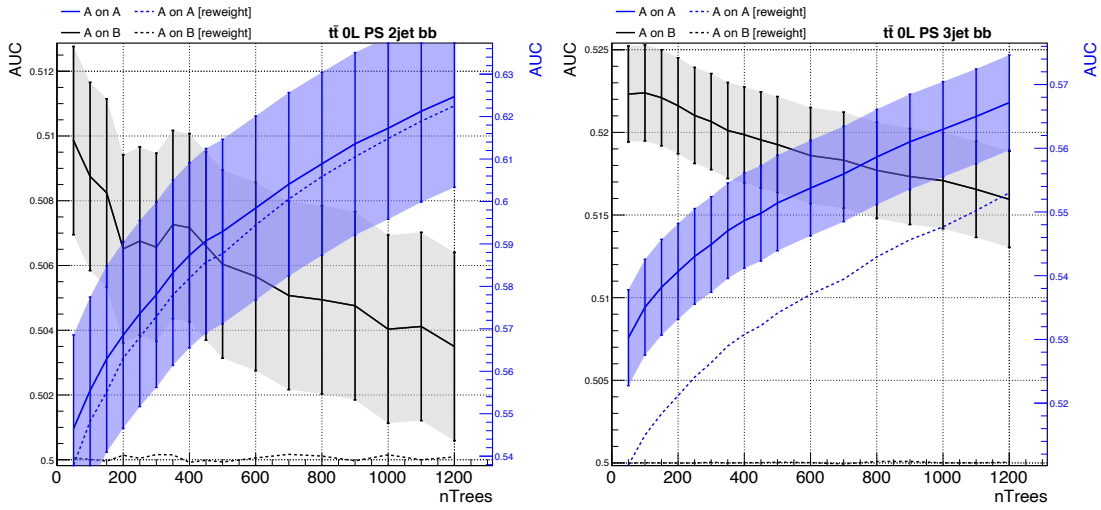


Figure 5.21 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level for the m_{bb} distributions. The bottom part of the plots shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The distributions are plotted in the bb (top), bc (middle) and other (bottom) flavours of the two b-tagged jets in the 2tag3jets region. Considered variations are the ME (left), and the PS (right).

variant mass of the 3jets in the 2tag3jets case. In the "bc" and "oth" categories, there is a high chance that the 2-tagged jets are coming from one of the top decay (with a hadronic decay of the W boson). So the reconstructed invariant mass is peaking at the top mass. Important mismodelling is observed in Figure 5.23, but is nicely accounted for by the BDT reweighting procedure.

Since most of the non-closures are covered by the method, the optimisation of the BDT hyper-parameters such as the number of trees is conducted. Adding more trees should improve the discrimination but could also lead to over-training and consequently, this eventuality should be checked carefully. The folding strategy already allows checking for this feature, since the BDT reweighting can also be applied to the events used in the training and not only to the folded events. Since two folds are used in the training the evaluation can be called A-on-B (usual folding) or A-on-A (unfolding method). The metric used to assess the overtraining is the AUC of the BDT score while varying the number of trees. By construction, the reweighted BDT distribution of the nominal should match exactly the one from the variation leading to an AUC of 0.5, while the non-reweighted AUC should increase with the number of trees up to the point where no new feature is learned. In the A-on-A case, the AUC is expected to increase in the same way as the A-on-B case up to the point where the overtraining effect arises, leading to a linear increase with respect to the number of trees. The reweighted distribution, however, is using the weight from the folded method and therefore is not expected to be at 0.5. Figure 5.22 is showing the evolution in case of the bb flavour events considering the PS variation. More plots can be found in Appendix B.

It appears that the expected behaviour of the A-on-B reweighted AUC is well observed with some small statistical fluctuations around 0.5. However, the A-on-A non-reweighted AUC is decreasing with the number of trees, illustrating the poor statistical power of the discrimination with the samples used in the training. Indeed by the addition of more trees, the events are simply moving from the well-classified category to the mis-classified one, making the distributions of BDT scores less separated. The unfolded AUC is showing the expected rise, illustrating the effect of the over-training. The difference between the reweighted and non-reweighted AUC is shown to be roughly conserved between the A-on-A and A-on-B evaluations since the reweighting power is the same between the two methods as they use the same ratio. All in all the setup used for the hyper-parameters seems reasonable in terms of performance and over-training.



(a) (b)

Figure 5.22 – Evolution of the AUC for the bb flavoured events with respect to the number of trees used in the training. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots. Only the PS variation is presented. The right (left) axis is presenting the AonA (AonB) scale. Reweighted AUC are shown in dashed lines while the non reweighted AUCs are in solid lines. Errors are shown with light areas for the non reweighted quantities following the recommendations of [182].

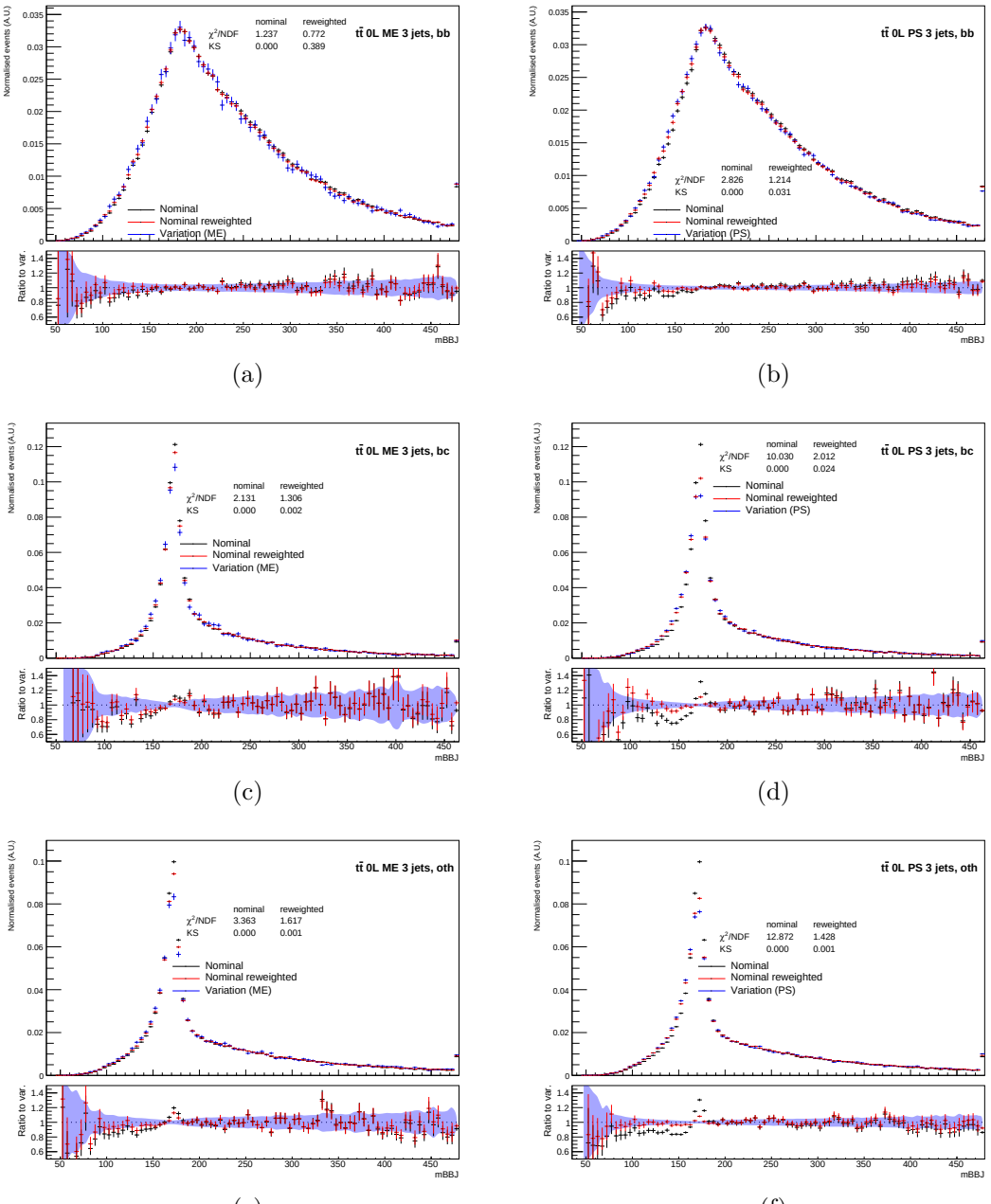


Figure 5.23 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level for the m_{bbJ} distributions. The bottom part of the plots shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with tthe light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The distributions are plotted in the bb (top), bc (middle) and other (bottom) flavours of the two b-tagged jets in the 2tag3jets region. Considered variations are the ME (left), and the PS (right).

6 Statistical analysis and results

The goal of this analysis is to answer the formal question: is the Higgs boson decaying into a pair of b-quarks in the way predicted by the Standard Model? This question is structured in two parts: is the Higgs boson coupled to b-quarks, and is the rate of production and decay observed in the data the same as predicted by theory. The answer is eminently linked to probabilities and should be treated with the appropriate statistical tools: how likely is the observed data consistent with the background+signal hypothesis. These tools will be described in this Section, together with the results obtained. An introduction to the Profiled Likelihood Ratio method is given in Section 6.1. Section 6.2 gives a global description of the systematic uncertainties. Two sets of results are then detailed in Section 6.3 and Section 6.4, the former being the result obtained in the nominal analysis and published in [21], the latter showing the effect of changing the anti-QCD cuts as described in Section 5.1.1.

6.1 The statistical analysis: The Profiled Likelihood Ratio fit

At the LHC the statistical framework used for the discoveries has been chosen to be fully frequentist contrary to the mixed Bayesian-frequentist approach used at Tevatron. Even if the two methods are different in their philosophy, the same quantity is measured

from the data:

$$\mu = \frac{\sigma \cdot BR}{\sigma_{SM} \cdot BR_{SM}}$$

where σ is the considered cross-section of production and BR the decay branching ratio. This quantity, called signal-strength, represents a scale factor, that would be 0 if no signal is observed and 1 if the observed amount of signal as predicted by a certain theory, in our case the Standard Model is observed. The procedure used for its measurement relies on the computation of Profiled Likelihood, which is used in a Profiled Likelihood Ratio as a test-statistics.

6.1.1 The Profiled Likelihood

For each possible value of the signal strength, the level of compatibility between this hypothesis and the measured dataset is assessed. The formalism introduces a Likelihood function $\mathcal{L}$ that is going to take into account the specifics of the analysis. The simplest definition that would correspond to a cut and count experiment is to use the Poissonian probability $\mathcal{P}$:

$$\mathcal{L}_{pois.}(\mu) = \mathcal{P}(n|\mu \cdot s + b) = \frac{(\mu \cdot s + b)^n}{n!} e^{-(\mu \cdot s + b)}$$

where s is the predicted amount of signal events by the model, b the predicted amount of background events, and n the number of observed events.

This rather simple expression can then be extended to the case of a discriminant based analysis, which can be considered as a multiple cut and count experiment:

$$\begin{aligned} \mathcal{L}_{disc.}(\mu) &= \prod_{i=1}^{N_{bins}} \mathcal{L}_{pois.,i}(\mu) \\ &= \prod_{i=1}^{N_{bins}} \mathcal{P}(n_i|\mu \cdot s_i + b_i) \end{aligned}$$

As described in Chapter 4, the $Vh(b\bar{b})$ analysis is further split into several regions (leptons, jets, SR/CR) which can share the same discriminating variable or not, but are going to be treated with the same formalism:

$$\begin{aligned} \mathcal{L}_{regions}(\mu) &= \prod_{j=1}^{N_{SR}} \mathcal{L}_{disc.,j}(\mu) \prod_{j=1}^{N_{CR}} \mathcal{L}_{disc.,j}(\mu) \\ &= \prod_{j=1}^{N_{SR}} \prod_{i=1}^{N^j_{bins}} \mathcal{P}(n^j_i|\mu \cdot s^j_i + b^j_i) \cdot \prod_{l=1}^{N_{CR}} \prod_{k=1}^{N^l_{bins}} \mathcal{P}(n^j_k|b^l_k) \end{aligned}$$

This basic expression, however, does not account for all the uncertainties that are impacting the measured data as well as the MC predictions of the signal and background yields as shown in Section 5.2. The inclusion of those parameters is well treated inside this formalism through the introduction of Nuisance Parameters (NPs), as described in Section 5.2.1. They are going to affect the number of signal events (in which case they

are denoted by the parameters θ_s), background events (in which case they are denoted by the parameters θ_b), or both (in which case they are denoted by the parameters θ_a). NPs are used to give the fit some freedom to adjust the number of events while penalising modifications using subsidiary measurements $f(\theta_s, \theta_b, \theta_a)$ that profile-out the knowledge of the NPs from the signal strength measurement. All this leads to the final Profile Likelihood expression:

$$\begin{aligned}\mathcal{L}_{profiled}(\mu,\theta) = \quad & \prod_{j=1}^{N_{SR}} \prod_{i=1}^{N^j_{bins}} \mathcal{P}(n_i^j | \mu \cdot s_i^j(\theta_s^{i,j}, \theta_a^{i,j}) + b_i^j(\theta_b^{i,j}, \theta_a^{i,j})) \\ & \cdot \prod_{l=1}^{N_{CR}} \prod_{k=1}^{N^l_{bins}} \mathcal{P}(n_k^j | b_k^l(\theta_b^{k,l}, \theta_a^{k,l})) \\ & \cdot f(\theta_s, \theta_b, \theta_a)\end{aligned}$$

The subsidiary measurements are the priors that can be inserted by the user in the statistical process which is similar to the Bayesian philosophy of statistics. The usual functions used in the process are log-normal or Gaussian distributions, the former having the advantage of always being positively defined. The width of the probability density functions (pdf) governs the characterisation of a deviation, with central value 0 (or 1 in the case of normalisation NPs). The determination of the various constraints is treated in Section 6.2. In a few cases, nuisance parameters do not have such subsidiary measurements and are therefore said to be free to float in the fit.

Due to the high number of NPs, it is important to assess their relations not only to the signal and background quantities but also amongst each other. Two NPs are said to be correlated if their variations are coherent. The correlations are effectively decreasing the number of degrees of freedom in the Likelihood. Dedicated studies are conducted to understand the correlation scheme, allowing a-posteriori modifications by the analysers. For instance, for important backgrounds, the normalisations NPs can be separated into sub-components that are then uncorrelated.

The last step is to consider the statistical uncertainties, that are incorporated as Poissonian distributions with dedicated NPs described later on.

6.1.2 The Profiled Likelihood Ratio as a test-statistics

Once the likelihood is built, the value of the signal strength can be extracted by maximising its logarithm with respect to all the NPs, getting the so-called fitted signal strength $\hat{\mu}$.
However, this estimator of the central value of the true signal strength does not reflect the knowledge obtained on the "significance" of the measurement. This result is obtained when comparing the hypothesis assuming the searched signal H_0 and the background-only hypothesis H_1. The term of significance (Z) in the frequentist statistic is linked to the p-value (p) defined as the probability, under the assumptions of the hypothesis H_0, of finding data of equal or larger incompatibility with the predictions of H_0:

$$Z = G^{-1}(1-p)$$

where G^{-1} is the inverse of the Gaussian function, providing a direct mapping between the p-value and the number of standard deviations from the Gaussian distribution as seen from Figure 6.1. This final significance is one of the key ingredients in searches for claiming a discovery or an observation. Two thresholds are historically defined in HEP for claims: the 1.35×10^{-2} % p-value level, corresponding to Z=3, is associated to an evidence while the 2.87×10^{-5} % p-value level, corresponding to Z=5, is associated to an observation. However, the level of significance is not always sufficient by itself to get credits from the community. For instance, the supraluminic neutrinos announcement was observed with a Z=6.0 significance [183], and more recently the DAMA/LIBRA experiment reported a Z=12.9 significance on a dark matter observation and annual oscillation. However, the first result got explained by a bad connection between a fibre-optic cable that connects to the GPS receiver used to correct the timing of the neutrinos' flight and an electronic card in a computer [184], and was corrected by the collaboration when published. In the second case several experiments based on similar and other detection methods ruled out the published result, and some theoretical and instrumental explanations have been issued [185]. These two examples, tend to show that the level of trust is also based on the social belief in the community upon the produced result, such as the reproducibility of the result by independent experiments. Other important criteria are, among others, related to the good comprehension of its detector and the modelling of the background, aspects which are fundamental to make a strong assessment.

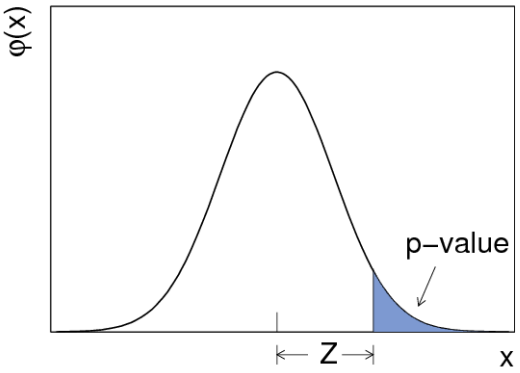


Figure 6.1 – The standard Gaussian distribution $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ showing the relation between the significance Z and the p-value p [186].

The p-value in the case of the two hypothesis described before is computed from a test-statistics called the Profiled Likelihood Ratio q_μ computed as:

$$q_\mu = -2 \log \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}$$

Here $\hat{\hat{\theta}}(\mu)$ denotes the conditional Maximum-Likelihood (ML) estimator of θ for the given value μ. The denominator is obtained with the unconditional maximising of the Likelihood given by the ML estimators $\hat{\theta}$ and $\hat{\mu}$. In the case of discovery searches, such as in the

$Vh(b\bar{b})$ analysis, the goal is to reject the background only $\mu = 0$ hypothesis. Hence the relevant test-statics is defined as $q_0 = \lim_{\mu \to 0} q_\mu$. This value is capped at 0 for negative $\hat{\mu}$ since the presence of a signal is only expected to give an enhancement of observed events. This said, a negative $\hat{\mu}$ value would indicates a disagreement with the background-only hypothesis while not agreeing with the initial signal hypothesis, pointing rather to systematic error or statical fluctuations. The p-value p can be directly extracted from this PLR value:

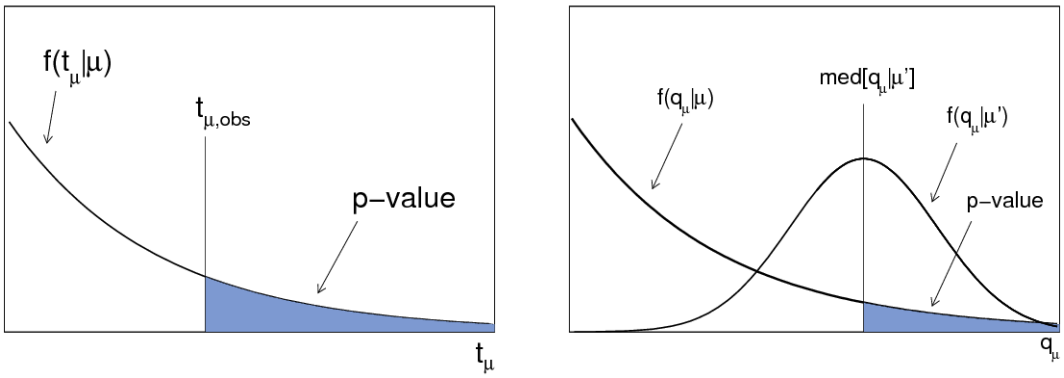
$$p = \int_{q_0}^{\infty} f(q|0)dq$$

where $f(q|0)$ is the probability density function of the test-statistics. This function can be computed using the Wald approximation [186, 187]:

$$f(q|0) = \frac{1}{2}\delta(q) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q}}e^{-q/2}$$

where $\delta(q)$ denotes the Dirac distribution at 0, and where the second term represents a chi-square distribution for one degree of freedom. This function is represented in Figure 6.2a. The integration allows then to get a simple relationship between the observed p-value and the significance:

$$Z = \sqrt{q_0}$$



(a) (b)

Figure 6.2 – Illustration of the link between the probability density function of the test-statistics, the measured (a) or the median (b) PLR and the p-value [186].

6.1.3 The Asimov dataset and expected significance

The statistical treatment of data leads to the extraction of the final results and this is generally the last step in the analysis. Indeed, since the observed significance can lead to important decisions, such as publishing groundbreaking results as mentioned earlier, the analyser should remain as unbiased as possible when designing and improving his analysis. Therefore, to get the most rigorous result a blinded procedure is applied.

This procedure states that the full chain of analysis should be understood and frozen before looking at data results in regions where a signal might be found. However, tuning of the fit procedure requires several studies, often based on data. In some cases the data sample is not large enough to guarantee the validity of the extracted results. All these reasons push to build an artificial dataset from the MC prediction and is called Asimov data-set [186].

This data-set must be constructed in such a way that when one uses it to evaluate the estimators for all parameters, the true parameter values are obtained :

$$\begin{cases} n_i^j & = \mu' \cdot s_i^j + b_i^j \\ m_k^l & = b_k^l \end{cases}$$

where μ' is the tested signal strength corresponding formally to the hypothesis H_0, hence $\mu' = 1$ here. The term Asimov was coined due to its predictive power that is based on a single entity representing the global sum of processes, such as in one of Asimov's short stories [188].

The strength of this procedure is to give a natural estimator of what would be the significance of the analysis to reject H_1 given the predicted signal and background, where, by construction, all the statistical fluctuations are suppressed. Since the significance function is monotonic under the p-value(p), and thus under the test-statistics q_0, computing the median of the significance is equivalent to computing the median of q_0:

$$\begin{aligned} med(Z|\mu' = 1) &= med(g(p)|\mu' = 1) \\ &= med(g(h(q_0))|\mu' = 1) \\ &= g(h(\underbrace{med(q_0|\mu' = 1)}_{q_{0,A}})) \end{aligned}$$

where g, h are the monotonic functions to go from the test-statistics to the significance, and $q_{0,A}$ the test-statistics obtained from the Asimov data-set as shown in Figure 6.2b. In the case of the Wald approximation we have $g(h(x)) = \sqrt{x}$.

6.1.4 Post-fit inquiries

Once the fit is performed based on the real or the Asimov data-set, any features need to be understood before looking at the significance. All the parameters are allowed to evolve in the procedure, and thus the analyser must check that the values obtained are in agreement with what was predicted, i.e. the prior set on the NPs.

The NPs parameters are given as input in the Likelihood with subsidiary measurements, which present two parameters: the central value, and the width of the distribution. The actual effect of the fit will be to adapt these values with the Maximum Log Likelihood procedure, resulting in two effects:

— constraints: sometimes the fit can reduce the width of the prior since the data can constraint it much better than the prediction. This could be caused by a mismatch between the NPs needed by the fit and the inputed ones, especially for the experimental NPs where a reduction scheme is applied on the number of NPs for time computation reasons. On the analysis specific quantities, it is the sign that the first computation may have been too conservative, or that the parametrisation is not good enough and that more degrees of freedom must be provided. In some rare cases, the systematics are under-constrained, showing some pathological problems of the fit itself.

— pulls: The central values are often shifted away from their initial value for good reasons. These pulls should be compared with the resulting constrained widths and the initial values/widths. For example, a NP with post-fit values of 0.4 ± 0.1 is still highly compatible with the initial guess that is $0. \pm 1.0$. However, the NPs should be checked if the central value is subsequently shifted away from zero without the width being constrained.

These checks must be conducted with the Asimov and the real data-set. In the first case the pulls must be aligned with their initial values. Some constraints can appear, and should be correlated with similar phenomena when fitting with the data. A detailed comparison of the pulls and constraints between the two data-sets is therefore crucial.

As we have more than one parameter acting on the results, they might share correlations reflected in the profiled likelihood. These effects can be physically genuine for instance in the case of the background normalisations when the factors are separated between regions but are still evolving coherently. The global effect of the correlation is to reduce the effective number of degrees of freedom in the fit, making the interpretation of results less straight-forward. For instance a pull or a constraint observed for one NP, could in fact be due to a second NP that is correlated to the first one. The decorrelation of parameters can be invoked in the case when a NP is applied in different regions undergoing the effects of independent constraints in each of these. In practice the correlation matrix is computed by the inversion of the Hessian matrix:

$$Corr = \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \mu^2} & \frac{\partial^2 \mathcal{L}}{\partial \mu \partial \theta_1} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial \mu \partial \theta_n} \\ \frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \mu} & \frac{\partial^2 \mathcal{L}}{\partial \theta_1^2} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{L}}{\partial \theta_n \partial \mu} & \frac{\partial^2 \mathcal{L}}{\partial \theta_n \partial \theta_1} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial \theta_n^2} \end{bmatrix}^{-1}$$

The next relevant point to be investigated is the impact of the introduced NP on the signal strength error. A global approach is conducted by looking at the profiled likelihood with respect to the signal strength μ near to the best-fit value. For each

value, the NPs are fitted again leading to the new likelihood $\mathcal{L}(\mu,\theta)$. A one standard deviation uncertainty is assigned corresponding to the test-statistics $q_\mu = 2\cdot(\log(\mathcal{L}(\hat{\mu},\hat{\theta})) - \log(\mathcal{L}(\mu,\hat{\hat{\theta}}(\mu))))$ variation by one unit. However, this value is inclusive in the various effects of the NPs. Two different tests are then conducted to investigate the effect of individual NP:

— **Breakdown method:** NPs are merged in categories to evaluate the impact of each group. The method used in the $Vh(b\bar{b})$ analysis is considering the difference in quadrature of the individual deviations and the fit done by fixing all the NPs in the category to their initial values. The advantage of this method is that the result is closer to the impact of an individual NP when approaching the limit of one NP per category. However, the sum in quadrature of all the effects will not result in the full uncertainty because of the correlations. A different option is introduced in Section 7.1, following the recommendations of the Higgs combination group as well as the CMS methodology.

— **Ranking method:** Each NP, one at the time, is scaled by its post-fit $\pm1\sigma$ variation and the signal strength best-fit value is recomputed, all the other NPs being allowed to float from their initial post-fit value to take into account correlations. The difference of signal strength $\Delta\mu^{\pm} = \mu_{\pm1\sigma} - \hat{\mu}$ for each NP is ranked by importance, and plotted. Due to the correlation effects, the positive and negative variations may not have the same magnitude resulting in asymmetrical representation.

The last element to be checked is the distributions used to perform the fit. The post-fit distributions are plotted using the best-fit values of the signal strength $\hat{\mu}$ as well as the NPs $\hat{\theta}$. The variables used in the BDT training are also plotted using the best-fit values of the NP, except for the MC statistical ones, which is hard to interpolate between different distributions. The data over MC ratio obtained is thus a good sign of the fit convergence and possible mismodelling.

6.2 The nuisance parameters

The previously introduced NPs can be classified according to their origins: the statistical uncertainties linked to the finite number of events, the experimental systematics arising from the known biases in reconstruction and identification as presented in the

Chapter 3 and the modelling effects already presented in Chapter 5. Finally, to reduce the number of degrees of freedom, a reduction and smoothing scheme of the NPs is presented.

6.2.1 The statistical uncertainties

The limited statistics of the MC generators is taken into account as a separate term of the likelihood following the recommendations of Beeston–Barlow technique [189]. The full approach would require to consider as many parameters as the number of bins times the number of processes, and could be minimised using parallel computation. However, a lighter approach, called Beeston–Barlow lite, considering all the backgrounds morphed into one component has led to a simpler calculation without loosing any precision in all the tests conducted so far. The resulting likelihood term can be written as Poisson form [190]:

$$\mathcal{L}_{stat}(\gamma) = \prod_{i=1}^{N_{bins}} \mathcal{P}(m_i(\theta_s^i, \theta_b^i, \theta_a^i)|\gamma_i\tau_i)$$

where the γs are the NPs associated to the statistical fluctuations of the MC samples, and $\tau_i = (m_i(\theta_s^i, \theta_b^i, \theta_a^i)/\delta_i)^2$ is the Poissonian predicted central value from the total statistical uncertainty δ_i of the bin i. The second effect in the likelihood is a bin-by-bin global reweighting:

$$\mathcal{P}(n_i^j|\mu \cdot s_i^j(\theta_s^{i,j}, \theta_a^{i,j}) + b_i^j(\theta_b^{i,j}, \theta_a^{i,j})) \rightarrow \mathcal{P}(n_i^j|\gamma_i \cdot (\mu \cdot s_i^j(\theta_s^{i,j}, \theta_a^{i,j}) + b_i^j(\theta_b^{i,j}, \theta_a^{i,j})))$$

This method is shown to give a significant improvement in case the number of MC events is lower than 10 times the number of data events per bin. In this case, to limit the impact of the low MC statistics, the BDT distribution is binned with an adaptive and iterative binning procedure, which would allow getting an optimal S/B ratio while maintaining a reasonable MC statistical uncertainty. The algorithm uses the Z function defined as:

$$Z = z_s n_s/N_s + z_b n_b/N_b$$

where $n_{s,b}$ is the number of signal/background events in the bin with respect to the total $N_{s,b}$ number of signal/background events while $z_{s,b}$ are two parameters which will be optimised by applying the following procedure:

1. aggregate the bins from the right to compute the value of the Z function.
2. if the Z function is above a threshold (here 1) and the MC statistical uncertainty is below 20%, rebin all the selected bins into a single bin.
3. repeat the same procedure until all the bins are grouped.

The values of $z_{s,b}$ were chosen to be (10,5) in the case of the $Vh(b\bar{b})$ BDT and (5,5) in the case of the di-boson cross-check due to the lower statistics of the signal MC sample resulting in 15 bins in the former case and 10 bins in the latter.

6.2.2 The experimental systematics

The measurements and calibrations of the objects described in Section 3.2 are associated with uncertainties given by the dedicated Performance Working Group. These uncertainties are computed per event in both the data and the MC samples and propagated to the fit through histogram variations with Gaussian priors. The following categories are used in the $Vh(b\bar{b})$ analysis.

Luminosity and pile-up

The luminosity is used to weight the MC samples to get the correct final number of events. The uncertainty on this value provided for 2015, 2016 and 2017 is respectively 2.1%, 2.6%, and 2.4%. A combined measurement is obtained using a method similar to [191] , from calibrations of the luminosity scale using x–y beam-separation scans performed in August 2015, May 2016 and July 2017, and yields a mean uncertainty of 2.0%. This value is applied as the width of a Gaussian prior in the fit.
The pile-up emulated in the MC samples is scaled up by a 1.03 factor to account for data/MC differences. By comparison, the MC samples used in previous studies with A2 tunes were reweighted by a 1.09 factor. This factor is varied with an up/down values of 1.00/1.18.

$E_{\mathbf{T}}^{\mathbf{miss}}$

The uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ come from various sources and can be classified into two categories, the trigger and the object related effects. Scale factors have been derived to account for the trigger turn-on curve and related uncertainties are obtained from that method. Three NPs are used to account for the statistical variation of the fit, the MC process used in the derivation, and the kinematical dependence in terms of the offline scalar sum of the p_{T} of the jets.
The second category is concerning the resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and its objects. The value is recomputed first while varying the soft-terms or hard-terms by $\pm 1\sigma$ of their resolution. Two resolution terms for the soft-terms coming from comparisons between data and various MC samples are also used. These resolutions use the perpendicular and parallel projections of the soft-terms calculated and binned as a function of the hard-objects p_{T}.

Leptons

For both the muons and the electrons three sets of systematics are considered. First, since these objects are entering at the trigger level in the 1- and 2-leptons cases, the trigger scale factors used to correct for data/MC differences are varied both for statistical and systematic sources, resulting in a global NP for electrons and two NPs for the muons. Second, the identification efficiency variations are used as set by the calibrations prescriptions. While the electrons only use three NPs for the identification, reconstruction

and isolation, the muons use eight variations dividing the NPs between statistical and systematical variation, as well as low- and high-p_{T} for the identification due to the differences of calibration technique used to extract the scale factors. Last, the object energy scale and resolution are used from the energy calibration. The electron scale is not the dominant source of uncertainty in the analysis, and thus all the variations provided (more than sixty) are classified into two NPs corresponding to scale and energy resolution. For the muons, the p_{T} resolution is separated into two components (inner-detector and muon system) while one NP is set for the momentum scale.

Jets

The calibration of the jet energy scale described in Section 3.2.2.2 requires a lot of different steps and various uncertainties can be applied: in-situ corrections, eta intercalibration, high-p_{T} jets, pile-up, flavour composition, flavour response, b-jets, and punch-through jets. This baseline uses roughly 50 parameters to allow for a full combination with other analyses. However, a reduced scheme of 23 parameters is also proposed. The reduction is undertaken through a principal component analysis, combining the resulting NPs to preserve correlations.
An uncertainty is also set on the jet energy resolution derived from data/MC comparison.

b-tagging

In order to correct for the data/MC discrepancies in flavour tagging, scale factors are derived, which have corresponding systematics variations. These scale factors are separated into different flavours (b, c, light and τ), and around 40 variations are considered for each. Those variations arise from experimental (e.g. JES modifications since the factors are binned in p_{T}), theoretical (e.g. p_{T} spectrum of $\mathrm{t\bar{t}}$ events in the calibration) and statistical uncertainties. Using the same approach as for the jet energy scale, a reduced scheme is created separated between b, c and light with τs being accounted in the last one. The remaining eleven NPs are completed by a high-p_{T} extrapolation NP to account for the limited p_{T} range of the calibration and a b-to-c extrapolation NP.

6.2.3 The multi-jet background modelling uncertainties

In the 1-lepton channel, a template is used to determine the shape of the QCD multi-jet background in the SR (see Section 5.1.2). Several sources of systematic uncertainties are derived to encompass the hypotheses made for this estimation, both on the shape and normalisation of the multi-jet distributions.

Shape Uncertainties

In order to evaluate the shape uncertainty of the MJ background estimate, several shape systematic uncertainties are considered:

— **Lepton trigger**: the impact of the choice of lepton trigger on the MJ estimate is evaluated, as this may introduce a bias in the inverted isolation region. This systematic affects only the electron channel, since only the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used in the muon channel. Instead of using the combination of triggers, listed in Table 4.1, simply the lowest p_{T} trigger is used.

— **Isolation requirements**: an evaluation of the uncertainty introduced by the isolation criteria to define the MJ-enriched CR is considered. A reduced inverted-isolation region is defined, with additional isolation cuts applied. In the electron channel, this is defined requiring in addition TopoEtCone20 $<$ 12 GeV , and in the muon channel, PtCone20 $<$ 2.9 GeV. The additional cuts are optimised for preserving about half of data events in the QCD enriched-CR retaining the fraction that is closer to the signal region and thus is subject to a smaller extrapolation uncertainty.

— **EW normalisation factors**: the impact of using the normalisation factors extracted in the SR m_T^W fit for the Top and W+jets processes for the EW background subtraction procedure in the inverted isolation region is assessed. The nominal MJ template shape is evaluated with and without applying the normalisation factors and the difference in shape taken as the systematic uncertainty.

These uncertainties are implemented as shape-only systematics by normalising the variation to the nominal MJ yield. Plots in Figure 6.3 show the shape comparison for the nominal BDT and the main shape systematics variations for both electron and muon channels in the 2tag2jets region.

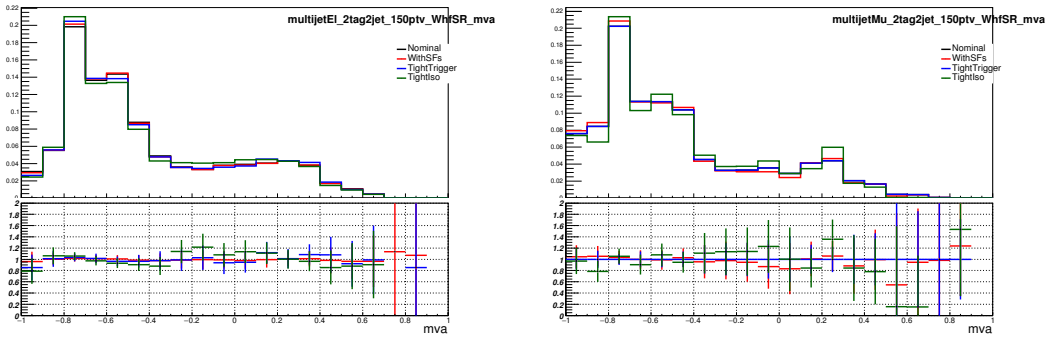


Figure 6.3 – The MJ BDT shape comparison for the nominal (in black) and some main shape variations in the 2tag2jets region, electron channel (left) and muon channel (right). The histograms in blue indicate the impact of using the lowest lepton p_{T} trigger (Lepton trigger systematics), the green histograms indicate the impact of using the reduced inverted isolation region (Isolation requirements systematics), the red histograms indicate the impact of using the Top and W+jets normalisation factors in the inverted isolation region (EW normalisation factors systematics).

Normalisation Uncertainty

The sources of systematic uncertainties that have an impact on the multi-jet distributions are also considered to derive an uncertainty on the estimated multi-jet normalisation. The impact on the multi-jet normalisation is indirectly driven by changes to the m_T^W template distributions, and to the relative yield in the signal and W+hf control regions. The individual contributions are added in quadrature to give the overall normalisation uncertainty, separately in the 2 and 3 jet regions, and for the electron and muon channels.

In addition to the same sources of uncertainties considered for the shape and previously described, a few more are added exclusively for the normalisation uncertainty:

— $E_{\mathbf{T}}^{\mathbf{miss}}$ **requirements**: an estimates of the uncertainty introduced by the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement in the electron channel selection is calculated. The $E_{\mathrm{T}}^{\mathrm{miss}}$ < 30 GeV region is included in the template fit, which induces a change to the m_T^W distribution both for the derived multi-jet component and for the MC EW background components.

— m_T^W **fit**: an estimate of the uncertainty introduced by fitting on the m_T^W variable is considered. An alternative distribution ($\Delta\Phi(l,h)$, where l is the selected lepton and h the reconstructed Higgs boson) is used in the template fit. This variable is selected due to the good discrimination between MJ and EW background.

All in all the fractions of the multi-jet contribution compared to the total background and their uncertainties are presented in Table 6.1.

Region	channel	MJ Fractions (%)	MJ norm. uncertainty
2tag2jets	electron	$1.91^{+1.96}_{-1.91}$	-100% / +105%
	muon	$2.76^{+2.06}_{-1.65}$	-60% / +75%
2tag3jets	electron	$0.15^{+0.24}_{-0.15}$	-100% / +160%
	muon	$0.43^{+1.10}_{-0.43}$	-100% / +260%

Table 6.1 – Summary of MJ fractions, along with their associated uncertainties in the 2tag2jets and 2tag3jets separately

6.2.4 Acting on systematics

As presented before, the systematics are either contributing to a global scale factor acting on all the bins at the same time or kept as histograms that have a bin-by-bin effect. Amongst the latter, the effect could come from a variation of the event weights (e.g. flavour tagging, or shape modelling), or from a modification of the event properties (e.g. the jet energy scale) that would lead to a different shape of a certain variable. However, when the population of a bin is low, the second category could induce statistical shape fluctuations. To mitigate this effect a smoothing procedure has been adopted to rebin the distribution of the systematic variation of the discriminant. An iterative algorithm merges adjacent bins with the two conditions that no local extrema can be found after rebinning while preserving at most a 5% statistical error in each bin of the nominal distribution. The resulting variation is computed from the difference of the rebinned histograms while being applied on the initial binning.

If some of the systematics are presented with a $\pm 1\sigma$ variation naturally such as in the jet energy scale, it is not the case for all of them. The so-called 1-sided systematics are typically obtained when only the positive variation is described (as in the 2-point systematics of the shape modelling), or when the distribution is symmetric by construction (as in the jet energy resolution). However, 2-sided systematics can present some differences in magnitude between the positive and negative regions. Such effects are treated accordingly in the fit procedure and are considered genuine. The 1-sided systematics are symmetrised around the origin to be evaluated in the same way as the 2-sided ones.

Since a large number of NPs is present, the fit stability could be questionable. Some of the systematics described earlier have a very limited impact on the distribution. Even worse, in limited statistical regions, they could induce some artificial fluctuation that would be hard to control. Therefore a pruning algorithm is designed to take out the problematic and low impact NPs. Several rules are set depending on the type of systematics:

- **normalisation:** if the variation is less than 0.5% or the two sides have the same sign.
- **shape:** all the bins have a variation smaller than 0.5%.

In addition, the signal shape and normalisation systematics are dropped in regions where the signal is below 2% of the total background in all the bins and the effect is lower than 0.5% of the total background.

6.3 Results

The analysis is performed using the $\mathcal{L} = 79.8\ fb^{-1}$ dataset from Run-2 and the MC described in Section 4.1 and presented in [21]. The results are presented in Section 6.3.1 for the MVA analysis, whilst the Cut Based Analysis fit results are given in Section 6.3.2. The results are combined with the Run-1 results obtained from a different analysis [13], and with the other channels to get the final Vh and $h \to b\bar{b}$ results.

6.3.1 MVA Analysis

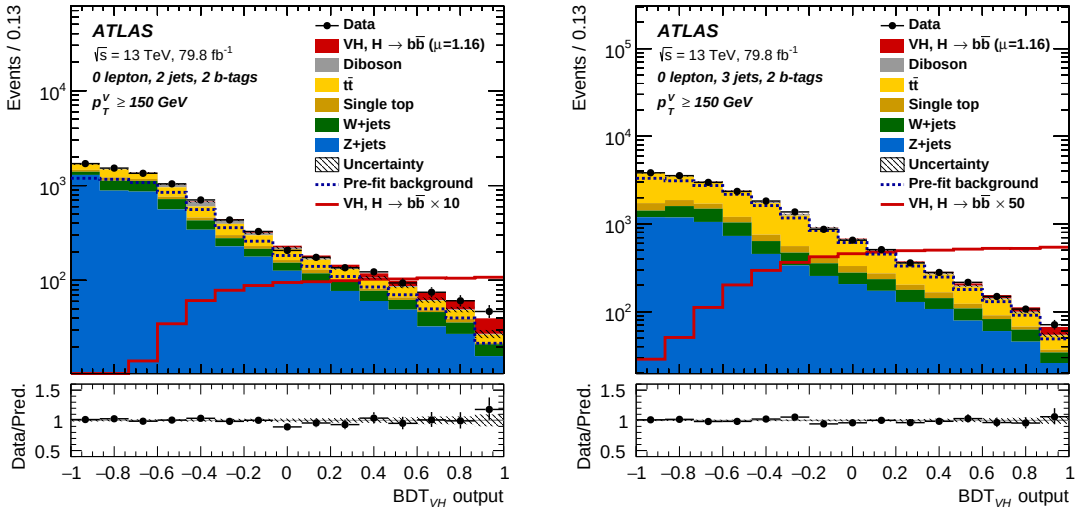
The main result in the $Vh(b\bar{b})$ analysis is driven by using the MVA BDT as a discriminant in the profile likelihood ratio fit, using the signal and control regions defined in the Section 4.2. However, since the control regions are not used to discriminate between the signal and the backgrounds but rather to model the latter, other distributions are used. The variables used are summarised in Table 6.2.

Channel	SR/CR	Categories			
		$75 < p_T^V < 150$ GeV		$p_T^V > 150$ GeV	
		2 jets	(≥) 3 jets	2 jets	(≥) 3 jets
0-lepton	SR	–	–	BDT	BDT
1-lepton	SR	–	–	BDT	BDT
2-leptons	SR	BDT	BDT	BDT	BDT
1-lepton	W+hf CR	–	–	Yield	Yield
2-leptons	Top emu CR	m_{bb}	m_{bb}	Yield	m_{bb}

Table 6.2 – Regions used in likelihood fit, and their associated distributions. In practice the yields are obtained from the m_{bb} distributions with one bin.

The post-fit plots of the most relevant distributions are summarised in Figures 6.4, 6.5 and 6.6. The background contributions after the global likelihood fit are shown in the filled histograms. The Higgs boson signal is shown as a red filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data (μ=1.16) and superimposed as a solid line in the plot, scaled by the factor indicated in the legend for visibility purposes. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal (μ=1.16) and background is shown in the lower panel. The

distributions are shown with the binning used in the global likelihood fit. The overall agreement between the data and the MC is good.



(a) (b)

Figure 6.4 – The BDT output post-fit distributions in the 0-lepton channel for 2-btag events, in the 2-jets (a) and exactly 3-jets (b) categories.

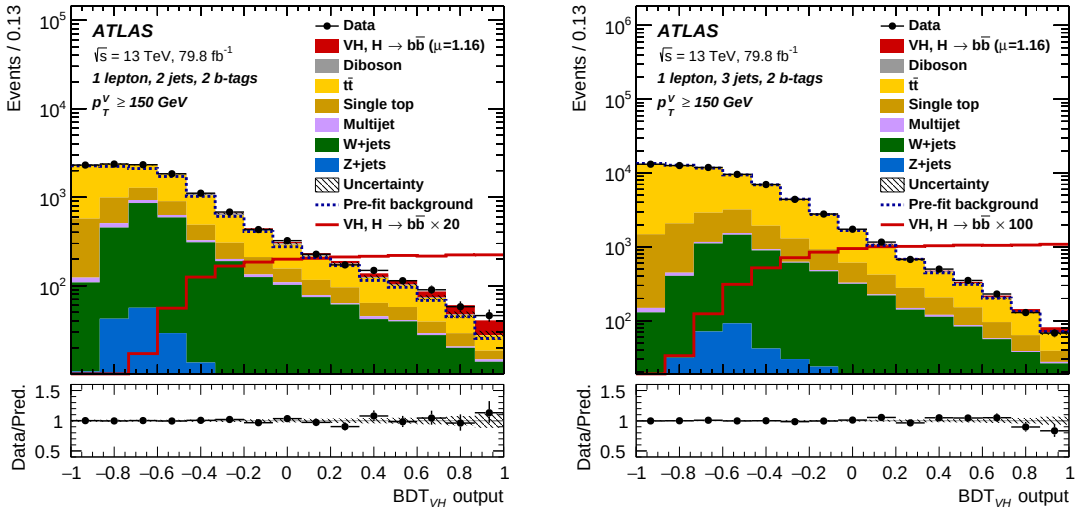
Pull values are investigated in Figure 6.7. The highest pulled NPs are concerning the Z+hf m_{bb} shape modelling. These effects have already been observed in the past and are linked to the large data sample in the 0- and 2-leptons mass sidebands that allow the fit to pull and constrain the nuisance parameter on the m_{bb} shape of the Z+hf background. The pull corrects a mismodelling, observed in Z+hf enriched sideband regions, of the m_{bb} distribution by the simulation. Some over-constraints are observed especially for modelling systematics, which could be explained by the large uncertainties associated to those variables at the reconstructed level, especially in the 2-leptons region where the fit has more constraining power. This could hopefully be improved with the new shape modelling proposed in Section 5.2.4.

