
**Search for the decay of the Higgs boson to a Z boson and a
long-lived axion-like particle decaying to two photons with
the ATLAS Detector**

Dissertation submitted
for the award of the title
"Doctor
of Natural Sciences"
to the Faculty of Physics, Mathematics, and Computer Science
of Johannes Gutenberg University Mainz
in Mainz

Georgios Lamprinoudis

born in Chios, Greece

Mainz, March 6, 2025

Declaration of Authorship

Doctoral examination date: February 25, 2025

I, Georgios Lamprinoudis, hereby declare that the dissertation presented is my own work and that I have not used any sources other than those listed in the bibliography. In addition, I affirm to have clearly marked and acknowledged all quotations or references that have been taken from the works of others. Finally, I am aware of the stipulations for the safeguarding of good academic practice in research and teaching and the consequences of academic misconduct.

Mainz, March 6, 2025 _____

Signature

Abstract The Standard Model of particle physics is one of the most successful theories, capable of describing a wide range of observed physical processes. However, some phenomena cannot be explained within the Standard Model, necessitating a theoretical extension. The Large Hadron Collider (LHC) at CERN aids in this effort through precision tests and searches for Beyond Standard Model particles using dedicated analyses. Several well-motivated extensions of the Standard Model include Axion-Like Particles (ALPs), which can couple to photons. This thesis presents two searches for ALPs.

The first search, described in the thesis, uses ATLAS Run 2 proton-proton (pp) data to constrain the exotic decay of the Higgs boson into a Z boson and an ALP, leading to a final state of two leptons and two photons. This measurement focuses on long-lived ALPs, which do not decay promptly, extending the excluded photon-ALP coupling phase space and providing limits for a range of couplings not previously covered by ATLAS.

The second search, described in the thesis, utilizes ATLAS Run 3 lead-lead (PbPb) data to probe the resonant production of an ALP in photon-photon scattering within the ATLAS detector. This is the first ALP measurement in ATLAS at a collision energy of $\sqrt{s} = 5.36$ TeV.

Kurzzusammenfassung Das Standardmodell der Teilchenphysik ist eine der erfolgreichsten Theorien, die in der Lage ist, ein breites Spektrum beobachteter physikalischer Prozesse zu beschreiben. Einige Phänomene können jedoch nicht mit dem Standardmodell erklärt werden, so dass eine theoretische Erweiterung erforderlich ist. Der Large Hadron Collider (LHC) am CERN unterstützt diese Bemühungen durch Präzisionstests und die Suche nach Teilchen jenseits des Standardmodells mit Hilfe spezieller Analysen. Mehrere gut begründete Erweiterungen des Standardmodells umfassen Axion-ähnliche Teilchen (ALPs), die an Photonen koppeln können. In dieser Arbeit werden zwei Suchläufe nach ALPs vorgestellt.

Die erste Suche verwendet ATLAS Run 2 Proton-Proton (pp)-Daten, um den exotischen Zerfall des Higgs-Bosons in ein Z-Boson und ein ALP einzugrenzen, was zu einem Endzustand aus zwei Leptonen und zwei Photonen führt. Diese Messung konzentriert sich auf langlebige ALPs, die nicht sofort zerfallen. Dadurch wird der Phasenraum der ausgeschlossenen Photon-ALP-Kopplung erweitert und es werden Grenzwerte für eine Reihe von Kopplungen ermittelt, die bisher von ATLAS nicht erfasst wurden.

Die zweite Suche nutzt ATLAS Run 3 Blei-Blei (PbPb) Daten, um die resonante Produktion eines ALP in der Photonen-Photonen-Streuung innerhalb des ATLAS-Detektors zu untersuchen. Dies ist die erste ALP-Messung in ATLAS bei einer Kollisionsenergie von $\sqrt{s} = 5.36$ TeV.

Acknowledgements

My PhD journey was filled with many interesting experiences and remarkable people. I would like to express my gratitude to some of them, without whom this thesis would not have been possible.

First, I would like to express my highest gratitude to my supervisor Prof. Matthias Schott, for his permanent support and his endless patience. He taught me the correct scientific approach on CERN particle physics and he provided the necessary stimuli for my curiosity and inquisitiveness. His deep level of physics knowledge and curiosity will always be an example to me. I am profoundly grateful to be given the opportunity to work with him and his work ethic and experience will be an example for me for my future life.

I would also like to express my gratitude to Dr. Kristof Schmieden for his endless patience and resourcefulness when my analysis was reaching dead-ends. Our discussions were essential for the completion of my PhD thesis. His deep understanding of the ATLAS detector and overall physics knowledge shows years of dedication and craftsmanship which I will be grateful to emulate in any area of my life.

I would also like to thank all the other members of the Schott group. First and foremost, Dr. Jack MacDonald whose guidance and expertise made the hZ reinterpretation possible. In addition, I would like to thank Peter Kraemer and Olivera Vujinovic for their excellent work in the $h \rightarrow aa$ analysis which was an essential part of my analysis and an inspiration. Furthermore, I would like to thank Particja Potepa, Radek Varvicka and particularly Dhruv Chouhan for their friendship.

I would like to thank the faculty at the University of Mainz which welcomed me in Germany during my internship and PhD. In particular, Prof. Lucia Masetti who offered me my internship position.

I would like to thank CERN and the ATLAS collaboration for the excellent detector operation which makes our work possible and allows us to observe physics at the elementary scale. In particular, I would like to thank the members of the LbyL analysis and the STG team.

Finally, I would like to express my profound gratitude and gracefulness to my parents and sister for all the support from Greece who taught me to aspire for greater heights in my life. I would also like to thank my friends who taught me to think outside the box. Finally, I would like to thank my nephew George who was born 18 days after the start of my PhD and has been a constant source of happiness.

Contents

1	Introduction	10
2	The Standard Model	11
2.1	Elementary particles	11
2.1.1	Fermions	11
2.1.2	Bosons	12
2.1.3	Symmetries	12
2.2	Gauge field theories	13
2.2.1	QED	13
2.2.2	QCD	14
2.2.3	Electroweak force	15
2.2.4	Higgs Mechanism	16
2.3	Higgs boson phenomenology at the LHC	20
2.3.1	Proton-Proton collisions and Higgs production	20
2.3.2	Higgs decays	22
2.3.3	$H \rightarrow Z\gamma$	23
3	Beyond the Standard Model theories	24
3.1	Introduction	24
3.2	ALPs(Axion Like Particles)	24
3.2.1	ALP decays to SM particles	25
3.2.2	Constraints on the ALP-photon coupling	27
3.2.3	Exotic decays of the Higgs boson into ALPs	28
3.2.4	ALP searches in $h \rightarrow Za$ decay	29
3.2.5	Constraining the ALP-photon coupling via $h \rightarrow Za$	30
4	CERN	32
4.1	Accelerator complex	32
4.2	Large Hadron Collider	33
4.3	Luminosity	34
4.4	Fiducial and total cross section	34
5	ATLAS	35
5.1	Inner detector	36

5.1.1	IBL	37
5.1.2	Pixel detector	37
5.1.3	SCT	38
5.1.4	TRT	38
5.2	Calorimeters	39
5.2.1	Electromagnetic calorimeter	39
5.2.2	Hadronic calorimeter	42
5.3	Muon Spectrometer	42
5.4	Object reconstruction	43
5.4.1	Charged particles tracks	43
5.4.2	Electromagnetic objects	44
5.4.3	Jets	46
5.5	Trigger system	46
5.6	Event simulation	47
5.6.1	Event generation	48
5.6.2	Detector simulation	48
5.6.3	Digitization	48
5.6.4	Reconstruction and Derivation	49
6	Search for Higgs boson decays into a Z boson and an ALP	50
6.1	Introduction	50
6.2	Dataset and MC simulated events	51
6.2.1	MC Signal sets	51
6.2.2	Monte Carlo backgrounds sets	51
6.2.3	Data	52
6.3	Displaced Vertices in experimental analysis	52
6.4	Object selection	53
6.4.1	Triggers	53
6.4.2	Object reconstruction	55
6.4.3	Analysis tools and framework	56
6.5	hZa ATLAS displaced event selection description and results	56
6.5.1	Z boson selection	56
6.5.2	Diphoton selection	57
6.5.3	Background estimation	61
6.5.4	Systematic Uncertainties	66
6.5.5	Results	72
6.6	Optimization of the existing hZa selection	76
6.6.1	Introduction	76
6.6.2	Event selection	76

6.6.3	Background estimation	81
6.6.4	Systematic uncertainties	91
6.6.5	Results	95
6.6.6	Comparison of prompt sensitivities in the two analyses	97
6.7	Summary of $h \rightarrow Za$	98
7	Light-by-Light scattering	99
7.1	Introduction	99
7.2	LbyL scattering in ATLAS	99
7.3	Data and MC samples	101
7.4	Event selection	102
7.4.1	Triggers	102
7.4.2	Object reconstruction	103
7.4.3	Analysis selection	104
7.5	Backgrounds	105
7.5.1	Electron-Positron background	105
7.5.2	Central exclusive production	106
7.5.3	Fake photons background	108
7.6	Photon identification	108
7.6.1	Motivation	108
7.6.2	Signal and background photons	109
7.6.3	Training MC and data	113
7.6.4	BDT PID	113
7.6.5	NN3 PID	114
7.6.6	Classifiers comparison	115
7.6.7	Signal efficiency and background rejection studies	117
7.6.8	Conclusions	119
7.7	Results	120
8	Conclusion	121
A	Appendix	123
A.1	Filenames of MC samples	123
A.2	Higgs boson theoretical elements	125
A.2.1	Higgs symmetry breaking mechanism	125
A.2.2	Weak bosons and fermion masses	126
A.2.3	Unitarity arguments for Higgs boson	130
A.3	Standard Model Lagrangian	133
A.4	Beyond the Standard Model theories	135
A.4.1	EFT essentials	135
A.4.2	QCD Axion	137

	A.4.3	ALPs	138
A.5		Machine Learning	142
	A.5.1	Introduction	142
	A.5.2	Deep neural networks	142
	A.5.3	Boosted decision trees	148
	A.5.4	DNNs and BDTs in ATLAS	151

1 Introduction

Particle physics is the study of fundamental particles and forces that constitute matter and radiation. The current prominent model for particle physics description is called the Standard Model, born out of a combination of special relativity and quantum mechanics. It is widely regarded as one of the most successful theories in the history of science. However, there are some phenomena it can not explain, highlighting the need for an extension to its particle and interaction content. This thesis explores one such extension with the introduction of a pseudoscalar particle called the Axion-Like-Particle (ALP).

The axion was initially proposed by Steven Weinberg to address the strong CP problem [1]. Axion-like particles (ALPs) are a broader class of particles with similar properties, which do not necessarily resolve the CP problem [2]. In several theoretical models, ALPs are light, long-lived, and couple to Standard Model photons. In recent years, ALPs have become the focus of numerous experiments and searches.

This thesis investigates ALPs using data from the A Toroidal LHC Apparatus (ATLAS) detector at the Large Hadron Collider (LHC), located at CERN. The LHC collides proton-proton and lead-lead beams, and the resulting interactions are detected and recorded by ATLAS. This thesis includes two separate ALP searches based on these data.

The first search, based on proton-proton (pp) collision data, investigates exotic Higgs boson decays into a Z boson and an ALP, which subsequently produce a final state containing two leptons and two photons. In this thesis, two selection methods are described: one based on the ATLAS paper [3] and another developed by the author. These methods constrain the ALP parameter space for masses in the range of 2–33 GeV. The analysis employs the displaced photon origin vertices method, targeting non-prompt ALP decays, which significantly extends the accessible phase space.

The second search, based on lead-lead (PbPb) collision data, focuses on light-by-light (LbyL) scattering. For specific ALP masses (6–100 GeV), a significant enhancement in the LbyL cross section is expected. This thesis details the current progress of the LbyL Run 3 analysis, with a particular emphasis on photon identification.

The thesis begins in chapter 2 with a brief description of the Standard Model and the phenomenology of the Higgs boson at the LHC. In chapter 3 the Beyond Standard Model extension involving ALPs is described using an Effective field theory to derive the phenomenological ALPs characteristics. Finally, chapter 4 is a concise introduction to CERN and its accelerator complex, including the LHC.

The main part of the thesis is divided into three parts. Chapter 5 describes the operation and structure of the ATLAS experiment. Following that, chapter 6 focuses on the $h \rightarrow Za$ analysis and presents the final results. Finally, chapter 7 offers a description of the LbyL scattering run 3 analysis with an expansion of the photon identification derivation.

2 The Standard Model

2.1 Elementary particles

The Standard Model of particle physics (SM) is the quantum field theory that describes the interactions of the known fundamental particles. It is the most successful model of particle physics currently known as it is able to describe phenomena with an accuracy of 10^{-12} [4]. Its particle content is split into three categories, the integer spin bosons, the $1/2$ spin fermions, and the scalar Higgs boson. These particles are elementary, which means they cannot be broken down to smaller particles. The particle content is shown in Figure 1 and is described in the following subsections.

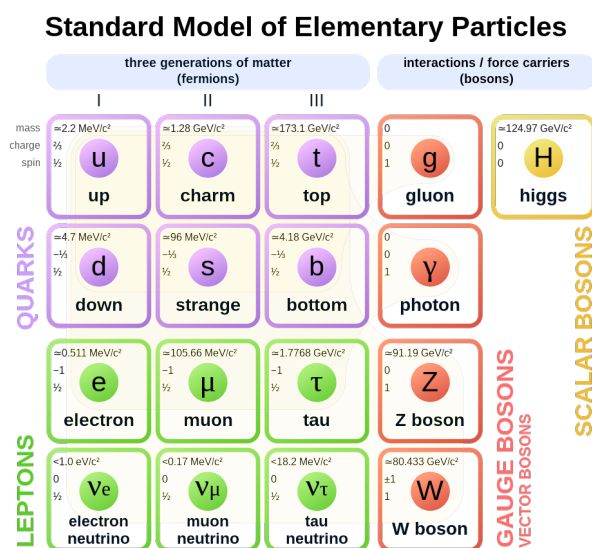


Figure 1: The Standard Model particles [5].

2.1.1 Fermions

Fermions are the particles that combine to form visible matter in the universe. They are split into two categories: quarks and leptons. In addition, each fermion has a corresponding anti-particle, with exactly the same properties but opposite charge. Fermions exist in three generations, each generation made up of two quarks or two leptons, as shown in Figure 1. Only the fermions in the first generation are stable, with the fermions of the other generations having shorter lifetimes and decaying into particles of the first generation.

Quarks are the particles that combine to form composite particles called hadrons, the most stable of which are protons and neutrons, the components of atomic nuclei. They interact with all bosons as they possess colour, weak and electric charge. They are split into two categories regarding the electric charge: up-type quarks and down-type quarks.

The up-type quarks have charge $+2/3$ and they are the up, charm and top quarks. The down-type quarks have charge $-1/3$ and they are the down, strange and bottom quarks.

Leptons are the particles that do not undergo strong interactions. They are comprised of three pairs that include a heavier charged particle called electron, muon and tau and a much lighter neutral particle called a neutrino (ν_e, ν_μ, ν_τ). The electron, muon and tau carry an electric charge of -1 and interact both electromagnetically and weakly. The neutrinos have no electric charge and only interact through the weak force. The neutrinos have extremely low mass compared to the other Standard Model particles.

2.1.2 Bosons

Bosons are the particles that mediate the interactions in the standard model. The gluon (g) is the mediator of the strong interaction, interacting with quarks to form baryons. The gluons carry a colour charge and can interact with each other. The photons (γ) are the mediators of the electromagnetic interaction. They interact with all particles that possess electric charge, including the W^\pm bosons. Photons do not possess an electric charge and do not interact with each other.

The three bosons associated with the weak interaction are W^\pm and Z . The W bosons carry an electric charge of -1 or $+1$, while the Z boson is electrically neutral. Both the W and Z bosons couple to all fermions. The name weak interaction can be explained by the limited range and weakness of the interaction which is due to the fact that W and Z are massive at 80.38 and 91.19 GeV [6].

The last boson in the standard model is the Higgs boson. It is a scalar boson with spin 0 and mass of 125.1 GeV. It was the last particle of the standard model formalism to be discovered, in 2012 in ATLAS [7] and CMS [8].

2.1.3 Symmetries

It is evident that the Standard Model encompasses various particle interactions and mass ranges, presenting a complex picture of dynamics. In modern physics the concept of symmetries is instrumental to the understanding of the forces and the dynamic description of the system. The Standard model is no exception as it is invariant under Lorentz transformations for fermions and bosons and three distinct local gauge symmetries for the fundamental forces. The three distinct symmetries are $U(1)_Y$, $SU(2)_W$ and $SU(3)_C$. These symmetries correspond to the electroweak force (which is split into the electromagnetic force and the weak force after symmetry breaking) and the strong nuclear force respectively. These are three of the four fundamental forces of nature with the fourth being gravity, which is not included in the Standard Model. The Standard Model Lagrangian is invariant under all these symmetries at the same time.

These forces interact with the fundamental representations of the elementary particles.

For example the strong nuclear force interacts with the (3) representation of the quarks signifying the three different colors. The mediators of the fundamental forces are called gauge bosons, which are described by Gauge field theories.

2.2 Gauge field theories

Gauge field theories are the culmination of the exploration of physics at the fundamental scale. Their predictions have been tested repeatedly by numerous experiments. The simplest gauge field theory is Quantum Electrodynamics commonly known as QED.