The values of the free-floating scale factors are presented in detail in Table 6.3. While the scale factors are compatible with the MC predictions for the $t\bar{t}$ and the V+jets in the ($\geq$) 3-jets channels, important corrections are needed in the 2-jets especially for the Z+hf, as observed in the previous result [16], as well as in the CMS result [192] where the scale factor is even higher. This is expected due to the poor modelling of the cross-section for these processes and is also observed in SM Z+bb analysis [193], in particular when requiring b-tagging while it is not observed in the general case [194]. The same observation holds to a less extent in the W+hf, which is more compatible with predicted MC values and can be explained by difficult modelling, with less pure regions for constraints, whereas the Z+hf can use the 2-leptons sidebands.

The correlation plot is shown in Figure 6.8a for the Asimov fit and Figure 6.8b for the fit to data. Important correlations are observed for the normalisation factors, as well

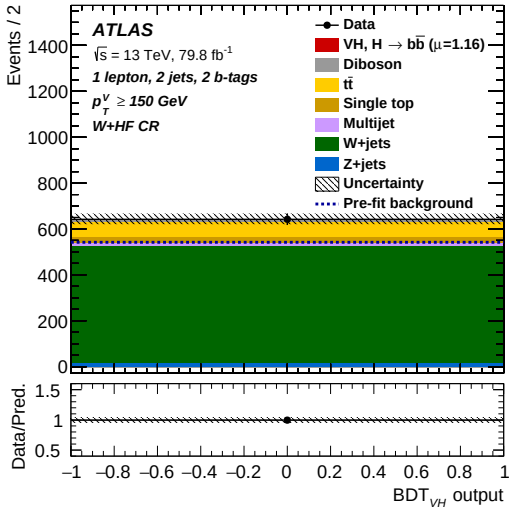
Process	Normalisation factor
$t\bar{t}$ 0- and 1-lepton	0.98 ± 0.08
$t\bar{t}$ 2-leptons 2-jets	1.06 ± 0.09
$t\bar{t}$ 2-leptons $\geq$3-jets	0.95 ± 0.06
W+hf 2-jets	1.19 ± 0.12
W+hf ($\geq$) 3-jets	1.05 ± 0.12
Z+hf 2-jets	1.37 ± 0.11
Z+hf ($\geq$) 3-jets	1.09 ± 0.09

Table 6.3 – Factors applied to the $t\bar{t}$, Z+hf, and W+hf backgrounds, as obtained from the global likelihood fit to the $\sqrt{s} = 13$ TeV data for the multivariate analysis, used to extract the Higgs boson signal. The errors represent the combined statistical and systematic uncertainties.

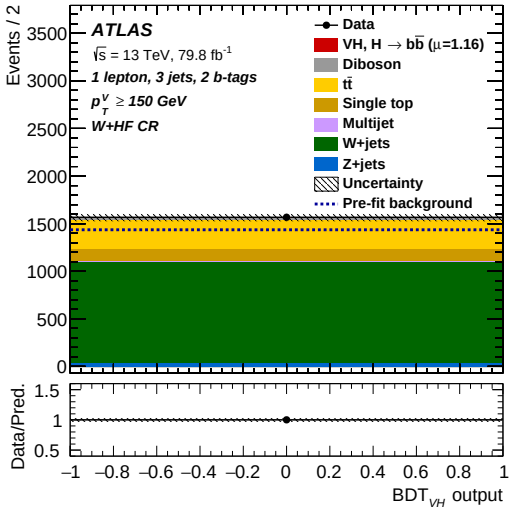


(a)

(b)



(c)



(d)

Figure 6.5 – The BDT output post-fit distributions in the 1-lepton channel for 2-btag events, in the 2-jets (a) and exactly 3-jets (b) categories. The W+hf control region plots are shown for 2-btag events, in the 2-jets (c) and exactly 3-jets (d) categories.

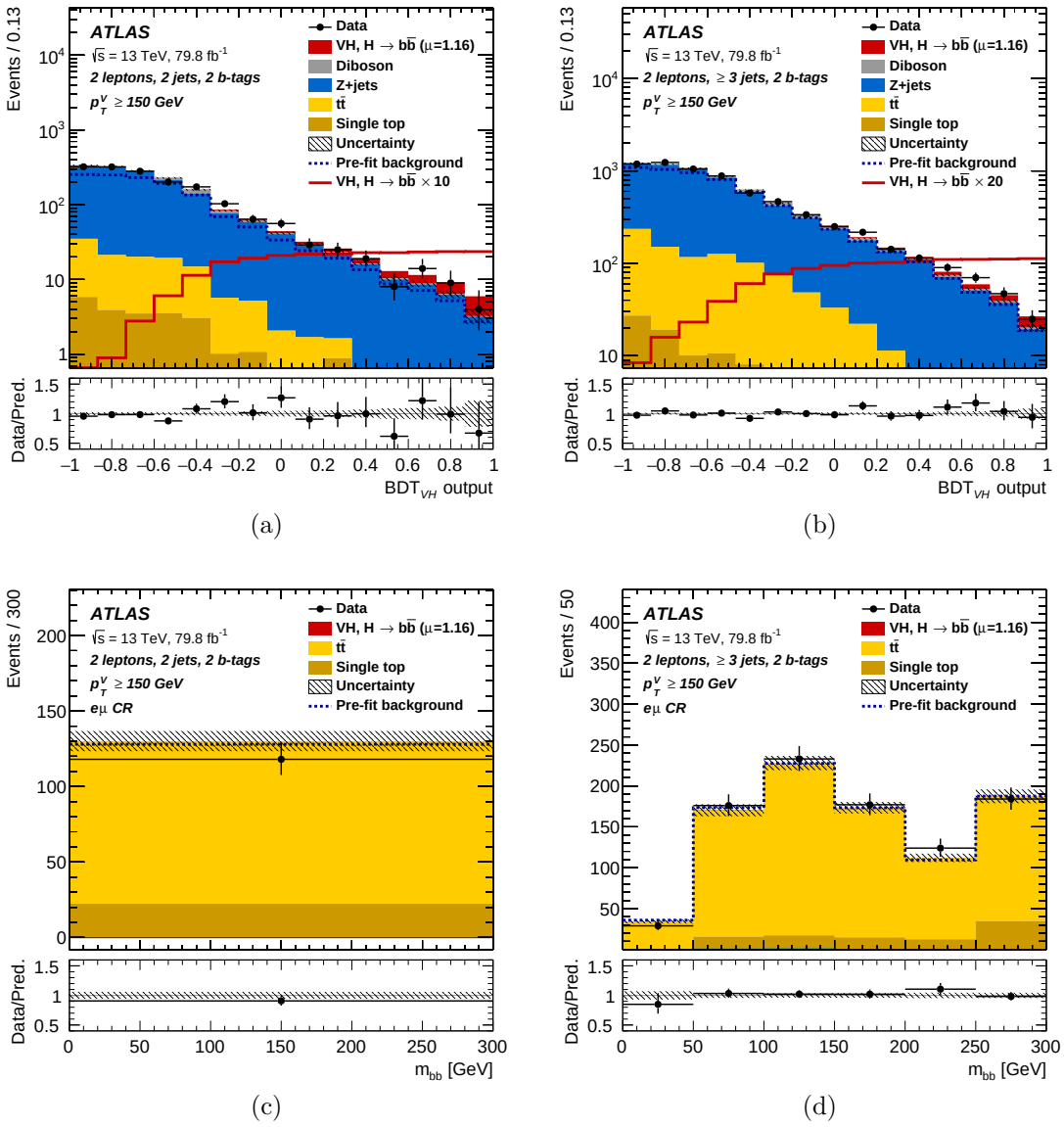


Figure 6.6 – The BDT output post-fit distributions in the 2-leptons channel for 2-btag events, in the 2-jets (a) and ≥ 3 jets (b) categories in the high-$p_{\mathrm{T}}^{\mathrm{V}}$ region. The Top emu control region plots are shown for 2-btag events, in the 2-jets (c) and exactly 3-jets (d) categories.

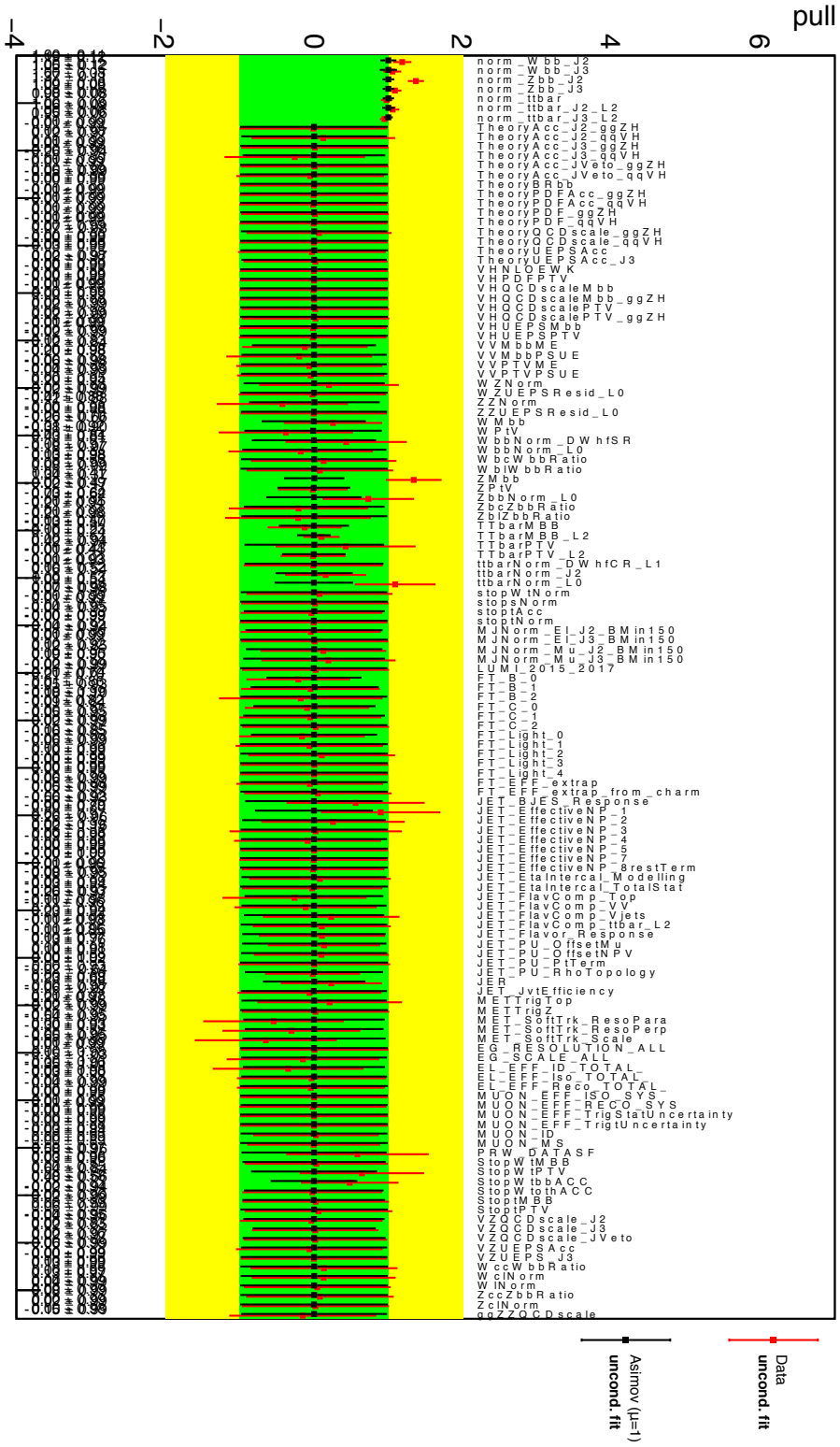
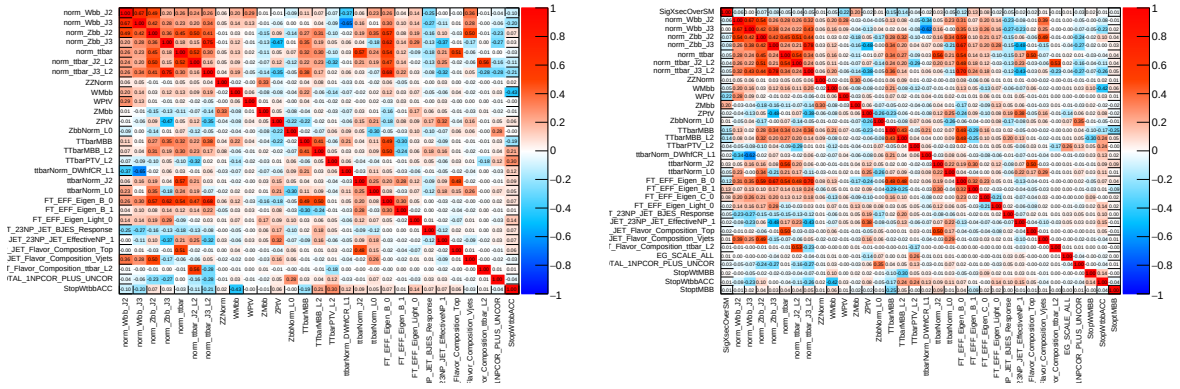


Figure 6.7 – Nuisance parameter pulls and the free parameter scale factors corresponding to a conditional combined fit performed to the Asimov dataset (black) and to the Run-2 data (red).

as with the b-tagging NP. This is due to the importance of b-tagging scale factors in the shapes of the distributions in this phase space and the normalisation of the backgrounds.



(a) (b)

Figure 6.8 – Correlation matrix from the Asimov fit (a) and fit to the data (b). Only variables with at least one correlation larger than 0.25 are shown. The order of the variables along the axis is arbitrary as it comes from the list of systematics provided to the framework by the analyser.

The ranking of the NPs is shown in Figure 6.9 for the fit to data. Due to improvements in the methodology used to estimate the modelling uncertainties for the single top-quark, VH signal and di-boson processes, the relative importance of the corresponding NPs is modified compared to the ranking presented in [16]. Modelling systematics are still clearly dominating while signal acceptance systematics and b-tagging are coming next.

The impact on the signal strength is given by the breakdown in Table 6.10. As shown in the table and the ranking plot in Figure 6.9, the systematic uncertainties due to the modelling of the signal play a dominant role, followed by the uncertainty due to the limited size of the simulated samples (MC statistical), the modelling of the backgrounds and the b-tagging uncertainties. The analysis is also systematics dominated, and in this sense, the next improvements will have to focus on the reduction of these contributions, starting with the modelling, as presented in Section 5.2.4. It is worth mentioning that the categories listed in this table are coming from the choice of the analysis strategy reflecting what seems to be a good compromise between the hundreds of NPs and the main categories.

The effects of the nuisance parameters in the fit are globally understood, and the result in terms of signal strength and significance can be extracted. A full combination of all the lepton channels yields to an observed p_0 probability of $5.3 \cdot 10^{-7}$ of a background fluctuation to explain the observed data, while the expected one is $7.3 \cdot 10^{-6}$. This result is translated in terms of significance to 4.9 standard deviations compared to 4.3 expected.

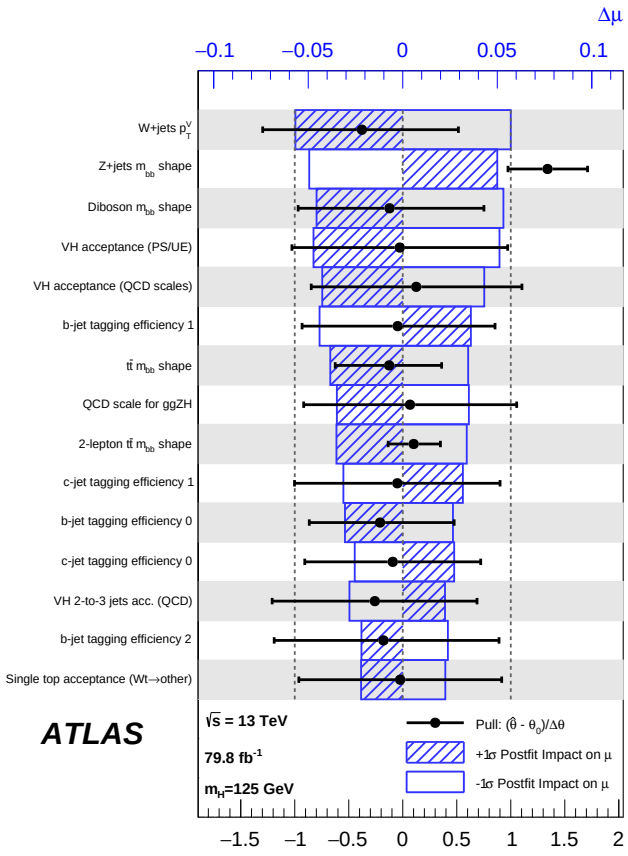


Figure 6.9 – Ranking of the nuisance parameters based on the impact on the best-fit value $\hat{\mu}$ for the fit to the data. The hatched and open areas correspond to the upwards and downwards variations, respectively. The filled circles are representing the corresponding pulls for the specific NP. Only the 15 firstly ranked NPs based on the global effect are shown.

Source of uncertainty		σ_μ
Total		0.259
Statistical		0.161
Systematic		0.203
Experimental uncertainties		
Jets		0.035
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.014
Leptons		0.009
b-tagging	b-jets	0.061
	c-jets	0.042
	light-flavour jets	0.009
	extrapolation	0.008
Pile-up		0.007
Luminosity		0.023
Theoretical and modelling uncertainties		
Signal		0.094
Floating normalisations		0.035
Z + jets		0.055
W + jets		0.060
$t\bar{t}$		0.050
Single top quark		0.028
Diboson		0.054
Multi-jet		0.005
MC statistical		0.070

Figure 6.10 – Breakdown of the contributions to the uncertainty in μ. The sum in quadrature of the systematic uncertainties attached to the categories differs from the total systematic uncertainty due to correlations.

The obtained value for the best-fitted signal strength is:

$$\mu_{VH}^{bb} = 1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$$

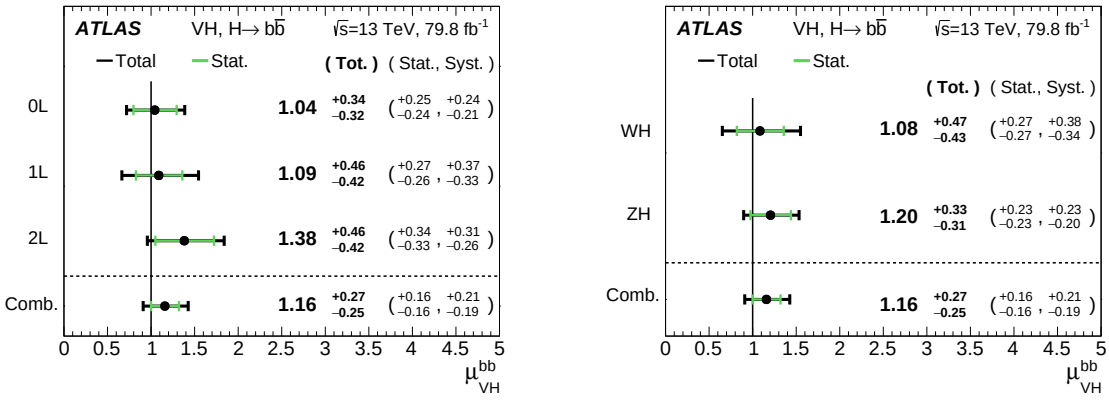
The decomposition of the signal strength uncertainties has already been presented in Table 6.10. The fit has also been computed considering each lepton channel signal strength independent and the results are reported in Table 6.4. The leading channels in the discovery are shown to be the Z-dominated 0- and 2-leptons modes, with the latter pushing-up the value of the signal strength. The comparison in terms of signal strength values is shown in Figure 6.11a. A compatibility test can be made separating the calculation into three independent μ and the global fit. It is assumed that the differences in the maximum likelihood values between fits follow a χ^2 distribution based on the difference of parameter

of interests, i.e. of the number of μ values. Performing a χ^2 between the single lepton channel fits and the global one lead to a measured compatibility of 80%

Channel	Signal strength	p_0		Significance	
		Exp.	Obs.	Exp.	Obs.
0-lepton	$1.04^{+0.34}_{-0.32}$	$9.5 \cdot 10^{-4}$	$5.1 \cdot 10^{-4}$	3.1	3.3
1-lepton	$1.09^{+0.46}_{-0.42}$	$8.7 \cdot 10^{-3}$	$4.9 \cdot 10^{-3}$	2.4	2.6
2-leptons	$1.38^{+0.46}_{-0.42}$	$4.0 \cdot 10^{-3}$	$3.3 \cdot 10^{-4}$	2.6	3.4
combination	$1.16^{+0.27}_{-0.25}$	$7.3 \cdot 10^{-6}$	$5.3 \cdot 10^{-7}$	4.3	4.9

Table 6.4 – Measured signal strengths with their combined statistical and systematic uncertainties, expected and observed p_0 and significance values (in standard deviations) from the combined fit with a single signal strength, and from a combined fit where each of the lepton channels has its own signal strength, using $\sqrt{s} = 13$ TeV data.

A two signal strength parametrisation separating the Z- and W-induced signals (referred to as ZH and WH) has been refitted. The correlation between the two signals has been found to be -1%. The observed (expected) significance of ZH mode is 4.0 (3.5) standard deviations, whilst 2.5 (2.3) for the WH that does not benefit from the 0- and 2-leptons channels. This result demonstrates that such a differential measurement, could be accessible with a full Run-2 analysis and provides valuable inputs to the theoretical communities. The extracted signal strengths are shown in Figure 6.11b. The level of statistical compatibility between the inclusive and differential fit is 84%.



(a) (b)

Figure 6.11 – Measured best-fit signal strengths for the combination of all three channels in a single μ and three μ fit (left) and for a two μ fit separating WH and ZH production (right).

6.3.2 Cut Based Analysis (CBA)

The cross-check performed for the cut based analysis differs from the multivariate one. The m_{bb} distribution is used as the discriminant, as described in Section 4.2.1. The modelling systematics were derived accordingly to the specific phase space, and distributions are binned with constant size contrary to the MVA distribtuions. The control region definition, as well as the $p_{\mathrm{T}}^{\mathrm{V}}$ binning, is modified as presented in Table 6.5.

Channel	SR/CR	Categories					
		$75 < p_{\mathrm{T}}^{\mathrm{V}} < 150$ GeV		$150 < p_{\mathrm{T}}^{\mathrm{V}} < 200$ GeV		$p_{\mathrm{T}}^{\mathrm{V}} > 200$ GeV	
		2 jets	$(\geq)$ 3 jets	2 jets	$(\geq)$ 3 jets	2 jets	$(\geq)$ 3 jets
0-lepton	SR	–	–	m_{bb}	m_{bb}	m_{bb}	m_{bb}
1-lepton	SR + CR	–	–	m_{bb}	m_{bb}	m_{bb}	m_{bb}
2-leptons	SR	m_{bb}	m_{bb}	m_{bb}	m_{bb}	m_{bb}	m_{bb}
2-leptons	Top emu CR	m_{bb}	m_{bb}	Yield*	$m_{bb}{}^{**}$	Yield*	$m_{bb}{}^{**}$

Table 6.5 – Regions used in likelihood fit, and their associated distributions. The medium- and high-$p_{\mathrm{T}}^{\mathrm{V}}$ bins are merged in the 2-leptons CR due to statistics issues.

The post-fit plots of the m_{bb} distributions are summarised in Appendix C. The distributions are shown with the binning used in the global likelihood fit. The overall agreement between the data and the MC is relatively good. As expected the fit is less smooth and some statistical fluctuations are observed. This is also reflected in the comparison of pulls and constraints from Figure C.5. The same constraints and pulls are observed compared to the MVA fit. The behaviours are overall fairly similar. The comparison of the correlation plots in Figure C.6 gives the same conclusions, despite small local differences, with the example of the Z m_{bb} shape modelling NP that is more strongly correlated to normalisations, probably due to its importance in the final fit.

The main difference arises from the NP's impacts on the signal strength. The ranking plots in Figure C.7 show the importance of the MC statistical NPs that appears in the first 15 ranked NP. This is also visible from the breakdown Table C.8. Floating normalisation and MC statistics are significantly impacting the final uncertainty on the signal strength, due to the stringent selections, and the higher number of bins. The jet systematics are also increased due to the relationship between the di-jet mass and the importance of the jet energy definition. However, most of the modelling systematics have a similar impact on the final result in the two approaches.

The significance and p-values of the CBA analysis are summarised in Table 6.6. The significance of the analysis is smaller than in the MVA analysis as expected. The global signal strength is found to be:

$$\mu_{VH}^{bb} = 1.06^{+0.36}_{-0.33} = 1.06 \pm 0.20(\mathrm{stat.})^{+0.30}_{-0.26}(\mathrm{syst.}),$$

The signal strength is derived also by channel and is also more prone to fluctuations with larger uncertainties than in the MVA case. A good agreement between the fitted signal strengths is shown in Figure 6.12.

Channel	Signal strength	Significance	
		Exp.	Obs.
0-lepton	$1.03^{+0.48}_{-0.45}$	2.6	2.6
1-lepton	$1.39^{+0.71}_{-0.62}$	2.4	2.6
2-leptons	$0.85^{+0.57}_{-0.51}$	2.5	2.4
combination	$1.06^{+0.36}_{-0.33}$	3.6	3.5

Table 6.6 – Measured signal strengths with their combined statistical and systematic uncertainties, expected and observed significance values (in standard deviations) from the combined fit with a single signal strength, and from a combined fit in which each of the lepton channel signal strengths is fitted independently of the others, using $\sqrt{s} = 13$ TeV data, for the CBA approach.

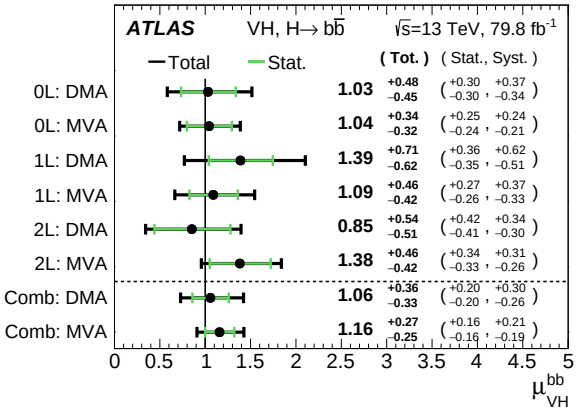


Figure 6.12 – The fitted values of the Higgs boson signal strength μ^{bb}_{VH} for the 0-, 1- and 2-leptons channels and their combination, using the $\sqrt{s} = 13$ TeV data. The results are shown both for the nominal multivariate analysis (MVA) and for the di-jet mass analysis (DMA), also-called CBA in this document. The individual μ^{bb}_{VH} values for the lepton channels are obtained from a simultaneous fit with the signal strength for each of the lepton channels floating independently.

The background-subtracted m_{bb} distribution showing both the di-boson and Higgs peak is shown in Figure 6.13. In order to increase the discrimination, the events are weighted by the S/B ratio in each of the categories defined in Table 6.5.

6.3.3 The combination

In order to benefit from all the possible measurements, the Run-2 result is combined with different analyses to improve its physics interpretation.

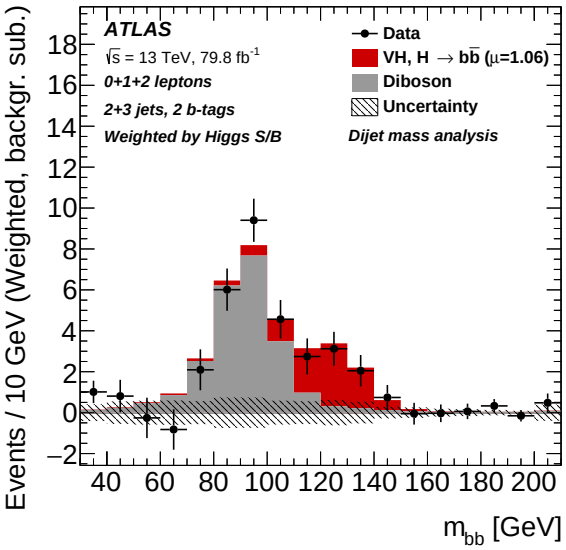


Figure 6.13 – The distribution of m_{bb} in data after subtraction of all backgrounds except for the WZ and ZZ di-boson processes, as obtained with the dijet-mass analysis (CBA). The contributions from all lepton channels, $p_{\mathrm{T}}^{\mathrm{V}}$ regions, and number-of-jets categories are summed and weighted by their respective S/B, with S being the total fitted signal and B the total fitted background in each region. The expected contribution of the associated VH production of a SM Higgs boson is shown scaled by the measured signal strength ($\mu = 1.06$).The size of the combined statistical and systematic uncertainty for the fitted background is indicated by the hatched band.

Combination with the Run-1 result

The released Run-1 analysis [13] showed a small excess for an Higgs boson at a mass of 125.36 GeV with an observed (expected) significance of 1.4 (2.6) standard deviations and a signal strength $\mu_{VH}^{bb} = 0.52 \pm 0.32(\text{stat.}) \pm 0.24(\text{syst.})$. A downward fluctuation was observed in the 0-lepton channel while the 2-leptons channel showed an excess of events. This study was already performed for the previous result [16], where the correlation scheme of the NPs was studied in detail. To be consistent with the other analysis in ATLAS, only the signal theory uncertainty and the b-jet energy scale are considered correlated, using a single signal strength in the fit. The correlations of other NPs have been found to have negligible impacts in the fit. The fitted signal was found to be:

$$\mu_{VH}^{bb} = 0.98^{+0.22}_{-0.21} = 0.98 \pm 0.14(\text{stat.})^{+0.17}_{-0.16}(\text{syst.}).$$

with an observed (expected) significance of 4.9 (5.1) standard deviations. While the expected significance increased thanks to the addition of data, due to the Run-1 measured rate being below the SM expectation in the 0-lepton channel the observed significance does not increase. Only a strong evidence can be officially claimed for the $Vh(b\bar{b})$ analysis, but with the assurance that the 5.0 σ significance threshold will be overcome with the inclusion of more data. The measurement of the ZH and WH components has been also conducted with the full data-set available and the result is presented in Figure 6.14

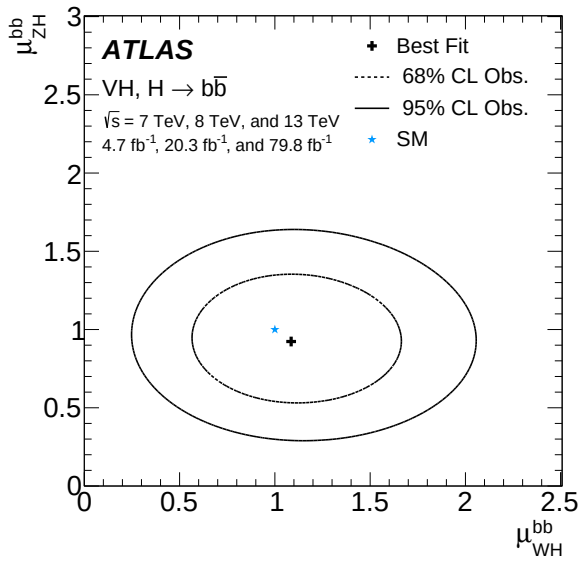


Figure 6.14 – The fitted values of the Higgs boson signal strength μ_{bb} from m_h =125GeV for the WH and ZH processes. The individual μ_{bb} values for the (W/Z)H processes are obtained from a simultaneous VH fit with the signal strength for each of the WH and ZH processes floating independently. The compatibility of the two signal strengths is 84%.

Combination with other production modes

The $Vh(b\bar{b})$ is not the only analysis targeting a $b\bar{b}$ decay mode for the Higgs boson. The $t\bar{t}$ associated production and the vector boson fusion modes are combined with it to provide enhanced sensitivity to the $h \to b\bar{b}$ decay. In all the analyses, the Run-1 (only considering the $\sqrt{s} = 8$ TeV, with $\mathcal{L} = 20.2\ fb^{-1}$) and Run-2 data (with a partial $\mathcal{L} = 36.1\ fb^{-1}$ for the ttH mode [195] and $\mathcal{L} = 24.5\ fb^{-1}$ and $\mathcal{L} = 30.6\ fb^{-1}$ due to trigger availability in the two sub-analysis of the VBF mode [51]) are considered. The latter mode is also sensitive to the gluon-gluon fusion mode through the all-hadronic decay channel. For the Run-2 analyses, the $t\bar{t}H$ and VBF are correlated through the luminosity, jet energy scales and resolutions, JVT and pile-up. The two analyses are further correlated to the $Vh(b\bar{b})$ through the e/γ NP, but most importantly through the branching ratios.

The resulting observed (expected) significance is found to be 5.4 (5.5) standard deviations, leading to the formal observation of the $h \to b\bar{b}$ decay mode. The signal strength for the decay to b-quarks is found to be:

$$\mu_{H\to bb} = \frac{BR_{bb}}{BR_{bb,SM}} = 1.01 \pm 0.20 = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.15}(\text{syst.}).$$

in agreement with the SM. The different values of the signal strengths and significances by production mode can be found respectively in Figure 6.15 and Table 6.16. As expected the $Vh(b\bar{b})$ dominates the result with a small contribution from the VBF+ggF channel. The same fit floating independently the signal strength per data period (Run-1 and Run-

2) has been performed resulting in a probability of compatibility of 54% between the two datasets due to the different values observed in Run-1 for some of the modes.

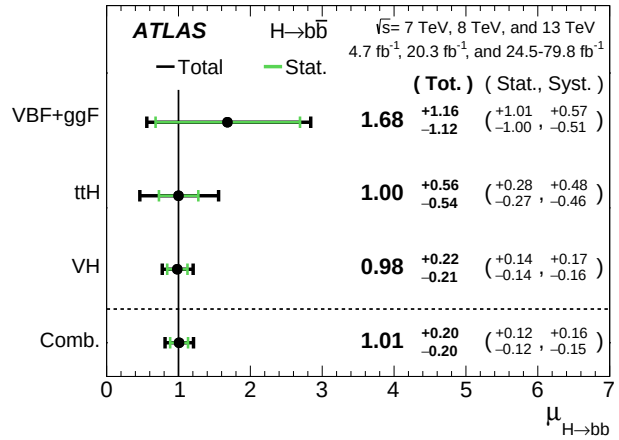


Figure 6.15 – The fitted values of the Higgs boson signal strength $\mu_{H\to bb}$ separately for the Vh, $t\bar{t}$ H and VBF+ggF analyses along with their combination, using the $\sqrt{s} = 7$ TeV, $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV data. The individual $\mu_{H\to bb}$ values for the different production modes are obtained from a simultaneous fit with the signal strengths for each of the processes floating independently.

Channel	Significance	
	Exp.	Obs.
VBF+ggF	0.9	1.5
$t\bar{t}H$	1.9	1.9
VH	5.1	4.9
$H \to b\bar{b}$ combination	5.5	5.4

Figure 6.16 – Expected and observed significance values (in standard deviations) for the $h \to b\bar{b}$ channels fitted independently and their combination using the $\sqrt{s} = 7$ TeV, $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV data.

Combination with other decay modes

In the same spirit, the Vh production can be reconstructed in several other decay modes of the Higgs, such as the di-photon or $ZZ^* \to 4$ leptons channels.

The di-photon channel analysis uses a multi-categorisation of the vector-boson decay in both the leptonic (5 categories) and hadronic (2 categories) channels. The $ZZ^* \to 4$ leptons analysis classifies the events in three categories: fully-hadronic, only charged-leptons (linked to the 2-leptons channel), or presence of neutrinos (linked to the 0- and 1-lepton channel). Both analyses are updated with $\mathcal{L} = 79.8\ fb^{-1}$ of data while the $\mathcal{L} = 36.1\ fb^{-1}$ were published in [196, 197]. They are considered orthogonal due to their different final state, and thus the combination was worked out as presented in [52], fixing the different Branching Ratio to the one predicted by the SM.

The resulting observed (expected) significance is found to be 5.3 (4.8) standard deviations, leading to the formal observation of the Vh production mode. The signal strength of the Vh cross-section is found to be:

$$\mu_{VH} = \frac{\sigma_{VH}}{\sigma_{VH,SM}} = 1.13^{+0.24}_{-0.23} = 1.13 \pm 0.15(\text{stat.})^{+0.18}_{-0.17}(\text{syst.}).$$

in agreement with the SM. The individual values of the signal strengths and significances by decay modes can be found respectively in Figure 6.17 and Table 6.18. As expected the $Vh(b\bar{b})$ analysis dominates the global significance. The fit compatibility between the three decay modes is 96%, showing the robustness of the analysis and the result.

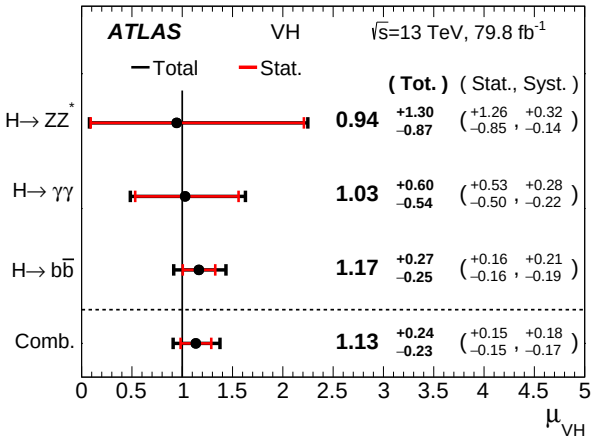


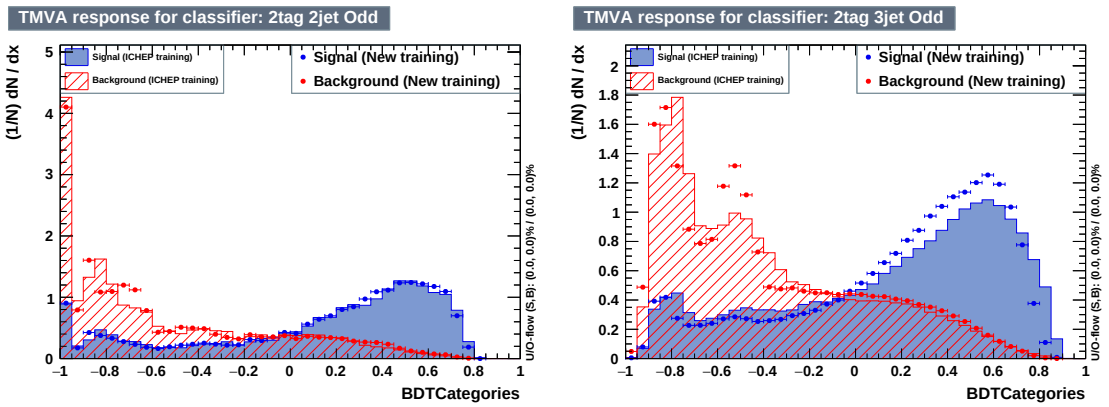
Figure 6.17 – The fitted values of the Higgs boson signal strength μ_{VH} separately for the $H \to b\bar{b}$, $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ decay modes, along with their combination. The individual μ_{VH} values for the different decay modes are obtained from a simultaneous fit with the signal strengths for each of the processes floating independently.

Channel	Significance	
	Exp.	Obs.
$H \to ZZ^* \to 4\ell$	1.1	1.1
$H \to \gamma\gamma$	1.9	1.9
$H \to b\bar{b}$	4.3	4.9
VH combined	4.8	5.3

Figure 6.18 – Expected and observed significance values (in standard deviations) for the VH production channels from the combined fit and from a combined fit where each of the lepton channels has its own signal strength, using $\sqrt{s} = 13$ TeV data.

6.4 Effects of changing the anti-QCD cuts

The evaluation of the effects of the new anti-QCD cuts presented in Section 5.1.1 has been studied up to the fit. In order to fully take into account the new selection, the BDT has been retrained using the same hyper-parameters as those described in Table 4.9. The performance of the retraining can be observed in Figure 6.19. From the ROC curve and a simple binned significance, a global similar performance is observed with respect to the official result presented in Figure 4.8. The structure of the background distribution is more peaked in the 2tag3jets channel. This could be explained by better intrinsic discrimination between the different backgrounds due to the $t\bar{t}$ reduction.



(a) (b)

Figure 6.19 – Distribution of the BDT scores after training with the odd events of the ICHEP training (plain), and the new training (dots), for the signal (blue) and the backgrounds (red) in the 2tag2jets (a) and 2tag3jets (b) regions.

This BDT is then used in the Profile Likelihood fit. All the experimental and modelling systematics have been kept the same as in the observation analysis. In the case of the modelling systematics it is not certain that the used NPs are still correct, given the new phase space. However, it was considered as a second-order issue and remains a caveat of this study. The pulls of the existing nuisance parameter reveal that only a few modifications in the individual constraints and pulls are observed as shown in Figure 6.20. The major change that is observed is a variation in the Z+jets normalisation factor in the 3-jets category and the $t\bar{t}$ normalisation factor. Most likely the two scale factors are compensating for one another due to the differential reduction obtained with the new cut scheme shown in Table 5.4. Some decorrelation studies have been performed, allowing the fit to recover the differences without success. This effect could be cured by designing specific control regions for the Z+jets background, as done in the 1- and 2-leptons channels.
When looking at the post-fit distribution of the BDT score in Figure 6.21, an overall good agreement is observed between the data and the MC from the ratio plot given. The main issue seems to appear at high BDT score where the MC under(over)-estimates the data in the 2(3)-jets category. The situation in the 2tag2jets is somewhat similar to what obtained in the observation fit result while the data-MC agreement is worse in the 2tag3jets.

In order to assess the effect of the fit on the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. variable used in the selection process, the distributions are plotted before and after applying the fitted nuisance parameters in Figure 6.22. If the situation is under control in the 2tag2jets region, a clear mismodelling, most likely coming from the important Z+jets scale factors, is observed in the 2tag3jets. Furthermore the resolution of the objects entering the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. definition might be non-optimal. For instance the resolution of the jets is known to have some data-MC mismodelling. While this is supposed to be accounted for in the official recommendations, it could still appear when requesting the extra jet in the 2tag3jets category.

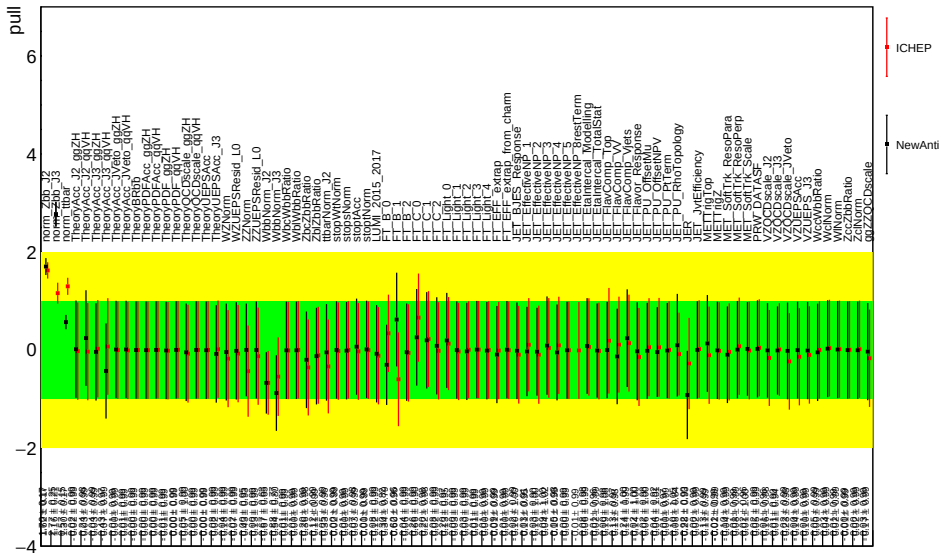
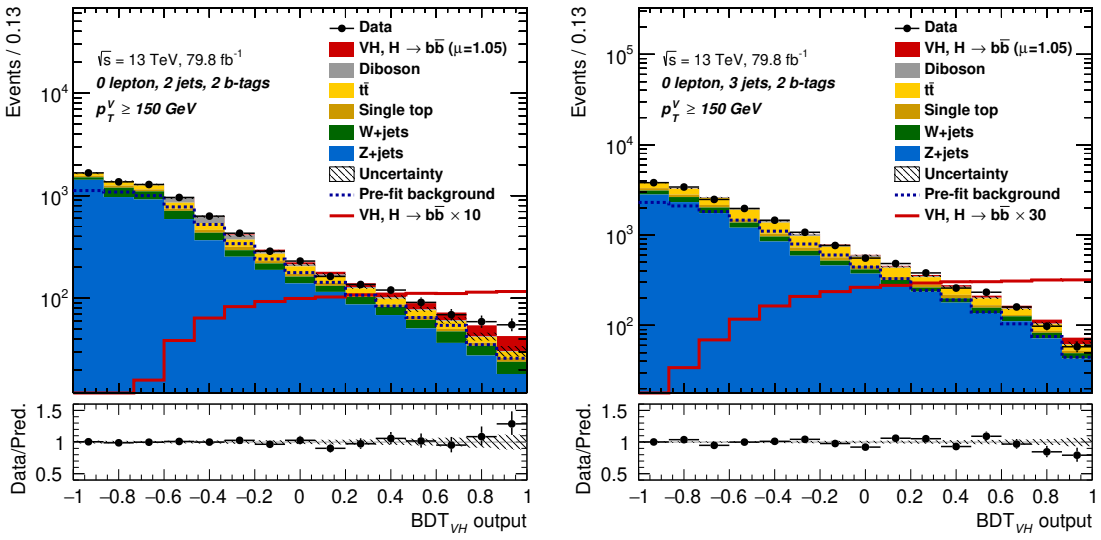


Figure 6.20 – Comparison of the pulls of the nuisance parameters between the observation result denoted as ICHEP and plotted in red and the result obtained with the new anti-QCD cuts denoted newAnti and plotted in black. The two fits are involving only the 0-lepton channel.