2.2.1 QED

The QED gauge theory uses the group U(1). The equation of motion of a free electron can be described by a Lagrangian which stems from the Dirac equation:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (1)$$

Here $\psi, \bar{\psi}$ are the fermion spinors, γ are the Dirac matrices and m is the fermion mass. The principle of gauge invariance allows the introduction of a local phase transformation or U(1) symmetry:

$$\psi(x) \rightarrow \psi(x)' = e^{ia(x)}\psi(x) \quad (2)$$

with $a(x)$ being a real scalar field. The Lagrangian itself is not gauge invariant but it can be made so by the addition of a covariant derivative transformation D_μ , with the fundamental electric charge e of the electron:

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu - ieA_\mu \quad (3)$$

where the A_μ is a vector field which transforms under a U(1) symmetry as:

$$A_\mu = A_\mu + \frac{1}{e}\partial_\mu a \quad (4)$$

Then the Lagrangian is gauge invariant.

$$\mathcal{L}_{QED} = i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi = i\bar{\psi}\gamma^\mu\partial_\mu\psi + e\bar{\psi}\gamma^\mu\psi A_\mu - m\bar{\psi}\psi \quad (5)$$

This additional vector field A_μ , required to preserve the theory's invariance under the U(1) transformation, is called the gauge field. It is coupled to the fermion field ψ via the $+e\bar{\psi}\gamma^\mu\psi A$ commonly known as the interaction term. This term is at the heart of QED as the theory in its simplest form is the interaction of the current terms $j^\mu \equiv \bar{\psi}\gamma^\mu\psi$ with the EM field potentials given by A_μ . This field can be linked with the electromagnetic field by introducing a dynamical term in the Lagrangian:

$$\mathcal{L}_A = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (6)$$

with the electromagnetic field tensor $F_{\mu\nu}$:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (7)$$

2.2.2 QCD

Quantum Chromodynamics (QCD) is more complex than Electromagnetic (EM) theory because, unlike EM, there are three distinct colors that define the quantum number of color. This number is described by the non-Abelian gauge group $SU(3)_C$. As a result, QCD is formulated as a Yang-Mills theory, a more intricate mathematical framework. Specifically, QCD is a Yang-Mills theory for the $SU(3)$ which has 8 generators (since $SU(N)$ has $N^2 - 1$ generators). These generators correspond to the 8 different types of gluons. To describe the strong interaction of the dynamic field F one has the following Lagrangian:

$$\mathcal{L}_{YM} = -\frac{1}{4}F_{\mu\nu}^\alpha F^{\mu\nu\alpha} \quad (8)$$

where $F^{\mu\nu\alpha}$ is the tensor of the chromodynamic field which is defined as:

$$F_{\mu\nu}^\alpha = \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha + g_s f^{\alpha\beta\gamma} G_\mu^\beta G_\nu^\gamma \quad (9)$$

where $f^{\alpha\beta\gamma}$ are the structure constants of the Lie algebra $SU(3)$. As evident by the colour factor term $g_s f^{\alpha\beta\gamma} G_\mu^\beta G_\nu^\gamma$ the local potential fields G have charge in QCD so they interact with each other in combinations of 3 and 4 gluons. Two important effects of QCD are asymptotic freedom and colour confinement [9]. Asymptotic freedom is a property of Yang-Mills gauge theories in which the interactions between particles become progressively weaker as the energy scale increases and the corresponding length scale decreases. Colour confinement is caused by the opposite effect where the interactions become extremely strong below a certain energy scale, $\Lambda_{QCD} = 200$ MeV.

The covariant derivative D_μ is:

$$D_\mu \equiv \partial_\mu - ig_s T^\alpha G_\mu^\alpha \quad (10)$$

where T^α are the generators of $SU(3)$. In a complete analogy with QED the following equations of motion for the F field can be written.

$$(D_\mu F^{\mu\nu})^\alpha = \mathcal{J}^{\nu\alpha}, (D_\mu \tilde{F}^{\mu\nu})^\alpha = 0 \quad (11)$$

where $\mathcal{J}^{\nu\alpha}$ is the vector current and $\tilde{F}^{\mu\nu}$ is the dual-strength tensor:

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F^{\rho\sigma} \quad (12)$$

Strong CP Problem In quantum physics an anomaly or quantum anomaly is the failure of a symmetry of a theory's classical action to be a symmetry of any regularization of the full quantum theory. One such symmetry is the axial symmetry generated by the vector current $j^{\mu 5} = \bar{\psi} \gamma^\mu \gamma^5 \psi$, for which classically $\partial_\mu j^{\mu 5} = 0$. However, the existence of massive fermions leads to the creation of an axial current from the Adler-Bell-Jackiw anomaly which violates the symmetry [2].

$$\partial_\mu j^{\mu 5} = \frac{g^2 N}{32\pi^2} F_{\mu\nu}^\alpha \tilde{F}^{\mu\nu\alpha} \quad (13)$$

where $N = 6$ is the number of quarks. To regulate this term, preventing it from creating divergences, and in recognition of the complex vacuum structure of QCD, a term has to be added in the Lagrangian:

$$\mathcal{L}_\theta = \theta \frac{g^2}{32\pi^2} F_{\mu\nu}^\alpha \tilde{F}^{\mu\nu\alpha} \quad (14)$$

This term violates Parity(P) and Time(T) reversal invariance, but conserves Charge(C) conjugation invariance, resulting in CP violation. The degree of CP violation is characterized by the angle θ . The θ term gives rise to a contribution to the neutron electric dipole moment, d_N . According to the strongest existing bound [10], $d_N < 10^{-26}$ e cm. The contribution of θ on d_N is $d_N = (5.2 \times 10^{-16})\theta$ e cm, making the value of θ approximately close to zero at $\theta < 10^{-10}$. The θ angle could take any value within the range $[0, 2\pi]$. Therefore, it is desirable to have an explanation as to why this parameter must be so small and fine tuned. This is called the strong CP problem. One possible solution to the strong CP problem is given by axions, described in section A.4.2.

2.2.3 Electroweak force

The electroweak force is the unified description of two of the fundamental interactions of nature: electromagnetism (electromagnetic interaction) and the weak interaction. The electroweak force in the Standard Model is described by the Glashow-Salam-Weinberg Model. The mediators of this force are the γ , $W^{-/+}$ and Z^0 (Z) bosons and the local gauge group of the theory is $SU(2)_L \times U(1)_Y$ where Y is the quantum number of hypercharge, a quantum number relating the electric charge and the third component of weak isospin [9]. The coupling of the SU(2) group is g and of U(1) g' as the two terms are independent of each other in the unbroken symmetry. This leads to the following Lagrangian for the field kinetic terms:

$$\mathcal{L} = -\frac{1}{4} W_{\mu\nu}^\alpha W^{\mu\nu\alpha} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (15)$$

where W^α is the tensor field associated with the SU(2) symmetry and B is the tensor field associated with the U(1) symmetry. The field strength tensor of the U(1) is defined like the Electromagnetic theory while the field strength tensor of the SU(2) field is non-Abelian and is defined like the QCD strength tensor only using the SU(2) structure elements. The covariant derivative of the electroweak force is :

$$D_\mu \equiv \partial_\mu - igA_\mu^\alpha \tau^\alpha - i\frac{1}{2}g'B_\mu \quad (16)$$

where A_μ^α and B_μ are the gauge fields of SU(2) and U(1) respectively, while τ^α are the Pauli matrices.

A unique feature of this theory is that it incorporates handedness or helicity as there is interaction only of the left handed quarks and leptons, creating $SU(2)_L$ doublets. The right handed quarks and leptons are singlets. So for each generation of quarks there are

four fields: one right handed u_R , one left handed u_L , one right handed d_R and one left handed d_L . The doublet Q_L is then defined as :

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad (17)$$

This creates a problem with the mass term in the Lagrangian. For example for the up quark mass term one can not write:

$$\Delta\mathcal{L} = m_u(\bar{u}_R u_L + \bar{u}_L u_R) \quad (18)$$

as the two fields u_L and u_R belong to different representations. As such, a term that is a doublet under $SU(2)$ should be introduced to contract with the u_L doublet, as discussed in section [A.2.2](#).

Furthermore, the Weak gauge bosons $W^{-/+}$ and Z have been experimentally confirmed to have masses. The generation of masses for the weak bosons can occur via a symmetry breaking mechanism, called the Higgs Mechanism.

2.2.4 Higgs Mechanism

Higgs Potential The Higgs mechanism was developed by Higgs, Kibble, Guralnik, Hagen, Brout, and Englert [[11](#), [12](#), [13](#)]. The main feature of the Higgs mechanism is the Higgs potential. The Higgs potential is described by the Lagrangian:

$$\mathcal{L} = |D_\mu\phi|^2 + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \quad (19)$$

This expression generates a potential with a local maximum at zero energy and a circular minimum at the value where symmetry breaking occurs, as shown in [Figure 2](#). The field ϕ is a complex scalar with four degrees of freedom. If $\mu < 0$ then the field has a total minimum at the value u , so it acquires a vacuum expectation value (VEV). Then a new field can be defined in the circle of value u which is the Higgs field with mass m_h . The remaining three degrees of freedom are 'eaten' by the weak bosons, which acquire mass, as detailed in section [A.2.2](#).

The values of the potential away from u are dominated by the constant λ . This value can be constrained by measurements of the 4-Higgs coupling at the LHC. The determination of the shape of the Higgs potential can shed light into phenomena like dark matter [[15](#)], Higgs inflation [[16](#)] and gravitational models [[17](#)].

The Higgs Boson The scalar field that causes spontaneous breaking of the gauge symmetry is an important ingredient in the structure of the Glashow-Weinberg-Salam (GWS) theory. The scalar field ϕ can be parametrized in the following way:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (20)$$

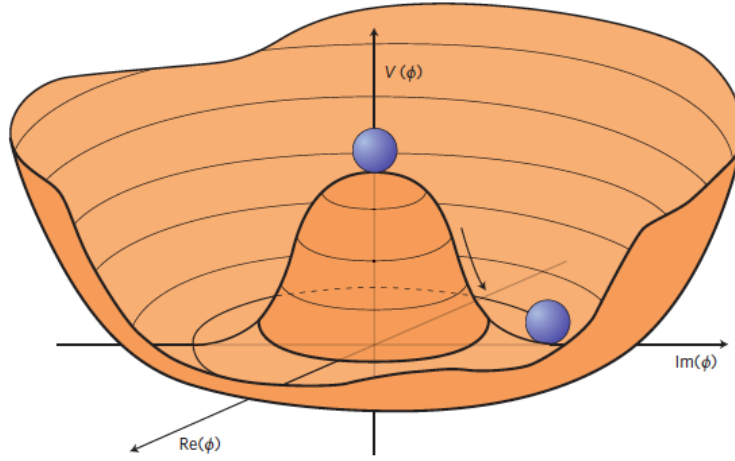


Figure 2: An illustration of the Higgs potential in the case that $\mu^2 < 0$, in which case the minimum is at $\phi^2 = \frac{\mu^2}{2\lambda}$. Choosing any of the points at the minimum of the potential leads to the symmetry breaking [14].

where $h(x)$ is a fluctuating scalar field corresponding to a small perturbation around the value of v .

An explicit renormalizable Lagrangian that leads to a vacuum expectation value for $\langle \phi \rangle$ is:

$$\mathcal{L} = |D_\mu \phi|^2 + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \quad (21)$$

The minimum of the potential energy is at:

$$v = \sqrt{\frac{\mu^2}{\lambda}} \quad (22)$$

So the potential energy terms take the following form, in unitarity gauge ¹:

$$\mathcal{L} = -\mu h^2 - \lambda v^2 h^3 - \frac{1}{4} \lambda h^4 = -\frac{1}{2} m_h^2 h^2 - \sqrt{\frac{\lambda}{2}} m_h h^3 - \frac{1}{4} \lambda h^4 \quad (23)$$

the quantum of the field $h(x)$ is a scalar particle with mass:

$$m_h = \sqrt{2}\mu = \sqrt{2\lambda}v. \quad (24)$$

This particle is known as the Higgs boson. As for the fermions in the GWS theory, the Higgs boson has a mass whose general magnitude is controlled by the vacuum expectation value v , but whose precise value is determined by a new, unspecified, renormalizable coupling constant, λ .

Expanding the kinetic term of the scalar field Lagrangian with the covariant derivative of the weak bosons for the broken symmetry, in unitarity gauge, yields the gauge boson

¹special frame of reference of bosons used to simplify the renormalization of the theory

mass term and allows us to determine the couplings of the Higgs boson to the weak bosons.

$$\mathcal{L}_{\mathcal{K}} = \frac{1}{2}(\partial_{\mu}h)^2 + [m_W^2 W^{\mu+} W_{\mu}^{-} + \frac{1}{2}m_Z^2 Z^{\mu} Z_{\mu}] \left(1 + \frac{h}{v}\right)^2 \quad (25)$$

Finally, the fermion mass terms lead to direct couplings with the Higgs boson to fermions, as described in section A.2.2. Evaluating these terms in unitarity gauge, we find that, for any quark or lepton flavor, the Higgs boson couples according to:

$$\mathcal{L}_f = -m_f f \bar{f} \left(1 + \frac{h}{u}\right) \quad (26)$$

Summarizing, the Higgs boson has the following couplings with various particles in the Standard Model:

$$\begin{aligned} g_{hf\bar{f}} &= \frac{\sqrt{2}m_f}{u} \\ g_{hW^+W^-} &= 2\frac{m_W^2}{u} \\ g_{hZZ} &= 2\frac{m_Z^2}{u} \\ g_{hhW^+W^-} &= 2\frac{m_W^2}{u^2} \\ g_{hhZZ} &= 2\frac{m_Z^2}{u^2} \\ g_{hhh} &= 3\frac{m_h^2}{u} \\ g_{hhhh} &= 3\frac{m_h^2}{u^2} \end{aligned} \quad (27)$$

Thus the Higgs boson interacts more strongly with heavier particles in the Standard Model.

Weak bosons decays The W and Z bosons decay to fermion pairs but neither the W nor the Z bosons can decay to top quarks due to energy conservation. Neglecting phase space effects and higher order corrections, simple estimates of their branching fractions can be calculated.

W decays W bosons can decay to a lepton and an antilepton (one charged and one neutral) or to a quark and antiquark of an allowed configuration. The decay width of the W boson to a quark–antiquark pair is proportional to the corresponding squared CKM matrix element times the number of quark colour charges, $N_C = 3$. The CKM matrix elements contribution is close to two due to the diagonal elements and the CKM unitarity. All leptonic decays are equiprobable due to lepton universality. Hence the relative factors shown in Table 1 are produced.

Therefore the leptonic branching ratio are $BR(W \rightarrow e\nu_e) = BR(W \rightarrow \mu\nu_{\mu}) = BR(W \rightarrow \tau\nu_{\tau}) = \frac{1}{9}$. The hadronic branching ratio is dominated by the modes $u\bar{b}$ and $c\bar{s}$ due to the CKM diagonal elements, with a tree level expectation of $\frac{2}{3}$. The sum of hadronic modes

Particles	Relative factor	Branching ratio expected (tree level)	Experimental measurements
hadrons	2/3	66.67 %	67.41 ± 0.27 %
$e\nu_e$	1/9	11.11 %	10.71 ± 0.16 %
$\mu\nu_\mu$	1/9	11.11 %	10.63 ± 0.15%
$\tau\nu_\tau$	1/9	11.11 %	11.38 ± 0.21%

Table 1: W bosons decays [6].

Particles	Relative factor ($g_{AV} + g_{AI}$)	Branching ratio expected (tree level)	Experimental measurements
Neutrinos	$3(1/2)^2$	20.00 %	20.00 ± 0.06 %
Leptons	$3(-1/2 + x)^2 + 3x^2$	10.10 %	10.097 ± 0.003 %
Electron	$(-1/2 + x)^2 + x^2$	3.33 %	3.363±0.004 %
Muon	$(-1/2 + x)^2 + x^2$	3.34 %	3.366±0.007 %
Tau	$(-1/2 + x)^2 + x^2$	3.36 %	3.367±0.008 %
Hadrons	-	69.90 %	69.91±0.06 %
Down type	$(-1/2 + x/3)^2 + (x/3)^2$	15.4 %	15.6±0.4 %
Up type (no t)	$(1/2 - 2x/3)^2 + (2x/3)^2$	11.6 %	11.6±0.6 %

Table 2: Z bosons decays ($x = \sin\theta_w$) [6].

has been experimentally measured to be 67.41 ± 0.27 % and the sum of the leptonic modes as 32.58 ± 0.27 %.

Z decays Z bosons decay into a fermion and its antiparticle. As the Z is a mixture of the W_μ^3 and B fields each vertex contains a factor $T_3 - Q\sin\theta_w$, as described in A.2.2. As the right handed fermions and the left handed fermions have different quantum numbers for the weak isospin T_3 the coupling is also different, creating what is called a axial-vector current and a vector current. Phenomenologically the width is given by:

$$\Gamma = 4 \frac{G_F m_Z^3}{24\sqrt{2}\pi} N (g_{AV}^2 + g_{AI}^2) \quad (28)$$

where N is 1 for leptons and 3 for quarks due to colour charge and G_F is the Fermi constant. The axial-vector coupling g_{AV} is given by $g_{AV} = T_3 + Q\sin\theta_w$ and the vector coupling g_{AI} is given by $g_{AI} = Q\sin\theta_w$. The Z decay modes can be calculated from the strengths of each coupling and they are shown in Table 2.