(a) (b)

Figure 6.21 – Distribution of the BDT scores using the nuisance parameters of the profile likelihood fit made with in the phase space using the newly designed anti-QCD cuts in the 2tag2jets (a) and 2tag3jets (b) regions.

Finally, the significance reflects the same observations. The expected and observed significances obtained in the 0-lepton only fit are presented in Table 6.7. Three different steps of nuisance parameters inclusion are considered. The first step is to consider only the data and the MC statistical nuisance parameters. A small increase in the expected significance is observed despite the smaller number of events in the fit and thus due to the better BDT separation obtained with the retraining. The normalisation nui-

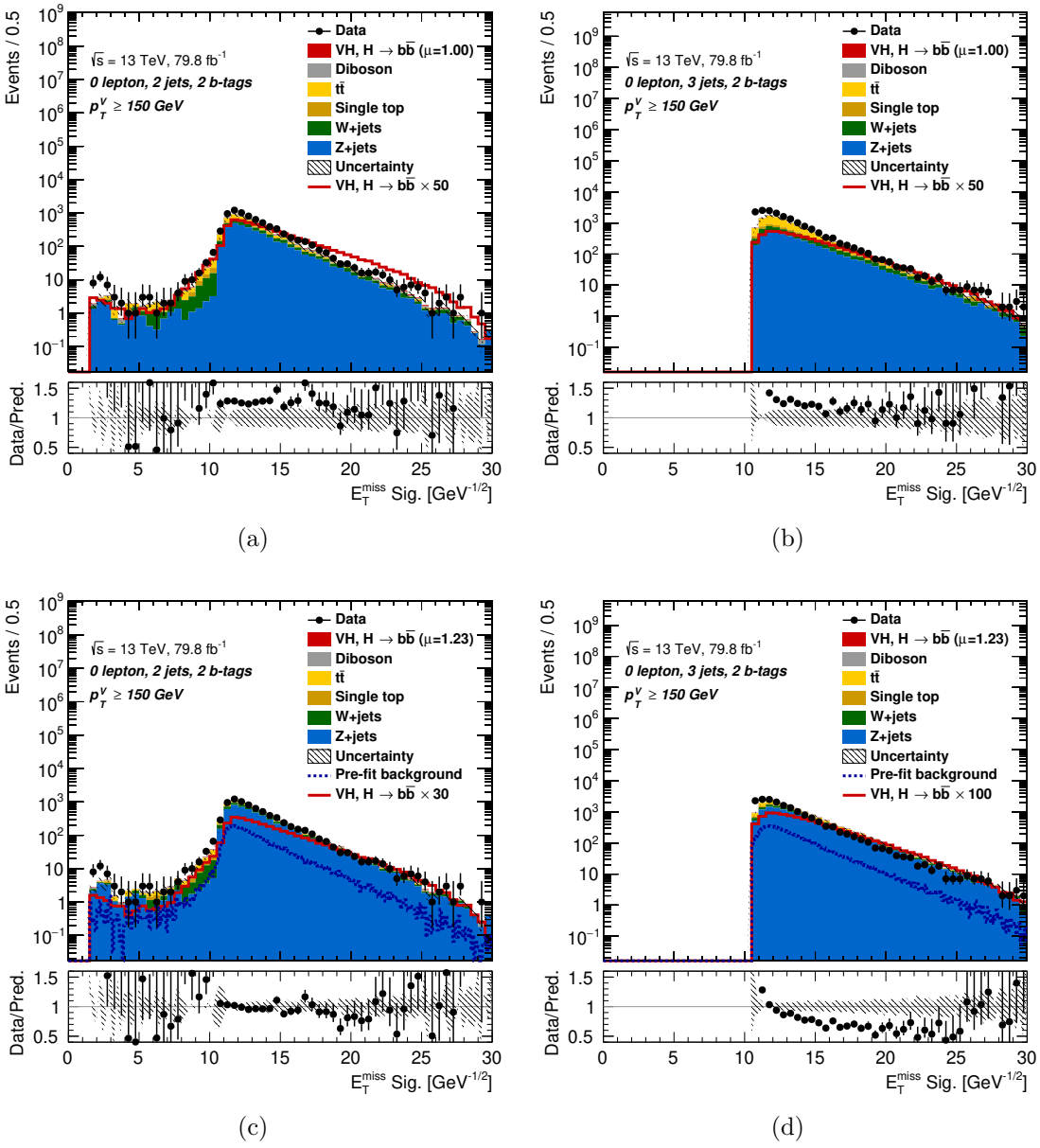


Figure 6.22 – Distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. prefit (top) and using the nuisance parameters of the profile likelihood fit (bottom) made with in the phase space using the newly designed anti-QCD cuts in the 2tag2jets (a,c) and 2tag3jets (b,d) regions.

sance parameters (Z+hf and $\mathrm{t\bar{t}}$ scale factors) are then considered. A small drop in the expected significance is however observed when including all the systematics.

All the observations and conclusions from this study, are postponing the re-designing of the anti-QCD cuts and the inclusion of the $E_{\mathrm{T}}^{\mathrm{miss}}$sig.. However, some other opportunities to use this variable in the analysis are under discussion for instance for the EW background separation in a MVA analysis that would separate not only the signal from the background but would resolve the EW backgrounds among themselves.

Anti-QCD cut scheme	Statistics	Normalisations	Full systematics
Nominal	4.21	3.89	3.02
New cuts	4.30	3.93	3.00
diff. (%)	2.1	1.0	-0.7

Table 6.7 – Expected significance with different NPs added: Statistics only includes the data and the MC statistical NP, Normalisations adds the Z+hf and $t\bar{t}$ normalisation scale factors, Full systematics includes all the nuisance parameters considered in the analysis. The relative difference is given with respect to the nominal cut scheme.

7 Physics and technical prospects for the High-Luminosity phase of the LHC

The LHC has been built as a staged accelerator, with a plan to increase the centre-of-mass energy and instantaneous luminosity with time as shown in Figure 7.1. While the first and second run exceeded the expected luminosity performance, the Run-3 (2021-2023) will be dedicated to the increase of the centre-of-mass energy, improving also marginally the peak luminosity. The total amount of data collected by the end of this run is expected to be at least $\mathcal{L} = 300\ fb^{-1}$, while the already collected data-set corresponds to almost $\mathcal{L} = 190\ fb^{-1}$. During this phase, the major Higgs physics program focussed on discovery, and some partial differential and coupling measurements. Constraints on some EFT models were set using the simplified template cross-section measurement [198].

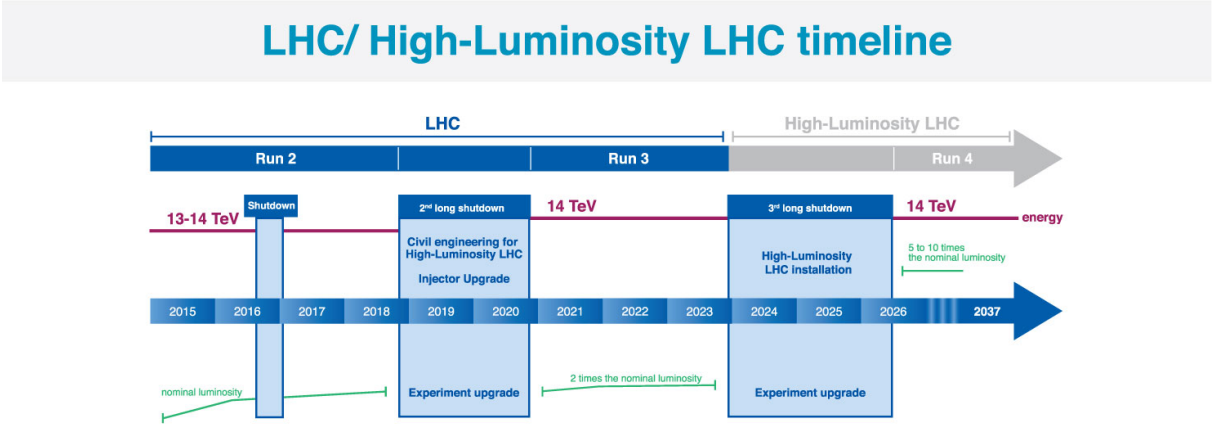


Figure 7.1 – Timeline for the LHC accelerator operation and planned upgrades [199].

However, to extract the full information of the Higgs sector, more data is needed. This will allow studying in detail statistically limited processes, such as the coupling to charm or muons but also the Higgs trilinear self-coupling. The precise measurements of the Higgs properties could also help to constrain some open questions such as the hierarchy problem or the nature of dark matter. Before and after the start of the LHC operation, several documents have summarised the physics cases for the luminosity upgrade. In the last ten years, these studies have been refined with the update of the European Strategy for Particle Physics [200] followed rapidly by the Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment [201]. In parallel with the development of the present analyses, predictions on the performance of the Higgs boson searches for the next few years of the LHC have been studied extensively. The extrapolation of the $Vh(b\bar{b})$ analysis in this context is described in Section 7.1. The last phase of the LHC program, called the High-Luminosity LHC (HL-LHC), aims at increasing the instantaneous luminosity by installing a better injection complex and a new focusing scheme. The expected integrated luminosity in each year should be equivalent to the total collected during the Run-1, -2 and -3, resulting in a final dataset of $\mathcal{L}$ = 3000-4000 fb^{-1}. To achieve this, not only do the accelerators need to be redesigned, but the experiments will have to face data-taking conditions. In this context, the design of the new tracking detector of the ATLAS experiment is discussed in Section 7.2.

7.1 Physics prospects on the $Vh(b\bar{b})$ measurements at HL-LHC

In ATLAS, a preliminary study on the physics discovery potential expected from the luminosity upgrade was conducted for the ECFA HL-LHC workshop in 2014 [202], with a special analysis focusing on $Vh(b\bar{b})$ [203] based on the Run-1 results. The analysis was performed by smearing the detector performance and the reconstruction efficiencies for the objects to account for the difference in the pile-up and data-taking conditions. In particular, a dedicated $Vh(b\bar{b})$ analysis study extrapolated the Run-1 analysis using a fit on the m_{bb} distribution, considering only the 1- and 2-leptons channels. The considered conditions for this study were using an effective pile-up of $<\mu>=140$. Experimental systematics were also updated. For instance, the jet energy scale uncertainties were increased by a factor 1.75 to account for the difference of centre-of-mass energy, but as well scaled down to take into account the increase of statistics. As a figure of merit, two scenarios were defined: in the first scenario the systematics were divided by a factor 10, and in the second by a factor 20. The results obtained are presented in Table 7.1. With the present experience, it should be noted that not only was this study relying on the

wrong pile-up profile but was missing the improvements obtained with the Run-2 analysis and the addition of the 0-lepton channel.

		1-lepton	2-leptons	1- & 2-leptons
Stat-only	Significance	7.7	7.5	10.7
	μ_{stat} error	±0.13	+0.14 -0.13	±0.09
Theory-only	+0.09 -0.07	+0.07 -0.08	±0.07	
Scenario I	Significance	1.8	5.6	5.9
	$\mu_{w/theory}$ error	+0.56 -0.54	+0.20 -0.19	±0.19
	$\mu_{wo/theory}$ error	±0.54	±0.18	+0.18 -0.17
Scenario II	Significance	3.2	–	6.4
	$\mu_{w/theory}$ error	+0.33 -0.32	–	+0.18 -0.17
	$\mu_{wo/theory}$ error	±0.32	–	±0.16

Table 7.1 – Expected signal sensitivity as well as the precision on the signal strength measurement for $m_H = 125$ GeV for the 1-lepton, 2-leptons and combined searches with $\mathcal{L} = 3000\ fb^{-1}$ with $< \mu > = 140$ in the two benchmark scenarios considered

In the scope of the definition of the new European Strategy for Particle Physics, an updated draft of the physics performance has been written in the form of a public document ("yellow report") [204]. The purpose of this document was to serve as a basis of the discussions held at the Granada open symposium during the 13^{th} to the 16^{th} of May 2019. It benefited also from all the latest improvements introduced in the analysis with partial Run-2 datasets, as well as the secured developments on the hardware of the experiments. Discussions were conducted between the CMS & ATLAS collaborations, as well as with theoreticians, to define a common treatment for the analyses. A specific document summarising the ATLAS results on the Higgs sector is presented in [205], which includes the results presented in the next subsections.

7.1.1 Method

The HL-LHC projections for the $Vh(b\bar{b})$ channel are performed using extrapolations based on the results of the analysis of $\mathcal{L} = 79.8\ fb^{-1}$ of pp collision data collected at $\sqrt{s} = 13$ TeV [21]. The same statistical framework and analysis strategies are used. In particular, the same selection and event categories, for both signal and control regions, are maintained. While the objects were updated in the previous study, it was decided that the object reconstruction efficiencies, resolutions and fake rates would be kept the same between the Run-2 and the HL-LHC environment. This hypothesis is based on the assumption that detector improvements for the HL-LHC phase would be compensated by the degradation from the increased pile-up and running conditions. This choice was made for the sake of simplicity, and is the result of an agreement between the analysis teams and the Combined Performance Groups of the two experiments. Any change in the coverage, especially for the tracking system, or the implementation of timing information

has been neglected resulting in a more conservative approach. The fit is conducted on the same BDT as the one trained in the Run-2 analysis. Concerning the parameters of interest (POI), two different measurements are performed: the signal-strength defined as in the Run-2 analysis, and the cross-section multiplied by the branching ratio that would serve as input for theoretical comparisons. In practice, the latter is normalised by its SM value and therefore obtained from the first result by removing any theory uncertainties on the cross-section and the branching ratio. Three types of POI with different interpretations of the results are provided: the 1-POI fit considers the full VH signal, the 2-POI separates the contributions between the Wh and Zh processes, while the 3-POI undergoes the breakdown of $q\bar{q}Z\mathrm{h}$, $\mathrm{gg}Z\mathrm{h}$ and Wh modes. The last one is motivated by the intrinsic difference of the signal modelling systematics as described in Section 5.2.2. Since no data is present, only Asimov datasets are considered assuming all the production cross-sections and branching fractions to be the ones from SM.

Moving from the $\sqrt{s} = 13$ TeV MC samples to the $\sqrt{s} = 14$ TeV conditions of the HL-LHC, cross-sections of the processes were reweighted to take into account the change of the centre-of-mass energy. The latest theoretical results gave the following scale factors: 1.102 for Wh, 1.105 for $q\bar{q}Z\mathrm{h}$, 1.176 for $\mathrm{gg}Z\mathrm{h}$ taken from [49]. The background processes scale factor is taken to be 1.1 inclusively as explained in [170]. The samples are furthermore scaled up by the ratio of luminosity ($\mathcal{L} = 3000\ fb^{-1}$ with respect to $\mathcal{L} = 79.8\ fb^{-1}$) directly for the inputs produced by the analysis framework.

The systematics scheme has been partially updated to take into account a few features from the HL-LHC conditions:

— the MC statistics NPs are not considered since it is thought that by the end of the HL-LHC they would be negligible compared to other source of systematics, considering the number of samples available by that time. This also allows the theoreticians to factor out this source of uncertainty that is not related to the physics.

— the Pythia 8 Herwig component of the signal normalisation NPs have been dropped, as this mismodelling is supposed to have disappeared with the new generation of MC generators.

— an extra systematic uncertainty is set when considering the cross-section measurement. Since an overall cross-section is applied on both the $q\bar{q}Z\mathrm{h}$ and $\mathrm{gg}Z\mathrm{h}$ processes, a residual QCD effect is inserted to control the known differences. This extra systematic is only needed in the 1- and 2-POI cases when measuring the cross-section since those effects are mitigated via the theoretical QCD uncertainty in the signal strength case. Used values are summarised in Table 7.2.

The systematics values have been also updated to take into account the various expected performance with the future analysis developments and decided in agreement

effect from	$q\bar{q}Zh$	$ggZh$
applied on $q\bar{q}Zh$	0.1 %	-3.7%
applied on $ggZh$	-0.5 %	21.3 %

Table 7.2 – Priors applied on the residual cross-section uncertainty in the 2-POI cross-section result.

between all the analysis channels and between the two experiments. Two scenarios are considered: a conservative one where all the systematics are kept with the same values (scenario S1) and a more optimistic one where reductions ranging from 0.6 to 1/3 are set (scenario S2), according to the Table 7.3. The baseline for the results is Scenario 2, Scenario 1 being considered as a conservative approach.

Nuisance parameter		Scenario 1	Scenario 2
Luminosity		–	0.5
JES		–	0.5
JER		–	0.5
tau		–	0.6
flavor tagging	b-jets	–	1/3
	c-jets	–	1/3
	light-jets	–	0.5
Shape syst.	Stop	–	0.5
	ttbar	–	0.5
	VV from ME	–	removed
	W+jets	–	0.5
	Z+jets	–	0.5
	Multijet (MJ)	–	0.5
	VH signal	–	0.5
Theory	backgrounds	–	0.5
	acceptance of signal from QCD scale	–	0.5
	acceptance of signal from PDF	–	0.5
	acceptance of signal from UEPS	P8-H7 diff. dropped	P8-H7 diff. dropped
	XS of $qqVH$ from QCD scale	–	0.5
	XS of $ggZH$ from QCD scale	–	27% $\rightarrow$ 15%
	XS of signal from PDF	–	0.5

Table 7.3 – Scale factors applied to reduce the prior systematic uncertainties for the HL-LHC $Vh(b\bar{b})$ from scenario S1 to scenario S2 projections. The electron and muon related systematic uncertainties are not reduced since they have negligible impacts on the μ measurement. The systematic uncertainties denoted with "removed" are considered not to be applicable to the HL-LHC analysis and they are not included in the fit. The systematic uncertainties denoted with "–" are kept the same as in Run-2 analysis.

The result has been combined with other decay modes: WW, ZZ, $\gamma\gamma$, Zγ, $t\bar{t}$H, and $\tau\tau$. The details about the correlation of the NPs is given in [206]. Several measurements such as the production cross-section (several decay modes are sensitive to the same production mode) or the decay branching ratios, are extracted and re-interpreted. Finally, the results are provided in the κ-framework formalism. Multiplicative factors for

the Higgs width (κ_H), the production (κ_i) and decay amplitudes (κ_f) in the zero-width approximation for the Higgs boson [170] are defined as:

$$\sigma_i \times \mathrm{BR}(H \to f) = \frac{\sigma_i \times \Gamma_f}{\Gamma_H} = \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2} \sigma_i^{\mathrm{SM}} \times \mathrm{BR}^{\mathrm{SM}}(H \to f)$$

A relationship is set between the Higgs width and decay amplitudes multipliers under the hypothesis of no trace of BSM couplings:

$$\kappa_H^2 = \sum_f \kappa_f^2 \mathrm{BR}^{\mathrm{SM}}(H \to f)$$

Different limits are also computed in the case of BSM but will not be discussed in this document.
The coupling modifiers can also be re-expressed with global couplings where all fermionic couplings and weak boson couplings would get the same values, respectively κ_f and κ_V. In this way the $Vh(b\bar{b})$ analysis is directly sensitive to both values.

7.1.2 Ranking of the Nuisance Parameters

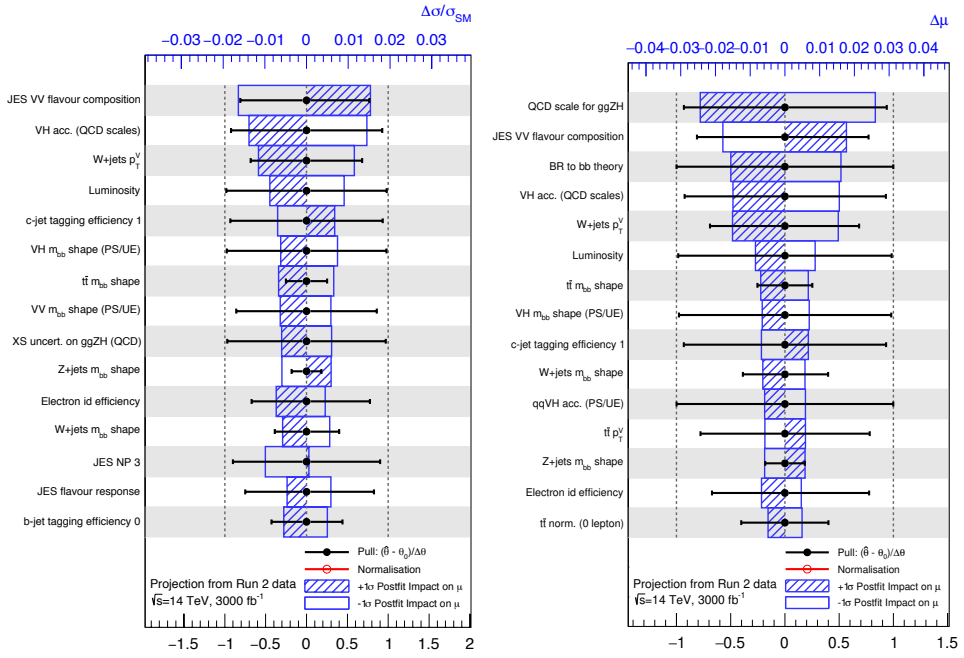
Figure 7.2 shows the ranking, for each production mode, of the systematic uncertainties with the largest impact on the cross-section times branching ratio and signal strength in the scenario S2 for the 1 POI fit, while the 3 POI result is shown in Figure 7.3.

The largest impact for the systematics comes from the flavour composition part of the jet energy scale assigned to the di-boson backgrounds, especially in the Zh channel. Theory uncertainties are also highly ranked thus highlighting the need for a strong theoretical development in the coming years.

7.1.3 Results on the cross-section and signal strength uncertainties

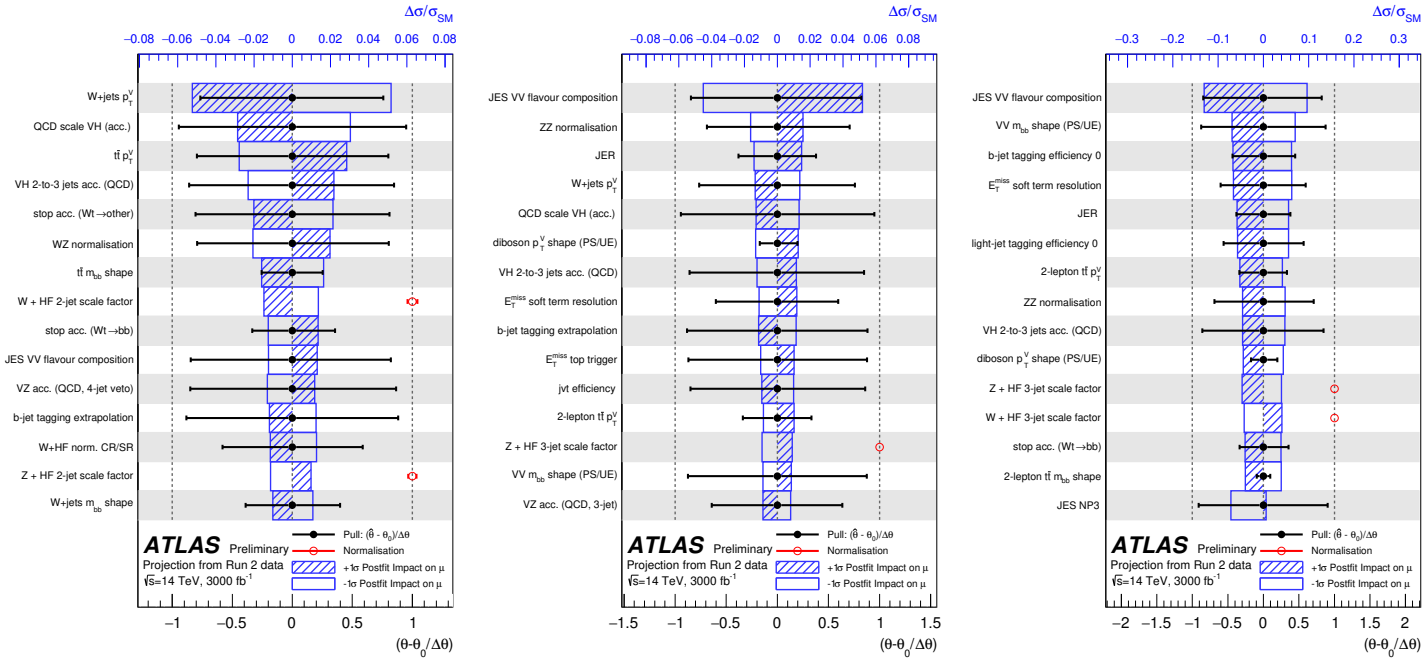
To understand the effect of the systematics on the results, a different scheme with respect to the Run-2 analysis [21] has been adopted to compute the breakdown of the effects. Following the recommendation of the Higgs Combination Group, as well as the CMS collaboration, four categories are defined: statistical, experimental, signal and background modelling impacts. Differences also arise from the technique to compute the values. In the Run-2 paper the impact of a systematic is computed by calculating the quadratic difference of the POI value X between a fit where all parameters are allowed to float, and the fit where the considered NPs are fixed to their nominal values:

$$\Delta \mathrm{X}_{\text{class n}} = \sqrt{|\Delta \mathrm{X}(\text{all floating})^2 - \Delta \mathrm{X}(\text{NP of class n frozen})^2|}$$

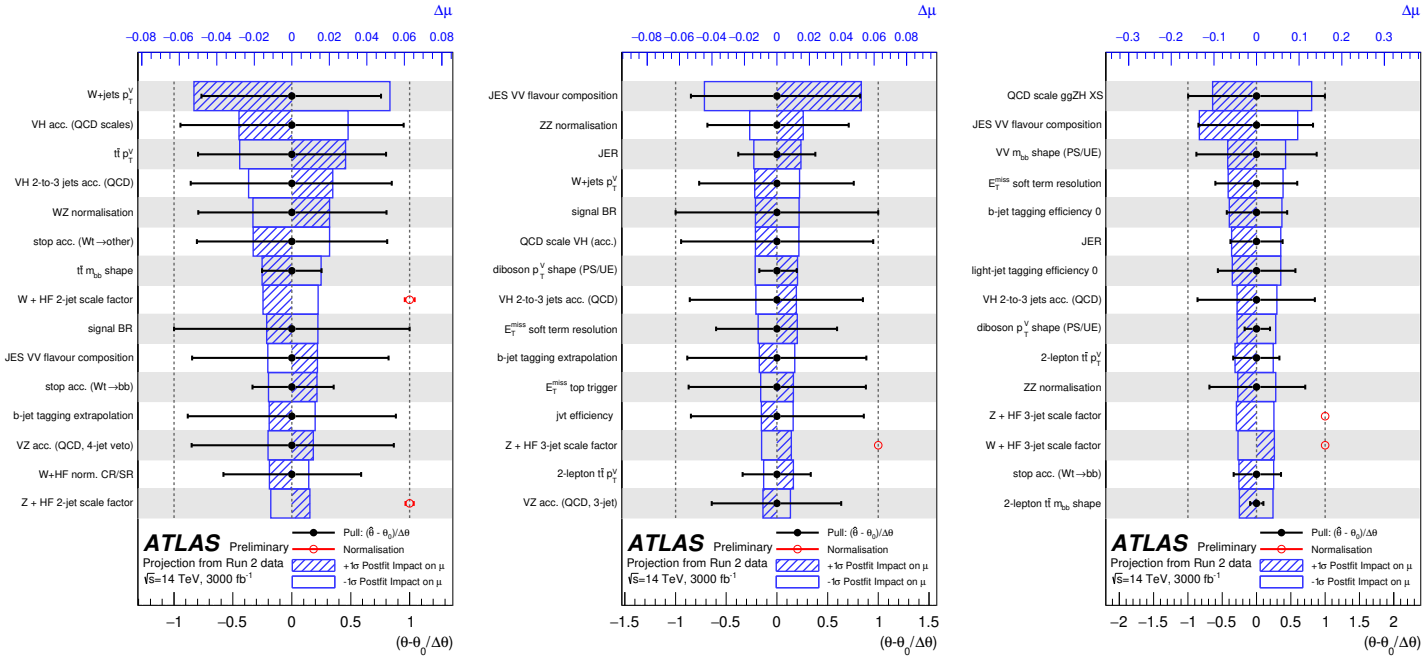


(a) σ_{VH} (b) μ_{VH}

Figure 7.2 – Ranking of the 15 systematic uncertainties (scenario S2) with the largest impact on the expected cross-section times branching ratio (left) and signal strength (right) measurement of the Vh inclusive production mode.



(a) σ_{WH} (b) $\sigma_{q\bar{q}\to ZH}$ (c) $\sigma_{gg\to ZH}$



(d) μ_{WH} (e) $\mu_{q\bar{q}\to ZH}$ (f) $\mu_{gg\to ZH}$

Figure 7.3 – Ranking of the 15 systematic uncertainties (scenario S2) with the largest impact on the expected cross-section times branching ratio (top) and signal strength (bottom) measurement of the $H \to b\bar{b}$ decay channel for the WH (left), $q\bar{q} \to ZH$ (middle), $gg \to ZH$ (right) production modes.

The positive aspect of this method is that it allows to group naturally NPs into a wide variety of classes, and in the limit of having one NP per group, it nicely converges to the expected impact of the NP on the POI. However, the other NPs are allowed to float to recover some of the impacts of the NPs in the group. Therefore the grouping should be implemented in a sub-phase space, where the NPs are not correlated with elements outside of the group. This causes the sum of the effects not to close as the total impact. To get perfect closure of the quadratic sum of the effects, a sequential algorithm can be used. This algorithm will assess fits by freezing NPs from the different groups one after the other and compute the difference in quadrature of the impact on the POI:

$$\Delta X_{\text{class n}} = \sqrt{|\Delta X(\text{NP of class n} - 1 \text{ frozen})^2 - \Delta X(\text{NP of class n} - 1 \text{ \& n frozen})^2|}$$

By construction, the sums of all the effects give the total impact of the systematics. However, the order of the fits is causing the individual values to change. This is illustrated using the Run-2 $\mathcal{L} = 79.8\ fb^{-1}$ results when applying the different types of breakdown and when changing the order of the categories. The main lesson learned is that background modelling and experimental systematics are found to be quite correlated since a difference of impact in one of the categories is compensated by the other one. The order chosen by the ATLAS Higgs combination group is the following: background, signal, experimental systematics. On the contrary, the CMS experiment chose to use the parallel breakdown algorithm [207].

	Background (B)		Signal (S)		Experimental (E)	
order	+	-	+	-	+	-
parallel breakdown						
	0.092	0.091	0.082	0.052	0.021	0.014
sequential breakdown						
B/S/E	0.092	0.091	0.082	0.056	0.040	0.035
B/E/S	0.092	0.091	0.082	0.056	0.041	0.035
S/B/E	0.093	0.093	0.082	0.052	0.040	0.035
S/E/B	0.099	0.098	0.082	0.052	0.022	0.014
E/S/B	0.099	0.098	0.082	0.052	0.021	0.014
E/B/S	0.099	0.096	0.082	0.056	0.021	0.014

Table 7.4 – Positive (+) and negative (-) impacts on the 1-POI signal strength fit using the Run-2 result as an illustration of the different results obtained with the two breakdown techniques. In the case of the sequential breakdown the order of the fits is indicated in the first column by the initial of the group used in 1/2/3

The measured products of the cross-section times the $H \to b\bar{b}$ branching ratio for each signal process, are the following, where the uncertainties outside (inside) the

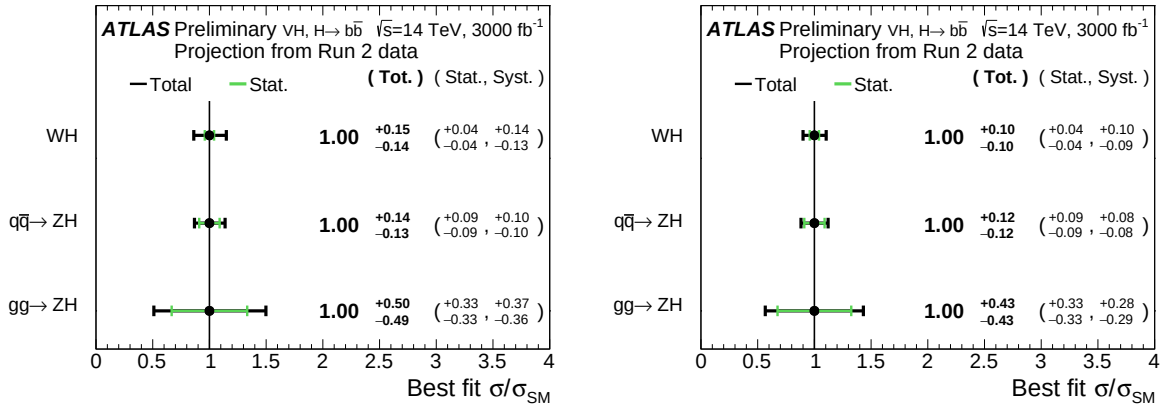
parentheses correspond to scenario S1 (S2):

$$\begin{aligned}
\sigma(WH)\cdot \mathrm{BR}(H\to b\bar{b}) &= 0.877\,^{+0.131}_{-0.121}(^{+0.091}_{-0.088})\,\mathrm{pb} \\
&= 0.877\,^{+0.036}_{-0.036}(^{+0.036}_{-0.036})\,(\mathrm{stat})\,^{+0.042}_{-0.041}(^{+0.039}_{-0.038})\,(\mathrm{exp}) \\
&\qquad ^{+0.070}_{-0.061}(^{+0.040}_{-0.036})\,(\mathrm{sig})\quad ^{+0.095}_{-0.088}(^{+0.063}_{-0.061})\,(\mathrm{bkg})\ \mathrm{pb} \\
\sigma(q\bar{q}\to ZH)\cdot \mathrm{BR}(H\to b\bar{b}) &= 0.488\,^{+0.067}_{-0.064}(^{+0.059}_{-0.058})\,\mathrm{pb} \\
&= 0.488\,^{+0.044}_{-0.043}(^{+0.044}_{-0.043})\,(\mathrm{stat})\,^{+0.032}_{-0.031}(^{+0.028}_{-0.027})\,(\mathrm{exp}) \\
&\qquad ^{+0.030}_{-0.027}(^{+0.015}_{-0.014})\,(\mathrm{sig})\quad ^{+0.026}_{-0.023}(^{+0.023}_{-0.022})\,(\mathrm{bkg})\ \mathrm{pb} \\
\sigma(gg\to ZH)\cdot \mathrm{BR}(H\to b\bar{b}) &= 0.084\,^{+0.042}_{-0.041}(^{+0.036}_{-0.036})\,\mathrm{pb} \\
&= 0.084\,^{+0.028}_{-0.028}(^{+0.028}_{-0.028})\,(\mathrm{stat})\,^{+0.021}_{-0.021}(^{+0.017}_{-0.017})\,(\mathrm{exp}) \\
&\qquad ^{+0.015}_{-0.012}(^{+0.008}_{-0.007})\,(\mathrm{sig})\quad ^{+0.017}_{-0.018}(^{+0.015}_{-0.015})\,(\mathrm{bkg})\ \mathrm{pb}
\end{aligned}$$

These numbers are translated into relative precision in Table 7.5. Figure 7.4 summarises the expected precision of the measured cross-sections for the three production modes. Figure 7.5 shows the expected precision of the measured cross-sections when the gg and $q\bar{q}$ to ZH production modes are combined. It's worthwhile to note that in this second fit, the uncertainty on the inclusive ZH signal process is much smaller than the uncertainties on the single $q\bar{q}\to ZH$ and $gg\to ZH$ processes, due to correlations between their measurements. This result shows that, in both the scenarios, the Wh and the $q\bar{q}Z$h processes will be systematically limited, while the rarer gg$Z$h signal would still have an important contribution from the statistical uncertainty.

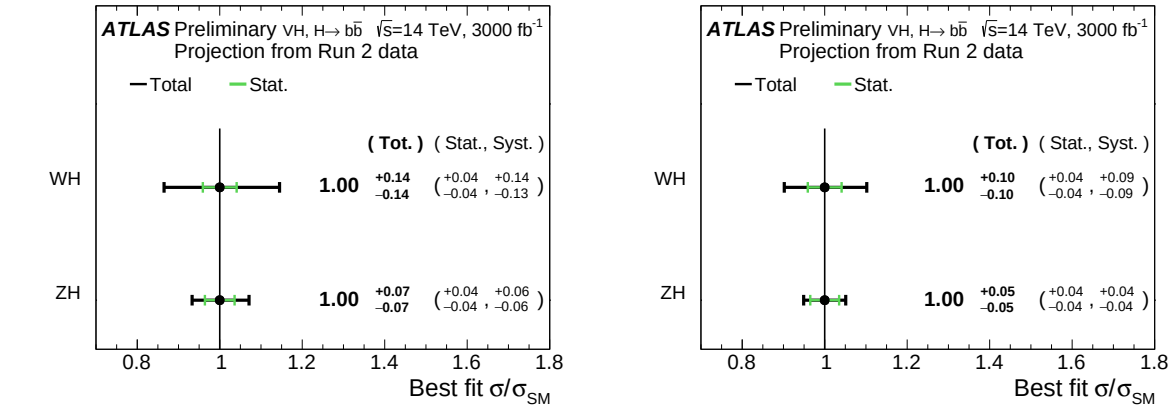
The results are then passed to the combination framework where a global fit is performed. The results are shown in Table 7.6, separated into the WH and ZH production cross-sections, the $\mathrm{b\bar{b}}$ Higgs branching ratio, and the product of the two. Uncertainties can be compared with the ones extracted with the $V\mathrm{h}(\mathrm{b\bar{b}})$ only fit, with fairly similar values at least in the Wh case, showing that the $V\mathrm{h}(\mathrm{b\bar{b}})$ is quite orthogonal to the other analyses and is therefore not impacted in the fit. The breakdown of uncertainties demonstrates the important contribution of the signal and background components in Scenario 1, while the improvements described in Scenario 2, lead to a more balanced contribution from the three categories.

The values of the kappa modifiers introduced earlier and related to the $V\mathrm{h}(\mathrm{b\bar{b}})$ analysis (W, Z, b) are gathered in Table 7.7. Similar conclusions as for the cross-section values are drawn. These results are obtained considering the photon and Zγ couplings and the gluon-fusion process to be point-like, avoiding the loop effects, and parametrised as single parameters.



(a) Scenario S1 (b) Scenario S2

Figure 7.4 – The fitted values of the Higgs boson cross-section divided by their SM values for the WH, $q\bar{q} \to ZH$ and $gg \to ZH$ processes expected with $\mathcal{L} = 3000\ fb^{-1}$ at the HL-LHC in the (a) scenario S1 and (b) S2 extrapolations. The individual cross-section values for the three processes are obtained from a simultaneous fit in which the cross-section parameters for the WH, $q\bar{q} \to ZH$ and $gg \to ZH$ processes are floating independently in a so-called 3 POI fit.



(a) Scenario S1 (b) Scenario S2

Figure 7.5 – The fitted values of the Higgs boson cross-section divided by their SM values for the WH and ZH processes expected with $\mathcal{L} = 3000\ fb^{-1}$ at the HL-LHC in the (a) scenario S1 and (b) S2 extrapolations. The individual cross-section values for the two processes are obtained from a simultaneous fit in which the cross-section parameters for the WH and ZH processes are floating independently in a so-called 2 POI fit.

Prod. mode	Scenario	$\Delta_{\text{tot}}/\sigma_{\text{SM}}$	$\Delta_{\text{stat}}/\sigma_{\text{SM}}$	$\Delta_{\text{exp}}/\sigma_{\text{SM}}$	$\Delta_{\text{sig}}/\sigma_{\text{SM}}$	$\Delta_{\text{bkg}}/\sigma_{\text{SM}}$	$\Delta\mu_{\text{sig}}$
WH	Run 2	$^{+0.462}_{-0.425}$	$^{+0.272}_{-0.265}$	$^{+0.157}_{-0.127}$	$^{+0.176}_{-0.075}$	$^{+0.224}_{-0.213}$	$^{+0.180}_{-0.077}$
	HL-LHC S1	$^{+0.149}_{-0.138}$	$^{+0.041}_{-0.041}$	$^{+0.048}_{-0.047}$	$^{+0.080}_{-0.070}$	$^{+0.108}_{-0.100}$	$^{+0.085}_{-0.074}$
	HL-LHC S2	$^{+0.104}_{-0.100}$	$^{+0.041}_{-0.041}$	$^{+0.044}_{-0.043}$	$^{+0.046}_{-0.041}$	$^{+0.072}_{-0.069}$	$^{+0.050}_{-0.045}$
$q\bar{q} \to ZH$	Run 2	$^{+0.667}_{-0.629}$	$^{+0.578}_{-0.562}$	$^{+0.129}_{-0.101}$	$^{+0.175}_{-0.105}$	$^{+0.143}_{-0.126}$	$^{+0.180}_{-0.105}$
	HL-LHC S1	$^{+0.138}_{-0.132}$	$^{+0.090}_{-0.089}$	$^{+0.065}_{-0.063}$	$^{+0.061}_{-0.055}$	$^{+0.054}_{-0.048}$	$^{+0.067}_{-0.059}$
	HL-LHC S2	$^{+0.121}_{-0.118}$	$^{+0.090}_{-0.089}$	$^{+0.057}_{-0.055}$	$^{+0.031}_{-0.028}$	$^{+0.048}_{-0.046}$	$^{+0.037}_{-0.033}$
$gg \to ZH$	Run 2	$^{+2.629}_{-2.608}$	$^{+2.105}_{-2.105}$	$^{+0.606}_{-0.677}$	$^{+0.658}_{-0.454}$	$^{+1.012}_{-1.037}$	$^{+1.269}_{-0.645}$
	HL-LHC S1	$^{+0.498}_{-0.490}$	$^{+0.333}_{-0.333}$	$^{+0.249}_{-0.250}$	$^{+0.181}_{-0.140}$	$^{+0.207}_{-0.218}$	$^{+0.495}_{-0.209}$
	HL-LHC S2	$^{+0.432}_{-0.433}$	$^{+0.333}_{-0.333}$	$^{+0.208}_{-0.204}$	$^{+0.096}_{-0.080}$	$^{+0.177}_{-0.181}$	$^{+0.222}_{-0.115}$

Table 7.5 – Expected precision of the production-mode cross-section measurements in the WH, $q\bar{q} \to ZH$ and $gg \to ZH$ production modes for the $H \to b\bar{b}$ decay channel with $\mathcal{L} = 79.8\ fb^{-1}$ of Run-2 data and at HL-LHC. Uncertainties are reported relative to the SM cross-section at the corresponding centre-of-mass energy. Both HL-LHC scenarios have been considered for the systematic uncertainties. The last column shows the theory uncertainty component when the measurement parameters are production mode signal strengths instead of cross-sections.

7.2 Technical prospects on the ATLAS experiment for the HL-LHC: the Inner Tracker

The experimental data-taking conditions during the HL-LHC phase will require a significant upgrade of the present detectors. The goal is to reach at least the same reconstruction performance of the present detector in the HL-LHC but with more difficult experimental conditions. The can be achieved thanks to the fast development of new technologies in the field of solid-state detectors. The two main considerations entering in the design are the higher number of interactions per collision, and the higher level of radiation in the detector area. Figure 7.6 shows the integrated radiation predicted by

POI	Scenario	Δ_{tot}	Δ_{stat}	Δ_{exp}	Δ_{sig}	Δ_{bkg}
$\sigma_{\text{WH}}/\sigma_{\text{WH}}^{\text{SM}}$	HL-LHC S1	$^{+0.095}_{-0.092}$	$^{+0.041}_{-0.040}$	$^{+0.041}_{-0.039}$	$^{+0.053}_{-0.048}$	$^{+0.055}_{-0.054}$
	HL-LHC S2	$^{+0.078}_{-0.076}$	$^{+0.041}_{-0.040}$	$^{+0.035}_{-0.034}$	$^{+0.034}_{-0.031}$	$^{+0.045}_{-0.045}$
$\sigma_{\text{ZH}}/\sigma_{\text{ZH}}^{\text{SM}}$	HL-LHC S1	$^{+0.063}_{-0.061}$	$^{+0.034}_{-0.034}$	$^{+0.025}_{-0.024}$	$^{+0.035}_{-0.033}$	$^{+0.031}_{-0.030}$
	HL-LHC S2	$^{+0.049}_{-0.048}$	$^{+0.034}_{-0.034}$	$^{+0.018}_{-0.018}$	$^{+0.020}_{-0.019}$	$^{+0.022}_{-0.021}$
$\text{BR}_{\text{bb}}/\text{BR}_{\text{bb,SM}}$	HL-LHC S1	$^{+0.079}_{-0.072}$	$^{+0.020}_{-0.020}$	$^{+0.025}_{-0.024}$	$^{+0.052}_{-0.047}$	$^{+0.050}_{-0.045}$
	HL-LHC S2	$^{+0.052}_{-0.049}$	$^{+0.020}_{-0.020}$	$^{+0.020}_{-0.019}$	$^{+0.029}_{-0.027}$	$^{+0.032}_{-0.031}$
$\sigma(\text{WH}, H \to bb)/\sigma_{\text{SM}}$	HL-LHC S1	$^{+0.146}_{-0.135}$	$^{+0.044}_{-0.043}$	$^{+0.050}_{-0.048}$	$^{+0.078}_{-0.068}$	$^{+0.104}_{-0.097}$
	HL-LHC S2	$^{+0.102}_{-0.099}$	$^{+0.044}_{-0.043}$	$^{+0.042}_{-0.040}$	$^{+0.044}_{-0.040}$	$^{+0.070}_{-0.068}$
$\sigma(\text{ZH}, H \to bb)/\sigma_{\text{SM}}$	HL-LHC S1	$^{+0.071}_{-0.068}$	$^{+0.035}_{-0.035}$	$^{+0.027}_{-0.026}$	$^{+0.042}_{-0.038}$	$^{+0.037}_{-0.035}$
	HL-LHC S2	$^{+0.052}_{-0.051}$	$^{+0.035}_{-0.035}$	$^{+0.020}_{-0.019}$	$^{+0.022}_{-0.021}$	$^{+0.024}_{-0.024}$

Table 7.6 – Expected uncertainties on the cross-sections, branching ratios and their products normalised to their SM predictions for both systematics scenarios S1 and S2, in the context of a global fit with all the processes.

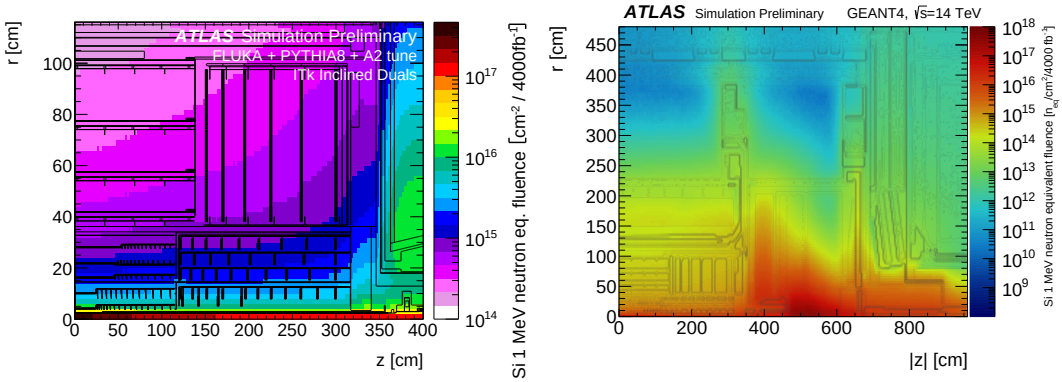
POI	Scenario	Δ_{tot}	Δ_{stat}	Δ_{exp}	Δ_{sig}	Δ_{bkg}
κ_W	HL-LHC S1	$^{+0.032}_{-0.031}$	$^{+0.008}_{-0.008}$	$^{+0.014}_{-0.013}$	$^{+0.019}_{-0.019}$	$^{+0.020}_{-0.019}$
	HL-LHC S2	$^{+0.022}_{-0.022}$	$^{+0.008}_{-0.008}$	$^{+0.012}_{-0.011}$	$^{+0.012}_{-0.011}$	$^{+0.013}_{-0.012}$
κ_Z	HL-LHC S1	$^{+0.026}_{-0.025}$	$^{+0.008}_{-0.009}$	$^{+0.011}_{-0.011}$	$^{+0.019}_{-0.017}$	$^{+0.012}_{-0.012}$
	HL-LHC S2	$^{+0.018}_{-0.018}$	$^{+0.008}_{-0.009}$	$^{+0.009}_{-0.009}$	$^{+0.010}_{-0.010}$	$^{+0.008}_{-0.008}$
κ_b	HL-LHC S1	$^{+0.064}_{-0.060}$	$^{+0.016}_{-0.016}$	$^{+0.023}_{-0.022}$	$^{+0.038}_{-0.036}$	$^{+0.043}_{-0.040}$
	HL-LHC S2	$^{+0.044}_{-0.043}$	$^{+0.016}_{-0.016}$	$^{+0.020}_{-0.020}$	$^{+0.022}_{-0.021}$	$^{+0.029}_{-0.028}$

Table 7.7 – Expected uncertainties on each Higgs boson coupling modifier sensitive to the $Vh(b\bar{b})$ analysis for scenarios S1 and S2. No BSM contribution to the Higgs boson total width is considered.

simulations in different parts of the detector, with the detail of the tracker region. The damage is parameterized in terms of 1-MeV equivalent neutron fluence, defined as the flux of 1-MeV neutrons which would produce the same displacement damage in Silicon lattice as the mixture of different particles with given energy distributions predicted by simulation over the detector lifetime.

Studies were made in Run-2 to compare the integrated fluence measured with dedicated sensors in different parts of the detector and predictions from simulations. Results, normalised by luminosity are shown in Figures 7.7 and 7.6. These studies indicate that the predicted rates at the end of the HL-LHC phase, despite being a bit conservative

compared to the real measurements obtained in Run-2, will involve the replacement of the actual tracker. The new system, so-called Inner Tracker (ITk) project [208], will have a better spatial resolution thanks to the smaller pixel pitch, and an increased rate readout capability and radiation hardness, to withstand a level of radiation one hundred times larger of what affects the present tracking detector. The expected fluence also indicates the need for the replacement of the readout electronics of the Liquid Argon and the Hadron calorimeters: the former one will have its readout channel upgraded, while the latter one will have its mechanical structure as well modified [201]. The muon systems are also going to be upgraded in order to cope with the higher rate (more recent electronics and smaller pitch of the detector), but also to extend the volume of detection up to $\eta < 4$ [209]. A new High Granularity Timing Detector (HGTD), a calorimeter whose sensitive layers are made with Low Gain Avalanche Diodes (LGAD) based sensors, should be inserted in the end-cap part ($2.4 < \eta < 4$). This detector would be able to resolve the collision sub-structure of the different beam bunches thanks to its good timing resolution $\sigma \simeq 30ps$, hence providing a good way to reduce the pile-up, as well as improving the object reconstruction in the forward region.

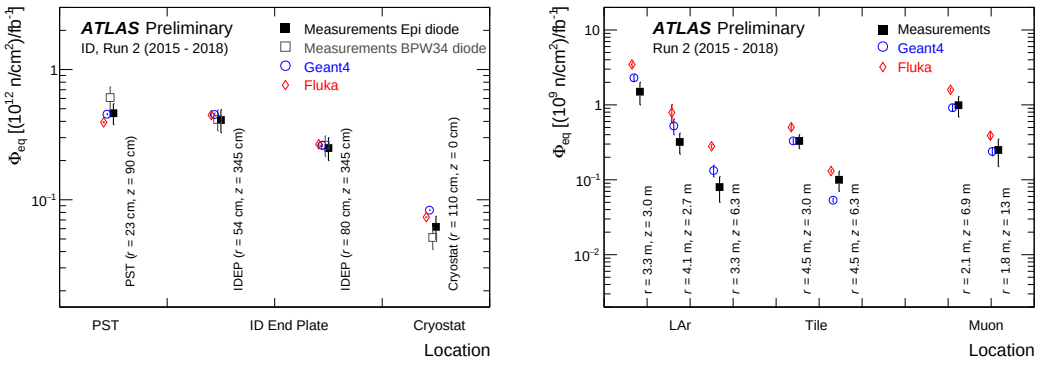


(a) (b)

Figure 7.6 – Prediction of the 1-MeV neutron equivalent fluence expected for $\mathcal{L} = 4000\ fb^{-1}$ of integrated luminosity in the ATLAS full (a) and Inner Tracker (b) detector. The minimum-bias pp events are simulated with PYTHIA 8 using A2 tune variation at $\sqrt{s} = 14$ TeV centre-of-mass energy and a predicted inelastic cross-section of 80 mb. Particle tracking and interactions with material are simulated with the GEANT4 [210] (a) and the FLUKA 2011 [211] code using the latest geometric description available [212].

7.2.1 The requirements and design of the ITk pixels

The current ATLAS tracking system is composed of Silicon- and Gas-based systems to detect the crossing of electrically charged particles. With the HL-LHC expected conditions, only a pure Silicon-based system is considered. The number of pixel layers will be increased to five, surrounded by four strip layers, and the full tracker will be approximately of the same size as the present one. The system is also expected to get



(a) (b)

Figure 7.7 – Summary of measurements and simulations of equivalent fluence per unit of integrated luminosity in the Inner Detector during Run-2. Measurements are averages from sensors at same (r, z) but at different azimuth angles. Error bars include variation of dose, integrated luminosity ratios during Run-2, variations between sensors and calibration. Neutron equivalent fluence is measured with two types of sensors at each location: BPW34 diodes (forward bias) and epitaxial diodes (reverse bias). Error bars on simulation (Geant4 and Fluka) points are standard deviations of simulated doses and fluences per fb^{-1} in intervals of coordinates around monitoring location [212].

extended up to a coverage of $\eta < 4$ as shown in Figure 7.8. The current modules are either placed parallel to the beam pipe in the barrel region or as a ring system in the end-caps. The novelty of the ITk design is that the outermost modules in η in the barrel region are inclined with a different tilting per layer, as can be seen in the lower part of Figure 7.8. This layout will lead to a reduction of the amount of material crossed by particles, being these modules essentially perpendicular to the tracks coming from the interaction point, even at high η in the detector. This will also lead to an optimisation of tracking performance.

To cope with a higher charged track density, the size of the Pixels are required to shrink as well. From an actual pixel size of $400\mu m \times 50\mu m$ (or $250\mu m \times 50\mu m$ for the IBL sensors) two options are evaluated: $50\mu m \times 50\mu m$ or $25\mu m \times 100\mu m$. The reduction of the pixel size allows reducing the cells capacitance, occupation (and thus dead-time), leakage current as well as improving the track resolution. However, it comes with read-out issues since the addition of channels would increase the power consumption and the bandwidth necessary to send all the data.

The effects of radiation

When a particle created during a collision is passing through the Silicon material composing one of the sensors, in addition to the production of electron-hole pairs which generate the signal, it can interact with the lattice structure of the crystal (microscopic effects), thus modifying the global properties of the sensor (macroscopic effects).
In the bulk of the sensor, massive particles such as neutrons, pions, electrons and protons

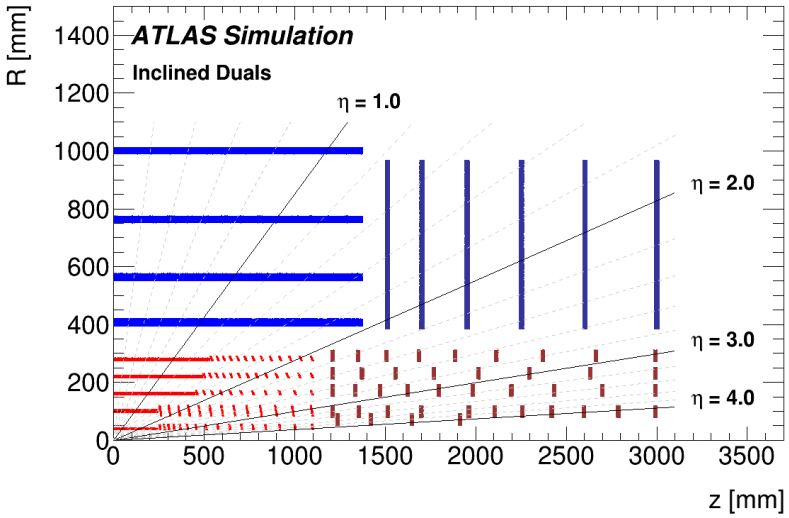


Figure 7.8 – Schematic layout of the ITk for the HL-LHC phase of ATLAS. The active elements of the barrel and end-cap Strip Detector are shown in blue, for the Pixel Detector the sensors are shown in red for the barrel layers and in dark red for the end-cap rings. Here only one quadrant and only active detector elements are shown. The horizontal axis is the axis along the beamline with zero being the interaction point. The vertical axis is the radius measured from the interaction region. The outer radius is set by the inner radius of the barrel cryostat that houses the solenoid and the electromagnetic calorimeter[208].

can displace atoms in the lattice, causing damages described under the Non-Ionising Energy Loss (NIEL) hypothesis. Such effects can be normalised for the energy or mass of the incoming particle using a hardness factor [213], resulting in a global flux ϕ_{eq} of equivalent 1-MeV neutrons as presented in Figures 7.6 and 7.7. The deformations of the lattice can be classified in different categories but generally speaking they create a vacancy at the initial position of the atom, and a new local structure in the inter lattice space at the final position of the atom, composing what is called a Frenkel defect pair. Depending on the energy and nature of the impinging particle, such effects can cluster, as the initial atom can produce a global displacement of about 100 atoms. These defects are evolving with time and temperature in a complex way called annealing. Beneficial annealing, where atoms can occupy back their original positions, is obtained by increasing the temperature for a short period of time (typically 60°C during 80 minutes [214]). Holding the exposition longer could lead to reverse annealing where the effective doping concentration would increase. Particles can also interact strongly changing the nature of the atoms in the lattice increasing the number of donors or acceptors.

The microscopic effects are changing the operational characteristics of the sensors in different ways:

— **leakage current:** in semiconductors, an electron-hole generation process happens all the time at room temperature through thermal excitation and is compensated by recombination effects either direct (an electron from the conduction band recombine with a hole of the valence band), or through re-

combination centres represented by the dopants or lattice defects. In the case of an np junction, if an inverse polarization voltage is applied, a significant fraction of the full junction region is depleted of free charge and an electric field is present. Under the effect of the field, the electrons and holes are separated before they can recombine and this creates a leakage current. In presence of radiation damage, new energy levels are created in the bandgap, resulting in an increased generation process that is emphasised by a reduction of the recombination time. This leads to an increase in the leakage current:

$$\Delta I(\Phi_{eq}, t, T) = \alpha(t, T) V_{dep} \Phi_{eq}$$

where V_{dep} is the depleted volume, $\alpha(t,T)$ the current related damage rate (typically $\alpha \sim 4 \times 10^{-17}$ A/cm [214]), and Φ_{eq} the 1-MeV neutron equivalent flux.

— **depletion voltage:** since a certain number of donors or acceptors can be trapped into the defects, the effective concentration N_{eff} is modified. Furthermore, space-charges can also be created. These two effects are leading to a modification of the width of the depletion region for a given applied reverse potential, or in other words to the necessary potential which is necessary to apply to fully deplete the device. Consequently, the voltage at which the sensor is operated V_{dep} has to be modified :

$$V_{dep} = \frac{e|N_{eff}(\Phi_{eq})|d^2}{2\epsilon_r \epsilon_0}$$

where d is the sensor's thickness, ϵ_r and ϵ_0 the electrical permittivity of respectively the sensor and the vacuum. This voltage increase must be controlled to limit the resulting leakage current. The second effect is a possible type inversion from an n-type bulk to a p-type, leading to changes in the depletion zone, and thus a reduced electrical signal. This is why the ATLAS and CMS collaboration are now using an n-on-p technology to mitigate this effect.

— **charge collection efficiency:** The electrical signal obtained through the Ramo's theorem is linked to the charge drifting in the medium. However, the defects introduced in the crystal can trap the charges for a time longer than the charge collection one. The time before a particle is getting trapped is given by:

$$\tau_{trap.}^{-1} = \beta \Phi_{eq}$$

where β is a constant of $4-6 \times 10^{-16}$ cm^2/ns for electrons and $6-8 \times 10^{-16}$ cm^2/ns for holes [215]. Combining this information with the mobility that is more important for electrons than holes, it is preferred to collect electrons

than holes. The trapping can have a side effect by releasing charges after the signal is processed creating spurious signals.

Surface effects can also be present in Silicon oxides which lead to the creation of a layer of trapped charges at the interface between the silicon and the passivation. This results in a perturbation of the electrical field inside the sensor or an increase of the strip or pixel capacitance, affecting the noise. This effects can also induce problems in the dedicated electronics used to process the signal.

The electronic read-out

To record the signal induced in the electrodes, a read-out chip is directly connected to the pixel sensor using a bump-bonding solution: indium or lead-tin alloy spheres are used to make the connection. The chip that will be used for the ATLAS and CMS pixel tracker upgrades is based on a Complementary Metal Oxide Semiconductor (CMOS) technology and is the result of a collaboration between the CMS and ATLAS experiments and called RD53 [216]. The signal is treated as follows:

1. **Charge amplification:** as close as possible to the sensor, an analogue amplification is provided. In the preliminary RD53 chip prototype, called RD53A, three different technologies are proposed: the linear case proposing a single-stage amplification, a differential case where the signal is separated into two single-stage amplification, and a synchronous case which uses a single-stage amplification as in the linear case. The signal is then shaped with a triangular analogue signal.
2. **Signal discrimination:** To give information to the system about the signal which has been induced in the pixel, the signal is compared to a threshold. This comparison is also different in the three front-end cases: the linear case compares the signal to a flat threshold decided by the user, the differential compares the two differentially amplified signals, while the synchronous front end uses a time-dependent threshold that is tuned by the chip itself based on the running conditions. The user is free to choose the various parameters for the first two FEs (threshold and gains), and the choice would depend on the level of noise of the considered chip. Some local variations of the gain are also required to get a uniform channel-by-channel response across the chip. They are set by evaluating the response of the chip while injecting the same signal in all the pixels. This calibration is referred to as "tuning" and is described in Section 7.2.2.
3. **Hit storage and retrieval:** from the comparison of the signal to the baseline, the Time over Threshold (ToT) is computed. This time is quantised in the LHC bunch crossing time and digitised to be subsequently processed by the digital logic to use the latest techniques of data storing and compression.
4. **Trigger comparison:** The data from different neighbouring cells are gathered, and sent to the data controller when the trigger signal is received.

Radiation effects can originate from the same physics described earlier, resulting in a smaller amplification, and thus lowering the discriminating ToT value, but also

creating Single Event Upsets. The latter effect is due to ionisation in the chip and results in memory corruption. This is mitigated by a clever redundancy of registers.

The main difference with the existing FE-I4 [87] chip used in the ATLAS IBL, apart from the amplification stage that is using the linear design, is the in-pixel size ($50 \times 250 \mu m^2$ vs $50 \times 50 \mu m^2$). The $50 \times 50 \mu m^2$ pattern allows both the pixel geometries described earlier by smart design of the connection to the bonding pads in the sensor. The FE-I4 was based on a 130 nm CMOS technology that is demonstrated to work to a dose of at least 250 Mrad, but is not enough to withstand the 500 Mrad expected at the HL-LHC. For this reason, a 65 nm technology is used for the RD53 chip, to reduce the effects of radiation damage. The main drawback is the initial lack of developed libraries and design blocks, common in HEP but which were not yet available in 65 nm.
The first sensors compatible with the RD53 design were produced in 2018 however this thesis will present results obtained with the FE-I4 chip, the results with the RD53-A chip being still under study.

The sensor design

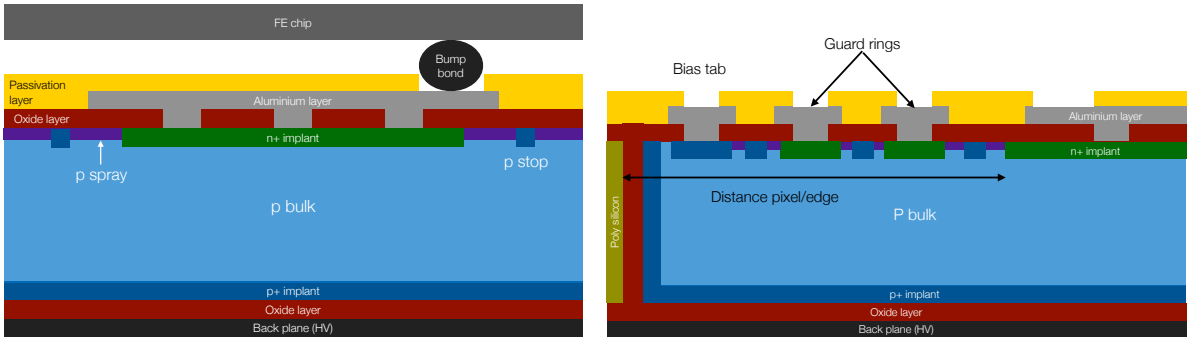
The conception of the sensors must keep into account the effects of the large expected radiation damage, as shown in Figure 7.6. The equivalent fluence is especially significant in the two innermost layers, given the proximity with the beam as shown in Table 7.8.

Luminosity $[fb^{-1}]$	Layer	Location	Fluence $[10^{15}\ n_{eq}/cm^2]$
2000	0	Barrel	13.1
		Inclined	12.3
		End-cap	6.8
2000	1	Barrel	2.7
		Inclined	3.5
		End-cap	3.8
4000	2-4	Barrel	2.8
		Inclined	3.0
		End-cap	3.8

Table 7.8 – The maximal 1-MeV neutron equivalent fluences for different parts of the Pixel Detector, for the baseline replacement scenario for the inner section. All values have been multiplied by a safety factor of 1.5 [208]

To moderate the radiation effects, a clever design of the sensor is required. The first solution that was proposed is to insert some electrically neutral elements that can combine to secondary defects and modify the radiation tolerance of the material. As an example, the inclusion of oxygen in the fabrication process [217] can lead to a better charged-hadrons tolerance, while no major impact has been observed for neutrons. The conception of the n- and p-implants also plays a major role in the measured tolerance.

The sensors designed and characterised at the LPNHE laboratory are based on the so-called planar pixel technology. An initial wafer obtained through the Float Zone technique (vertical growth that results in high purity ingots), is successively doped to obtain the result presented in the Figure 7.9a. From the bottom, a p+ implantation acts as a good ohmic contact and collects the holes. The p-doped bulk acts as the place for electron/hole creation. At the top, a complex structure is created through a lithography process using ion implantation. The n-doped layer acting as the counterpart of the p-type bulk in the n-p junction where the electrons are going to be collected is implanted using a complex sequence of masks to shape the pixels. However, at the interface between the oxide and the bulk between two pixels, a layer of charges can appear. This result in a possible short cut between the electrodes degrading the position resolution of the detector. Therefore additional thin implantation with p-type dopants could get obtained between the pixels (p-spray). A second solution called p-stop a p-type implantation between the pixels up to much higher concentration with respect to p-spray, but this requires an additional lithographic mask. The n-doped implants are later on connected to aluminium structures to collect the electrical signal. A passivation layer is covering the sensor to protect the sensor from dust and humidity, but also to avoid any electrical shortcuts between other parts the sensors and the read-out chip. The connection to the chip is obtained thanks to a bump-bonding technique.



(a) (b)

Figure 7.9 – Section (not to scale) of a planar sensor as designed of the ITk project. A zoom on one pixel is shown in (a), while (b) is focusing on the edge of the sensor.

The edges of the sensors are of particular interest being the un-instrumented area of the sensor, which could reduce the geometrical efficiency and the performance of the detector. Even if some overlap between neighbouring sensors is obtained by construction in the Φ direction, it is worth extending as much as possible the active area of the sensors. The separation between devices in the wafer is done with the help a diamond saw, which is creating microscopic cracks in the silicon, resulting in an important damaged area as illustrated in Figure 7.10a. The defects act as electron-hole generators as already described. If the electrical field is extended in this region, this charge generation gives rise to a large current (edge current). For this reason, the standard design of sensors tries to keep the electric field away from the edge, but this corresponds to having a dead region at the border. One way to avoid this problem is to perform the cut with a deep reactive ion etching (DRIE) technique. The trench reaching the bottom of the wafer, a mechanical

support is needed. It is further on doped with boron to reach the same potential on the side than at the bottom and filled with polysilicon as shown in Figure 7.9b. The technique, called "Active Edge", also implies a smaller number of generation centres at the edge, which are also prone to recombine quickly due to the small electrical field, reducing de-facto the induced leakage current. The designed trench can be straight or have more complicated shapes, such as the staggered trench presented in Figure 7.19b. The latter has the advantage not to require a priori a support wafer. To avoid an important voltage drop between the edge and the last pixel, free-floating guard rings are used to reduce smoothly the voltage difference, while maintaining a good efficiency even in this un-instrumented area. This effect is less significant with the active-edge technology though.

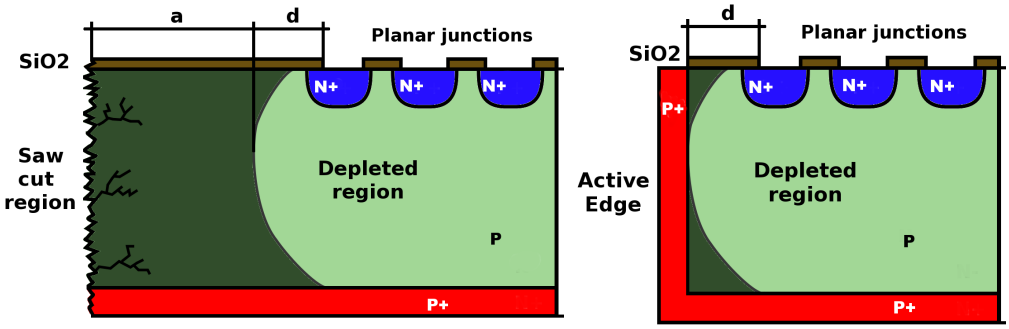


Figure 7.10 – Section of the edge of a sensor without (a) and with (b) active edge. The size of the dead area is parametrised by a, while the distance between the last pixel and the end of the active area is shown by the parameter d.

Before bump-bonding the sensors, tests need to be performed to ensure the good quality of the devices. Foundries where they are fabricated, are measuring the sensors to ensure good and homogenous properties amongst the production. The pixels are not yet connected to the readout electronics and so they cannot be brought to the correct potential to deplete the sensor. Anyway, the measurement can be achieved with two techniques. The punch-through mechanism has been widely used for that process. A punch-through dot is used on every pixel corner and connected to a metal line running across all the pixels to be kept at ground. However, this method requires to add permanent structures in the sensor, degrading locally the efficiency of charge collection. A less invasive approach, called temporary metal line, provided by some foundry, consists in shortening all the pixel together with a metal grid, which is used to bias the sensor in order to measure the electrical properties. This metal line can be removed with chemical etching afterwards, leaving the sensor intact.

Amongst the other techniques used to build a pixel detector, the 3-D approach leads to better radiation resistance. The sensors are equipped with heavily doped n+ and p+ pillars amongst which the difference of tension will be applied as shown in Figure 7.11. In this case, the charge produced by the crossing of a particle drifts horizontally between the pillars instead of vertically. The inter-electrode distance, given by the distance between columns, can be much smaller than in the case of planar pixels, where it is fixed to the

sensor thickness (reaching roughly 30 μm [218] compared to a thickness of 100 μm for planars). Thanks to this feature the leakage current is reduced, as well as the trapping effects. These sensors can thus be operated at a smaller tension and in harder radiation conditions. The technology was first introduced in the ATLAS experiment with the IBL project [85], where part of the layer was equipped with such detectors. For what concerns the ITk project, the innermost layer is going to be instrumented with sensors using this technology, while planars are used in the four remaining layers.

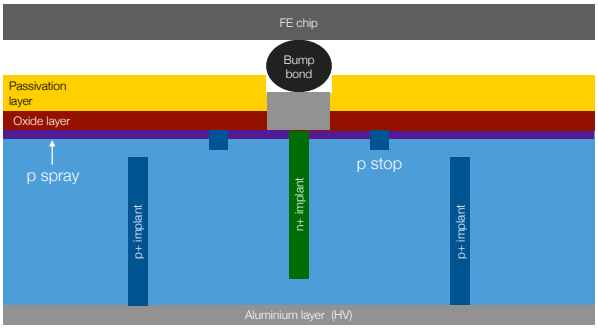


Figure 7.11 – Section (not to scale) of a 3d sensor as designed of the ITk project.

With the present technology, it is not possible to build a module which can resist the radiation conditions in the first two layers for the full length of the HL-LHC run while providing the expected performance. For this reason, a replacement of these modules is foreseen at half the HL-LHC run.

7.2.2 Measurements of the sensors' properties during test-beams

Amongst the technical solutions detailed previously, the LPNHE laboratory has collaborated with the Fundazione Bruno Kessler (FBK [1]) to design and produce several sensors with different technologies. Production 1 released the first Active Edge sensor produced at FBK, with a thickness of 200 μm. Production 2 aimed at reducing the thickness obtaining sensors at 100 and 130 μm, but without an active edge. Production 3 mixed the two features by producing 100 and 130 μm thin Active Edge sensors. The various productions are summarised with their intrinsic properties in Table 7.9. All the listed samples have been bump-bonded at the IZM Berlin to FE-I4B readout using $50 \times 250 \mu m^2$ pitched pixels.

To study the effect of radiation, the sensors from production 2 and 3 have been irradiated. The second production underwent staged irradiation campaigns at the CERN IRRAD [2] facility using a 24 GeV gaussian shaped proton beam. The beam characteristics are obtained with monitors during the irradiation and are summarized in Table 7.10. Since the width of the beam was smaller than the size of the sensor, different points across the

1. FBK-CMM (Trento, Italy): https://cmm.fbk.eu/
2. http://ps-irrad.web.cern.ch

	Production 1		Production 2		Production 3
Module name	LPNHE 5	LPNHE 7	W30	W80	M1.4
Thickness (μm)	200	200	100	130	130
Number of GR	0	2	5	2	0
Biasing	TM		PT		TM
Active Edge: · presence	✓(straight trench)		–		✓(staggered trench)
· Trench distance (μm)	100				37/52
Charge stopping: · p-spray	Yes				
· p-stop	Yes	No	No		Yes

Table 7.9 – Characteristics of the three sensor productions of the FBK/LPNHE collaboration. GR refers to Guard Rings. The biasing refers to the biasing process used to polarize the sensor before bump-bonding: TM stands for Temporary Metal and PT for Punch-Through.

beam spot corresponding to different values of the integrated fluence, thus allowing to perform fluence dependency measurements with a single sensor. Only the average value of the fluence is given in the table. The third production has been irradiated at the KIT [3] with a 25 MeV proton beam. In this facility, a narrow beam is scanning the sensor to get final uniform irradiation.

Module name	Beam spot size FWHM [mm^2]	fluence Φ_{eq} [10^{15} n_{eq}/cm^2]	cumulative fluence [10^{15} n_{eq}/cm^2]
Campaign 1			
W80	20×20	3	–
W30	12×12	4	–
Campaign 2			
W80	20×20	7	10
W30	20×20	7	11
Campaign 3			
M1.4	uniform	2.7	–

Table 7.10 – Irradiation programs for the different sensors tested at the test-beam

Electrical properties of the sensors

The electrical properties of the sensors are measured in a cleanroom environment at the LPNHE or during the test-beam campaigns for the irradiated modules. The temperature is controlled either by a thermal chuck in the laboratory or with a dedicated cooling system at the test-beam, always checked with platinum thermal probes. The result of the electrical characterization is presented in Figure 7.12 for the first production[219], and in Figure 7.13 for the second and third ones. For the first production, the sensors tested are coming from the same wafer but are not the sensors whose properties are reported afterwards. Even though the measurements have not been carried out at the same temperature, the effects of adding Guard Rings can be seen on the breakdown voltage that

3. https://www.etp.kit.edu/english/irradiation_center.php.

increased from 70 to 120 V. The only effect of the temperature can be seen from the value of the plateau for the leakage current. It is worth mentioning also that the depletion voltage is roughly stable at 20 V for the three solutions.

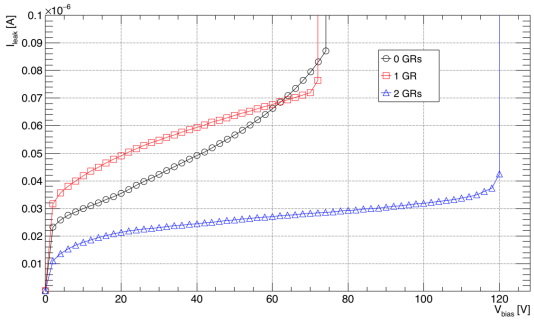


Figure 7.12 – Current-Voltage curves for test structures featuring different number of GRs. The innermost GR, if present, was kept at ground voltage. The shortest distance from the pixels to the trench is 100 μm. The measurement for the test structure with 2 GRs was taken at a lower temperature with respect to the other two samples.

For the production 2 and 3, the radiation effects on the sensors are generating both a shift in the depletion voltage and an increase in the leakage current as described earlier. For this purpose, the measurements presented are done with the readout chip turned on, taking into account the heat dissipation. For the last production, the sensor is shown to reach an early breakdown at 90 V which is suspicious when compared to the depletion and breakdown voltage of the thin sensors of the production 2. From the measurements the alpha coefficient relating the leakage current to the fluence can be derived:

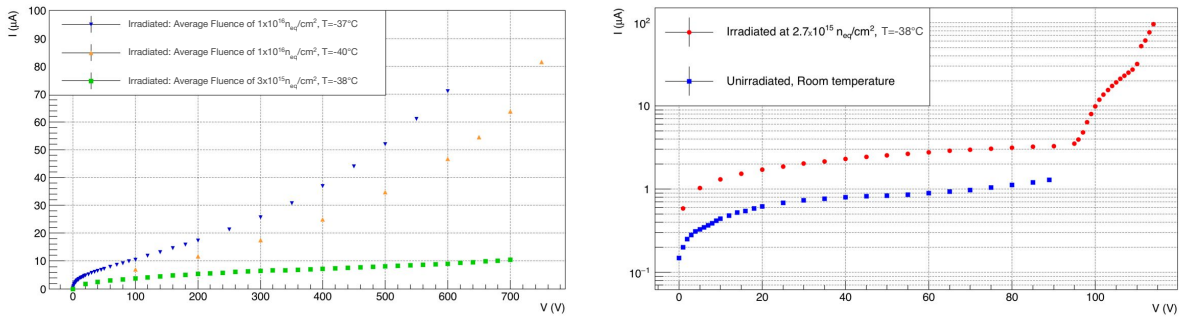
$$\Delta I(\Phi_{eq}, t, T) = \alpha(t, T) V_{dep} \Phi_{eq}$$

The leakage current has an important variation with the temperature as can be seen from Figure 7.13a, where for the same fluence of 1×10^{16} n_{eq}/cm^2 and at a bias of 600 V, the leakage current goes from 47 μA at -40°C to 71 μA at -37°C. This can be explained by the distribution dependence of charge carriers with temperature, as well as the self-heating of the sensor. The I-V curve is also not flat in the depleted regime. Therefore the measurement is performed at a fixed bias of 600 V for production 2 and 90 V for the production 3. The value of the temperature of -38°C is also considered in the following computations.

The literature provides an α coefficient at 20°C measured with a 60 minutes 80°C annealing of $(3.99 \pm 0.03) \cdot 10^{-17}$ A.cm^{-1}[214]. To calibrate the measurements of the leakage current at the desired temperature, a parametrisation is given in [220]:

$$\alpha(t, T_0) = \alpha(t, T) \cdot \left(\frac{T_0}{T}\right)^2 e^{\left(E_{eff} \frac{T_0 - T}{2 \cdot k_B \cdot T \cdot T_0}\right)} = \frac{\Delta I(\Phi_{eq}, t, T)}{V \cdot \Phi_{eq}} \cdot \left(\frac{T_0}{T}\right)^2 e^{\left(E_{eff} \frac{T_0 - T}{2 \cdot k_B \cdot T \cdot T_0}\right)}$$

The effective activation energy is close to the gap energy but gets modified by the trap energies of the bulk, and some temperature effects that result in different values found for



(a) (b)

Figure 7.13 – Current-Voltage curves of (a) W80 sensor after a fluence of 3×10^{15} n_{eq}/cm^2 (green markers) and after an cumulative fluence of 1×10^{16} n_{eq}/cm^2 (blue and yellow markers) (b) the M1.4 sensor before and after irradiation. The temperatures are specified in the legends.

this parameter. Thus two hypotheses are conducted choosing $E_{eff} = 1.21$ eV from [220] or the more recent value that has been measured $E_{eff} = 1.12$ eV [221]. The detector was considered fully depleted, such that the active volume has a depth 130μm and size of 2 × 1.68 cm^2. The borders of the module are not considered here, knowing that the effect is quite small in magnitude (~ 0.1) compared to the other sources of uncertainties. An uncertainty coming from the temperature measurement is quoted by varying the real temperature by $\Delta T = \pm 1^{\circ}$C. The values of the coefficients are summarised in the Table 7.11. The values obtained for W80 are highly compatible with the one measured in the reference, however an increased value for a higher dose could be linked to some not-well monitored annealing processes. The value for the third production is indicating that the sensor is not fully depleted at this bias voltage as corroborated by the previous IV curve measurements.

Sensor	$\Phi_{eq}[10^{15} n_{eq}/cm^2]$	$E_{eff} = 1.12$ eV	$E_{eff} = 1.21$ eV
W80	3	2.6 ± 0.4	4.0 ± 0.6
	10	5.3 ± 0.7	8.2 ± 1.2
M1.4	2.7	0.94 ± 0.1	1.46 ± 0.2

Table 7.11 – Values of the α coefficient for several modules at exposed at different fluences and rescaled at the temperature of 20°C. Two hypotheses for the effective energy are proposed, and the uncertainty comes from the variation of the temperature at which the current is measured.

The power dissipation of the sensor and the chip can be also measured, as it is a crucial figure of merit to design the required cooling power of the system [208]. This dissipation is obtained from the product of the leakage current and the bias voltage. The measurement can be rescaled to the desired temperature using the same formula as for the α coefficient, to allow comparisons at an operating temperature similar to the current IBL of -25°C. For W80, at 600 V, the power dissipation is 6 mW/cm^2 when irradiated at 3×10^{15} n_{eq}/cm^2 and 40 mW/cm^2 when irradiated at 1×10^{16} n_{eq}/cm^2, while M1.4

has a power dissipation at 90 V of 0.4 mW/cm^2 when irradiated at 2.7×10^{15} n_{eq}/cm^2. The requirement for the maximal power consumption of the sensor is 100 mW/cm^2 [208] with a current per pixel of 10 nA. Therefore the modules from production 2 are meeting the requirements. The power drained by the module from the third production is way under-estimated due to the low value of the bias voltage, thus no significant conclusion can be drawn from this measurement.

Test-beam measurement of the sensors' properties

In order to understand the performance of the sensors designed, designed sensors, they have to be tested in the same conditions as during the data-taking, with a controlled environment. This is usually done with a beam of highly-energetic particles obtained with secondary emissions from known accelerators. The results presented here are obtained at the CERN SPS and DESY II [222] beam facilities.
The SPS accelerator is providing a 400 GeV proton beam, that is sent to three primary targets to extract a final beam of 120 GeV pions. The CERN North Area site of the accelerator is providing up to four beamlines on which several experiments are located, and two dedicated lines for the NA62 [223] and COMPASS [224] experiments. As the priority is given to the operation of the LHC, the beam can be quite unstable but is optimised to deliver an average of 400 to 500 Hz rate.
The DESY II accelerator is a derived line of the PETRA III injector providing highly energetic electrons. Thanks to the bremsstrahlung effects, those electrons and positrons can radiate photons. Using a target, those photons can generate a quasi-continuous flow of electrons and positrons. A system of magnets can select the sign of the charge and the energy of the beam ranging from 3 to 6 GeV. A compromise between the energy of the electrons and the rate has to be found, such that most of the presented results are using a 4 GeV beam. At this energy however the particles are subject to multiple scattering degrading the spatial resolution of the measurements.
In order to measure the properties of the sensors, the trajectories of the incoming particles need to be known to extrapolate their hit positions in the devices under test (DUT). EUDET/AIDA telescopes are therefore used up- and down-stream of the position of the DUTs for that purpose. Each arm of the telescope is composed of three planes equipped with Mimosa26 monolithic active pixel sensors, with a fine pitch of 18.4 μm by 18.4 μm. These fine granularity sensors come with an important intrinsic time resolution (of the order of 200 μs, compared to the 25 ns clock of the FEI4 and RD53 FE chips). Therefore the time coincidence of the tracks is obtained by placing a reference FEI4 sensor next to the DUTs to reject out-of-time tracks. To avoid the continuous read-out of the sensors, a trigger unit is set up thanks to plastic scintillators whose coincidence signal is sent to the DUTs, the reference and the telescopes. In order to control the temperature, two systems are used depending on the test-beam facility. A chiller is used at CERN, allowing a fine control at the degree level, while dry ice (carbon dioxide, -78°C) is used at DESY, with a thermal sensor attached to the DUTs to measure the temperature over time. In order to provide a good thermal and optical sealing of the environment containing the sensors,

two different boxes are conceived as well. In both situations, nitrogen gas is flown in the box to avoid water to condensate on the module which could create shortcuts.

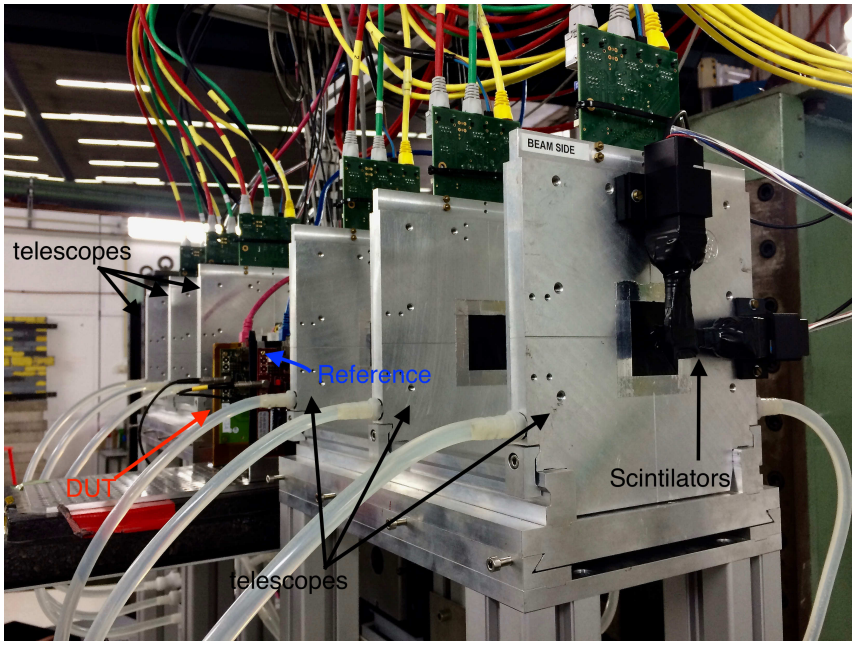


Figure 7.14 – Photography of the setup used in the test-beam area located at DESY. The six EUDET/AIDA telescope planes as well as the trigger plastic scintillators are shown in black, the DUT in red and the reference plane in blue.

In order to control the acquisition system, a remote PC is used. It interfaces the user with the readout boards connected to the modules under test via a graphical interface, called USBPix, controlling the trigger as well as the data acquisition. Before taking data, the FE chips have to be tuned to give an optimized and uniform performance across the sensor for the desired threshold and ToT/delivered charge relationship. The threshold dispersion in the sensor can be corrected using a global register for the full sensor with a 5-bits local register called TDAC to correct for pixel-by-pixel effects. Once the threshold is fixed, the ToT distribution can get shifted such that its most probable value corresponding to the charge left by a MIP is well centred inside the dynamic range of the chip. This operation is performed by modifying the charge to analogue ToT conversion at the pre-amplifier stage, using a 4-bits local register called FDAC.
Once a particle is passing through the six planes, a track can be reconstructed, by combining the information of the telescopes and the reference. This reconstruction is done offline using the EUTelescope framework [225]. Several steps are needed:

1. First a bank of noisy pixels is created for the DUT as well as the telescopes and reference, the last two being generally less affected by noise. All pixels showing hits at a frequency above a certain threshold (chosen by the user, and typically kept at 0.5% of the trigger frequency) are discarded for later analysis.

2. Hit pixels are then grouped to form clusters where the local centroid position is computed from the ToT values.

3. An alignment step is performed using the Millipede algorithm [226], which considers all the planes independent from each other and minimise the global χ^2 of the tracks.
4. A final fit is performed with a Kalman-filter based technique to take into account the multiple scattering in the DUTs and the materials crossed by the tracks.