2.3 Higgs boson phenomenology at the LHC

The Higgs boson is currently being exclusively researched experimentally by the Large Hadron Collider (LHC) located at CERN.

2.3.1 Proton-Proton collisions and Higgs production

The Large Hadron Collider is a proton-proton collider that, during its second operation run, achieved a center-of-mass energy of 13 TeV. At these energies, protons are composed of three valence quarks (uud), which carry the majority of the proton's momentum, along with a set of virtual sea quarks and gluons, all of which can participate in the interaction. The distributions of these particles, called parton distribution functions, depend on the so-called Bjorken scale, which is the fraction x of the total proton energy carried by a particle participating in the interaction compared to the energy Q^2 , the squared 4-momentum-transfer vector q of the exchanged virtual boson. Two examples of parton distribution functions are shown in Figure 3.

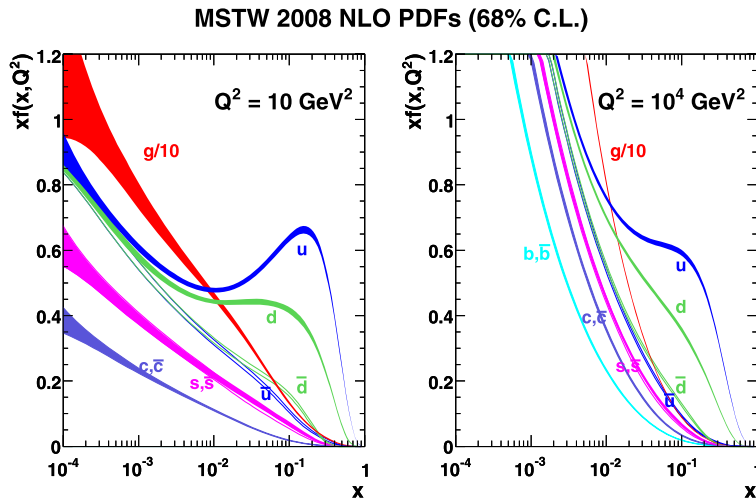


Figure 3: Parton distribution functions at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$ [18].

Scattering processes at high energy hadron colliders can be classified as either hard or soft. Quantum Chromodynamics (QCD) is the underlying theory for all such processes. The cross section of a process is given by the calculation of a matrix element, which is a quantum mechanical operator that describes the evolution from the initial to the final state. For most processes, hard interactions are accompanied by additional underlying interactions of the two particle beams at a collision point, away from the primary vertex, beyond the main collision under study. To understand the rates and characteristics of predictions for hard processes the underlying interactions should be taken into account.

For an inelastic process, the cross-section σ of a process $p + p \rightarrow X$ is given by:

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int_0^1 dx_1 f^{a,p}(x_1|\mu_F^2) \int_0^1 dx_2 f^{b,p}(x_2|\mu_F^2) \tilde{\sigma}_{ab \rightarrow X}(x_1 p_1, x_2 p_2, a_s(\mu_R)) \quad (29)$$

where a and b are the interacting partons inside protons, x_i is the fraction of the total momentum of the proton carried out by the corresponding parton and $\mu_R(\mu_F)$ is the renormalization (factorisation) scale.

Hadronization is the process by which hadrons are formed from quarks and gluons. This occurs after high energy collisions in which free quarks or gluons are produced. Due to colour confinement, these particles cannot exist individually. They combine with quarks and antiquarks, generated from the vacuum, to form hadrons. The resulting cascade of particles is called a jet. The process of hadronization in hadronic colliders is complex and not entirely well understood.

A small subset of these interactions produces Higgs bosons in multiple modes. The Higgs boson in proton-proton collisions is produced by the following modes: ggF (gluon-gluon fusion), VBF (vector boson fusion), VH (associated production with a vector boson, either W or Z), ttH or bbH (associated production with a pair of top quarks or b quarks) and tH (production in association with one top quark). The Feynman diagrams of these processes are shown below:

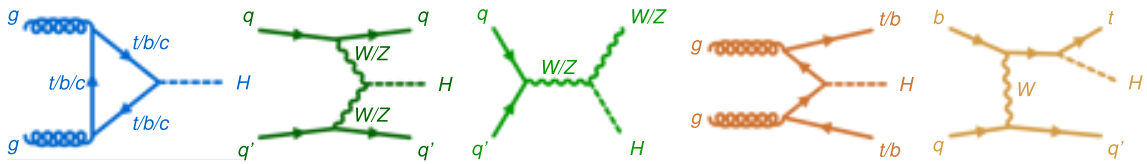


Figure 4: Feynman diagrams for the Higgs production modes at the LHC (from left to right): ggF, VBF, VH, ttH/bbH and tH/bH [19].

Illustration of the cross sections of these modes at the energy of 13 TeV:

Production mode	cross section
ggh	48.58 pb
VBF	3.78 pb
WH	1.37 pb
ZH	0.88 pb
ttH	0.51 pb
bbH	0.49 pb

Table 3: Higgs production cross sections at 13 TeV. The best-known measured value of the mass is 125.11 ± 0.11 [20],[21].

2.3.2 Higgs decays

The Higgs boson couplings scale with the masses of the weak bosons and the fermions as discussed in section 2.2.4. The mass of the Higgs influences the decay branching ratios. The Higgs decay modes are shown in Figure 5 with the relative branching ratios in Figure 6.

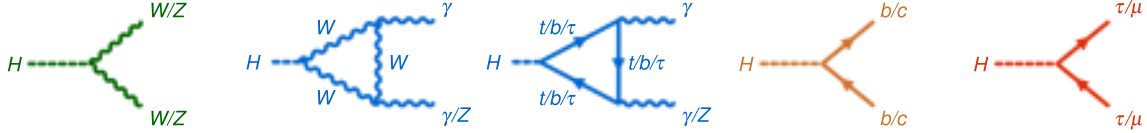


Figure 5: Feynman diagrams for the Higgs decay modes (from left to right): VV (vector boson), $Z\gamma$ or $\gamma\gamma$ (both with fermion and weak boson loop), qq (quark pair - mostly b or c) and ll (lepton pair - mostly τ or μ) [19].

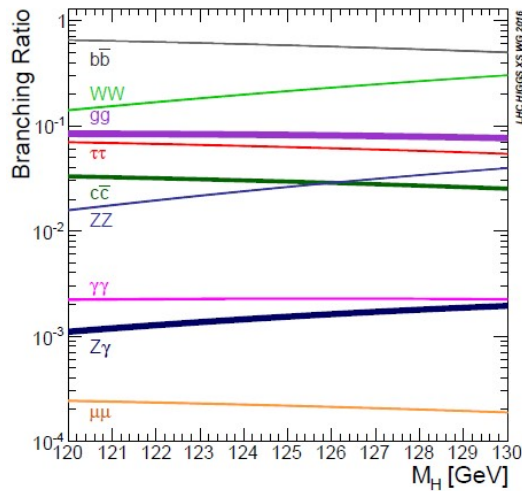


Figure 6: Higgs Branching ratios to various Standard model particles as a function of its mass. The best-known measured value of the mass is 125.11 ± 0.11 GeV [20], [19], [21].

As evident by Figure 6 the dominant Higgs decay mode is bb followed by WW . The Higgs decays to gluons and photons are mediated by a loop process. This process, in the fermionic loop is dominated by the top quark, due to the Yukawa coupling, while the bosonic loop is only possible for W . It can be replaced with an effective coupling treating it as a point interaction [22].

The $H \rightarrow \gamma\gamma$ and $H \rightarrow 4l$ channels are particularly interesting as they are the Higgs discovery channels [7] due to fewer backgrounds and superior mass resolution. Currently, they are used for precision mass measurements by ATLAS and CMS. The most recent and most precise measurement is done with a combination of the channels for 139 fb^{-1} data

[20], [19] and gives:

$$m_H = 125.11 \pm 0.11 = 125.11 \pm 0.09(\text{stat}) \pm 0.06(\text{syst}) \text{ GeV} \quad (30)$$

2.3.3 $H \rightarrow Z\gamma$

Measurements of the Higgs boson width by ATLAS [23], which compared the on-shell and off-shell $ZZ \rightarrow 4l$ production, reached the following value:

$$\Gamma_H = 4.5^{+3.3}_{-2.5} \text{ MeV} \quad (31)$$

As the value predicted for the Higgs boson width at 125 GeV is $4.07^{+0.16}_{-0.16}$ [6], it is evident that BSM decays that can increase the Higgs width compared to the SM value are not excluded yet. This makes the Higgs an excellent portal for the search of BSM effects.

In the SM, the $h \rightarrow Z\gamma$ decay is expected to have a relatively small branching fraction of $(1.5 \pm 0.1) \times 10^{-3}$ for a Higgs boson mass close to 125 GeV. The Feynman diagrams that contribute in the process are shown in Figure 7. As the decay occurs via loop graphs it is sensitive to modifications by BSM models that would cause the branching ratio to be enhanced compared with the SM value. Example include models where the Higgs boson is a composite state [24], a pseudo Nambu-Goldstone boson [25] or the Higgs decays to an Axion-Like Particle [26]. The latest scenario underlines one of the main projects of the thesis, the hZa analysis.

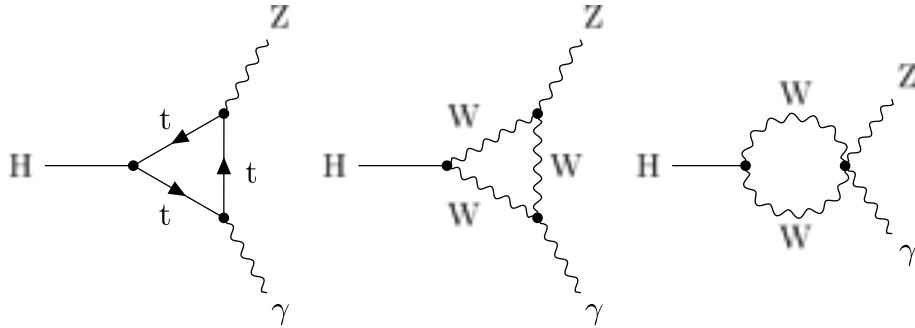


Figure 7: Examples of Feynman diagrams for $H \rightarrow Z\gamma$ decay [27].

3 Beyond the Standard Model theories

3.1 Introduction

There are many reasons to motivate an extension of the Standard Model. One shortcoming of the Standard Model is its inability to explain dark matter and dark energy. Dark matter and dark energy constitute 96% of the observable universe, so named due to their lack of interaction with the electromagnetic force. None of the particles in the Standard Model can explain dark matter, with the exception of neutrinos, which lack the necessary mass and abundance to account for all of it.

In addition, in the observable universe there is a massive asymmetry between baryonic matter and antibaryonic matter. The Standard Model does not have enough CP violation to explain this baryon asymmetry as the big bang should have produced matter and antimatter in equal quantities. This suggests the need for an ultraviolet (UV) complete theory, with more CP violation compared to the Standard Model. However, there is still no consensus on a theory to explain the phenomenon, as current theories, such as string theory [28], lack experimental data.

Finally, there are various discrepancies between the Standard Model and precision experiments. Examples include the muon $g-2$ magnetic moment [29] and the excess in excited Beryllium decays $Be^* \rightarrow Be + e^+e^-$ observed by the Atomki collaboration [30]. An elegant solution to some of these discrepancies is the introduction of a light pseudoscalar particle ² with a mass less than the electroweak scale. This particle is called an axion. In many models, the axion is a long-lived particle (LLP). Displaced vertices are associated with LLPs, which creates a need for theoretical studies of displaced vertices.

The LLP is defined as a particle that travels a macroscopic distance inside the detector before it decays to SM particles, which are subsequently detected. There are numerous theoretical models that contain LLPs like supersymmetry [31], Models with dark photons or ALPs [32] and models which naturally predict particles with low couplings, like weakly interacting massive particles (WIMPs) [33]. WIMPs are particularly interesting because they can explain dark matter and most models are motivated by this purpose.

3.2 ALPs(Axion Like Particles)

This section presents an EFT on ALPs and the phenomenology expected to be seen at particle colliders. The relevant information is reproduced from the paper Collider Probes of Axion-Like Particles [26]. This paper acts as the theoretical motivation of the ALP searches in hadronic colliders. ALPs, in general, is an umbrella term used to describe light pseudoscalar particles, compared to the electroweak bosons, that couple weakly to the

²Particles with spin 0 and odd parity, that is, a particle with no intrinsic spin with wave function that changes sign under parity inversion

Standard Model, beyond their initial relevance in the context of the strong CP problem, described in section 2.2.2.

Various models that include ALPs are supersymmetric models or Higgs doublet models. Pseudo Nambu-Goldstone bosons arise generically in models with spontaneous breaking of a global symmetry. Due to an (approximate) shift symmetry [34] they can naturally be light with respect to the electroweak or even the QCD scale [26], [35]. Low-energy observables, cosmological constraints and ALP searches with helioscopes probe a significant region of the parameter space in terms of the mass of the ALP and its couplings to photons and electrons.

Collider experiments have searched directly and indirectly for ALPs [36]. Besides ALP production in association with photons, jets and electroweak gauge bosons [37, 38, 39], searches for decays of the Z boson into a pseudoscalar a and a photon at LEP and the LHC provide limits for ALPs with up to electroweak scale masses [40]. Utilizing Higgs decays to search for light pseudoscalars has been proposed in [41]. Several experimental searches looking for the decay $h \rightarrow aa$ have been performed, constraining various final states [42, 43, 44]. Fewer experimental searches exist for the state $h \rightarrow Za$ [45].

3.2.1 ALP decays to SM particles

The effective Lagrangian, described in A.4.3, leads to decays of the ALP to SM particles. The decays can be split into three categories: decays to photons, decays to leptons and decays to hadrons.

ALP decays to photons In many scenarios, the diphoton decay is the dominant decay mode of a light ALP. On leading order it is mediated by a simple tree level diagram. The decay rate has been calculated to loop-level [26]. The relevant Feynman loop diagrams are shown in Figure 8. An effective coefficient $C_{\gamma\gamma}^{eff}$ is defined so that

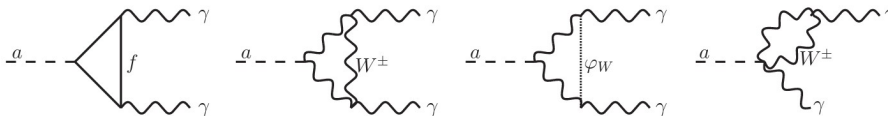


Figure 8: One loop-diagrams contributing to the decay $a \rightarrow \gamma\gamma$. The internal boson lines represent W bosons and the associated charged Goldstone fields in unitary gauge [26].

$$\Gamma(a \rightarrow \gamma\gamma) = \frac{4\pi\alpha^2 m_a^3}{\Lambda^2} |C_{\gamma\gamma}^{eff}|^2. \quad (32)$$

To an excellent approximation (with a slight difference due to mass-dependent loop effects), the decay rate scales with the third power of the ALP mass. This is expected, as it also occurs in the pion decay channel $\pi^0 \rightarrow \gamma\gamma$ and it is due to the properties and identical

quantum numbers of pseudoscalar particles [46]. For very light ALP masses $m_a < 2m_e$ this is the only decay mode and with decreasing mass the ALP will become long-lived.

Charged leptons If the ALP mass is larger than $2m_e \approx 1,022$ MeV, the leptonic decay $a \rightarrow e^+e^-$ or decays into heavier leptons (if kinematically allowed) can be the dominant ALP decay modes in some regions of parameter space. The corresponding decay rates can be written as follows with $l = e, \mu, \tau$ [26]

$$\Gamma(a \rightarrow l^+l^-) = \frac{m_a m_l^2}{8\pi\Lambda^2} |C_{ll}^{eff}|^2 \sqrt{1 - \frac{4m_l^2}{m_a^2}} \quad (33)$$

which scales linearly with the ALP mass. The ALP-lepton coupling receives contributions from loop effects by the operators $C_{\gamma\gamma}, C_{\gamma Z}, C_{ZZ}$ and C_{WW} showing that even if the tree level ALP-lepton coupling is zero it starts existing from the one loop-level. Due to the different scaling with the ALP mass, if ALPs only decay leptonically they are more long-lived compared to the decay into $\gamma\gamma$.

Charged hadrons The pseudoscalar a can decay into colored particles. At tree-level the relevant modes are $a \rightarrow gg$ and $a \rightarrow q\bar{q}$. Hadronic decays are only allowed if the ALP mass is larger than the pion mass m_π . For the decay to two gluon jets the decay rate is as follows

$$\Gamma(a \rightarrow \text{hadrons}) = \frac{32\pi a_s^2(m_a) m_a^3}{\Lambda^2} \left[1 + \frac{83}{4} \frac{a_s(m_a)}{\pi} \right] |C_{GG}^{eff}|^2. \quad (34)$$

Compared to the $a \rightarrow \gamma\gamma$ this decay rate is much larger on tree level and in addition receives larger 1-loop corrections. For the $a \rightarrow q\bar{q}$ decay, the decay rate is, very similar to the $a \rightarrow l^+l^-$ decay rate:

$$\Gamma(a \rightarrow q\bar{q}) = \frac{3m_a m_Q^2(m_a)}{8\pi\Lambda^2} |C_{QQ}^{eff}|^2 \sqrt{1 - \frac{4m_Q^2}{m_a^2}} \quad (35)$$

where the strong running coupling is taken at the value of the ALP mass.