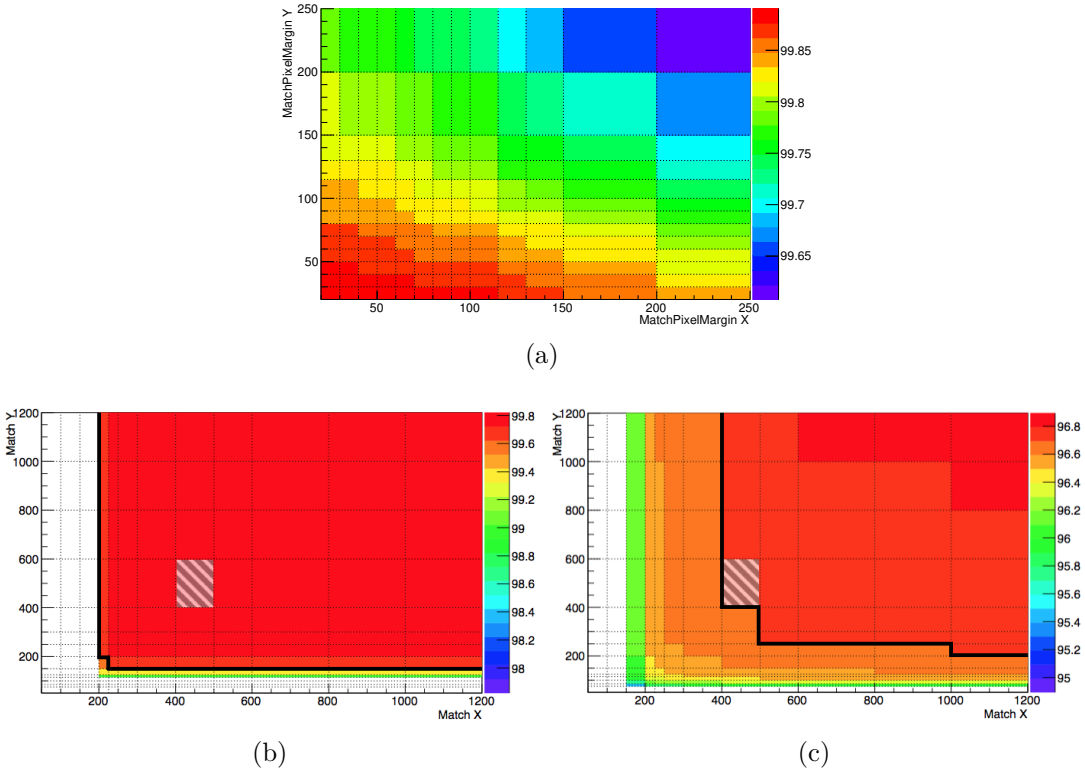
The result is obtained in a root interpreted file and analysed with the TBMon2 [227] framework to extract the useful information to the analyser. Amongst all the variables of particular interest from the development point of view, only the hit efficiency is going to be described here. This quantity is one of the sensor specifications defined for the HL-LHC tracker, where a 97% efficiency should be reached during all the HL-LHC run. It is defined by the fraction of good reconstructed tracks being matched to a hit. Two criteria are defined for a proper matching:

— To define the fiducial regions where tracks are accepted, windows of the pixel size are defined by a quantity called MatchPixelMargin. Having a window size bigger than the pixel size does not affect the results inside the sensor, but is crucial for edge studies where tracks can be located much further away from the pixel area.

— The hits considered in the analysis are weighted clusters of fired pixels. A window of a given size is opened around the cluster and the closest track to the centroid in the window is matched.

Those parameters can be optimised by looking at the evolution of the efficiency with respects to their values as shown in Figure 7.15. A small degradation ($\sim (O)(0.1\%)$) of the hit efficiency is observed when increasing the fiducial volume, as expected by the increase of non-matched tracks. This should be accounted for when considering the edge effects. As expected the global efficiency increases with respect to the matching window size. The effect of the detector resolution and to the multiple scattering which makes the tracking more problematic, are possible interpretations. A plateau is reached but with different values of the parameters for the two devices. After the turning-point, the efficiency is still increasing mainly driven by bad matching. The effect ($\sim (O)(0.1\%)$) however is under the statistical fluctuation of the data-taking but is considered genuine. All those elements are considered in the measurement by taking a conservative error (from statistics with an added systematic 0.4% term).

The hit performance of the first production has been documented in [219]. The studies have been focussed on the effect of the threshold and the active edge technology on the efficiency. The Figure 7.16a is showing the evolution of the global efficiency over the sensor as a function of the bias voltage. Different ToT to charge calibrations are tried, not affecting the observed measurements. Biasing the detector above a threshold of 25 V leads to efficiencies greater than 98%. The behaviour of the evolution of the efficiency with the position from the last pixel is obtained by focusing the beam on this part of the sensor and shown in Figure 7.16b. The global efficiencies are shown to be compatible from Figure 7.16a. The bias voltage was set to 40 V and the threshold and ToT to charge calibration as the one reported from Figure 7.16a. The effect of the active



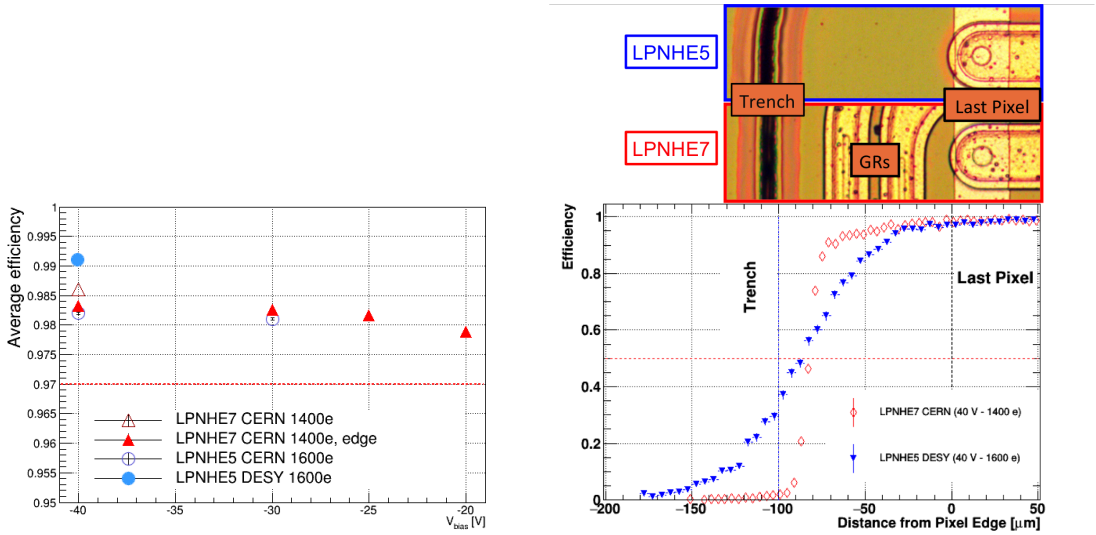
(a)

(b) (c)

Figure 7.15 – Evolution of the global hit efficiency with respect to the MatchPixelMargin (a) and the track to cluster window size (b,c). The X,Y index are referring to respectively the long and short side of the pixel. The results were obtained at CERN for the non-irradiated LPNHE7 sensor biased at 40 V (a,b) and the irradiated W80 sensor biased at 600 V (c).

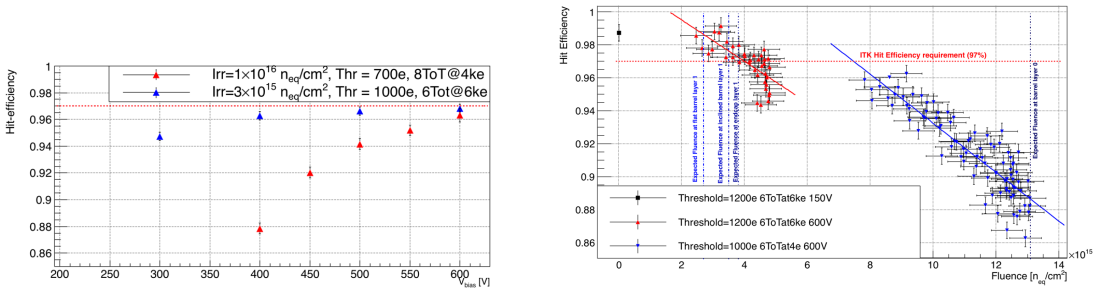
edge technology allows having an efficiency greater than 50% up to 90 μm away from the last pixel. Even though the data-taking conditions are not optimal, similar behaviour is observed at DESY. The sharpest transition for the LPNHE7 module is explained by the higher energy of the beam, resulting in less multiple scattering. To better interpret the result Technology Computer-Aided Design (TCAD) simulations[228] have been performed to visualise the electrical field in the sensor and the effects of the GRs.

The second production has been more intensively studied leading to several communications [229, 230]. The purpose of this production was to prove that the thin planar sensors were capable of resisting the high doses required to be qualified for the HL-LHC phase. Measurements of the global efficiency have been performed at the CERN SPS, varying the bias voltage as shown in Figure 7.17a. Only the W80 sensor quantities have been investigated as the results for the W30 were not considered as enough understood. At a voltage of 600 V, the observed global efficiency is reaching 97% for a fluence of $3\times10^{15}\,n_{eq}/cm^2$ while the value of 96.3 ± 0.5% is obtained for a fluence of $1\times10^{16}\,n_{eq}/cm^2$, almost reaching the ITk requirements at a value of the fluence which is much larger than the expected one for the layers in which the planar pixel sensors are supposed to be used.



(a) (b)

Figure 7.16 – Evolution of the global hit efficiency with respect to the bias voltage (a), or the local hit efficiency with respect to the local position of the hit from the last pixel (b). Both LPNHE5 and LPNHE7 are analysed with data taken either at the CERN SPS or DESY.



(a) (b)

Figure 7.17 – Evolution of the global hit efficiency with respect to the bias voltage (a), or the local hit efficiency with respect to the local fluence (b) of the W80 sensor. Two scenarios depending on the fluence are presented, red triangles are for the 3×10^{15} n_{eq}/cm^2 fluence while the blue ones are for the 1×10^{16} n_{eq}/cm^2. The measurement performed without irradiation is shown with a black triangle.

Since the irradiation has not been uniform, some regions corresponding to different fluence values can be identified and in this way, the evolution of the efficiency can be studied using data points corresponding to many different values of fluence. The values of the local fluence are determined from the fluence map given by the radiation facility which is then corrected by ToT in-situ measurement. Indeed the position of the ToT distribution is linked to the number of charges generated, and thus to the radiation dependent charge trapping effect. Several ToT-to-charge calibrations and thresholds have been used resulting in a global correction, reducing the dispersion of the values, as shown from Figure 7.18. The corrections obtained at CERN were used on DESY data, since the

aperture of the beam is larger there, allowing to expose a larger range of fluences with the same conditions. Results are shown in Figure 7.17b, where the points are fitted with a linear curve. Vertical lines representing the limits of fluences expected for different parts of the detectors are shown in dotted blue, while a horizontal line is drawn for the required efficiency of 97%. This goal is reached for most of the points corresponding to the fluence predicted in Layer 1 of the ITk system. The requirements could be met even for fluences of 7×10^{15} n_{eq}/cm^2 by lowering the threshold and improving the tuning.

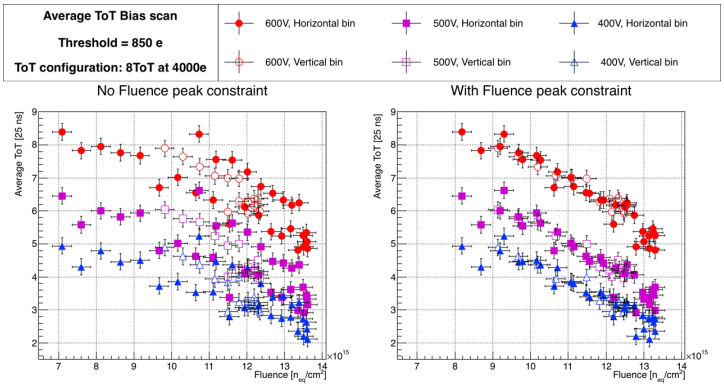
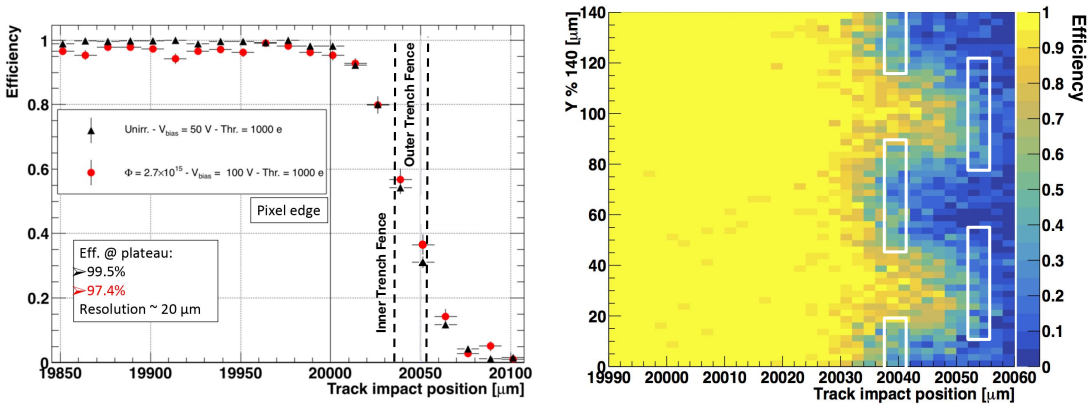


Figure 7.18 – Average ToT distribution vs fluence for 3 different bias voltages. The left/right plot is without/with fluence peak constraints. The horizontal/vertical bin label in the legend means that the fluence and average ToT have been extracted from an horizontal/vertical profile of the region of interest [230].

The last production has also been tested, even if the early breakdown observed after irradiation is weakening the conclusions [231]. Performance has been measured at DESY focusing on the staggered edge which is shown in Figure 7.19b as white rectangles. The efficiency evolution with respect to the position of the track is given in Figure 7.19a. The two positions of the trenches are given by the dotted black lines. Comparing the results with the first production, similar conclusions can be drawn about the extended lateral efficiency. The 20μm resolution due to the multiple scattering at DESY smears the strength of the results and explains the non zero efficiency obtained after the edge. In order to see the effect of the staggered design, a 2-D map is shown in Figue 7.19b, with data taken at CERN. A folding is operated in order to increase the statistic, representing only 140μm of the sensor, corresponding to two pitches of the structure. The efficiency is shown to be extended further in the area where there is the second edge fence but not the first one, the 50% efficiency point being reached for a distance of 50μm for the second edge compared to a distance of 35μm for the first edge. As for the first production, those effects have been corroborated by TCAD simulations [231], obtained without considering the radiation effects. However, the poor performance due to the early breakdown has conducted the collaboration to remain on a standard implementation in the design of the pixel sensors for the ITk.



(a) (b)

Figure 7.19 – Comparison of the 1-D (a) and 2-D (b) edge efficiency profile after irradiation of one sensor of the third production (M1.4). In the case of the 1-D profile, the comparison with the situation before irradiation is given. The data from the left figure are obtained at DESY, while the ones from the right are obtained at the CERN SPS facility [231].

Conclusions and outlook

On fait comme on a dit.

J. Bayou

I have presented, in this document, two analyses looking for the coupling of the Higgs boson to the b-quark in the associated production mode with a vector boson. The first one is based on a subset of Run-2 data ($\mathcal{L} = 79.8\ fb^{-1}$) obtained with the ATLAS detector while the second one extrapolates this result to the end of the HL-LHC phase with $\mathcal{L} = 3000\ fb^{-1}$. In that context, the development of a new tracking detector is crucial, especially in relationship to the b-tagging performance under the HL-LHC running conditions. Hence, the developments in view of the upgrade of the ATLAS pixel detector for the HL-LHC are described in the last part of this document.

The combination of the partial Run-2 dataset with the Run-1 result allowed strong evidence of the Higgs boson decay into a $b\bar{b}$ pair in the vector boson associated production mode to be obtained, with a 4.9 and 5.1 observed and expected significance respectively. This result has been further combined with other analyses, sensitive either to the same decay or production mechanism, leading to the first observation of the $b\bar{b}$ decay of the Higgs boson and the vector boson associated production mode with observing significances of 5.4 and 5.3 standard deviations respectively. The analyses have shown that all the results are in good agreement with each other, providing a coherent picture to the community. My work has been dedicated to the production and interpretation of the preliminary results in the 0-lepton channel while providing the combined results for the Cut Based Analysis approach in the final phase. I have also developed new strategies to improve the analysis for the future. A new event selection, to reject the QCD backgrounds using the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. variable, has shown promising results but currently suffers from fit instabilities. The background modelling techniques have been studied, introducing a new multivariate approach to better map the differences originating from the MC generators.

With time, this analysis could also be sensitive to specific BSM models which would result in a deviation between the measurement and prediction. Therefore even with the observation achieved, an effort should be conducted to further reduce the uncertainties to reach an even better sensitivity. A group of theorists and experimentalists from the ATLAS and CMS collaborations proposed a common framework called Simplified Template Cross-Section (STXS) to ease the interpretation of results and the handling of predictions. In the context of the $Vh(b\bar{b})$ analysis, this separates the processes involved in the production (WH, $q\bar{q}Zh$ and $ggZh$) in a first stage ("stage 0"), and using particle level quantities such as the number of jets, the p_T of the vector boson and the rapidity range of the Higgs boson ("stage 1"). A first set of results has been produced with the dataset used for the observation [198] using a reduced "stage 1" description adapted to the available statistics but is not described in this work. The result was further interpreted in terms of constraints on parameters of an Effective Field Theory (EFT) model, extending the scope of the analysis to BSM extensions. Such a framework will now become the standard procedure for this analysis, and a new binning will be used in p_T^V for the next publication: a resolved regime in the [150,250] p_T^V bin using the method described in this document, and a boosted regime using merged b-jets for the events with $p_T^V > 250$ GeV.

The "stage 0" of the STXS framework has been used to derive the extrapolated measurements for the end of the high luminosity phase of the LHC. A distinction between the signal strength and the cross-section results which would allow theorists to compare their predictions to the extrapolations has been implemented for the given results. Two systematics reduction schemes have also been introduced, giving the reader the possibility to choose a more conservative or aggressive scenario. The uncertainties on the parameter of interests range from 10 (15) % for the WH mode to 43 (50) % for $ggZh$ for scenario 2 (1). The first mode is already systematically limited, while the second and mode still has a large statistical uncertainty. These numbers are to be compared to the uncertainties obtained in the Run-2 analysis for the same processes ranging from 46 to 260 %.

Those results will be influenced by the technological improvements that need to be implemented due to the much more challenging data-taking conditions of the HL-LHC. The developments of the next pixel detector presented in this thesis aim at meeting this expectation. Several designs have been considered to resist to the high expected fluences and to improve the geometrical active area of the sensors. The three sensor productions detailed have been irradiated to simulate the conditions corresponding to the end of the data-taking period. Their efficiency and electrical behaviour have been measured and tested on beams at the CERN SPS and DESY II accelerators. The active-edge technology (used in production 1 and 3) has shown an excellent lateral extension of the depletion region. However, the early breakdown observed in the thin sensors, probably due to the very aggressive design with no or a limited residual guard rings, and the lack of maturity of the technique lead the collaboration to the decision to use for now the conservative standard design for the ITk sensors. A new radiation-hard chip, referred to as RD53, has also been developed, but no preliminary results on irradiated samples are included in this manuscript.

Likewise, the design of a new trigger system is one of the main goals for the upgrade of the detector for the high luminosity phase. The FTK project was created to compute the tracking information at an early stage of the ATLAS trigger workflow during Run-3 of the LHC. It relies on Associative Memory chips and simplified linear fits in a segmented design. The work presented in this thesis has improved the optimisation used to reduce the power consumption of the chip and the design of the pattern banks. While this project is conceived for the Run-3 of the LHC, it is thought to be superseded by the Hardware Tracking Trigger (HTT) [232] for the high-luminosity phase. Indeed the addition of the tracking information will be critical to maintain the performance of the trigger system during that phase and will also benefit rarer searches like the di-Higgs searches.

In conclusion, the work detailed in this thesis has spanned various areas from detector design, trigger optimisation to physics analysis. The main result obtained is the observation of the Higgs boson in the $b\bar{b}$ channel using a partial Run-2 dataset which will be further improved with the inclusion of the full Run-2 data-set and with the introduction of the developments presented in this document. Concerning the future high luminosity phase of the LHC, the expectations in terms of physics are encouraging but ultimately rely on the technical solutions currently being developed.

A Linearisation of helix parameters

A.1 The linearisation of constants

The constants presented in equation 3.2 can be extracted as presented in 3.3 through the following computation:

$$A = \left\langle (\tilde{p}_i - p_i)^2 \right\rangle = \left\langle (\sum_{l=1}^{N} C_{il} x_l + q_i - p_i)^2 \right\rangle \qquad \forall i \in [1,5] \tag{A.1}$$

$$= \left\langle \sum_{k=1}^{N} \sum_{l=1}^{N} C_{il} C_{ik} x_l x_k + q_i^2 + p_i^2 + 2((q_i - p_i) \sum_{l=1}^{N} C_{il} x_l - q_i p_i) \right\rangle \tag{A.2}$$

Since C_{ik}, q_i are constant with respect to the muon samples:

$$= \left\langle \sum_{k=1}^{N} \sum_{l=1}^{N} C_{il} C_{ik} x_l x_k \right\rangle + q_i^2 + \langle p_i^2 \rangle + 2(\sum_{l=1}^{N} C_{il} (q_i \langle x_l \rangle - \langle p_i x_l \rangle) - q_i \langle p_i \rangle) \tag{A.3}$$

In order to find the constants that minimise the distance between the true parameters and the linearised ones, one needs to derive A with respect to the constants:

$$\frac{\partial A}{\partial q_i} = 2q_i + 2(\sum_{l=1}^{N} C_{il}\langle x_l\rangle - \langle p_i\rangle) \quad \forall i \in [1,5] \tag{A.4}$$

$$\frac{\partial A}{\partial q_i} = 0 \rightarrow \langle p_i\rangle = \sum_{l=1}^{N} C_{il}\langle x_l\rangle + q_i \tag{A.5}$$

$$\frac{\partial A}{\partial C_{il}} = 2\langle x_l \sum_{k=1}^{N} C_{ik}x_k\rangle + 2(q_i\langle x_l\rangle - \langle p_i x_l\rangle) \qquad \forall i \in [1,5], l \in [1,N] \tag{A.6}$$

$$\frac{\partial A}{\partial C_{il}} = 0 \rightarrow \langle x_l \sum_{k=1}^{N} C_{ik}x_k\rangle + q_i\langle x_l\rangle - \langle p_i x_l\rangle = 0 \tag{A.7}$$

Identifying from equation A.5 the constant:

$$q_i = \langle p_i\rangle - \sum_{l=1}^{N} C_{il}\langle x_l\rangle \tag{A.8}$$

can be eliminated from equation A.7.

$$\sum_{k=1}^{N} C_{ik}(\underbrace{\langle x_l x_k\rangle - \langle x_l\rangle\langle x_k\rangle}_{V_{lk}}) + \langle p_i\rangle\langle x_l\rangle - \langle p_i x_l\rangle = 0 \tag{A.9}$$

Where V_{lk} represents the element of the position covariance matrix.

$$\sum_{k=1}^{N} C_{ik}V_{lk} + \langle p_i\rangle\langle x_l\rangle - \langle p_i x_l\rangle = 0 \tag{A.10}$$

$$\sum_{k=1}^{N} C_{ik}V_{lk} = \langle p_i x_l\rangle - \langle p_i\rangle\langle x_l\rangle \tag{A.11}$$

And therefore by inverting the matrix V:

$$C_{il} = \sum_{m=1}^{N} V_{lm}^{-1}\langle p_i x_m\rangle - \langle p_i\rangle\langle x_l m\rangle \tag{A.12}$$

Re-injecting the C_{il} in equation A.8, the constants q_i can be extracted. Hence only measuring the covariance matrix and the true helix parameters from simulation provide the constants.

A.2 The χ^2 determination trough PCA

$$\chi^2 = \sum_{i,j=1}^{N} (x_i - \langle x_i \rangle) V_{ij}^{-1} (x_j - \langle x_j \rangle) \tag{A.13}$$

but $V_{ij}^{-1} = \sum_{k=1}^{N} \frac{U_{ik}U_{kj}}{e_k}$, where e_k is the k^{th} eigen values and U_{ik} is the i^{th} component of the k^{th} eigen vector. The Principal Component Analysis (PCA) allows to express a matrix in a new basis in order to get as much dominating eigen values as there is degree of freedom. In this case 5 eigen values are dominating the N measurements, and therefore have negligible contribution to V_{ij}^{-1}, that could be expressed has $V_{ij}^{-1} = \sum_{k=1}^{N-5} \frac{U'_{ik}U'_{kj}}{e_k}$, where U'_{ik} is the i^{th} component of the k^{th} eigen vector in this new basis. For later consideration the prime symbol will be dropped. Equation A.13 can be rewritten:

$$\chi^2 = \sum_{i,j=1}^{N} (x_i - \langle x_i \rangle) \sum_{k=1}^{N-5} \frac{U_{ik}U_{kj}}{e_k} (x_j - \langle x_j \rangle) \tag{A.14}$$

$$= \sum_{k=1}^{N-5} \sum_{i,j=1}^{N} \left((x_i - \langle x_i \rangle) \frac{U_{ik}}{\sqrt{e_k}} \right) \left(\frac{U_{kj}}{\sqrt{e_k}} (x_j - \langle x_j \rangle) \right) \tag{A.15}$$

calling $\frac{U_{ik}}{\sqrt{e_k}} = a_{ik}$

$$= \sum_{k=1}^{N-5} \sum_{i,j=1}^{N} \left((x_i - \langle x_i \rangle) a_{ik} \right) \left(a_{kj} (x_j - \langle x_j \rangle) \right) \tag{A.16}$$

$$= \sum_{k=1}^{N-5} \left(\sum_{i=1}^{N} (x_i - \langle x_i \rangle) a_{ik} \right)^2 \tag{A.17}$$

$$= \sum_{k=1}^{N-5} \left(\sum_{i=1}^{N} x_i a_{ik} - l_k \right)^2 \text{ with } l_k = \sum_{i=1}^{N} a_{ik} \langle x_i \rangle \tag{A.18}$$

B BDT ratio method applied on the tt sample in the 0-lepton channel

B.1 bb flavor

B.1.1 BDT ratio

B.1.2 ME variation

2-jets region

var	KS before	KS after
Ht	0.217	0.337
MET	0.235	0.335
dEtaBB	0.0	0.001
dPhiVBB	0.954	0.982
dRBB	0.0	0.001
mBB	0.001	0.002
pTB1	0.062	0.104
pTB2	0.999	0.999

var	Chi2 before	Chi2 after
Ht	1.34	1.29
MET	2.1	2.05
dEtaBB	1.7	1.53
dPhiVBB	0.82	0.83
dRBB	2.1	1.92
mBB	1.63	1.54
pTB1	1.5	1.47
pTB2	0.65	0.66

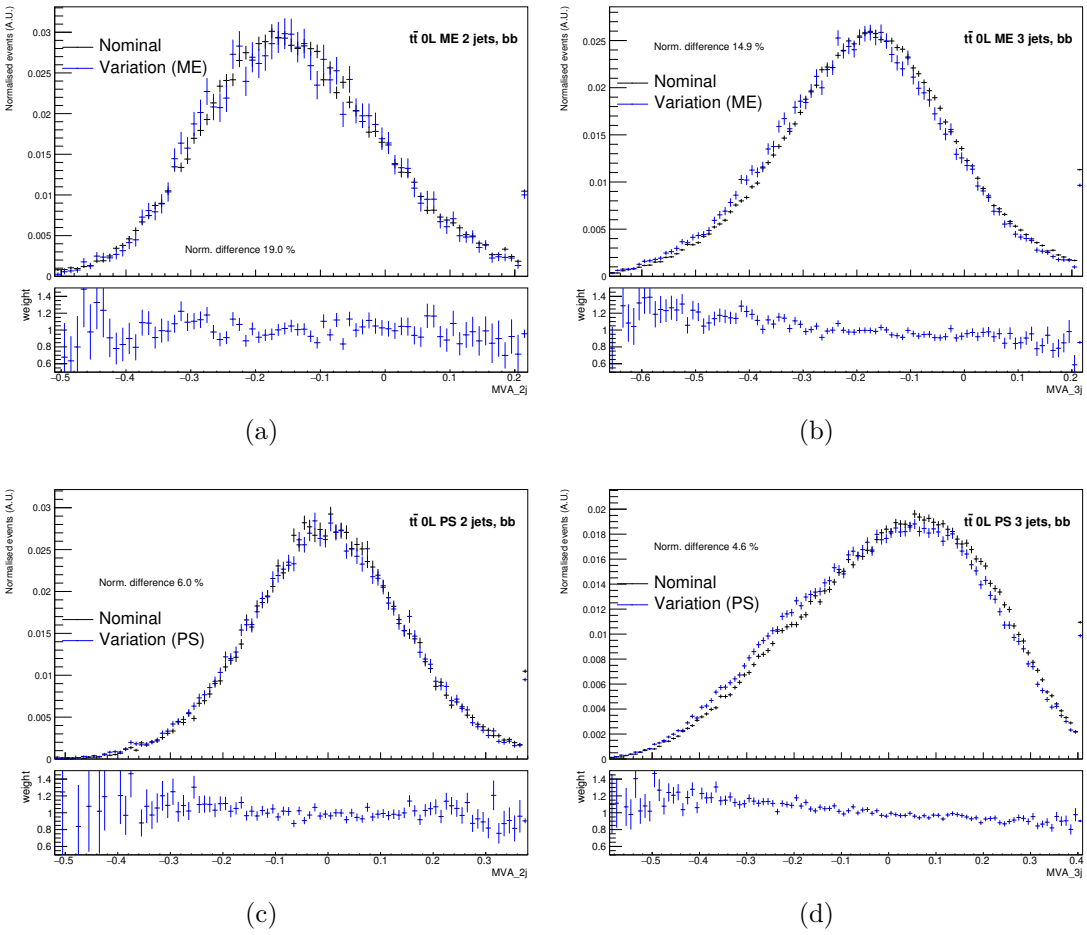


Figure B.1 – Distribution of the BDT scores for the bb flavoured events. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots, while the ME (PS) variation is shown on the top (bottom) part of the Figure. On each plot the ratio between the variation and the nominal is shown on the bottom pad.

3-jets region

var	KS before	KS after
Ht	0.0	0.178
MET	0.0	0.892
dEtaBB	0.0	0.254
dPhiVBB	0.043	0.605
dRBB	0.008	0.893
mBB	0.071	1.0
mBBJ	0.0	0.389
pTB1	0.263	0.926
pTB2	0.075	0.785
pTJ3	0.0	0.09

var	Chi2 before	Chi2 after
Ht	1.51	1.08
MET	1.74	1.33
dEtaBB	2.04	1.42
dPhiVBB	0.88	0.72
dRBB	1.66	1.24
mBB	0.88	0.78
mBBJ	1.24	0.77
pTB1	1.4	1.29
pTB2	1.76	1.33
pTJ3	1.96	1.29

B.1.3 PS variation

2-jets region

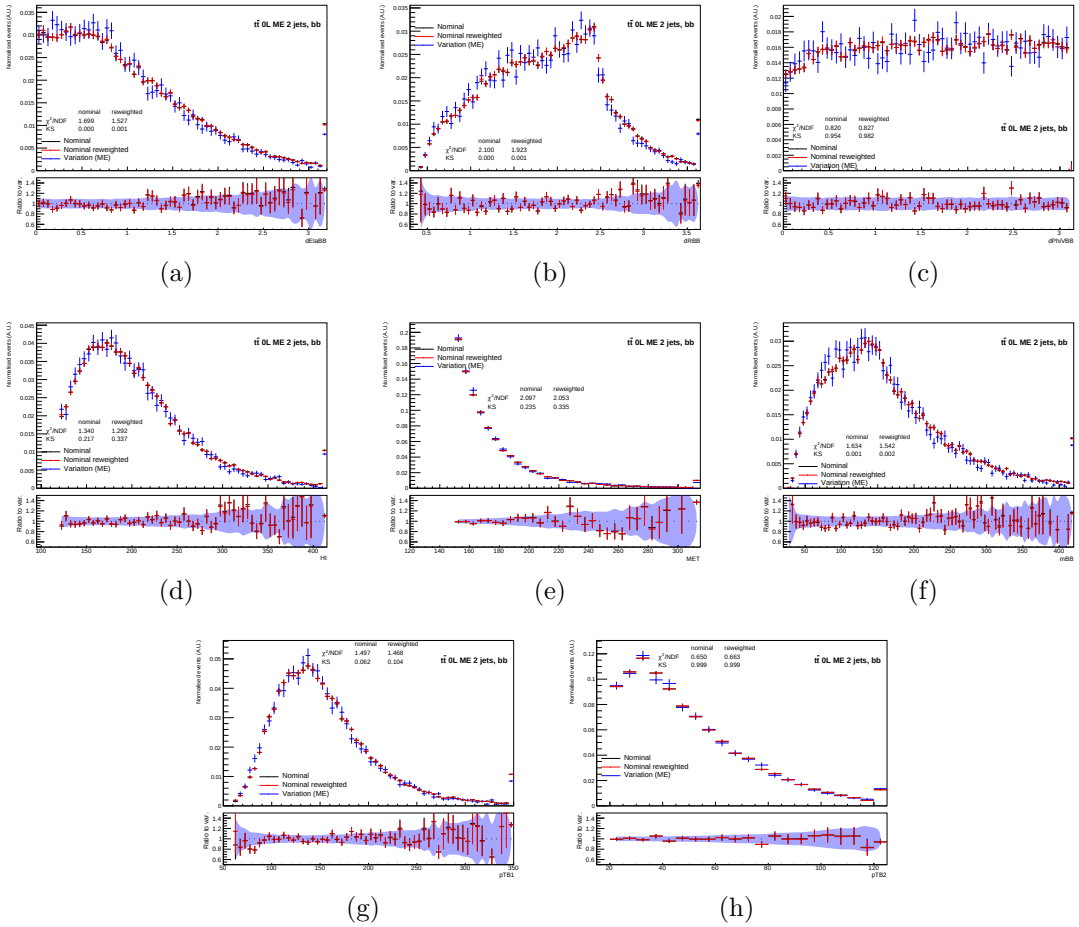
var	KS before	KS after
Ht	0.0	0.0
MET	0.874	0.911
dEtaBB	0.617	0.691
dPhiVBB	0.954	0.977
dRBB	0.088	0.137
mBB	0.002	0.011
pTB1	0.0	0.001
pTB2	0.238	0.397

var	Chi2 before	Chi2 after
Ht	1.71	1.52
MET	1.58	1.52
dEtaBB	1.15	1.1
dPhiVBB	0.72	0.67
dRBB	1.11	1.06
mBB	1.08	0.98
pTB1	1.63	1.43
pTB2	1.44	1.3

3-jets region

var	KS before	KS after
Ht	0.0	0.0
MET	0.719	0.999
dEtaBB	0.006	0.893
dPhiVBB	0.002	0.81
dRBB	0.088	0.969
mBB	0.0	0.42
mBBJ	0.0	0.031
pTB1	0.0	0.04
pTB2	0.0	0.721
pTJ3	0.0	0.033

var	Chi2 before	Chi2 after
Ht	3.89	1.46
MET	1.44	1.15
dEtaBB	0.98	0.81
dPhiVBB	0.86	0.49
dRBB	1.4	0.96
mBB	1.41	1.0
mBBJ	2.83	1.21
pTB1	2.2	1.32
pTB2	1.53	0.82
pTJ3	2.37	1.04

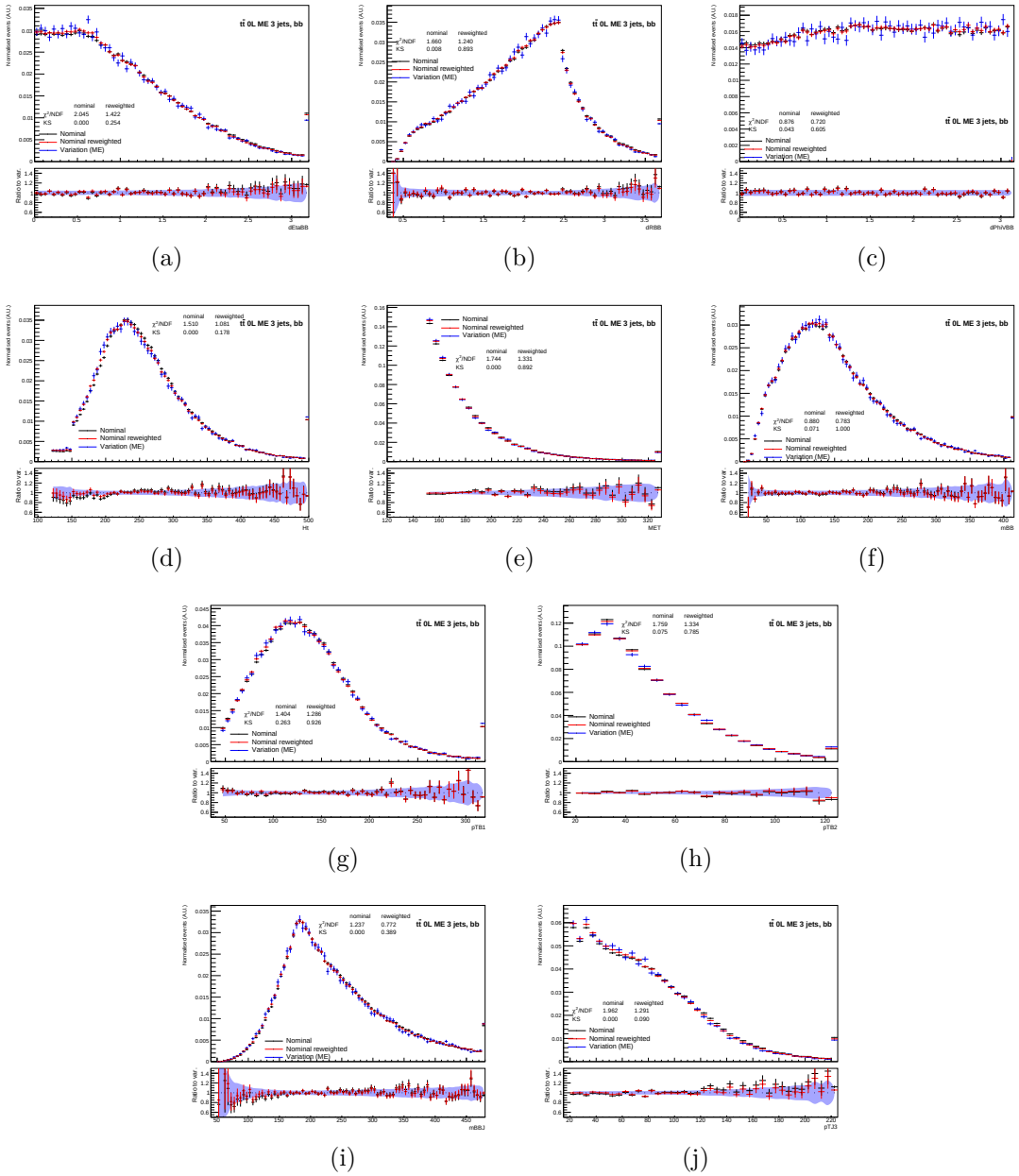


(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure B.2 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag2jets events with where the tagged jets are bb flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f) and the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h).



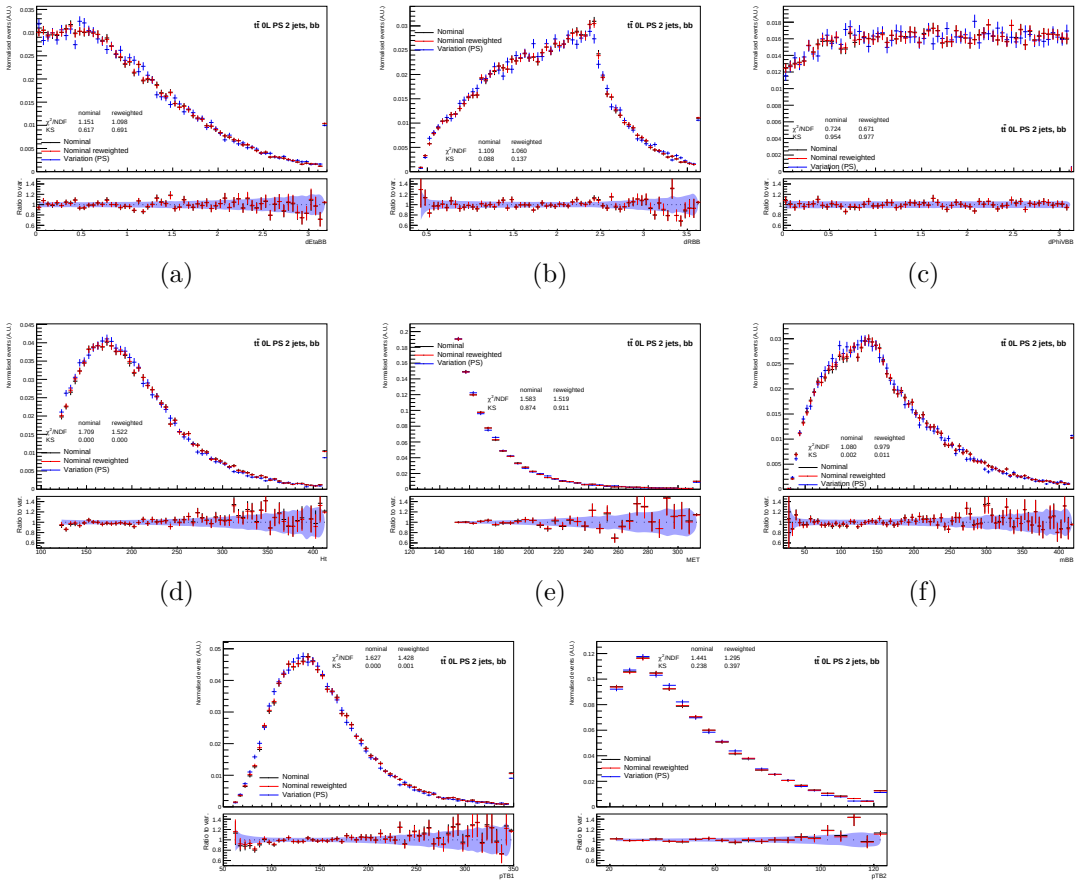
(a) (b) (c)

(d) (e) (f)

(g) (h)

(i) (j)

Figure B.3 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag3jets events with where the tagged jets are bb flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f), the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h), the invariant mass of the three jets $m_{BB,j3}$ (i) and the third jet momentum p_{T}^{j3} (j)



(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure B.4 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag2jets events with where the tagged jets are bb flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f) and the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h).