Summary of decay modes Combining all the above formulas, setting $\Lambda = 1$ TeV and all the relevant decay coupling coefficients to 1, to enable comparisons, a range of decay rates for different ALP masses can be generated which is shown in Figure 9.

From the Figure 9, taken from [26], it is evident that the $a \rightarrow \gamma\gamma$ decay mode is the only one which is available for every possible ALP mass motivating the search of ALP decays to a photon pair. If the total decay rate of the ALP is too small, the ALP will leave the detector before it decays. If the ALP is produced in the decays of heavier particles, the Lorentz boost can significantly increase its lifetime, compensating for the ALP mass.

It is also a possibility that the ALP decays invisibly into light particles of a hidden sector. In this case the decay products cannot be reconstructed, and hence the ALP signature would be that of missing energy and momentum.

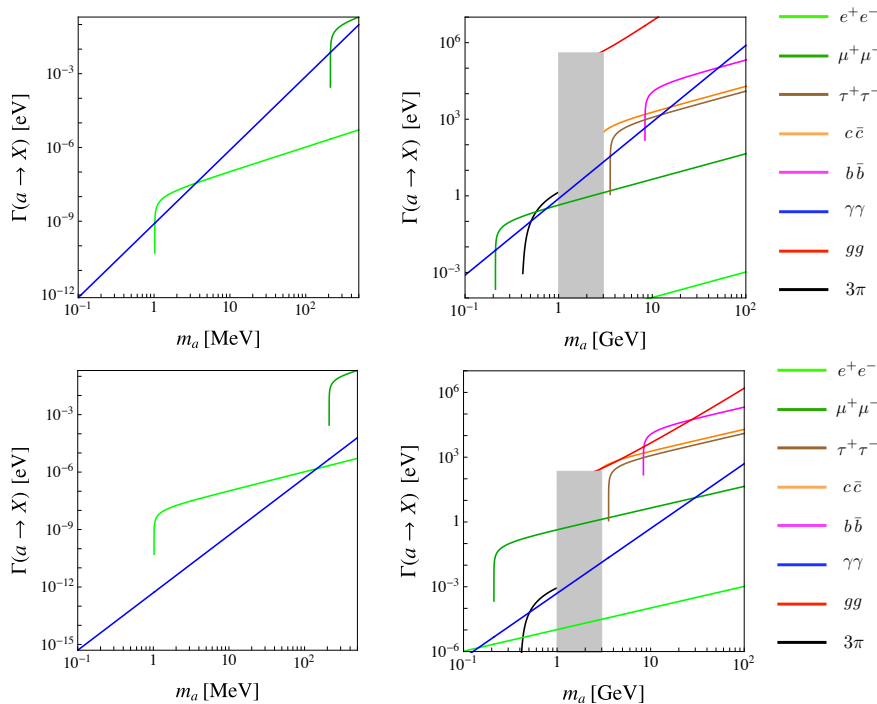


Figure 9: ALP decay rates into pairs of SM particles obtained by setting the relevant effective Wilson coefficients to 1 (top), or by setting the ALP-fermion couplings to 1 and the ALP-boson couplings to $1/(4\pi^2)$ (bottom). The gray area between 1 and 3 GeV shows the region in which various exclusive hadronic (and difficult to calculate) decay channels such as $a \rightarrow \rho\rho$ open up. In this interval the rate $\Gamma(a \rightarrow \text{hadrons})$ is expected to interpolate between the black and red lines, reproduced from [26].

3.2.2 Constraints on the ALP-photon coupling

The couplings of ALPs to photons have been constrained over vast regions of parameter space using a variety of experiments in particle physics, astro-particle physics and cosmology. The limits of ALP experiments can be reinterpreted to probe the couplings in the range of the EFT, with more details on the limit estimation available on [26]. One particularly interesting element for this thesis is the Light-by-Light ALP searches in ATLAS using Pb-Pb collisions which provide limits in the mass region 6-100 GeV [47].

The latest limits are shown in Figure 10. The Figure shows that above 30 MeV a window opens for the effective ALP-photon coupling $C_{\gamma\gamma}^{eff}/\Lambda \approx 1 \text{ TeV}^{-1}$ and above 400 MeV the ALP-photon coupling is unconstrained for $10^{-6} \text{ TeV}^{-1} \leq C_{\gamma\gamma}^{eff}/\Lambda \leq 1 \text{ TeV}^{-1}$. The mass range $m_a > 30 \text{ MeV}$ is thus well motivated to search for ALPs at high-energy particle colliders. The coupling-mass phase that can be probed by the high-energy colliders is fairly extensive and depends on the model and the detector sensitivity.

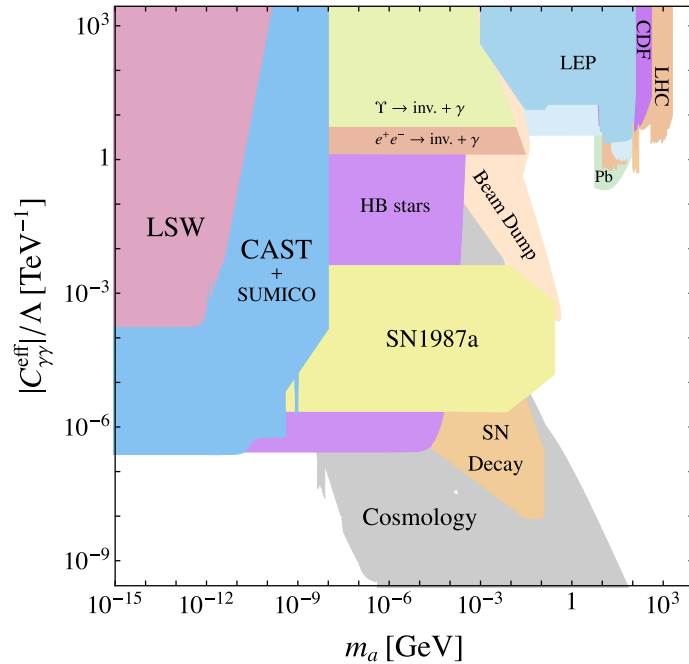


Figure 10: Existing constraints on the ALP-photon coupling derived from a variety of particle physics, astro-particle physics and cosmological observations, reproduced from [26]. Several of these bounds are model dependent.

3.2.3 Exotic decays of the Higgs boson into ALPs

The presence of ALP couplings to SM particles introduces the possibility of various exotic decay modes for the Higgs boson. As mentioned earlier, the relevant decay modes are $h \rightarrow aa$ and $h \rightarrow Za$. These modes offer a variety of interesting search channels for ALPs, depending on the specific decay modes of the ALP and the Z boson. In some regions of the parameter spaces, when the ALPs are very light, the decay $h \rightarrow Za$ may produce an enhancement in the $h \rightarrow Z\gamma$ search channel and appear as a new physics contribution. The $h \rightarrow Z\gamma$ channel is currently constrained to $\sigma(pp \rightarrow h \rightarrow Z\gamma) < 2.2 \times SM_{pred.}$ [27]. ALP searches in these two decay modes can constrain the ALP-photon coupling in a currently unexplored region of phase space.

The lifetime of ALPs and their Lorentz boost factor play a crucial role in determining their detectability. For very light ALPs or very weak couplings, the decay length can become macroscopic and hence only a small fraction of ALPs decay inside the detector. The relevant boost factors in the Higgs rest frame for the ALPs are given by $\gamma_a = (m_h^2 - m_Z^2 + m_a^2)/(2m_a m_h)$ for $h \rightarrow Za$ and $\gamma_a = m_h/(2m_a)$ for $h \rightarrow aa$. As expected lighter ALPs have bigger boost factors. This thesis will from now on focus on the $h \rightarrow Za$ decay.

3.2.4 ALP searches in $h \rightarrow Za$ decay

The $h \rightarrow Za$ decay proceeds via the following Feynman diagrams up to one-loop, shown in Figure 11. The effective Lagrangian does not contain a dimension-5 operator at tree level.