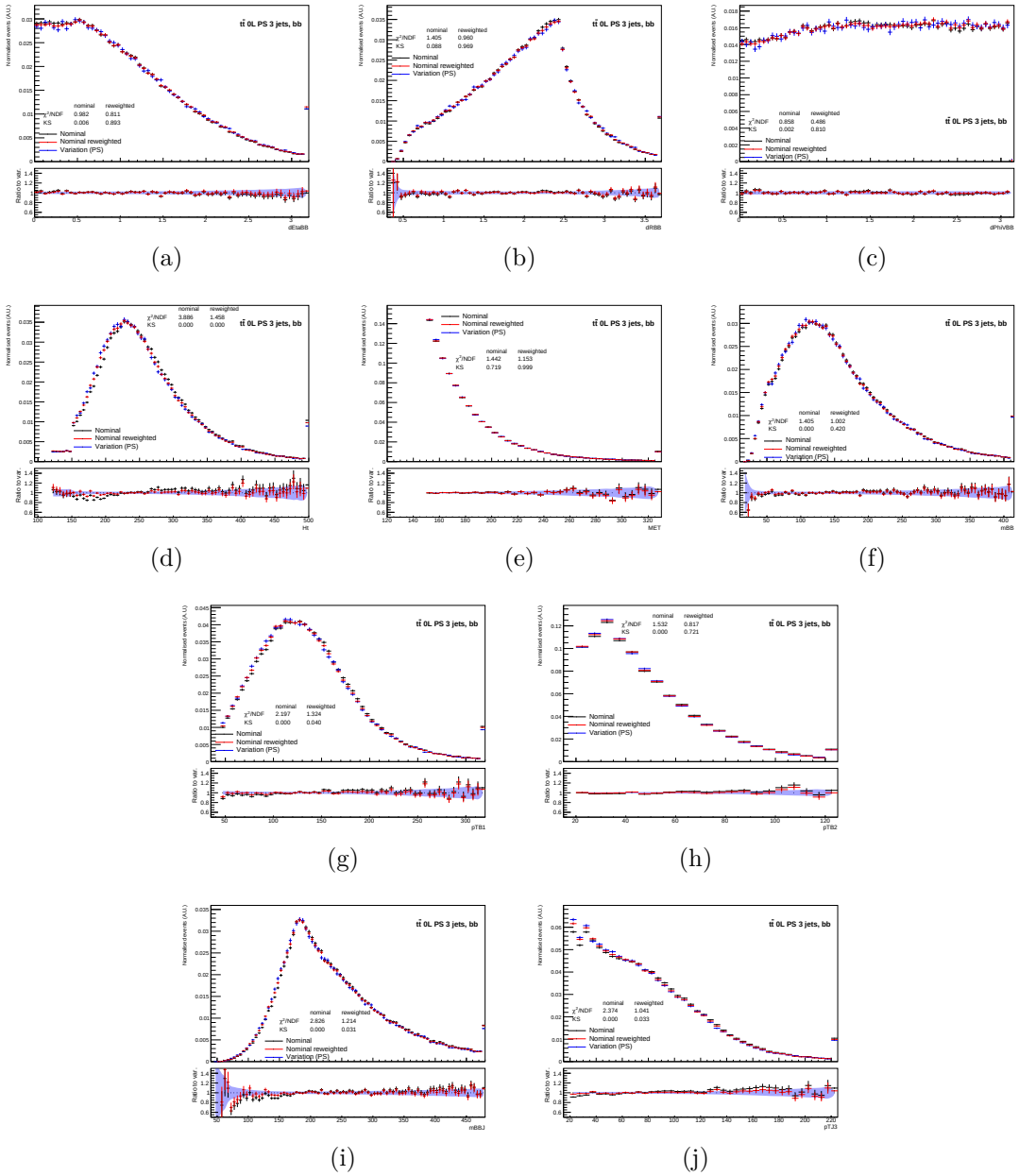


Figure B.5 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag3jets events with where the tagged jets are bb flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f), the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h), the invariant mass of the three jets $m_{BB,j3}$ (i) and the third jet momentum p_{T}^{j3} (j)

B.1.4 Hyper-parameter optimization

B.2 bc flavor

B.2.1 BDT ratio

B.2.2 ME variation

2-jets region

var	KS before	KS after
Ht	0.042	0.397
MET	0.001	0.031
dEtaBB	0.196	0.48
dPhiVBB	0.792	0.905
dRBB	0.716	0.968
mBB	0.778	0.955
pTB1	0.005	0.071
pTB2	0.272	0.891

var	Chi2 before	Chi2 after
Ht	2.92	2.34
MET	1.44	1.24
dEtaBB	1.48	1.41
dPhiVBB	1.24	1.12
dRBB	1.19	1.12
mBB	0.95	0.88
pTB1	1.79	1.57
pTB2	1.6	1.12

3-jets region

var	KS before	KS after
Ht	0.0	0.025
MET	0.022	0.25
dEtaBB	0.001	0.211
dPhiVBB	0.602	0.97
dRBB	0.005	0.231
mBB	0.262	0.525
mBBJ	0.0	0.002
pTB1	0.301	0.764
pTB2	0.289	0.996
pTJ3	0.021	0.198

var	Chi2 before	Chi2 after
Ht	1.37	1.06
MET	1.24	1.12
dEtaBB	1.06	0.93
dPhiVBB	0.7	0.57
dRBB	1.37	1.08
mBB	1.18	1.05
mBBJ	2.13	1.31
pTB1	1.22	1.16
pTB2	1.12	0.92
pTJ3	1.31	1.18

B.2.3 PS variation

2-jets region

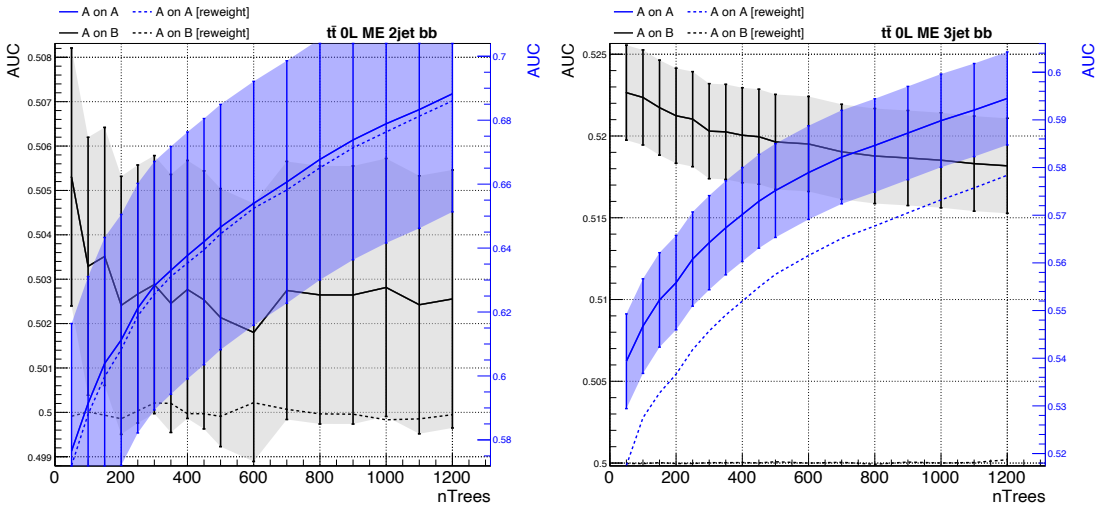
var	KS before	KS after
Ht	0.0	0.0
MET	0.085	0.518
dEtaBB	0.213	0.682
dPhiVBB	0.123	0.615
dRBB	0.702	0.993
mBB	0.0	0.001
pTB1	0.0	0.004
pTB2	0.0	0.018

var	Chi2 before	Chi2 after
Ht	3.47	2.45
MET	1.54	1.23
dEtaBB	0.91	0.8
dPhiVBB	1.3	1.01
dRBB	0.82	0.8
mBB	1.44	1.06
pTB1	2.65	2.12
pTB2	1.67	1.25

3-jets region

var	KS before	KS after
Ht	0.0	0.022
MET	0.001	0.989
dEtaBB	0.403	1.0
dPhiVBB	0.716	0.999
dRBB	0.03	0.852
mBB	0.0	0.703
mBBJ	0.0	0.024
pTB1	0.0	0.76
pTB2	0.0	0.635
pTJ3	0.0	0.671

var	Chi2 before	Chi2 after
Ht	2.36	1.07
MET	1.58	1.02
dEtaBB	1.13	0.89
dPhiVBB	1.19	0.84
dRBB	1.2	0.9
mBB	1.58	0.95
mBBJ	10.03	2.01
pTB1	1.7	1.08
pTB2	2.44	1.3
pTJ3	1.72	0.93



(a) (b)

(c) (d)

Figure B.6 – Evolution of the AUC for the bb flavoured events with respect to the number of trees used in the training. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots, while the ME (PS) variation is shown on the top (bottom) part of the Figure. The right (left) axis is presenting the AonA (AonB) scale. Errors are shown for the non reweighted quantities following the recommendations of [182].

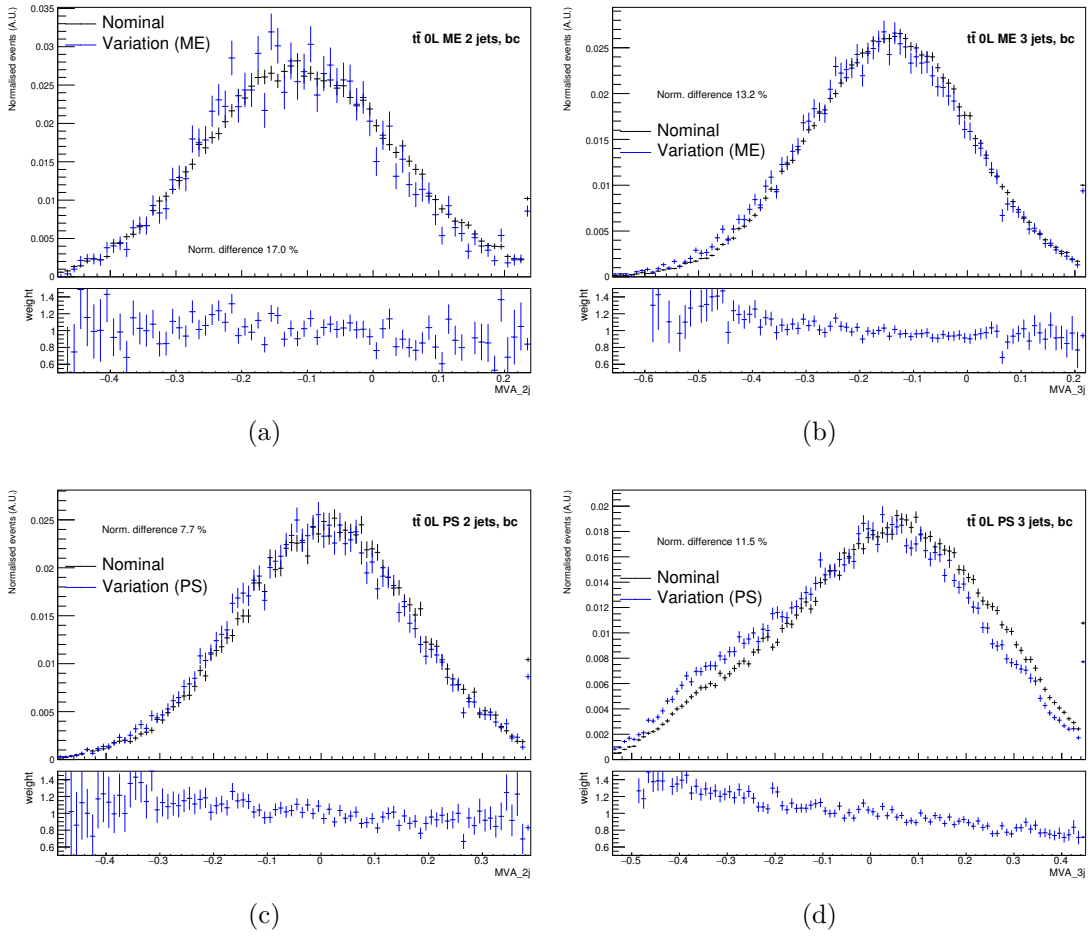
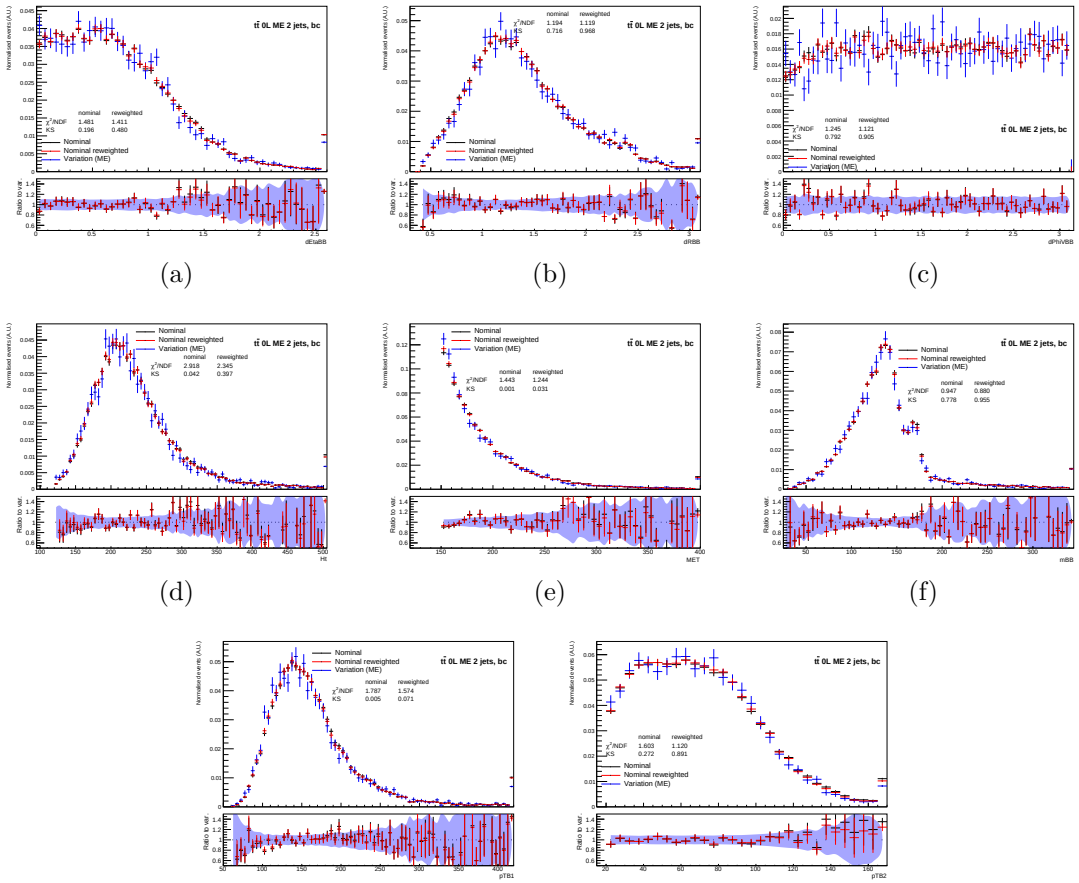


Figure B.7 – Distribution of the BDT scores for the bc flavoured events. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots, while the ME (PS) variation is shown on the top (bottom) part of the Figure. On each plot the ratio between the variation and the nominal is shown on the bottom pad.



(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure B.8 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag2jets events with where the tagged jets are bc flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f) and the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h).

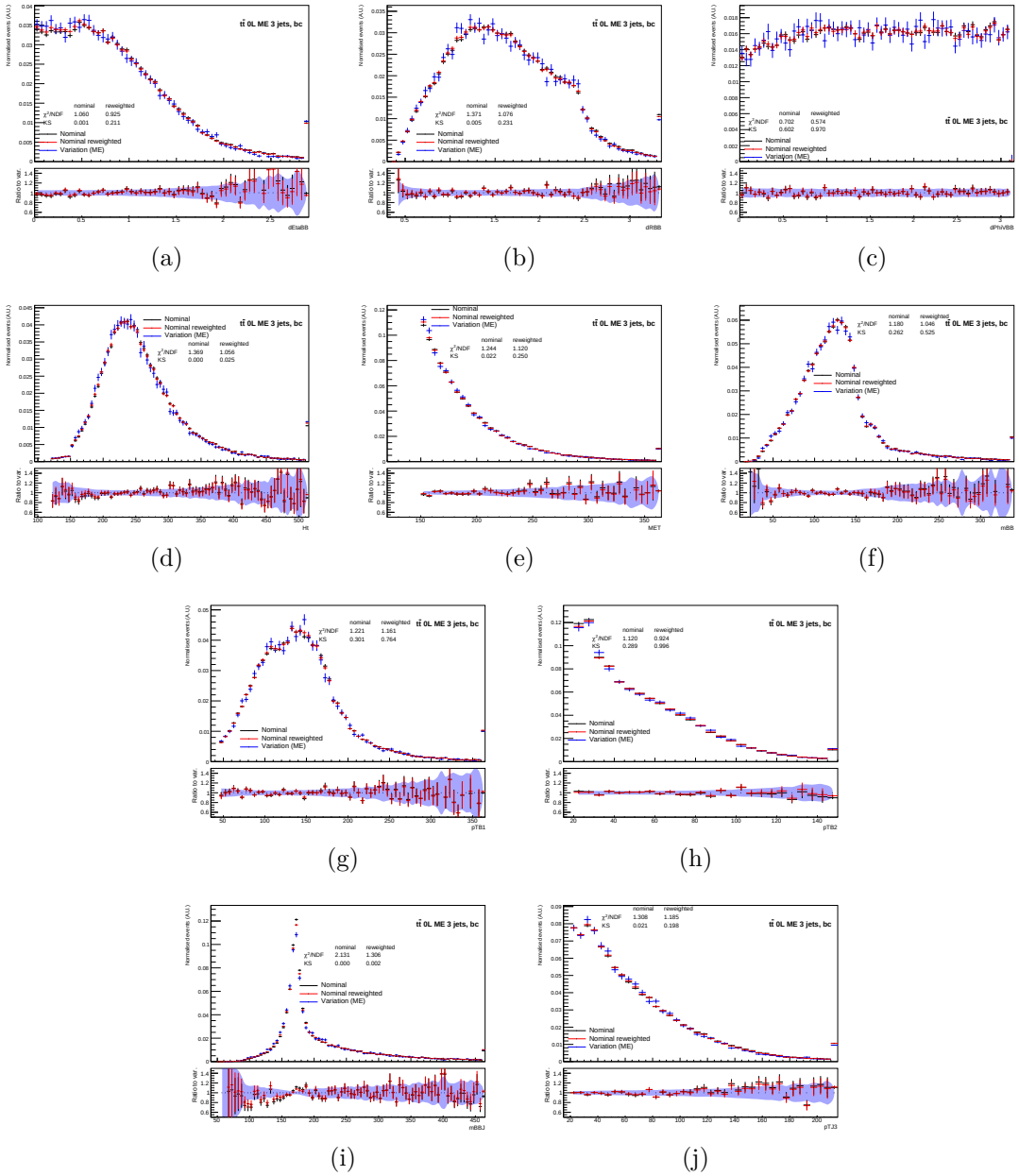
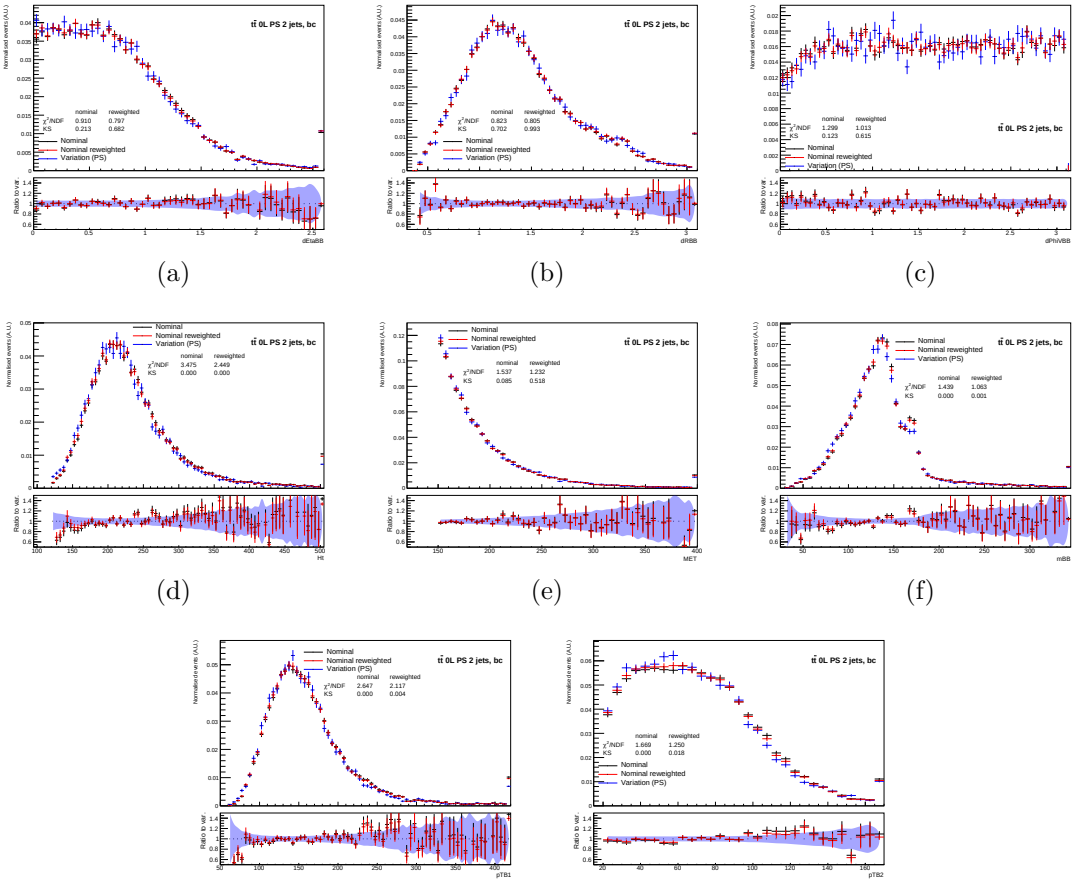


Figure B.9 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag3jets events with where the tagged jets are bc flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h), the invariant mass of the three jets $m_{BB,j3}$ (i) and the third jet momentum p_{T}^{j3} (j)

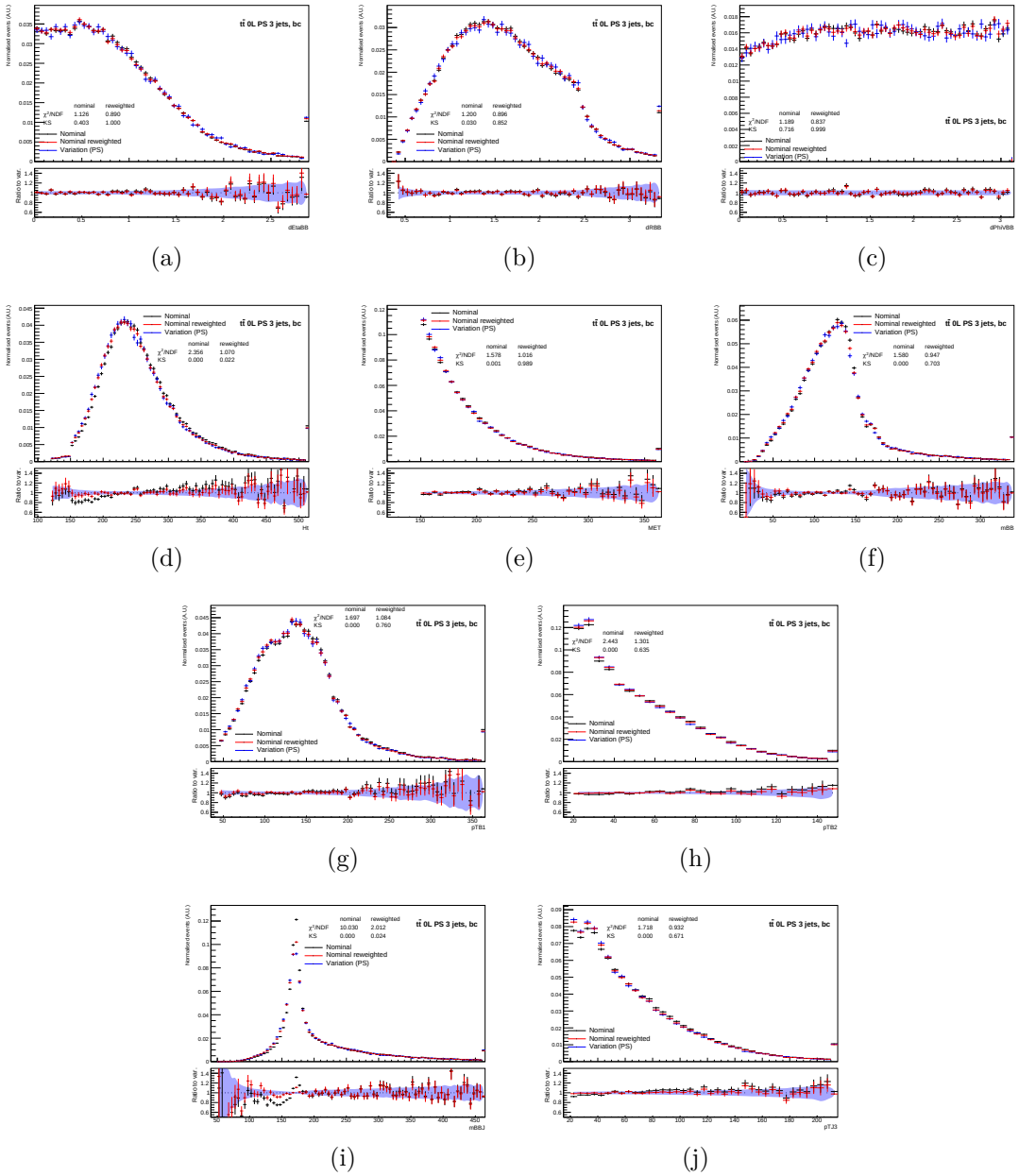


(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure B.10 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag2jets events with where the tagged jets are bc flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f) and the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h).



(a) (b) (c)

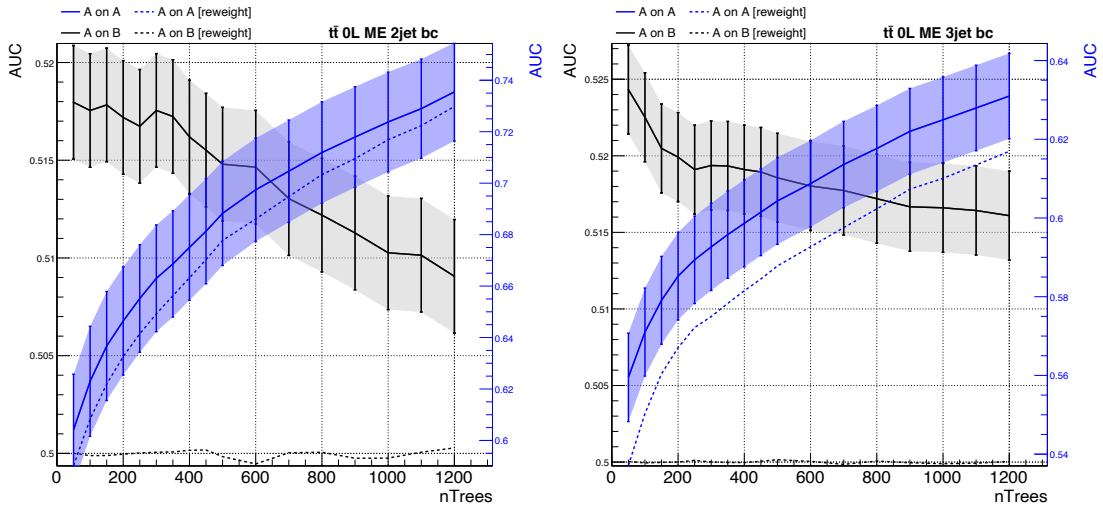
(d) (e) (f)

(g) (h)

(i) (j)

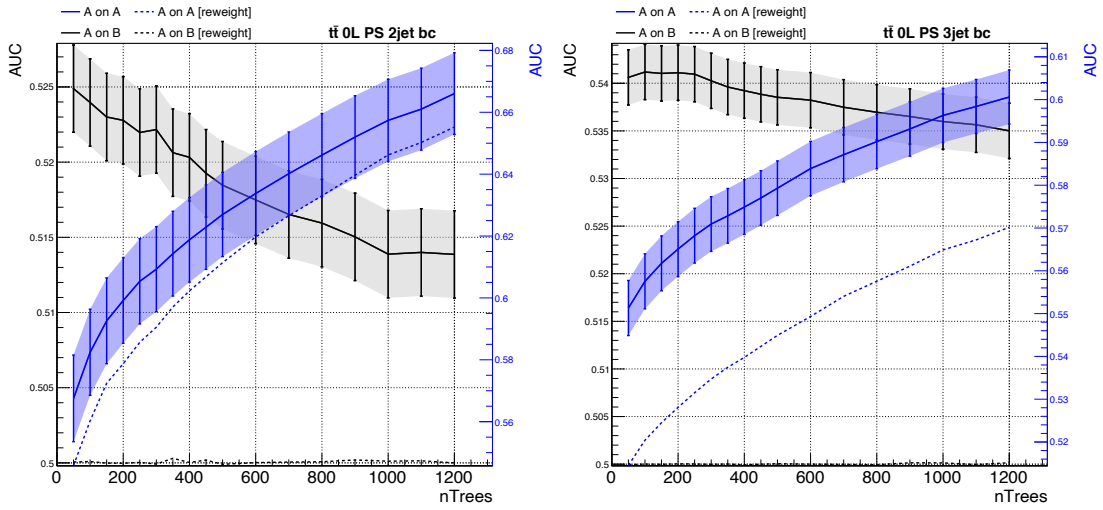
Figure B.11 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag3jets events with where the tagged jets are bc flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h), the invariant mass of the three jets $m_{BB,j3}$ (i) and the third jet momentum p_{T}^{j3} (j)

B.2.4 Hyper-parameter optimization



(a)

(b)



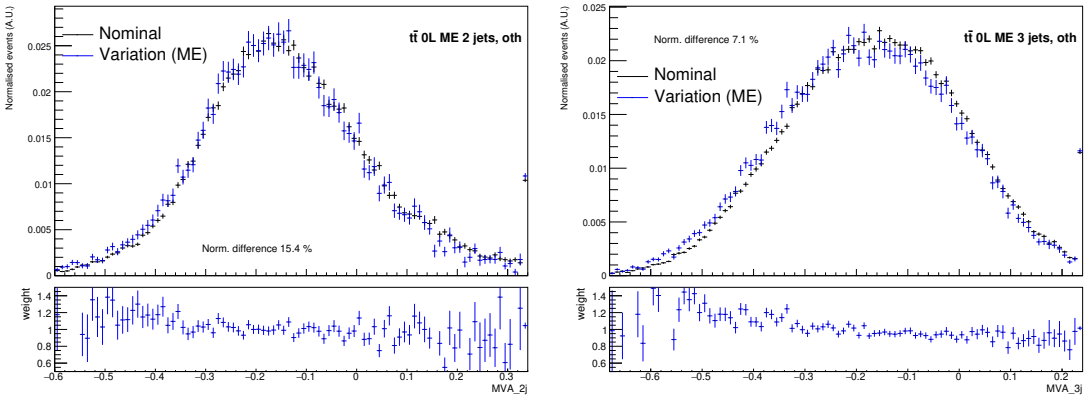
(c)

(d)

Figure B.12 – Evolution of the AUC for the bc flavoured events with respect to the number of trees used in the training. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots, while the ME (PS) variation is shown on the top (bottom) part of the Figure. The right (left) axis is presenting the AonA (AonB) scale. Errors are shown for the non reweighted quantities following the recommendations of [182].

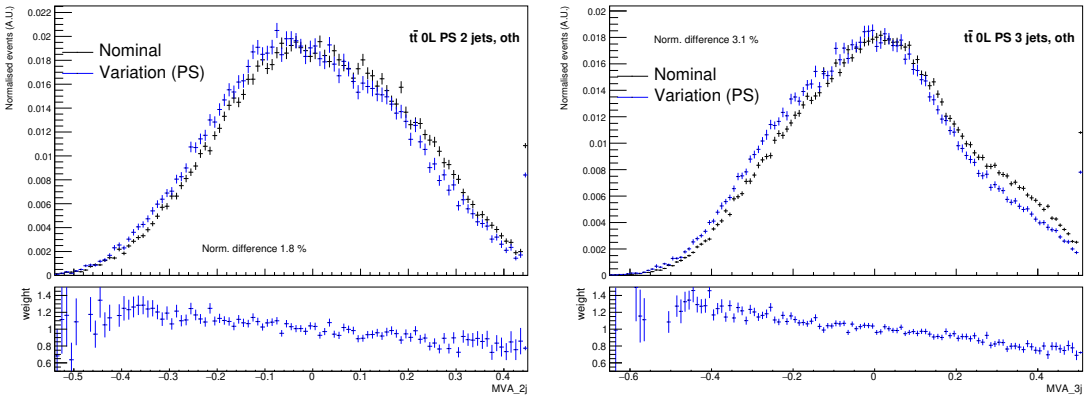
B.3 other flavors

B.3.1 BDT ratio



(a)

(b)



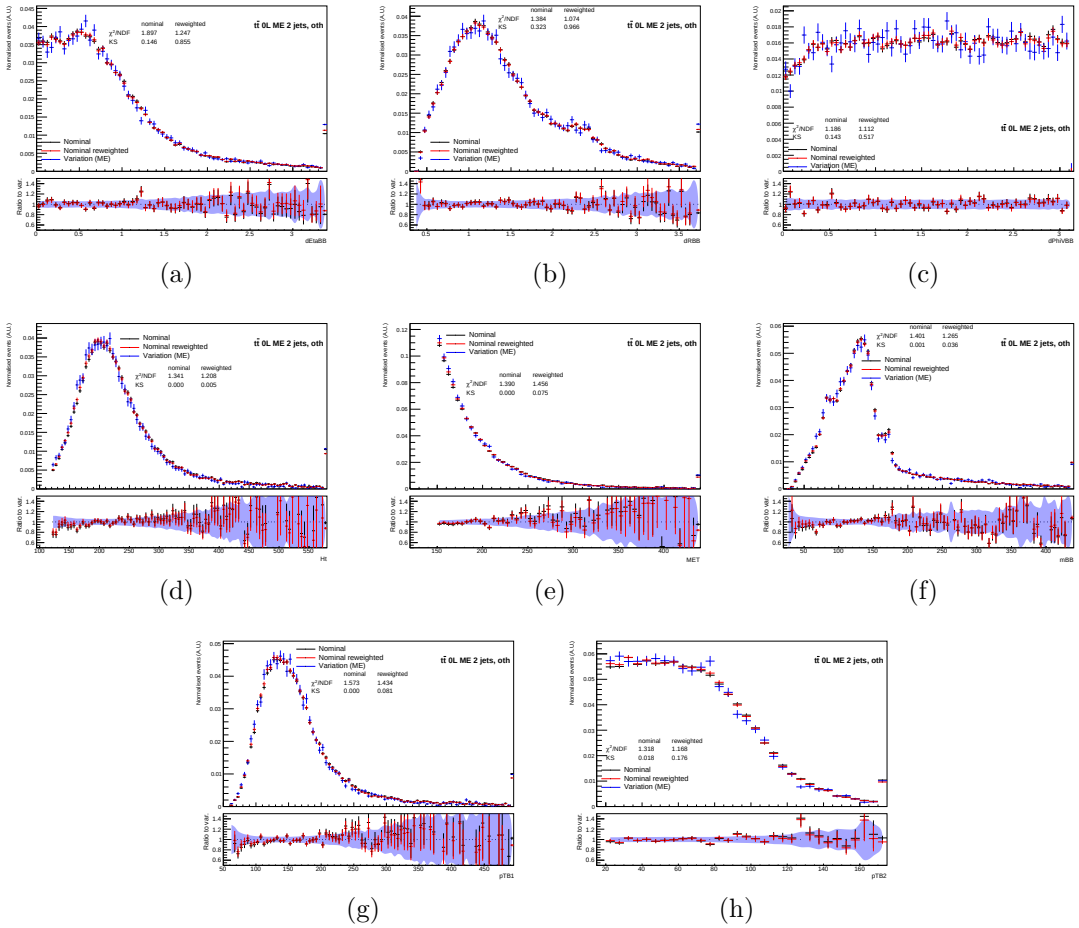
(c)

(d)

Figure B.13 – Distribution of the BDT scores for the oth flavoured events. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots, while the ME (PS) variation is shown on the top (bottom) part of the Figure. On each plot the ratio between the variation and the nominal is shown on the bottom pad.

B.3.2 ME variation

2-jets region



(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure B.14 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag2jets events with where the tagged jets are oth flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f) and the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h).

var	KS before	KS after
Ht	0.0	0.005
MET	0.0	0.075
dEtaBB	0.146	0.855
dPhiVBB	0.143	0.517
dRBB	0.323	0.966
mBB	0.001	0.036
pTB1	0.0	0.081
pTB2	0.018	0.176

var	Chi2 before	Chi2 after
Ht	1.34	1.21
MET	1.39	1.46
dEtaBB	1.9	1.25
dPhiVBB	1.19	1.11
dRBB	1.38	1.07
mBB	1.4	1.26
pTB1	1.57	1.43
pTB2	1.32	1.17

3-jets region

var	KS before	KS after
Ht	0.0	0.215
MET	0.0	0.523
dEtaBB	0.483	0.986
dPhiVBB	0.963	0.967
dRBB	0.002	0.119
mBB	0.0	0.033
mBBJ	0.0	0.001
pTB1	0.001	0.485
pTB2	0.752	0.902
pTJ3	0.246	0.688

var	Chi2 before	Chi2 after
Ht	1.18	0.99
MET	1.29	1.13
dEtaBB	1.62	1.07
dPhiVBB	0.82	0.74
dRBB	1.59	0.96
mBB	1.64	1.28
mBBJ	3.36	1.62
pTB1	1.32	1.11
pTB2	1.25	1.2
pTJ3	1.08	1.01

B.3.3 PS variation

2-jets region

var	KS before	KS after
Ht	0.0	0.023
MET	0.003	0.523
dEtaBB	0.0	0.057
dPhiVBB	0.844	1.0
dRBB	0.0	0.021
mBB	0.0	0.845
pTB1	0.0	0.049
pTB2	0.0	0.159

var	Chi2 before	Chi2 after
Ht	2.26	1.26
MET	1.67	1.11
dEtaBB	2.94	1.57
dPhiVBB	0.88	0.63
dRBB	1.8	0.88
mBB	1.9	1.11
pTB1	1.94	1.25
pTB2	2.29	1.06

3-jets region

var	KS before	KS after
Ht	0.0	0.28
MET	0.553	1.0
dEtaBB	0.0	0.423
dPhiVBB	0.824	1.0
dRBB	0.0	0.378
mBB	0.0	0.671
mBBJ	0.0	0.001
pTB1	0.0	0.737
pTB2	0.0	0.453
pTJ3	0.0	0.599

var	Chi2 before	Chi2 after
Ht	2.47	1.18
MET	1.2	1.06
dEtaBB	4.37	1.17
dPhiVBB	0.77	0.53
dRBB	2.9	0.85
mBB	2.77	1.27
mBBJ	12.87	1.43
pTB1	1.56	1.01
pTB2	1.6	1.07
pTJ3	2.85	1.48

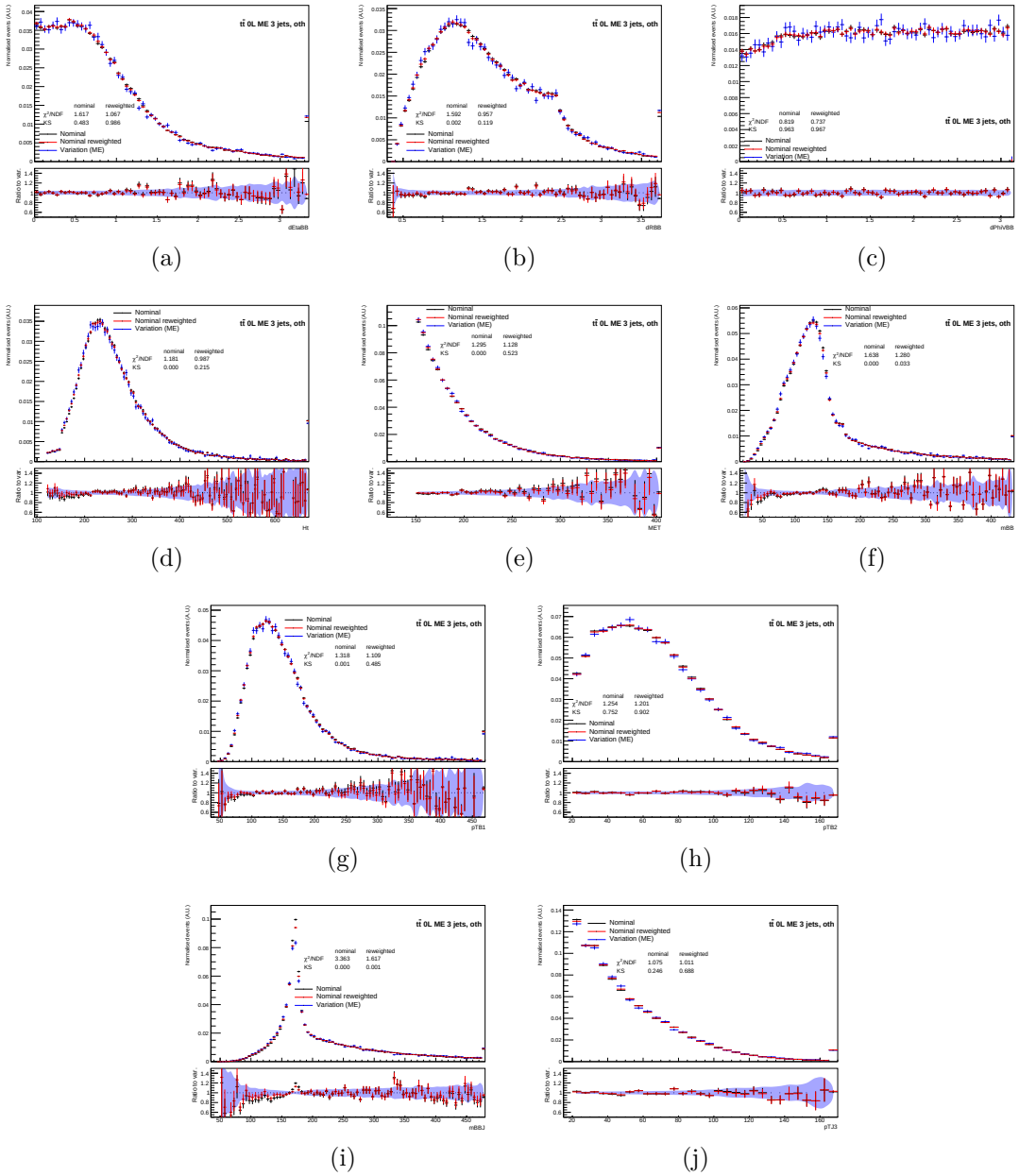
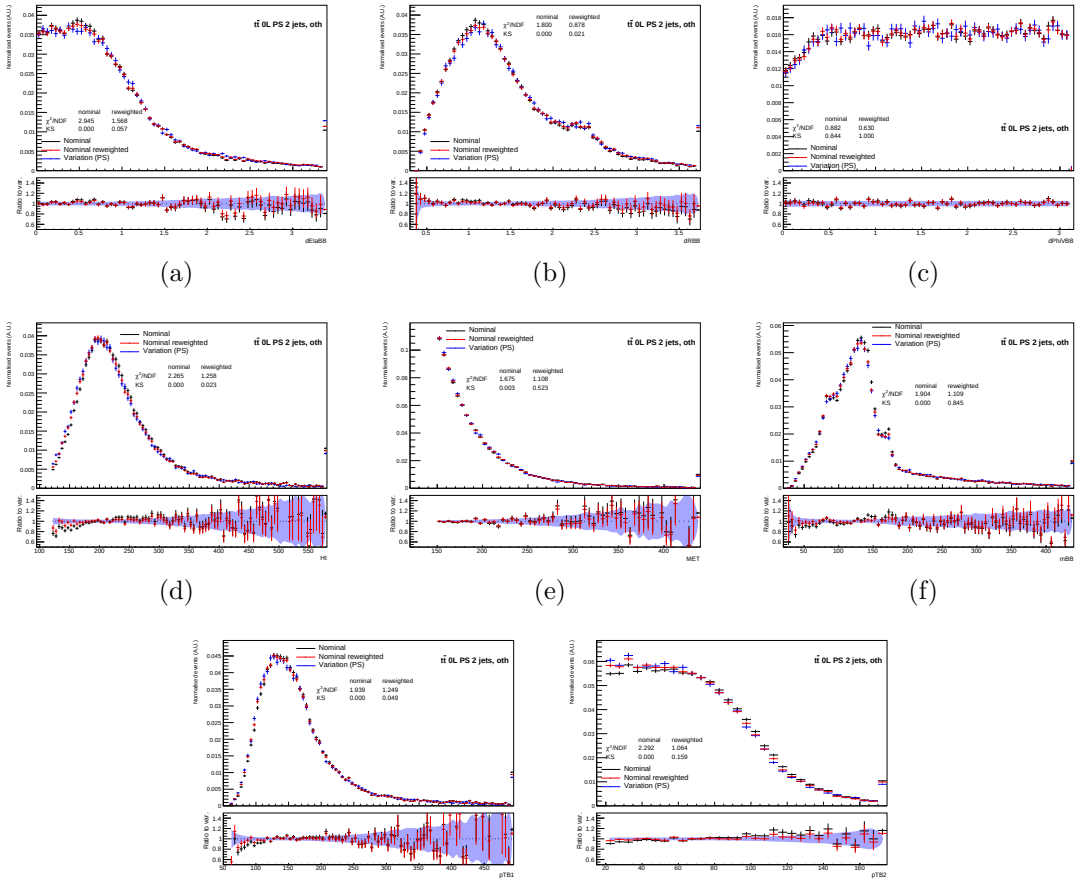


Figure B.15 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag3jets events with where the tagged jets are oth flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h), the invariant mass of the three jets $m_{BB,j3}$ (i) and the third jet momentum p_{T}^{j3} (j)



(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure B.16 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag2jets events with where the tagged jets are oth flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the di-jet invariant mass m_{bb} (f) and the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h).

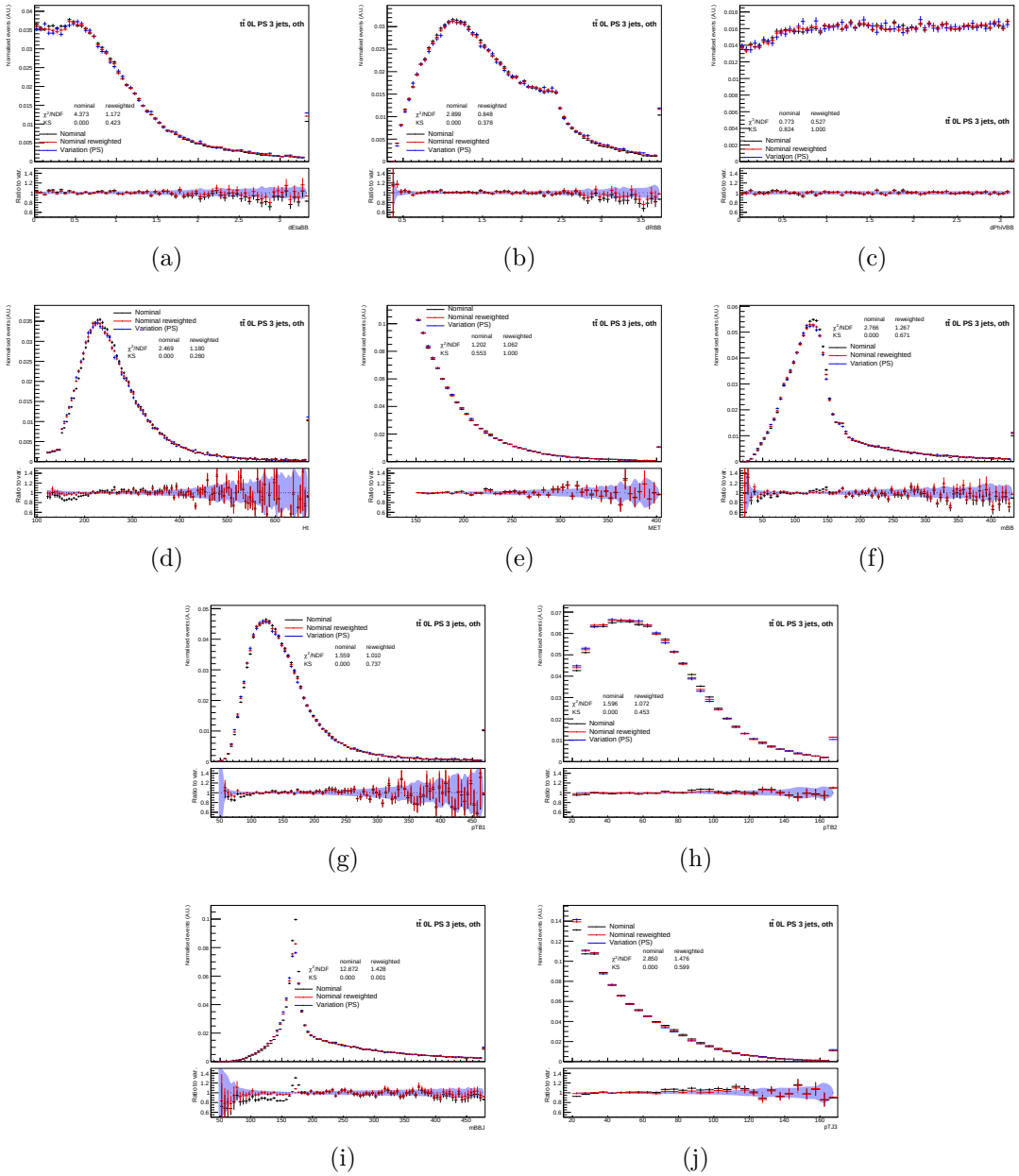
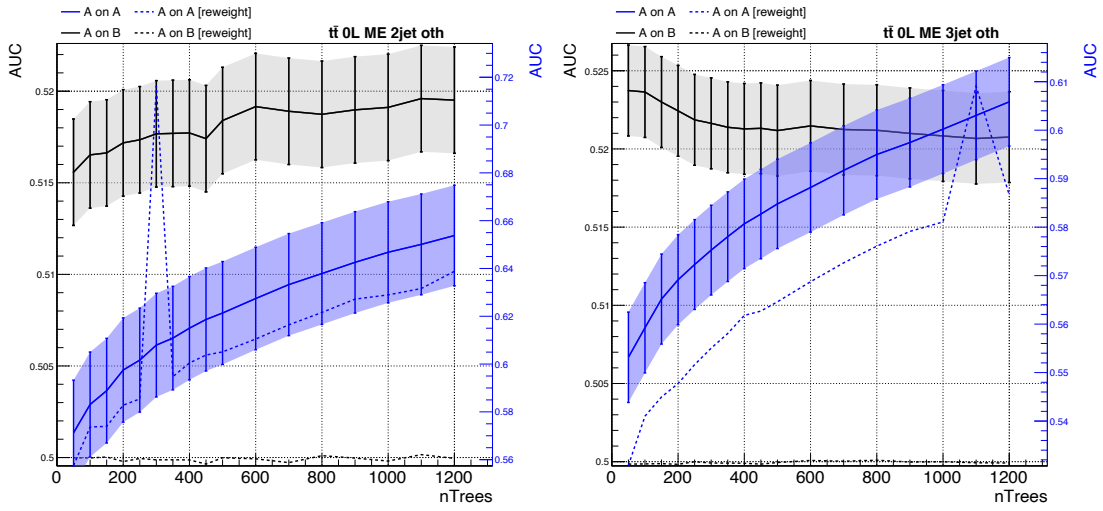


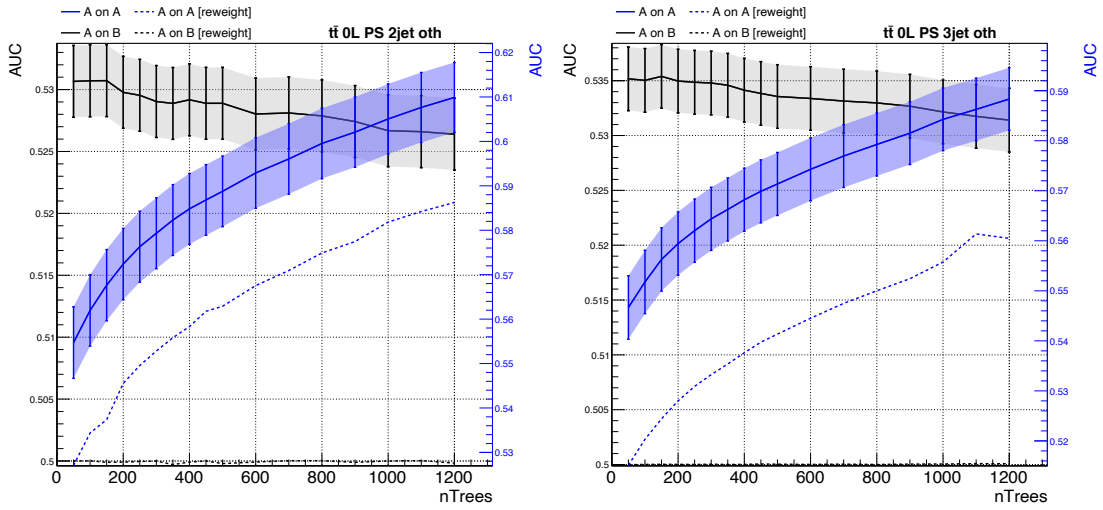
Figure B.17 – Comparison of shapes of the nominal MC generator before (black) and after reweighting (red) compared to the variation (blue) at truth level. 2tag3jets events with where the tagged jets are oth flavoured at truth level. The bottom plot shows the ratio of the nominal with respect to the variation. The statistical uncertainty from the variation only is represented with the light blue background. The χ^2 and KS scores are indicated before and after reweighting between the nominal and variation. The presented variables are $\Delta\eta(B,B)$ (a), $\Delta R(B,B)$ (b), $\Delta\phi(V,h)$ (c), the hard object scalar sum of p_{T} H_t (d), the $E_{\mathrm{T}}^{\mathrm{miss}}$ (e), the two b-jets p_{T} p_{T}^{b1} (g) and p_{T} p_{T}^{b2} (h), the invariant mass of the three jets $m_{BB,j3}$ (i) and the third jet momentum p_{T}^{j3} (j)

B.3.4 Hyper-parameter optimization



(a)

(b)



(c)

(d)

Figure B.18 – Evolution of the AUC for the oth flavoured events with respect to the number of trees used in the training. 2tag2jets (2tag3jets) regions events are shown on the left (right) plots, while the ME (PS) variation is shown on the top (bottom) part of the Figure. The right (left) axis is presenting the AonA (AonB) scale. Errors are shown for the non reweighted quantities following the recommendations of [182].

C CBA fit results

This chapter presents the fit results obtained with the CBA analysis, comparing the main results with the one obtained in the MVA analysis.

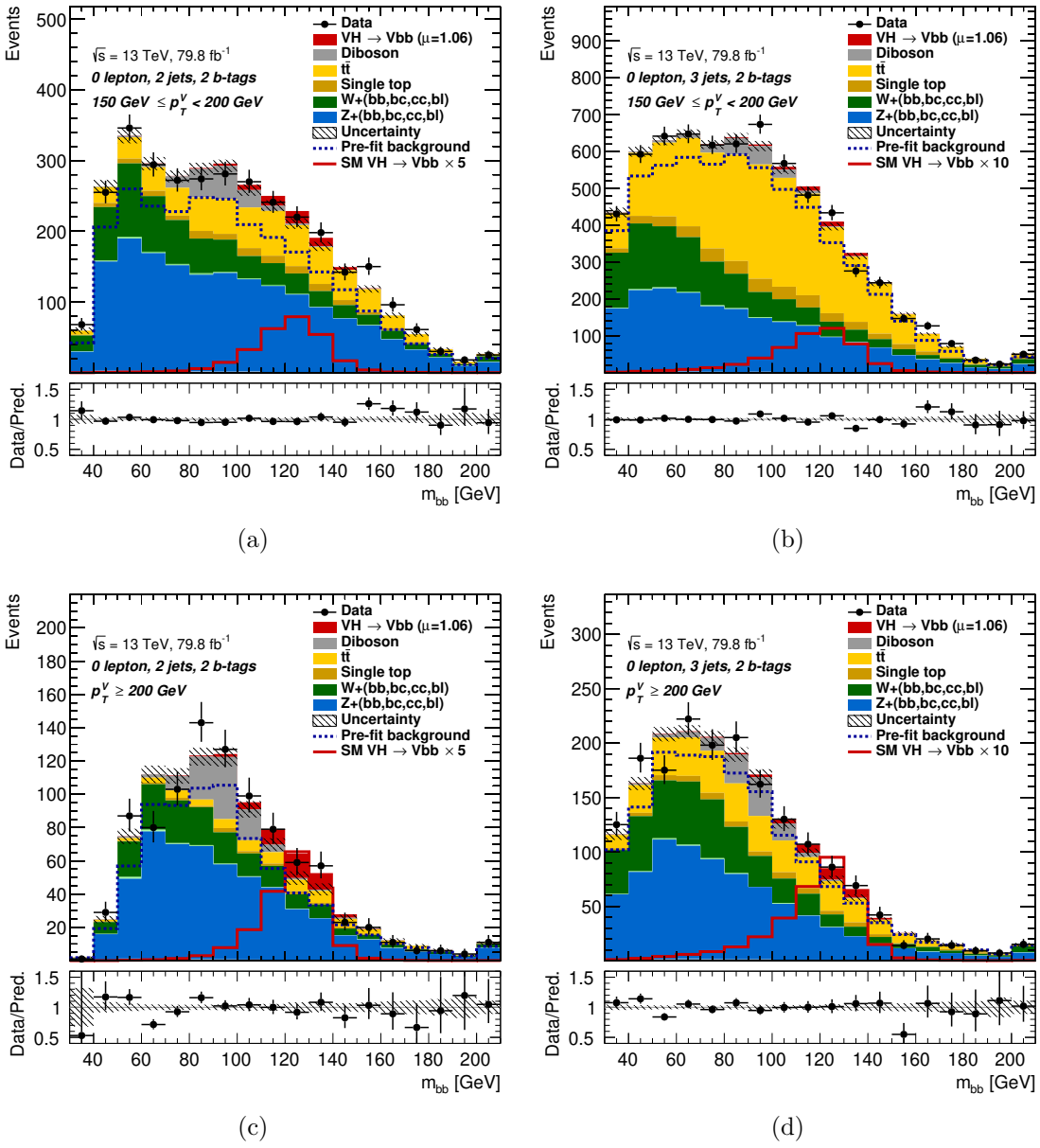


Figure C.1 – The m_{bb} post-fit distributions in the 0-lepton channel for 2-btag events, in the 2-jets (left) and exactly 3-jets (right) low-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (top) and high-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (bottom) categories.

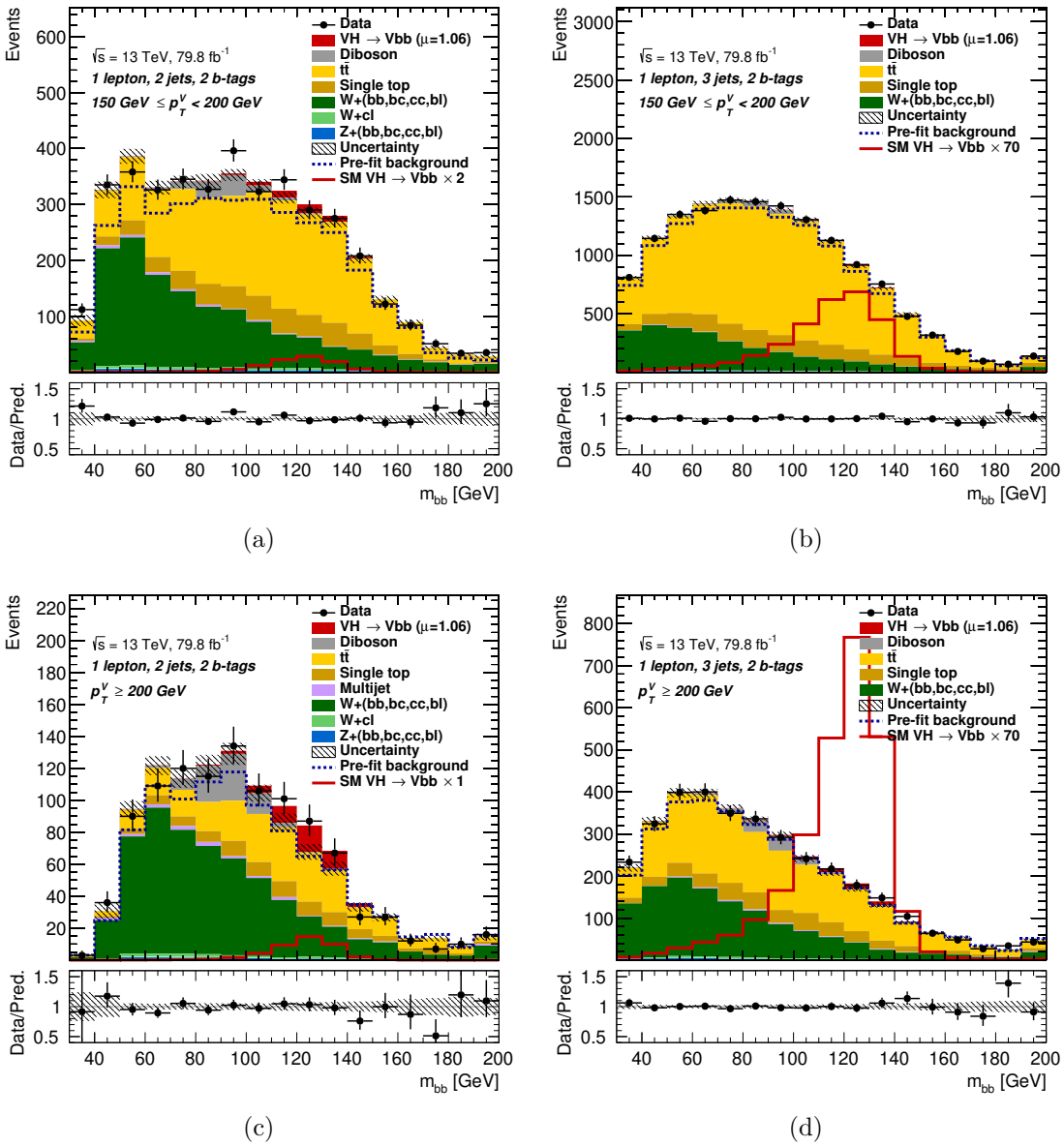


Figure C.2 – The m_{bb} post-fit distributions in the 1-lepton channel for 2-btag events, in the 2-jets (left) and exactly 3-jets (right) low-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (top) and high-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (bottom) categories in the signal region.

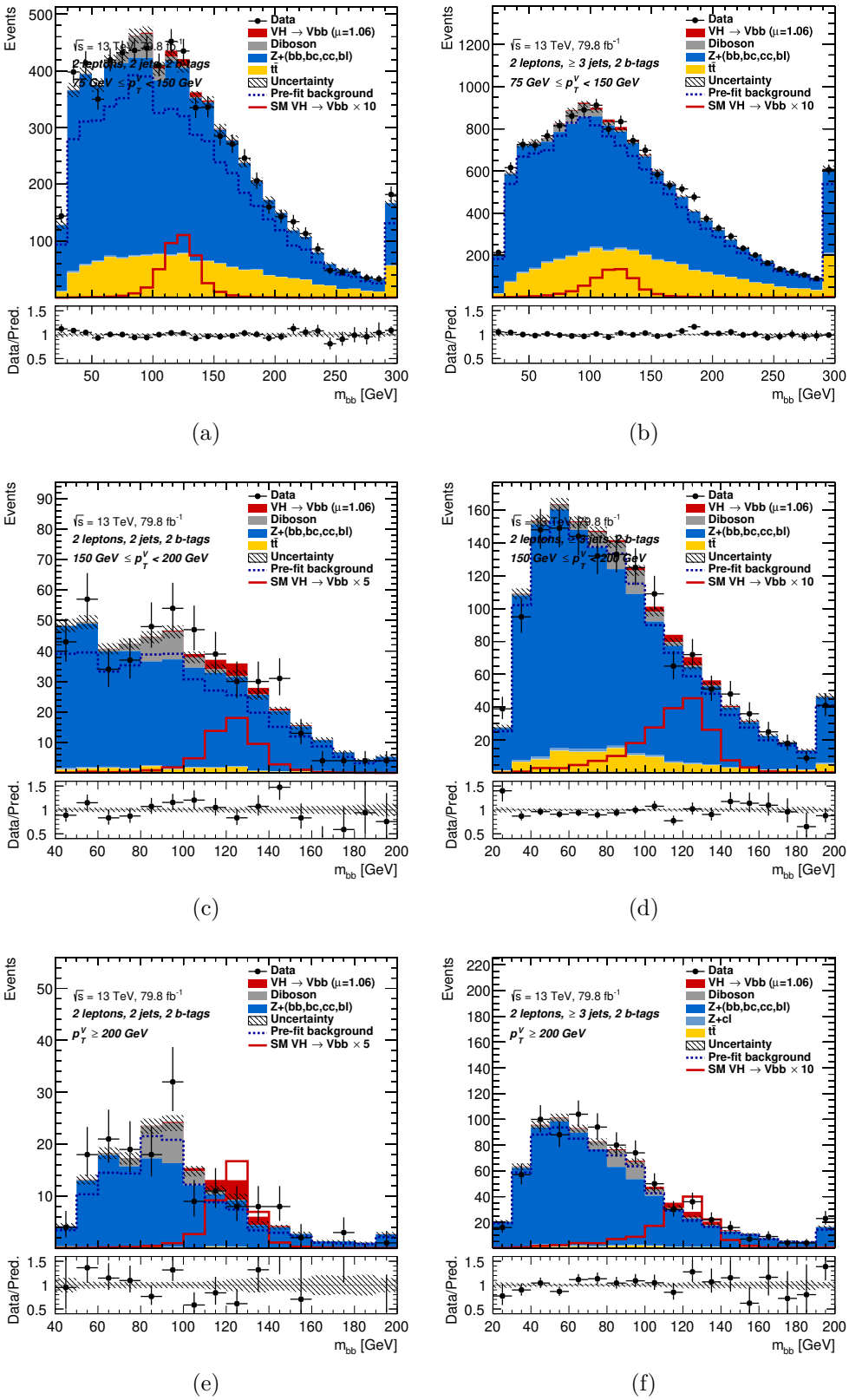
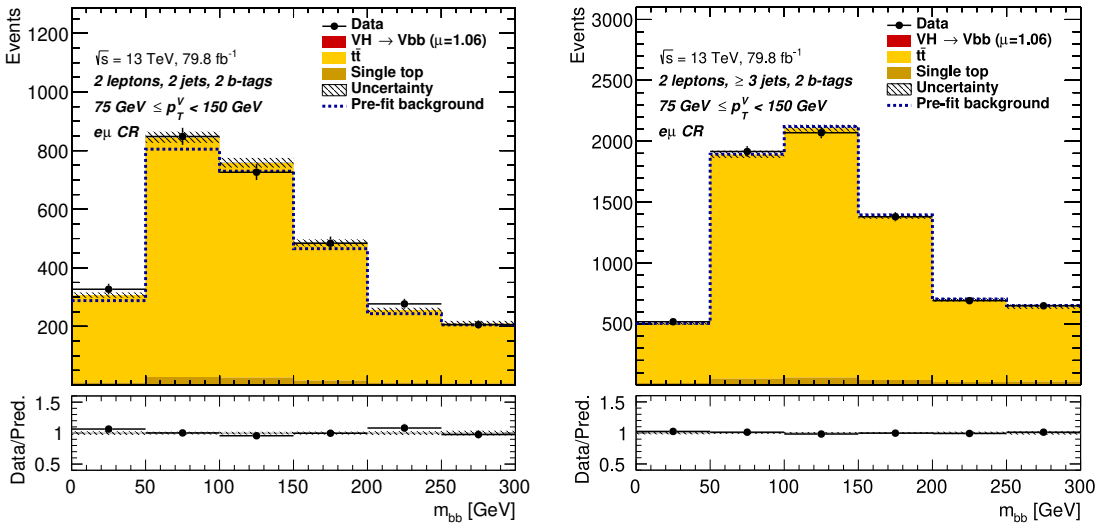
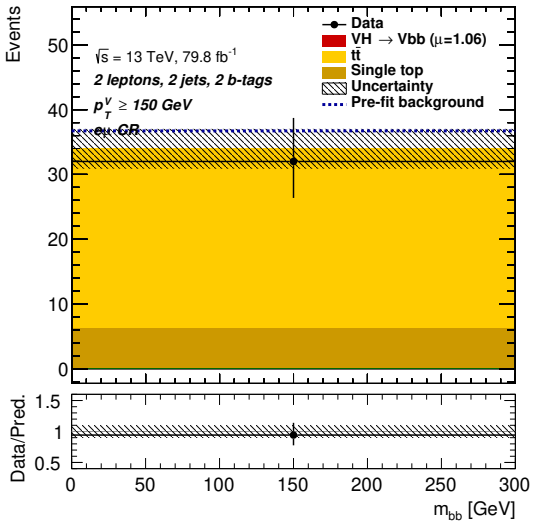


Figure C.3 – The m_{bb} post-fit distributions in the 2-leptons channel for 2-btag events, in the 2-jets (left) and exactly 3-jets (right) low-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (top), medium (middle) and high-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (bottom) categories in the signal region.

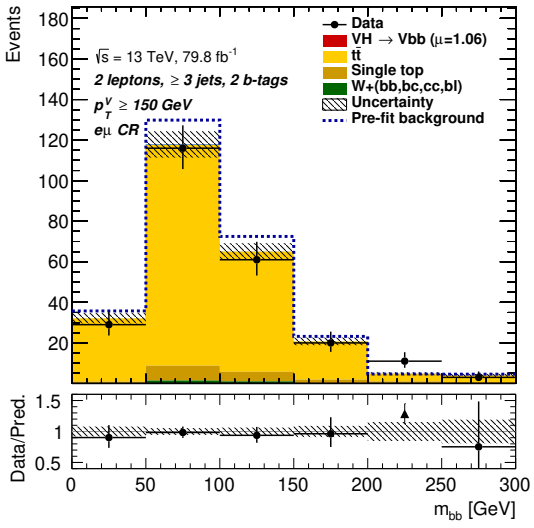


(a)

(b)

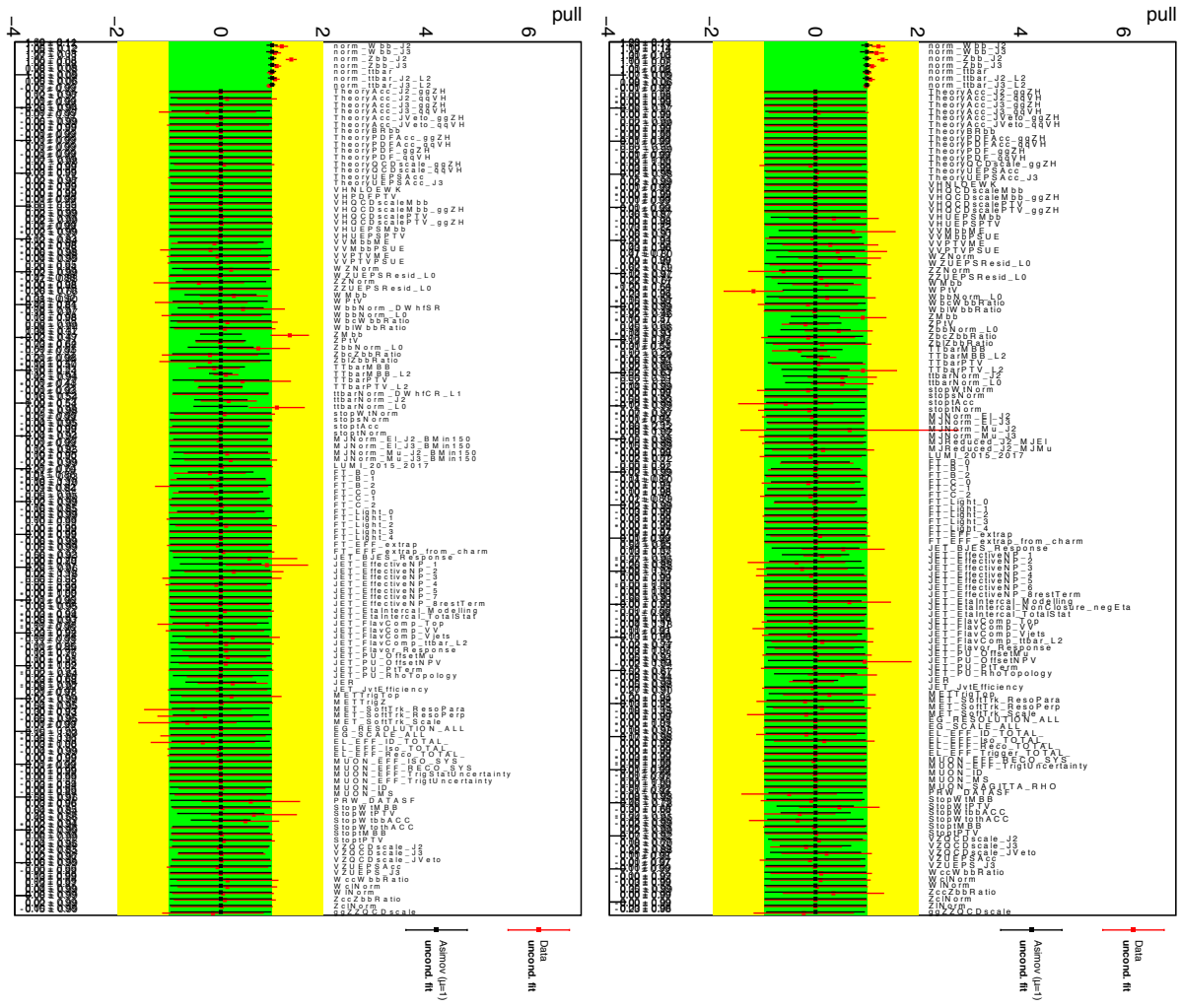


(c)



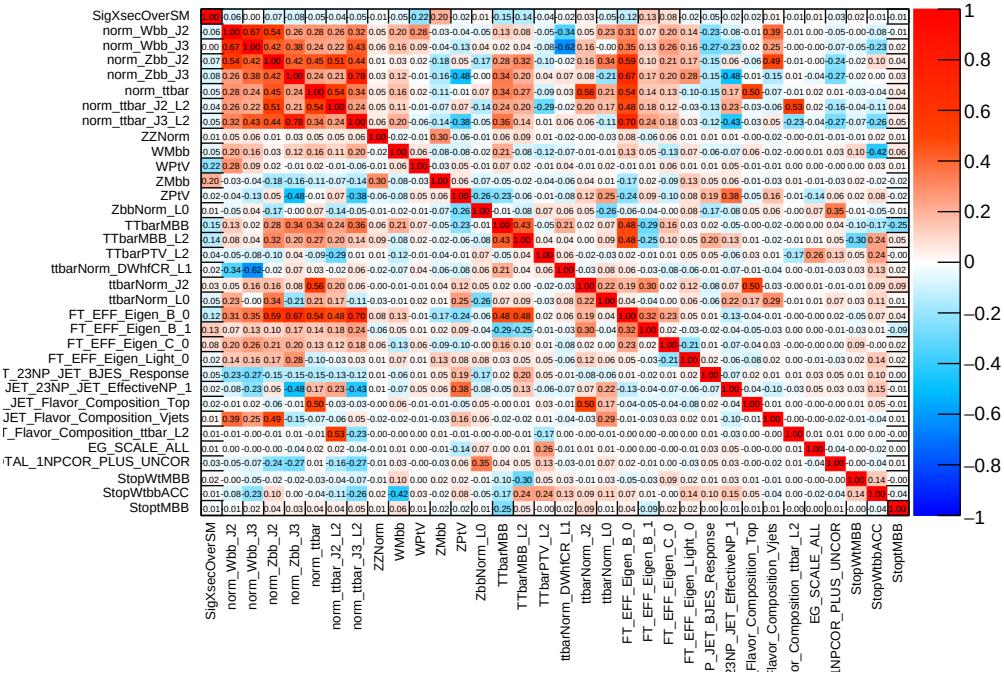
(d)

Figure C.4 – The m_{bb} post-fit distributions in the 2-leptons channel for 2-btag events, in the 2-jets (left) and exactly 3-jets (right) low-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (top), high-$p_{\mathrm{T}}^{\mathrm{V}}$ bin (bottom) categories in the top emu control region.

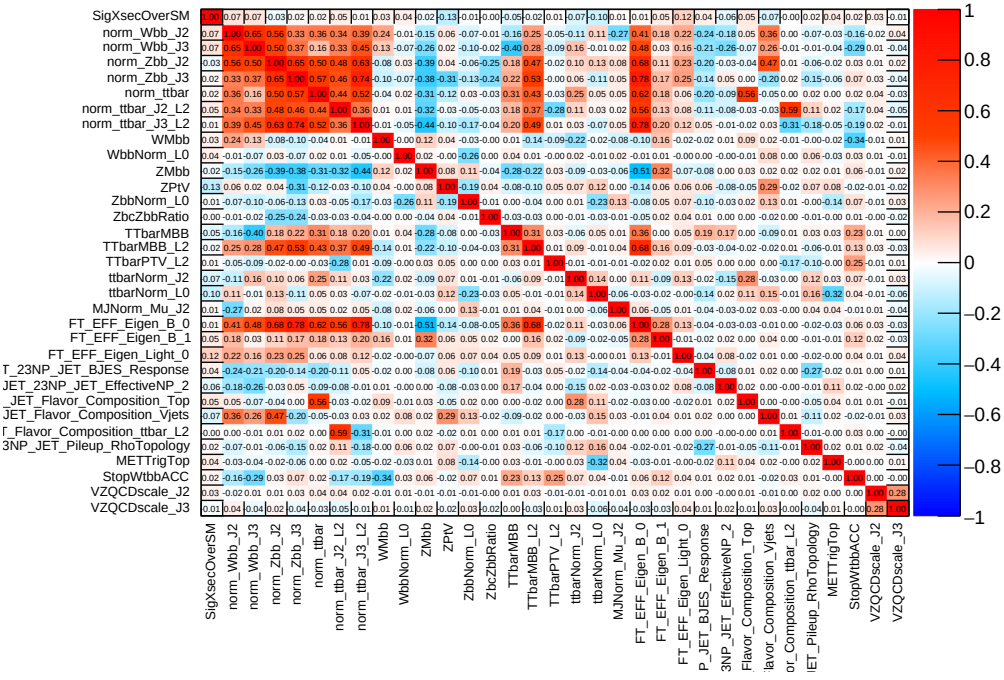


(a) (b)

Figure C.5 – Nuisance parameter pulls and the free parameter scale factors corresponding to a conditional combined fit performed to the Asimov dataset (black) and to the Run-2 data (red) for the MVA fit (a) and the CBA fit (b)

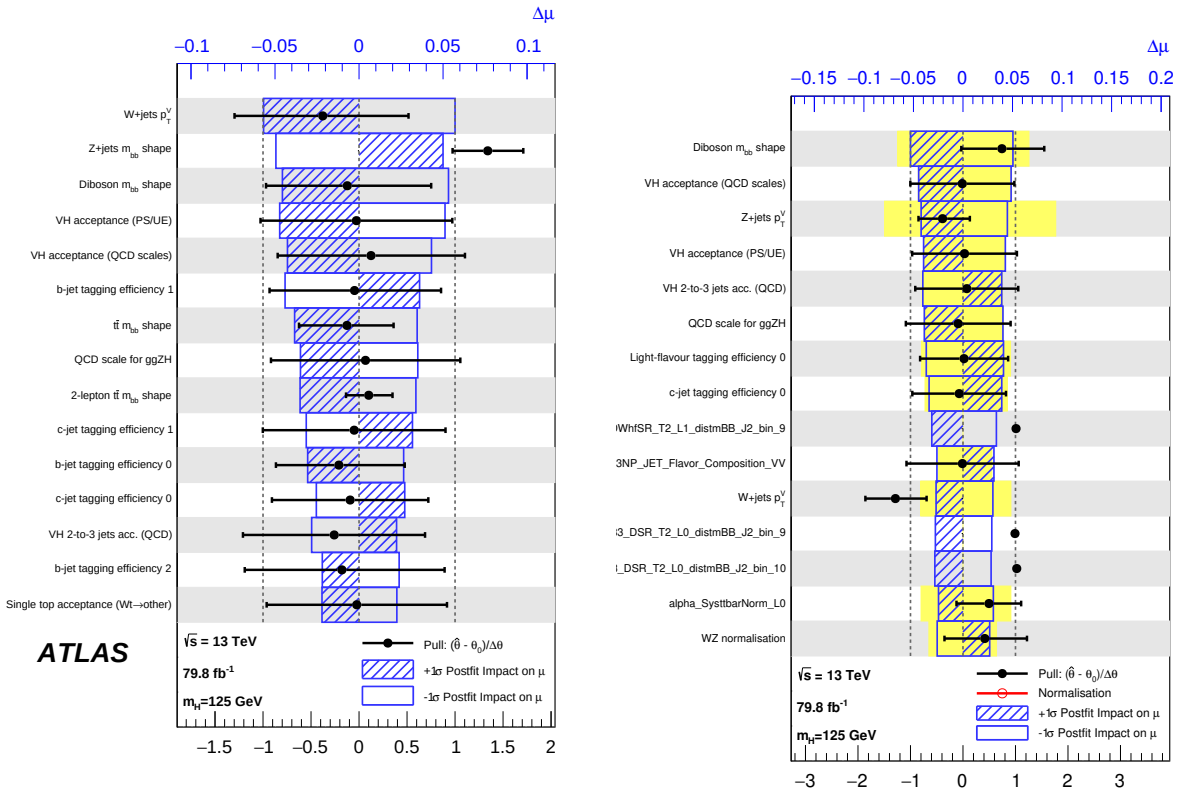


(a)



(b)

Figure C.6 – Correlation matrix from the MVA (a) and CBA (b) fit to the data. Only variables with at least one correlation with a magnitude greater than 0.25 are shown. The order of the axis is arbitrary as it comes from the list of systematics inputed by the analyzer in the framework.



(a) (b)

Figure C.7 – Ranking of the nuisance parameters based on the impact on the best fit value $\hat{\mu}$ for the MVA (a) and CBA (b) fit to the data. The hatched and open areas correspond to the upwards and downwards variations, respectively. The filled circles are representing the corresponding pulls for the specific NP. Only the 15 firstly ranked NPs based on the global effect are shown.

Source of uncertainty		σ_μ
Total		0.259
Statistical		0.161
Systematic		0.203
Experimental uncertainties		
Jets		0.035
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.014
Leptons		0.009
b-tagging	b-jets	0.061
	c-jets	0.042
	light-flavor jets	0.009
	extrapolation	0.008
Pile-up		0.007
Luminosity		0.023
Theoretical and modeling uncertainties		
Signal		0.094
Floating normalizations		0.035
Z + jets		0.055
W + jets		0.060
$\mathrm{t\bar{t}}$		0.050
Single top quark		0.028
Diboson		0.054
Multi-jet		0.005
MC statistical		0.070

Source of uncertainty		σ_μ
Total		0.347
Statistical		0.200
Systematic		0.283
Experimental uncertainties		
Jets		0.159
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.023
Leptons		0.009
b-tagging	b-jets	0.035
	c-jets	0.059
	light-flavor jets	0.040
	extrapolation	0.006
Pile-up		0.012
Luminosity		0.020
Theoretical and modeling uncertainties		
Signal		0.105
Floating normalizations		0.155
Z + jets		0.057
W + jets		0.040
$\mathrm{t\bar{t}}$		0.062
Single top quark		0.035
Diboson		0.160
Multi-jet		0.016
MC statistical		0.118

Figure C.8 – Breakdown of the contributions to the uncertainty in μ for the MVA (left) and CBA (right) fit. The sum in quadrature of the systematic uncertainties attached to the categories differs from the total systematic uncertainty due to correlations.

Bibliography

[1] Xavier Cortada. « In search of the Higgs boson: H → bottom bottom ». Apr. 2013. URL: https://cds.cern.ch/record/1541975 (cit. on pp. iii, 297).

[2] Daniel Stoljar. « Physicalism ». In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University, 2017 (cit. on p. 1).

[3] F. Englert and R. Brout. « Broken Symmetry and the Mass of Gauge Vector Mesons ». In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on pp. 1, 16).

[4] Peter W. Higgs. « Broken Symmetries and the Masses of Gauge Bosons ». In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on pp. 1, 16).

[5] Lyndon Evans and Philip Bryant. « LHC Machine ». In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001 (cit. on p. 1).

[6] G. Aad et al. « The ATLAS Experiment at the CERN Large Hadron Collider ». In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (cit. on pp. 1, 36, 46–53).

[7] S. Chatrchyan et al. « The CMS Experiment at the CERN LHC ». In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004 (cit. on pp. 1, 36).

[8] ATLAS Collaboration. « Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC ». In: *Physics Letters B* 716.1 (2012), pp. 1–29. DOI: https://doi.org/10.1016/j.physletb.

2012.08.020. URL: http://www.sciencedirect.com/science/article/pii/S037026931200857X (cit. on p. 1).

[9] Serguei Chatrchyan et al. « Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC ». In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex] (cit. on p. 1).

[10] Robert Rathbun Wilson. « The Tevatron ». In: *Phys. Today* 30N10 (1977), pp. 23–30. DOI: 10.1063/1.3037746 (cit. on p. 2).

[11] T. Aaltonen et al. « Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron ». In: *Phys. Rev. Lett.* 109 (2012), p. 071804. DOI: 10.1103/PhysRevLett.109.071804. arXiv: 1207.6436 [hep-ex] (cit. on p. 2).

[12] Serguei Chatrchyan et al. « Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks ». In: *Phys. Rev.* D89.1 (2014), p. 012003. DOI: 10.1103/PhysRevD.89.012003. arXiv: 1310.3687 [hep-ex] (cit. on p. 2).

[13] Georges Aad et al. « Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector ». In: *JHEP* 01 (2015), p. 069. DOI: 10.1007/JHEP01(2015)069. arXiv: 1409.6212 [hep-ex] (cit. on pp. 2, 100, 175, 186).

[14] Georges Aad et al. « Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV ». In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv: 1606.02266 [hep-ex] (cit. on p. 2).

[15] Georges Aad et al. *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a $b\bar{b}$ pair in pp collisions at 13 TeV using the ATLAS detector*. Tech. rep. ATLAS-CONF-2016-091. Geneva: CERN, Aug. 2016. URL: http://cds.cern.ch/record/2206813 (cit. on p. 2).

[16] M. Aaboud et al. « Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector ». In: *JHEP* 12 (2017), p. 024. DOI: 10.1007/JHEP12(2017)024. arXiv: 1708.03299 [hep-ex] (cit. on pp. 2, 176, 181, 186).

[17] Albert M Sirunyan et al. « Evidence for the Higgs boson decay to a bottom quark–antiquark pair ». In: *Phys. Lett.* B780 (2018), pp. 501–532. DOI: 10.1016/j.physletb.2018.02.050. arXiv: 1709.07497 [hep-ex] (cit. on p. 2).

[18] Roel Aaij et al. « Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity ». In: *JHEP* 02 (2016), p. 104. DOI: 10.1007/JHEP02(2016)104. arXiv: 1512.04442 [hep-ex] (cit. on p. 2).

[19] A. Abdesselam et al. « Angular analysis of $B^0 \to K^*(892)^0\ell^+\ell^-$ ». In: *Proceedings, LHCSki 2016 - A First Discussion of 13 TeV Results: Obergurgl, Austria, April 10-15, 2016*. 2016. arXiv: 1604.04042 [hep-ex] (cit. on p. 2).

[20] *Measurements and interpretations of Higgs-boson fiducial cross sections in the diphoton decay channel using 139 fb¯1 of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2019-029. Geneva: CERN, July 2019. URL: http://cds.cern.ch/record/2682800 (cit. on p. 2).

[21] Morad Aaboud et al. « Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector ». In: *Phys. Lett.* B786 (2018), pp. 59–86. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex] (cit. on pp. 3, 161, 175, 197, 200).

[22] J. J. Thomson. « Cathode rays ». In: *Phil. Mag. Ser.5* 44 (1897), pp. 293–316. DOI: 10.1080/14786449708621070 (cit. on p. 5).

[23] J. Chadwick. « Possible Existence of a Neutron ». In: *Nature* 129 (1932), p. 312. DOI: 10.1038/129312a0 (cit. on p. 5).

[24] Murray Gell-Mann. « A Schematic Model of Baryons and Mesons ». In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3 (cit. on p. 6).

[25] Martin Breidenbach, Jerome I. Friedman, Henry W. Kendall, Elliott D. Bloom, D. H. Coward, H. C. DeStaebler, J. Drees, Luke W. Mo, and Richard E. Taylor. « Observed Behavior of Highly Inelastic electron-Proton Scattering ». In: *Phys. Rev. Lett.* 23 (1969), pp. 935–939. DOI: 10.1103/PhysRevLett.23.935 (cit. on p. 6).

[26] Wolfgang Pauli. « Pauli letter collection: letter to Lise Meitner ». Typed copy. URL: https://cds.cern.ch/record/83282 (cit. on p. 6).

[27] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. « Detection of the free neutrino: A Confirmation ». In: *Science* 124 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103 (cit. on p. 6).

[28] M. Tanabashi et al. « Review of Particle Physics ». In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001 (cit. on pp. 6, 7, 15, 23, 24, 75, 82).

[29] Albert Einstein. « Concerning an heuristic point of view toward the emission and transformation of light ». In: *Annalen Phys.* 17 (1905), pp. 132–148 (cit. on p. 6).

[30] R. A. Millikan. « A Direct Determination of "h." » In: *Phys. Rev.* 4 (1 July 1914), pp. 73–75. DOI: 10.1103/PhysRev.4.73.2. URL: https://link.aps.org/doi/10.1103/PhysRev.4.73.2 (cit. on p. 6).

[31] Arthur H. Compton. « A Quantum Theory of the Scattering of X-rays by Light Elements ». In: *Phys. Rev.* 21 (1923), pp. 483–502. DOI: 10.1103/PhysRev.21.483 (cit. on p. 7).

[32] D. P. Barber et al. « Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies ». In: *Phys. Rev. Lett.* 43 (1979), p. 830. DOI: 10.1103/PhysRevLett.43.830 (cit. on p. 7).

[33] G. Arnison et al. « Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV ». In: *Phys. Lett.* B122 (1983), pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2 (cit. on p. 7).

[34] G. Arnison et al. « Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c^2 at the CERN SPS Collider ». In: *Phys. Lett.* B126 (1983), pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0 (cit. on p. 7).

[35] M. Banner et al. « Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider ». In: *Phys. Lett.* B122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2 (cit. on p. 7).

[36] P. Bagnaia et al. « Evidence for $Z_0 \to e^+e^-$ at the CERN anti-p p Collider ». In: *Phys. Lett.* B129 (1983), pp. 130–140. DOI: 10.1016/0370-2693(83)90744-X (cit. on p. 7).

[37] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. URL: http://www.slac.stanford.edu/~mpeskin/QFT.html (cit. on p. 9).

[38] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. « Experimental Test of Parity Conservation in Beta Decay ». In: *Phys. Rev.* 105 (1957), pp. 1413–1414. DOI: 10.1103/PhysRev.105.1413 (cit. on p. 10).

[39] Sidney A. Bludman. « On the universal Fermi interaction ». In: *Nuovo Cim.* 9 (1958), pp. 433–445. DOI: 10.1007/BF02725099 (cit. on p. 10).

[40] Chen-Ning Yang and Robert L. Mills. « Conservation of Isotopic Spin and Isotopic Gauge Invariance ». In: *Phys. Rev.* 96 (1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191 (cit. on p. 11).

[41] S. L. Glashow. « Partial Symmetries of Weak Interactions ». In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2 (cit. on p. 11).

[42] Roel Aaij et al. « Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$ ». In: *Phys. Rev. Lett.* 122.22 (2019), p. 222001. DOI: 10.1103/PhysRevLett.122.222001. arXiv: 1904.03947 [hep-ex] (cit. on p. 13).

[43] L. D. Faddeev and V. N. Popov. « Feynman Diagrams for the Yang-Mills Field ». In: *Phys. Lett.* B25 (1967), pp. 29–30. DOI: 10.1016/0370-2693(67)90067-6 (cit. on p. 14).

[44] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. « Global Conservation Laws and Massless Particles ». In: *Phys. Rev. Lett.* 13 (1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585 (cit. on p. 16).

[45] Steven Weinberg. « A Model of Leptons ». In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264 (cit. on p. 16).

[46] Abdus Salam. « Weak and Electromagnetic Interactions ». In: *Conf. Proc.* C680519 (1968), pp. 367–377 (cit. on p. 16).

[47] Nicola Cabibbo. « Unitary Symmetry and Leptonic Decays ». In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531 (cit. on p. 21).

[48] Makoto Kobayashi and Toshihide Maskawa. « CP Violation in the Renormalizable Theory of Weak Interaction ». In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652 (cit. on p. 21).

[49] S Heinemeyer, C Mariotti, G Passarino, and R Tanaka, eds. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. Oct. 2016. URL: https://cds.cern.ch/record/2227475 (cit. on pp. 22, 25, 26, 93, 96, 198).

[50] U. Aglietti et al. « Tevatron for LHC report: Higgs ». In: 2006. arXiv: hep-ph/0612172 [hep-ph]. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-06-467-E-T (cit. on p. 22).

[51] Morad Aaboud et al. « Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector ». In: *Phys. Rev.* D98.5 (2018), p. 052003. DOI: 10.1103/PhysRevD.98.052003. arXiv: 1807.08639 [hep-ex] (cit. on pp. 23, 187).

[52] Georges Aad et al. « Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment ». In: *Eur. Phys. J.* C76.1 (2016), p. 6. DOI: 10.1140/epjc/s10052-015-3769-y. arXiv: 1507.04548 [hep-ex] (cit. on pp. 26, 188).

[53] M. Aaboud et al. « Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment ». In: *Phys. Rev. Lett.* 120.21 (2018), p. 211802. DOI: 10.1103/PhysRevLett.120.211802. arXiv: 1802.04329 [hep-ex] (cit. on p. 26).

[54] *Search for the standard model Higgs boson decaying to charm quarks*. Tech. rep. CMS-PAS-HIG-18-031. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2682638 (cit. on p. 26).

[55] Morad Aaboud et al. « Searches for exclusive Higgs and Z boson decays into $J/\psi\gamma$, $\psi(2S)\gamma$, and $\Upsilon(nS)\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector ». In: *Phys. Lett.* B786 (2018), pp. 134–155. DOI: 10.1016/j.physletb.2018.09.024. arXiv: 1807.00802 [hep-ex] (cit. on p. 26).

[56] Thomas Sven Pettersson and P Lefèvre. *The Large Hadron Collider: conceptual design*. Tech. rep. CERN-AC-95-05-LHC. Oct. 1995. URL: https://cds.cern.ch/record/291782 (cit. on p. 30).

[57] Julie Haffner. « The CERN accelerator complex. Complexe des accélérateurs du CERN ». In: (Oct. 2013). URL: http://cds.cern.ch/record/1621894 (cit. on p. 30).

[58] L Arnaudon, P Baudrenghien, M Baylac, G Bellodi, Y Body, J Borburgh, P Bourquin, J Broere, O Brunner, L Bruno, C Carli, Friedhelm Caspers, S M Cousineau, Y Cuvet, C De Almeida Martins, T Dobers, T Fowler, R Garoby, F Gerigk, B Goddard, K Hanke, M Hori, M Jones, K Kahle, Willi Kalbreier, T Kroyer, D Küchler, A M Lombardi, L A López-Hernandez, M Magistris, M Martini, S Maury, E Page, M Paoluzzi, M Pasini, U Raich, C Rossi, J P Royer, E Sargsyan, J Serrano, R Scrivens, M Silari, M Timmins, W Venturini-Delsolaro, M Vretenar, R Wegner, W Weterings, and T Zickler. *Linac4 Technical Design Report*. Tech. rep. CERN-AB-2006-084. CARE-Note-2006-022-HIPPI. Geneva: CERN, Dec. 2006. URL: http://cds.cern.ch/record/1004186 (cit. on p. 31).

[59] *CERN Annual report 2017*. Tech. rep. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2624296 (cit. on p. 31).

[60] *LHC closer*. URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_running (cit. on p. 31).

[61] AC Team. « Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC ». June 1999. URL: https://cds.cern.ch/record/40524 (cit. on p. 33).

[62] Jean-Luc Caron. « LHC quadrupole cross section. » May 1998. URL: https://cds.cern.ch/record/841485 (cit. on p. 33).

[63] R. W. Assmann. « LEP luminosity revisited: Design and reality ». In: *Particle accelerator. Proceedings, 2nd Asian Conference, APAC'01, Beijing, P.R. China, September 17-21. 2001*. 2001, pp. 74–78. URL: http://accelconf.web.cern.ch/AccelConf/a01/PDF/WEAU01.pdf (cit. on p. 34).

[64] V. Papadimitriou. « Luminosity determination at the Tevatron ». In: *Proceedings, LHC Lumi Days, LHC Workshop on LHC Luminosity Calibration: Geneva, Switzerland, 13-14 Jan, 2011*. 2011. arXiv: 1106.5182 [physics.ins-det]. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-11-199 (cit. on p. 34).

[65] *Website of ATLAS Public Luminosity Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (cit. on p. 34).

[66] Frederick Bordry. « LHC status ». In: Rencontres de Moriond QCD 2019 (2019). 2019. URL: http://moriond.in2p3.fr/2019/QCD/Program.html (cit. on p. 35).

[67] C. Bernardini, G. F. Corazza, G. Di Giugno, G. Ghigo, R. Querzoli, J. Haissinski, P. Marin, and B. Touschek. « Lifetime and beam size in a storage ring ». In: *Phys. Rev. Lett.* 10 (1963), pp. 407–409. DOI: 10.1103/PhysRevLett.10.407 (cit. on p. 35).

[68] Andrew Presland, Brennan Goddard, J M Jiménez, D Ramos, and Raymond Veness. « A Large Diameter Entrance Window for the LHC Beam Dump Line ». In: LHC-Project-Report-823. CERN-LHC-Project-Report-823 (July 2005), 4 p. URL: http://cds.cern.ch/record/858492 (cit. on p. 35).

[69] A. Augusto Alves Jr. et al. « The LHCb Detector at the LHC ». In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005 (cit. on p. 36).

[70] K. Aamodt et al. « The ALICE experiment at the CERN LHC ». In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002 (cit. on p. 36).

[71] M Albrow, M Arneodo, V Avati, J Baechler, N Cartiglia, M Deile, M Gallinaro, J Hollar, M Lo Vetere, K Oesterberg, N Turini, J Varela, D Wright, and Collaboration CMS-TOTEM. *CMS-TOTEM Precision Proton Spectrometer*. Tech. rep. CERN-LHCC-2014-021. TOTEM-TDR-003. CMS-TDR-13. Sept. 2014. URL: https://cds.cern.ch/record/1753795 (cit. on pp. 36, 52).

[72] O Adriani, L Bonechi, M Bongi, R D'ALessandro, D A Faus, M Haguenauer, Y Itow, K Kasahara, K Masuda, Y Matsubara, H Menjo, Y Muraki, P Papini, T Sako, T Tamura, S Torii, A Tricomi, W C Turner, J Velasco, and K Yoshida. *LHCf experiment: Technical Design Report*. Technical Design Report LHCf. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/926196 (cit. on p. 36).

[73] « The Pierre Auger Project Design Report ». In: (1996) (cit. on p. 36).

[74] Hermann G. « The HESS array: a new system of 100 GeV IACTs for stereoscopic observations ». In: *Proceedings, 32nd Rencontres de Moriond 17th atrophysics meeting extragalactic astronomy in the infrared: Les Arcs, France, Mar 15-22, 1997*. Edition Frontieres. Paris: Edition Frontieres, 1997 (cit. on p. 36).

[75] E. Pare, T. Doke, M. Haguenauer, V. Innocente, K. Kasahara, T. Kashiwagi, J. Kikuchi, S. Lanzano, K. Masuda, H. Murakami, Y. Muraki, T. Nakada, A. Nakamoto, and T. Yuda. « Inclusive production of π_0'S in the fragmentation region at the SppS collider ». In: *Physics Letters B* 242.3 (1990), pp. 531–535. DOI: https://doi.org/10.1016/0370-2693(90)91807-N. URL: http://www.sciencedirect.com/science/article/pii/037026939091807N (cit. on p. 36).

[76] James Pinfold et al. « Technical Design Report of the MoEDAL Experiment ». In: (2009) (cit. on p. 36).

[77] Joao Pequenao. « Computer generated image of the whole ATLAS detector ». Mar. 2008. URL: https://cds.cern.ch/record/1095924 (cit. on p. 38).

[78] Matthias Schott and Monica Dunford. « Review of single vector boson production in pp collisions at $\sqrt{s}$=7 TeV ». In: *The European Physical Journal C* 74.7 (July 2014), p. 2916. DOI: 10.1140/epjc/s10052-014-2916-1. URL: https://doi.org/10.1140/epjc/s10052-014-2916-1 (cit. on p. 38).

[79] Joao Pequenao and Paul Schaffner. « How ATLAS detects particles: diagram of particle paths in the detector ». Jan. 2013. URL: https://cds.cern.ch/record/1505342 (cit. on p. 39).

[80] G. Aad et al. « The ATLAS Inner Detector commissioning and calibration ». In: *Eur. Phys. J.* C70 (2010), pp. 787–821. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293 [physics.ins-det] (cit. on p. 39).

[81] Joao Pequenao. « Computer generated image of the ATLAS inner detector ». Mar. 2008. URL: http://cds.cern.ch/record/1095926 (cit. on p. 40).

[82] Simon Ramo. « Currents induced by electron motion ». In: *Proc. Ire.* 27 (1939), pp. 584–585. DOI: 10.1109/JRPROC.1939.228757 (cit. on p. 41).

[83] Manfred Krammer. « Silicon Detectors ». URL: http://www.hephy.at/fileadmin/user_upload/Lehre/Unterlagen/Praktikum/Halbleiterdetektoren.pdf (cit. on p. 41).

[84] I. Peric, L. Blanquart, G. Comes, P. Denes, K. Einsweiler, P. Fischer, E. Mandelli, and Gerrit Jan Meddeler. « The FEI3 readout chip for the ATLAS pixel detector ». In: *Nucl. Instrum. Meth.* A565 (2006), pp. 178–187. DOI: 10.1016/j.nima.2006.05.032 (cit. on p. 41).

[85] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. ATLAS-TDR-19. 2010. URL: %7Bhttps://cds.cern.ch/record/1291633%7D (cit. on pp. 41, 42, 216).

[86] *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. ATL-PHYS-PUB-2015-022. Geneva: CERN, July 2015. URL: http://cds.cern.ch/record/2037697 (cit. on p. 42).

[87] M. Garcia-Sciveres et al. « The FE-I4 pixel readout integrated circuit ». In: *Nucl. Instr. and Meth. A* 636 (2011), S155–S159 (cit. on pp. 42, 213).

[88] The ATLAS TRT collaboration. « The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance ». In: *JINST* (), 3.02 (2008), P02013. (Cit. on p. 43).

[89] Klaus Pretzl. « TOPICAL REVIEW: Calorimeters in astro and particle physics ». In: *Journal of Physics G Nuclear Physics* 31.7 (July 2005), R133–R149. DOI: 10.1088/0954-3899/31/7/R01. arXiv: physics/0502065 [physics.ins-det] (cit. on p. 45).

[90] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375 (cit. on p. 45).

[91] G. Aad et al. « Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons ». In: *Eur. Phys. J.* C70 (2010), pp. 755–785. DOI: 10.1140/epjc/s10052-010-1403-6. arXiv: 1002.4189 [physics.ins-det] (cit. on p. 47).

[92] M. Aharrouche, J. Colas, L. Di Ciaccio, M. El Kacimi, O. Gaumer, M. Gouanere, D. Goujdami, R. Lafaye, S. Laplace, C. Le Maner, L. Neukermans, P. Perrodo, L. Poggioli, D. Prieur, H. Przysiezniak, G. Sauvage, F. Tarrade, I. Wingerter-Seez, R. Zitoun, F. Lanni, H. Ma, S. Rajagopalan, S. Rescia, H. Takai, A. Belymam, D. Benchekroun, M. Hakimi, A. Hoummada, E. Barberio, Y.S. Gao, L. Lu, R. Stroynowski, M. Aleksa, J. Beck Hansen, T. Carli, I. Efthymiopoulos, P. Fassnacht, F. Follin, F. Gianotti, L. Hervas, W. Lampl, Johann Collot, J.Y. Hostachy, Fabienne Ledroit, P. Martin, Fairouz Malek, S. Saboumazrag, M. Leltchouk, J.A. Parsons, M. Seman, S. Simion, D. Banfi, L. Carminati, D. Cavalli, G. Costa, M. Delmastro, M. Fanti, L. Mandelli, M. Mazzanti, G.F. Tartarelli, C. Bourdarios, L. Fayard, D. Fournier, G. Graziani, S. Hassani, L. Iconomidou-Fayard, M. Kado, M. Lechowski, M. Lelas, G. Parrour, P. Puzo, D. Rousseau, R. Sacco, L. Serin, G. Unal, D. Zerwas, A. Camard, D. Lacour, B. Laforge, I. Nikolic-Audit, Ph. Schwemling, H. Ghazlane, R. Cherkaoui El Moursli, A. Idrissi Fakhr-Eddine, M. Boonekamp, N.

Kerschen, B. Mansoulie, P. Meyer, J. Schwindling, B. Lund-Jensen, and Y. Tayalati. « Energy Linearity and Resolution of the ATLAS Electromagnetic Barrel Calorimeter in an Electron Test-Beam ». In: *Nucl. Instrum. Methods Phys. Res., A* 568.physics/0608012 (Aug. 2006), 601–623. 48 p. DOI: 10.1016/j.nima.2006.07.053. URL: http://cds.cern.ch/record/976098 (cit. on p. 47).

[93] Morad Aaboud et al. « Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data ». In: *Submitted to: JINST* (2018). arXiv: 1812.03848 [hep-ex] (cit. on pp. 47, 72, 73).

[94] E Khramov, N Rusakovich, T Carli, A Henriques, V Giangiobbe, Z Liang, C Santoni, and M Simonyan. *Study of the Response of the Hadronic Barrel Calorimeter in the ATLAS Combined Test-beam to Pions of Energies from 20 to 350 GeV for Beam Impact Points from 0.2 to 0.65*. Tech. rep. ATL-TILECAL-PUB-2009-007. ATL-COM-TILECAL-2009-006. Geneva: CERN, Apr. 2009. URL: http://cds.cern.ch/record/1172156 (cit. on p. 48).

[95] A. Artamonov et al. « The ATLAS forward calorimeters ». In: *JINST* 3 (2008), P02010. DOI: 10.1088/1748-0221/3/02/P02010 (cit. on p. 49).

[96] A. Yamamoto et al. « The ATLAS central solenoid ». In: *Nucl. Instrum. Meth.* A584 (2008), pp. 53–74. DOI: 10.1016/j.nima.2007.09.047 (cit. on p. 51).

[97] « ATLAS barrel toroid: Technical design report ». In: (1997) (cit. on p. 51).

[98] « ATLAS endcap toroids: Technical design report ». In: (1997) (cit. on p. 51).

[99] A Hervé, Bertrand Blau, P Brédy, D Campi, P Cannarsa, B Curé, T F Dupont, P Fabbricatore, S Farinon, F Feyzi, P Fazilleau, A Gaddi, H Gerwig, Michela Greco, J P Grillet, V Kaftanov, F Kircher, V Klyukhin, B Levesy, R Loveless, G Maire, R Musenich, Y Pabot, A Payn, G Perinic, P Petiot, F Rondeaux, H Rykaczewski, E Sbrissa, S Sequeira-Lopes-Tavares, Stefano Sgobba, R P Smith, L Veillet, and G Waurick. « Status of the construction of the CMS magnet ». In: *IEEE Trans. Appl. Supercond.* 14.2 (2004), 542–547. 6 p. DOI: 10.1109/TASC.2004.829715. URL: https://cds.cern.ch/record/806410 (cit. on p. 52).

[100] Peter Jenni, Marzio Nessi, and Markus Nordberg. *Zero Degree Calorimeters for ATLAS*. Tech. rep. CERN-LHCC-2007-001. LHCC-I-016. Geneva: CERN, Jan. 2007. URL: https://cds.cern.ch/record/1009649 (cit. on p. 53).

[101] S. Abdel Khalek et al. « The ALFA Roman Pot Detectors of ATLAS ». In: *JINST* 11.11 (2016), P11013. DOI: 10.1088/1748-0221/11/11/P11013. arXiv: 1609.00249 [physics.ins-det] (cit. on p. 54).

[102] ATLAS collaboration. « Performance of the ATLAS trigger system in 2015 ». In: *The European Physical Journal C* 77.5 (May 2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. URL: https://doi.org/10.1140/epjc/s10052-017-4852-3 (cit. on pp. 56, 58, 59, 71).

[103] *ATLAS level-1 trigger: Technical Design Report*. Tech. rep. Geneva, 1998. URL: https://cds.cern.ch/record/381429 (cit. on p. 56).

[104] ATLAS Collaboration. « The ATLAS Level-1 Calorimeter Trigger ». In: *Journal of Instrumentation* 3.03 (2008), P03001. URL: http://stacks.iop.org/1748-0221/3/i=03/a=P03001 (cit. on p. 57).

[105] ATLAS Collaboration. « The Level-1 Trigger Muon Barrel System of the ATLAS experiment at CERN ». In: *Journal of Instrumentation* 4.04 (2009), P04010. URL: http://stacks.iop.org/1748-0221/4/i=04/a=P04010 (cit. on pp. 57, 58).

[106] *Website of ATLAS Public Trigger Operation Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults (cit. on p. 58).

[107] ATLAS Collaboration. « The ATLAS central level-1 trigger logic and TTC system ». In: *Journal of Instrumentation* 3 (Aug. 2008), P08002. DOI: 10.1088/1748-0221/3/08/P08002 (cit. on p. 58).

[108] *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. Tech. rep. ATLAS-CONF-2012-042. Geneva: CERN, Mar. 2012. URL: https://cds.cern.ch/record/1435196 (cit. on p. 59).

[109] M Shochet, L Tompkins, V Cavaliere, P Giannetti, A Annovi, and G Volpi. *Fast TracKer (FTK) Technical Design Report*. Tech. rep. CERN-LHCC-2013-007. ATLAS-TDR-021. June 2013. URL: https://cds.cern.ch/record/1552953 (cit. on pp. 59, 60, 62, 63, 68).

[110] W Ashmanskas, A Bardi, M Bari, S Belforte, J Berryhill, M Bogdan, A Cerri, A.G Clark, G Chlachidze, R Condorelli, R Culberston, M Dell'Orso, S Donati, H.J Frisch, S Galeotti, P Giannetti, V Glagolev, A Leger, E Meschi, F Morsani, T Nakaya, G Punzi, L Ristori, H Sanders, A Semenov, G Signorelli, M Shochet, T Speer, F Spinella, P Wilson, X Wu, and A Zanetti. « The CDF silicon vertex tracker ». In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 477.1 (2002). 5th Int. Conf. on Position-Sensitive Detectors, pp. 451–455. DOI: https://doi.org/10.1016/S0168-9002(01)01830-7. URL: http://www.sciencedirect.com/science/article/pii/S0168900201018307 (cit. on p. 60).

[111] Alexander Schrijver. « On the History of Combinatorial Optimization (Till 1960) ». In: *Discrete Optimization*. Ed. by K. Aardal, G.L. Nemhauser, and R. Weismantel. Vol. 12. Handbooks in Operations Research and Management Science. Elsevier, 2005, pp. 1–68. DOI: https://doi.org/10.1016/S0927-0507(05)12001-5. URL: http://www.sciencedirect.com/science/article/pii/S0927050705120015 (cit. on p. 63).

[112] *Website of ATLAS Public FTK Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FTKPublicResults (cit. on pp. 67, 69).

[113] The ATLAS TDAQ Collaboration. « The ATLAS Data Acquisition and High Level Trigger system ». In: *Journal of Instrumentation* 11.06 (2016), P06008. URL: http://stacks.iop.org/1748-0221/11/i=06/a=P06008 (cit. on p. 69).

[114] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735 (cit. on p. 72).

[115] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012. URL: https://cds.cern.ch/record/1449796 (cit. on p. 72).

[116] *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2016-024. Geneva: CERN, June 2016. URL: https://cds.cern.ch/record/2157687 (cit. on p. 72).

[117] Morad Aaboud et al. « Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV ». In: *Eur. Phys. J.* C79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det] (cit. on pp. 73, 99).

[118] Georges Aad et al. « Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV ». In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex] (cit. on pp. 73–75).

[119] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. ATL-PHYS-PUB-2015-045. Geneva: CERN, Nov. 2015. URL: http://cds.cern.ch/record/2064383 (cit. on pp. 76, 77).

[120] T Barillari, E Bergeaas Kuutmann, T Carli, J Erdmann, P Giovannini, K J Grahn, C Issever, A Jantsch, A Kiryunin, K Lohwasser, A Maslennikov, S Menke, H Oberlack, G Pospelov, E Rauter, P Schacht, F Spanó, P Speckmayer, P Stavina, and P Strízenec. *Local Hadronic Calibration*. Tech. rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001. Geneva: CERN, June 2008. URL: https://cds.cern.ch/record/1112035 (cit. on p. 76).

[121] Georges Aad et al. « Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in *pp* collisions at $\sqrt{s}$=8 TeV ». In: *Eur. Phys. J.* C75.7 (2015), p. 303. DOI: 10.1140/epjc/s10052-015-3500-z. arXiv: 1412.7086 [hep-ex] (cit. on p. 76).

[122] Georges Aad et al. « Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1 ». In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex] (cit. on p. 77).

[123] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. « The anti-k_t jet clustering algorithm ». In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph] (cit. on p. 78).

[124] Stephen D. Ellis and Davison E. Soper. « Successive combination jet algorithm for hadron collisions ». In: *Phys. Rev.* D48 (1993), pp. 3160–3166. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266 [hep-ph] (cit. on p. 78).

[125] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. « Better jet clustering algorithms ». In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph] (cit. on p. 78).

[126] L Asquith, B Brelier, J M Butterworth, M Campanelli, T Carli, G Choudalakis, P A Delsart, S De Cecco, P O Deviveiros, M D'Onofrio, S Eckweiler, E Feng, P Francavilla, S Grinstein, I La Plante, J Huston, N Ghodbane, D Lopez Mateos, B Martin, N Makovec, S Majewsky, M Martinez, D W Miller, J Monk, K Perez, C Roda, J Robinson, A Schwartzmann, F Spano, K Terashi, F Vives, P Weber, and S Zenz. *Performance of Jet Algorithms in the ATLAS Detector*. Tech. rep. ATL-PHYS-INT-2010-129. Geneva: CERN, Dec. 2010. URL: https://cds.cern.ch/record/1311867 (cit. on p. 78).

[127] M. Aaboud et al. « Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector ». In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex] (cit. on pp. 78–80).

[128] Matteo Cacciari and Gavin P. Salam. « Pileup subtraction using jet areas ». In: *Phys. Lett.* B659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph] (cit. on p. 79).

[129] Georges Aad et al. « Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector ». In: *Eur. Phys. J.* C75 (2015), p. 17. DOI: 10.1140/epjc/s10052-014-3190-y. arXiv: 1406.0076 [hep-ex] (cit. on p. 79).

[130] ATLAS Collaboration. *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*. Tech. rep. ATLAS-COM-CONF-2015-024. Geneva: CERN, May 2015. URL: https://cds.cern.ch/record/2016323 (cit. on p. 80).

[131] *Tagging and suppression of pileup jets*. Tech. rep. ATL-PHYS-PUB-2014-001. Geneva: CERN, Jan. 2014. URL: http://cds.cern.ch/record/1643929 (cit. on pp. 80–82).

[132] Georges Aad et al. « Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector ». In: *Eur. Phys. J.* C76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex] (cit. on pp. 80, 81).

[133] N. S. Altman. « An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression ». In: *The American Statistician* 46.3 (1992), pp. 175–185. DOI: 10.1080/00031305.1992.10475879. eprint: https://www.tandfonline.com/doi/pdf/10.1080/00031305.1992.10475879. URL: https://www.tandfonline.com/doi/abs/10.1080/00031305.1992.10475879 (cit. on p. 81).

[134] Georges Aad et al. « Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector ». In: *Eur. Phys. J.* C74.8 (2014), p. 3023. DOI: 10.1140/epjc/s10052-014-3023-z. arXiv: 1405.6583 [hep-ex] (cit. on p. 82).

[135] Georges Aad et al. « Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC ». In: *The European Physical Journal C* 77.5 (May 2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [hep-ex] (cit. on pp. 82, 83).

[136] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. ATL-PHYS-PUB-2018-025. Geneva: CERN, Oct. 2018. URL: https://cds.cern.ch/record/2645405 (cit. on p. 83).

[137] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. Tech. rep. ATL-PHYS-PUB-2016-012. Geneva: CERN, June 2016. URL: https://cds.cern.ch/record/2160731 (cit. on pp. 82, 83).

[138] *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. ATL-PHYS-PUB-2017-011. Geneva: CERN, June 2017. URL: https://cds.cern.ch/record/2270366 (cit. on p. 83).

[139] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. Tech. rep. ATL-PHYS-PUB-2016-012. Geneva: CERN, June 2016. URL: https://cds.cern.ch/record/2160731 (cit. on p. 84).

[140] Morad Aaboud et al. « Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV ». In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex] (cit. on pp. 84, 85).

[141] *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*. Tech. rep. ATLAS-CONF-2018-006. Geneva: CERN, Apr. 2018. URL: https://cds.cern.ch/record/2314418 (cit. on pp. 84, 85).

[142] *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2018-001. Geneva: CERN, Mar. 2018. URL: https://cds.cern.ch/record/2306649 (cit. on pp. 84, 85).

[143] Morad Aaboud et al. « Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV ». In: *Eur. Phys. J.* C78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex] (cit. on pp. 87–89, 91).

[144] *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2015-023. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037700 (cit. on p. 89).

[145] *Object-based missing transverse momentum significance in the ATLAS detector*. Tech. rep. ATLAS-CONF-2018-038. Geneva: CERN, July 2018. URL: http://cds.cern.ch/record/2630948 (cit. on p. 90).

[146] M. Aaboud et al. « Measurement of b-hadron pair production with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV ». In: *JHEP* 11 (2017), p. 062. DOI: 10.1007/JHEP11(2017)062. arXiv: 1705.03374 [hep-ex] (cit. on p. 93).

[147] Serguei Chatrchyan et al. « Measurement of the cross section for production of bb^- bar X, decaying to muons in pp collisions at $\sqrt{s} = 7$ TeV ». In: *JHEP* 06 (2012), p. 110. DOI: 10.1007/JHEP06(2012)110. arXiv: 1203.3458 [hep-ex] (cit. on p. 93).

[148] Andy Buckley et al. « General-purpose event generators for LHC physics ». In: *Phys. Rept.* 504 (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph] (cit. on p. 95).

[149] William Buttinger. *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*. Tech. rep. ATL-COM-SOFT-2015-119. Geneva: CERN, May 2015. URL: https://cds.cern.ch/record/2014726 (cit. on p. 95).

[150] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. « $HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO ». In: *JHEP* 10 (2013), p. 083. DOI: 10.1007/JHEP10(2013)083. arXiv: 1306.2542 [hep-ph] (cit. on p. 96).

[151] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. « A Brief Introduction to PYTHIA 8.1 ». In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph] (cit. on p. 96).

[152] Georges Aad et al. « Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector ». In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex] (cit. on p. 96).

[153] Richard D. Ball et al. « Parton distributions for the LHC Run II ». In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph] (cit. on p. 96).

[154] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. « NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders ». In: *Phys. Lett.* B579 (2004), pp. 149–156. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206 [hep-ph] (cit. on p. 96).

[155] J. Ohnemus and W. J. Stirling. « Order-α_s corrections to the differential cross section for the WH intermediate-mass Higgs-boson signal ». In: *Phys. Rev. D* 47 (7 Apr. 1993), pp. 2722–2729. DOI: 10.1103/PhysRevD.47.2722. URL: https://link.aps.org/doi/10.1103/PhysRevD.47.2722 (cit. on p. 96).

[156] Lukas Altenkamp, Stefan Dittmaier, Robert V. Harlander, Heidi Rzehak, and Tom J.E. Zirke. « Gluon-induced Higgs-strahlung at next-to-leading order QCD ». In: *JHEP* 02.arXiv:1211.5015. CERN-PH-TH-2012-312. FR-PHENO-2012-023. WUB-12-21 (Nov. 2012), 078. 30 p. URL: https://cds.cern.ch/record/1495529 (cit. on p. 96).

[157] Robert V. Harlander, Anna Kulesza, Vincent Theeuwes, and Tom Zirke. « Soft gluon resummation for gluon-induced Higgs Strahlung ». In: *JHEP* 11 (2014), p. 082. DOI: 10.1007/JHEP11(2014)082. arXiv: 1410.0217 [hep-ph] (cit. on p. 96).

[158] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. « Event generation with SHERPA 1.1 ». In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph] (cit. on p. 96).

[159] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber. « QCD matrix elements + parton showers ». In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231 [hep-ph] (cit. on p. 96).

[160] Nils Lavesson and Leif Lonnblad. « W+jets matrix elements and the dipole cascade ». In: *JHEP* 07 (2005), p. 054. DOI: 10.1088/1126-6708/2005/07/054. arXiv: hep-ph/0503293 [hep-ph] (cit. on p. 96).

[161] Michał Czakon, Paul Fiedler, and Alexander Mitov. « Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha\frac{4}{S})$ ». In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph] (cit. on p. 97).

[162] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann. « HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR ». In: *Comput. Phys. Commun.* 182 (2011), pp. 1034–1046. DOI: 10.1016/j.cpc.2010.12.040. arXiv: 1007.1327 [hep-ph] (cit. on p. 97).

[163] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer. « HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions ». In: *Comput. Phys. Commun.* 191 (2015), pp. 74–89. DOI: 10.1016/j.cpc.2015.02.001. arXiv: 1406.4403 [hep-ph] (cit. on p. 97).

[164] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. « TMVA: Toolkit for Multivariate Data Analysis ». In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039 (cit. on pp. 107, 108, 148).

[165] *ROOT Data Analysis Framework, https://root.cern.ch/*. URL: https://root.cern.ch/ (cit. on p. 107).

[166] Yoav Freund and Robert E. Schapire. « A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting ». In: *J. Comput. Syst. Sci.* 55.1 (1997), pp. 119–139. DOI: 10.1006/jcss.1997.1504 (cit. on pp. 109, 149).

[167] M. Aaboud et al. « Measurement of the $b\bar{b}$ dijet cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector.. Measurement of the $b\bar{b}$ dijet cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector ». In: *Eur. Phys. J. C* 76.CERN-EP-2016-091. 12 (July 2016), 670. 36 p. DOI: 10.1140/epjc/s10052-016-4521-y. URL: https://cds.cern.ch/record/2202397 (cit. on pp. 123, 126).

[168] Andrew Stuart Bell and Francesco Lo Sterzo. *Signal and Background Modelling Studies for the Standard Model $VH, H \to b\bar{b}$ Analysis*. Tech. rep. ATL-COM-PHYS-2018-505. Geneva: CERN, May 2018. URL: https://cds.cern.ch/record/2316951 (cit. on pp. 130, 131).

[169] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. « Rivet user manual ». In: *Comput. Phys. Commun.* 184 (2013), pp. 2803–2819. DOI: 10.1016/j.cpc.2013.05.021. arXiv: 1003.0694 [hep-ph] (cit. on p. 135).

[170] J R Andersen et al. « Handbook of LHC Higgs Cross Sections: 3. Higgs Properties ». In: (2013). Ed. by S Heinemeyer, C Mariotti, G Passarino, and R Tanaka. DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph] (cit. on pp. 135, 198, 200).

[171] Bruce Mellado Garcia, Pasquale Musella, Massimiliano Grazzini, and Robert Harlander. « CERN Report 4: Part I Standard Model Predictions ». In: (May 2016). URL: https://cds.cern.ch/record/2150771 (cit. on p. 135).

[172] Iain W. Stewart and Frank J. Tackmann. « Theory Uncertainties for Higgs and Other Searches Using Jet Bins ». In: *Phys. Rev.* D85 (2012), p. 034011. DOI: 10.1103/PhysRevD.85.034011. arXiv: 1107.2117 [hep-ph] (cit. on pp. 136, 137).

[173] Roel Aaij et al. « Observation of CP violation in charm decays ». In: (2019). arXiv: 1903.08726 [hep-ex] (cit. on pp. 146, 147).

[174] D. Martschei, M. Feindt, S. Honc, and J. Wagner-Kuhr. « Advanced event reweighting using multivariate analysis ». In: *J. Phys. Conf. Ser.* 368 (2012), p. 012028. DOI: 10.1088/1742-6596/368/1/012028 (cit. on p. 148).

[175] A. Rogozhnikov. « Reweighting with Boosted Decision Trees ». In: *J. Phys. Conf. Ser.* 762.1 (2016), p. 012036. DOI: 10.1088/1742-6596/762/1/012036. arXiv: 1608.05806 [physics.data-an] (cit. on pp. 149, 150).

[176] *scikit-learn*. URL: https://scikit-learn.org/stable/index.html (cit. on p. 150).

[177] *hep_ml git repository*. URL: https://github.com/arogozhnikov/hep_ml (cit. on p. 150).

[178] Karl Pearson F.R.S. « On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling ». In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50.302 (1900), pp. 157–175. DOI: 10.1080/14786440009463897. eprint: https://doi.org/10.1080/14786440009463897. URL: https://doi.org/10.1080/14786440009463897 (cit. on p. 150).

[179] N. D. Gagunashvili. « Comparison of weighted and unweighted histograms ». In: *arXiv e-prints*, physics/0605123 (May 2006), physics/0605123. DOI: 10.1142/9781860948985_0010. arXiv: physics/0605123 [physics.data-an] (cit. on p. 151).

[180] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. « The Earth Mover's Distance as a Metric for Image Retrieval ». In: *International Journal of Computer Vision* 40.2 (Nov. 2000), pp. 99–121. DOI: 10.1023/A:1026543900054. URL: https://doi.org/10.1023/A:1026543900054 (cit. on p. 151).

[181] Monge G. « Mémoire sur la théorie des déblais et des remblais. » In: *Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année,* (1781), pp. 666–704. eprint: "ark:/12148/bpt6k35800" (cit. on p. 151).

[182] Hunyong Cho, Gregory J. Matthews, and Ofer Harel. « Confidence intervals for the area under the receiver operating characteristic curve in the presence of ignorable missing data ». In: *arXiv e-prints*, arXiv:1804.05882 (Apr. 2018), arXiv:1804.05882. arXiv: 1804.05882 [stat.AP] (cit. on pp. 152, 159, 244, 251, 260).

[183] T. Adam et al. « Measurement of the neutrino velocity with the OPERA detector in the CNGS beam ». In: *JHEP* 10 (2012), p. 093. DOI: 10.1007/JHEP10(2012)093. arXiv: 1109.4897 [hep-ex] (cit. on p. 164).

[184] Edwin Cartlidge. « Error Undoes Faster-Than-Light Neutrino Results ». In: *Science* (2012) (cit. on p. 164).

[185] Daniel Ferenc, Dan Ferenc Šegedin, Ivan Ferenc Šegedin, and Marija Šegedin Ferenc. « Helium Migration through Photomultiplier Tubes – The Probable Cause of the DAMA Seasonal Variation Effect ». In: (2019). arXiv: 1901.02139 [physics.ins-det] (cit. on p. 164).

[186] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. « Asymptotic formulae for likelihood-based tests of new physics ». In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an] (cit. on pp. 164–166).

[187] Abraham Wald. « Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large ». In: *Transactions of the American Mathematical Society* 54.3 (1943), pp. 426–482. URL: http://www.jstor.org/stable/1990256 (cit. on p. 165).

[188] I. Asimov and D. Shannon. *Franchise*. Creative Classic. Creative Education, 1989 (cit. on p. 166).

[189] Roger Barlow and Christine Beeston. « Fitting using finite Monte Carlo samples ». In: *Computer Physics Communications* 77.2 (1993), pp. 219–228. DOI: https://doi.org/10.1016/0010-4655(93)90005-W. URL: http://www.sciencedirect.com/science/article/pii/001046559390005W (cit. on p. 169).

[190] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. CERN-OPEN-2012-016. New York: New York U., Jan. 2012. URL: https://cds.cern.ch/record/1456844 (cit. on p. 169).

[191] Morad Aaboud et al. « Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC ». In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex] (cit. on p. 170).

[192] A. M. Sirunyan et al. « Observation of Higgs boson decay to bottom quarks ». In: *Phys. Rev. Lett.* 121.12 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex] (cit. on p. 176).

[193] Georges Aad et al. « Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector ». In: *JHEP* 10 (2014), p. 141. DOI: 10.1007/JHEP10(2014)141. arXiv: 1407.3643 [hep-ex] (cit. on p. 176).

[194] Morad Aaboud et al. « Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector ». In: *Eur. Phys. J.* C77.6 (2017), p. 361. DOI: 10.1140/epjc/s10052-017-4900-z. arXiv: 1702.05725 [hep-ex] (cit. on p. 176).

[195] Morad Aaboud et al. « Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector ». In: *Phys. Rev.* D97.7 (2018), p. 072016. DOI: 10.1103/PhysRevD.97.072016. arXiv: 1712.08895 [hep-ex] (cit. on p. 187).

[196] Morad Aaboud et al. « Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector ». In: *Phys. Rev.* D98 (2018), p. 052005. DOI: 10.1103/PhysRevD.98.052005. arXiv: 1802.04146 [hep-ex] (cit. on p. 188).

[197] Morad Aaboud et al. « Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector ». In: *JHEP* 03 (2018), p. 095. DOI: 10.1007/JHEP03(2018)095. arXiv: 1712.02304 [hep-ex] (cit. on p. 188).

[198] *Measurements of VH, $H \to b\bar{b}$ production as a function of the vector boson transverse momentum in 13 TeV pp collisions with the ATLAS detector*. Tech. rep. ATLAS-CONF-2018-053. Geneva: CERN, Nov. 2018. URL: http://cds.cern.ch/record/2649082 (cit. on pp. 195, 228).

[199] Cinzia De Melis. « Timeline for the LHC and High-Luminosity LHC. Frise chronologique du LHC et du LHC haute luminosite ». In: (Oct. 2015). General Photo. URL: https://cds.cern.ch/record/2063307 (cit. on p. 195).

[200] The ATLAS-collaboration. *Physics at a High-Luminosity LHC with ATLAS (Update)*. Tech. rep. ATL-PHYS-PUB-2012-004. Geneva: CERN, Oct. 2012. URL: https://cds.cern.ch/record/1484890 (cit. on p. 196).

[201] Collaboration ATLAS. *Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment*. Tech. rep. CERN-LHCC-2012-022. LHCC-I-023. Geneva: CERN, Dec. 2012. URL: https://cds.cern.ch/record/1502664 (cit. on pp. 196, 208).

[202] *Projections for measurements of Higgs boson signal strengths and coupling parameters with the ATLAS detector at a HL-LHC*. Tech. rep. ATL-PHYS-PUB-2014-016. Geneva: CERN, Oct. 2014. URL: https://cds.cern.ch/record/1956710 (cit. on p. 196).

[203] *Prospects for the study of the Higgs boson in the VH(bb) channel at HL-LHC*. Tech. rep. ATL-PHYS-PUB-2014-011. Geneva: CERN, July 2014. URL: https://cds.cern.ch/record/1740962 (cit. on p. 196).

[204] M. Cepeda et al. « Higgs Physics at the HL-LHC and HE-LHC ». In: (2019). arXiv: 1902.00134 [hep-ph] (cit. on p. 197).

[205] *Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC*. Tech. rep. ATL-PHYS-PUB-2018-054. Geneva: CERN, Dec. 2018. URL: http://cds.cern.ch/record/2652762 (cit. on p. 197).

[206] *Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*. Tech. rep. ATLAS-CONF-2018-031. Geneva: CERN, July 2018. URL: https://cds.cern.ch/record/2629412 (cit. on p. 199).

[207] *Sensitivity projections for Higgs boson properties measurements at the HL-LHC*. Tech. rep. CMS-PAS-FTR-18-011. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2647699 (cit. on p. 203).

[208] ATLAS Collaboration. *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*. Tech. rep. CERN-LHCC-2017-021. ATLAS-TDR-030. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2285585 (cit. on pp. 208, 210, 213, 219, 220).

[209] ATLAS Collaboration. *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*. Tech. rep. CERN-LHCC-2017-017. ATLAS-TDR-026. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2285580 (cit. on p. 208).

[210] S. Agostinelli et al. « GEANT4: A Simulation toolkit ». In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 208).

[211] T.T. Böhlen, F. Cerutti, M.P.W. Chin, A. Fassò, A. Ferrari, P.G. Ortega, A. Mairani, P.R. Sala, G. Smirnov, and V. Vlachoudis. « The FLUKA Code: Developments and Challenges for High Energy and Medical Applications ». In: *Nuclear Data Sheets* 120 (2014), pp. 211–214. DOI: https://doi.org/10.1016/j.nds.2014.07.049. URL: http://www.sciencedirect.com/science/article/pii/S0090375214005018 (cit. on p. 208).

[212] *Website of ATLAS Public Radiation Simulation Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RadiationSimulationPublicResults (cit. on pp. 208, 209).

[213] *Standard practice for characterizing neutron fluence spectra in terms of an equivalent monoenergetic neutron fluence for radiation-hardness testing of electronics*. West Conshohocken, PA: ASTM, 2009. URL: https://cds.cern.ch/record/2103968 (cit. on p. 210).

[214] Michael Moll. « Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties ». PhD thesis. Hamburg U., 1999. URL: http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis99-040 (cit. on pp. 210, 211, 218).

[215] G. Kramberger, V. Cindro, I. Mandic, M. Mikuz, and M. Zavrtanik. « Effective trapping time of electrons and holes in different silicon materials irradiated with neutrons, protons and pions ». In: *Nucl. Instrum. Meth.* A481 (2002), pp. 297–305. DOI: 10.1016/S0168-9002(01)01263-3 (cit. on p. 211).

[216] *RD-53 Collaboration Home Page, http://rd53.web.cern.ch/RD53/*. URL: http://rd53.web.cern.ch/RD53/ (cit. on p. 212).

[217] G Lindström, S Watts, and F Lemeilleur. *3rd RD48 status report: the ROSE collaboration (R & D on silicon for future experiments)*. Tech. rep. CERN-LHCC-2000-009. Geneva: CERN, Dec. 1999. URL: https://cds.cern.ch/record/421210 (cit. on p. 213).

[218] J. Lange, G. Giannini, S. Grinstein, M. Manna, G. Pellegrini, D. Quirion, S. Terzo, and D. Vázquez Furelos. « Radiation hardness of small-pitch 3D pixel sensors up to a fluence of 3×10^{16} n_{eq}/cm^2 ». In: *JINST* 13.09 (2018), P09009. DOI: 10.1088/1748-0221/13/09/P09009. arXiv: 1805.10208 [physics.ins-det] (cit. on p. 216).

[219] M. Bomben, A. Ducourthial, A. Bagolini, M. Boscardin, L. Bosisio, G. Calderini, L. D'Eramo, G. Giacomini, G. Marchiori, N. Zorzi, A. Rummler, and J. Weingarten. « Performance of active edge pixel sensors ». In: *Journal of Instrumentation* 12.05 (May 2017), P05006–P05006. DOI: 10.1088/1748-0221/12/05/p05006. arXiv: 1702.01709 [physics.ins-det]. URL: https://doi.org/10.1088%2F1748-0221%2F12%2F05%2Fp05006 (cit. on pp. 217, 222).

[220] A Chilingarov. « Temperature dependence of the current generated in Si bulk ». In: *Journal of Instrumentation* 8.10 (Oct. 2013), P10003–P10003. DOI: 10.1088/1748-0221/8/10/p10003. URL: https://doi.org/10.1088%2F1748-0221%2F8%2F10%2Fp10003 (cit. on pp. 218, 219).

[221] H. Spieler. *Semiconductor Detector Systems*. Series on Semiconductor Science and Technology. OUP Oxford, 2005. URL: https://books.google.fr/books?id=MUMb3y37yqYC (cit. on p. 219).

[222] R. Diener, J. Dreyling-Eschweiler, H. Ehrlichmann, I.M. Gregor, U. Kötz, U. Krämer, N. Meyners, N. Potylitsina-Kube, A. Schütz, P. Schütze, and M. Stanitzki. « The DESY II test beam facility ». In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 922 (2019), pp. 265–286. DOI: https://doi.org/10.1016/j.nima.2018.11.133. URL: http://www.sciencedirect.com/science/article/pii/S0168900218317868 (cit. on p. 220).

[223] Collaboration NA62. *2019 NA62 Status Report to the CERN SPSC*. Tech. rep. CERN-SPSC-2019-012. SPSC-SR-249. Geneva: CERN, Mar. 2019. URL: https://cds.cern.ch/record/2668548 (cit. on p. 220).

[224] *COMPASS-II Proposal*. Tech. rep. CERN-SPSC-2010-014. SPSC-P-340. Geneva: CERN, May 2010. URL: https://cds.cern.ch/record/1265628 (cit. on p. 220).

[225] *http://eutelescope.web.cern.ch/*. URL: http://eutelescope.web.cern.ch/ (cit. on p. 221).

[226] V. Blobel. « Millepede II: Linear Least Squares Fits with a Large Number of Parameters ». In: *Institut fur Experimentalphysik Universitat Hamburg 2007* (cit. on p. 222).

[227] *https://bitbucket.org/TBmon2/tbmon2/overview*. URL: https://bitbucket.org/TBmon2/tbmon2/overview (cit. on p. 222).

[228] M. Bomben, A. Bagolini, M. Boscardin, L. Bosisio, G. Calderini, J. Chauveau, G. Giacomini, A. La Rosa, G. Marchori, and N. Zorzi. « Development of Edgeless n-on-p Planar Pixel Sensors for future ATLAS Upgrades ». In: *Nucl. Instrum. Meth.* A712 (2013), pp. 41–47. DOI: 10.1016/j.nima.2013.02.010. arXiv: 1211.5229 [physics.ins-det] (cit. on p. 223).

[229] Audrey Ducourthial et al. « Thin and edgeless sensors for ATLAS pixel detector upgrade ». In: *JINST* 12.12 (2017), p. C12038. DOI: 10.1088/1748-0221/12/12/C12038. arXiv: 1710.03557 [physics.ins-det] (cit. on p. 223).

[230] A. Ducourthial et al. « Performance of thin planar $n-on-p$ silicon pixels after HL-LHC radiation fluences ». In: *Nucl. Instrum. Meth.* A927 (2019), pp. 219–229. DOI: 10.1016/j.nima.2019.02.033. arXiv: 1810.07279 [physics.ins-det] (cit. on pp. 223, 225).

[231] Audrey Ducourthial. « Upgrade of the ATLAS experiment Inner Tracker and related physics perspectives of the Higgs boson decay into two b quarks. Amélioration du trajectographe de l'expérience ATLAS et impact sur l'étude de la désintégration

du boson de Higgs en deux quarks b ». Nov. 2018. URL: https://cds.cern.ch/record/2652024 (cit. on pp. 225, 226).

[232] Todd Seiss. « ATLAS Hardware Based Track-Finding: Present and Future ». In: (July 2018). URL: http://cds.cern.ch/record/2632042 (cit. on p. 229).

List of figures

List of tables