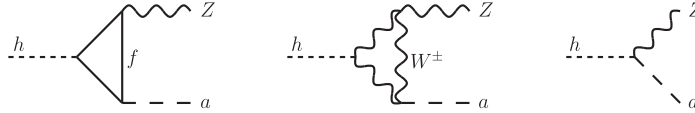


Figure 11: Feynman diagrams contributing to the decay $h \rightarrow Za$. [26]

The only contribution at this order comes from fermion loop graphs. Because both the Higgs boson and the ALP coupling to fermions is proportional to the fermion mass, due to the need to keep the diagonalization of the Higgs Yukawa couplings, the only relevant effects comes from the top quark. A tree-level contribution to the $h \rightarrow Za$ decay amplitude (third graph of Figure 11) arises first at dimension-7 order. Evaluating we obtain

$$\Gamma(h \rightarrow Za) = \frac{m_h^3}{16\pi\Lambda^2} |C_{Zh}^{eff}|^2 \lambda^{3/2} \left(\frac{m_Z^2}{m_h^2}, \frac{m_a^2}{m_h^2} \right) \quad (36)$$

where $\lambda(x, y) = (1 - x - y)^2 - 4xy$ and the effective coupling is defined as

$$C_{Zh}^{eff} = C_{Zh}^{(5)} - \frac{N_c y_t^2}{8\pi^2} T_3^t c_{tt} F + \frac{v^2}{2\Lambda^2} C_{Zh}^{(7)} \quad (37)$$

where y_t and T_3^t are the top Yukawa coupling and weak isospin and $C_{Zh}^{(5)} = 0$. After evaluation of the F top loop element, numerically one obtains [26]

$$C_{Zh}^{eff} \approx C_{Zh}^{(5)} - 0.016 c_{tt} + 0.030 C_{Zh}^{(7)} \left[\frac{1 \text{TeV}}{\Lambda} \right]^2 \quad (38)$$

The decay $h \rightarrow Za$ is kinematically allowed provided $m_a < m_h - m_Z \approx 33.9$ GeV. The LHC collaborations have reported the 95% CL upper limit $\text{BR}(h \rightarrow \text{BSM}) < 0.34$ on decays of the Higgs boson into non-SM final states, obtained from a combined analysis of the Higgs boson production and decay rates [48]. This implies the bound $\text{BR}(h \rightarrow \text{BSM}) < 2.1$, on branching ratio, on any decay rate involving new particles. So the effective coupling of $|C_{Zh}^{eff}| < 0.72 / \Lambda \text{ TeV}^{-1}$ is obtained. Depending on the dominant branching ratio of the ALP, the decay $h \rightarrow Za$ can give rise to various interesting experimental signatures. The main focus of this thesis will be on ALP decays into photons in the $h \rightarrow Za \rightarrow l^+ l^- + \gamma\gamma$ final state.

However, for strongly boosted ALPs the two photons would be reconstructed as a single photon jet and the decay $h \rightarrow Za$ could be comparable to the $pp \rightarrow h \rightarrow Z\gamma$. This interesting channel has been probed in an ATLAS analysis [3].

3.2.5 Constraining the ALP-photon coupling via $h \rightarrow Za$

Present and future searches for $h \rightarrow Za$ decays at the LHC can probe a wide range of ALP-photon couplings. The following estimates focus on Run 2 of LHC with an estimated luminosity of 300 fb^{-1} at $\sqrt{s} = 13 \text{ TeV}$. In the two photon channel it is required that the ALPs decay before the electromagnetic calorimeter, which is typically located at a distance of approximately 1.97 m from the beam axis. The candidate events are produced by Higgs bosons which have been generated by gluon-gluon fusion with cross section $\sigma(pp \rightarrow h) = 48.58 \text{ pb}$ [49]. The analysis reveals several distinct signatures. Light ALPs can enhance the observed $h \rightarrow Z\gamma$ rate, while heavier ALPs produce a clear diphoton signature. Additionally, ALPs with small couplings and longer lifetimes can result in displaced vertex signatures. Experimental strategies to isolate the signal and suppress the background vary significantly across these searches.

The parameter regions that can be constrained by $h \rightarrow Za$ are shown in Figure 12, taken from [26].

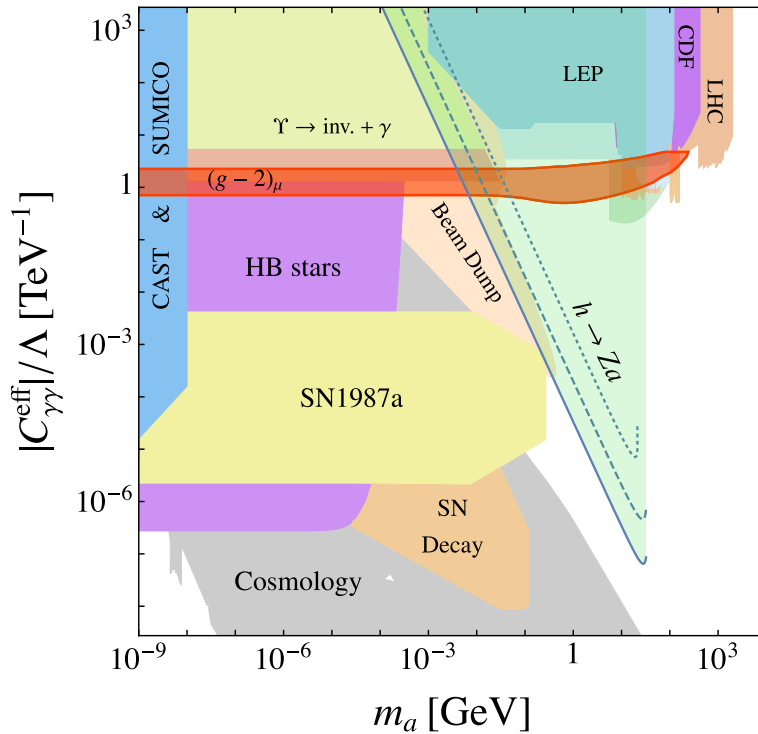


Figure 12: Constraints on the ALP mass and coupling to photons derived from various experiments [50], along with the parameter regions that can be probed using the Higgs decays $h \rightarrow Za \rightarrow l^+l^- + \gamma\gamma$. The contours correspond to $|C_{Zh}^{eff}|/\Lambda = 0.72 \text{ TeV}^{-1}$ (solid), 0.1 TeV^{-1} (dashed) and 0.015 TeV^{-1} (dotted). The red band shows the preferred parameter space where the $(g-2)_\mu$ anomaly can be explained at 95% CL, reproduced from [26].

The triangular shape of the region of the projected reach is a consequence of the fact that ALPs with either small masses or small couplings, which fall beyond the left boundary of the region of sensitivity, live long enough (on average) to leave the detector. The line in the $m_a - C_{\gamma\gamma}^{eff}$ plane where this happens depends on the width $\Gamma(a \rightarrow \gamma\gamma) \propto m_a^3 |C_{\gamma\gamma}^{eff}|^2 / \Lambda^2$. The region of parameter space accessible through exotic Higgs decays into ALPs complements the regions covered by existing searches almost perfectly. While existing searches focus on signatures of long-lived ALPs, our analysis targets ALPs that are so short-lived their decays can be fully reconstructed within the detector [26]. The red band shows the parameter space in which the anomalous magnetic moment of the muon can be explained in terms of loop corrections involving a virtual ALP exchange, assuming $|C_{\gamma\gamma}^{eff}| \Lambda < |c_{\mu\mu}| \Lambda < 5 \text{ TeV}^{-1}$, described in A.4.3.

While the graphical displays in Figure 12 accurately depict the parameter space, it is important to note that they do not account for experimental effects. As a result, the predictions are overly optimistic and likely unrealistic for LHC searches. In the $h \rightarrow Za$ channel, omitted experimental effects include reduced photon and lepton reconstruction efficiencies, limited luminosity, various systematic uncertainties, and challenges in background estimation. Nonetheless, these plots are meant to act as phenomenology and feasibility studies to motivate experimental searches, a goal they achieve successfully.

Finally, although the graphical displays in Figure 12 accurately depict the regions in the $m_a - C_{\gamma\gamma}^{eff}$ parameter space which can be probed using exotic Higgs decays, it is important to emphasize that finding a signal in these search regions requires sufficiently large ALP-Higgs couplings, as indicated by the solid, dashed and dotted contour lines in the plots. Consequently, not finding a signal in any of these searches would not necessarily exclude the existence of an ALP in this parameter space. The displays should be seen as the product of the $C_{\gamma\gamma}^{eff} * C_{Zh}^{eff}$ couplings for a range of ALP masses.

4 CERN

CERN is a research center located near Geneva, Switzerland, hosting the largest particle physics laboratory in the world. It was established in 1954 by twelve countries near the border of France and Switzerland. Currently it comprises of 23 member states and 12000 members [51] and provides the infrastructure and environment for high energy physics experiments.

4.1 Accelerator complex

CERN maintains and operates a complex of nine accelerators and two decelerators. These accelerators either supply experiments or serve as injectors, accelerating particles for larger accelerators. Some, such as the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), perform both roles—supplying particles for experiments while also injecting them into larger accelerators. Each accelerator is designed to operate at a specific energy spectrum. Some of the older accelerators, such as the PS (the oldest one), were used for landmark particle physics discoveries and are still in use today.

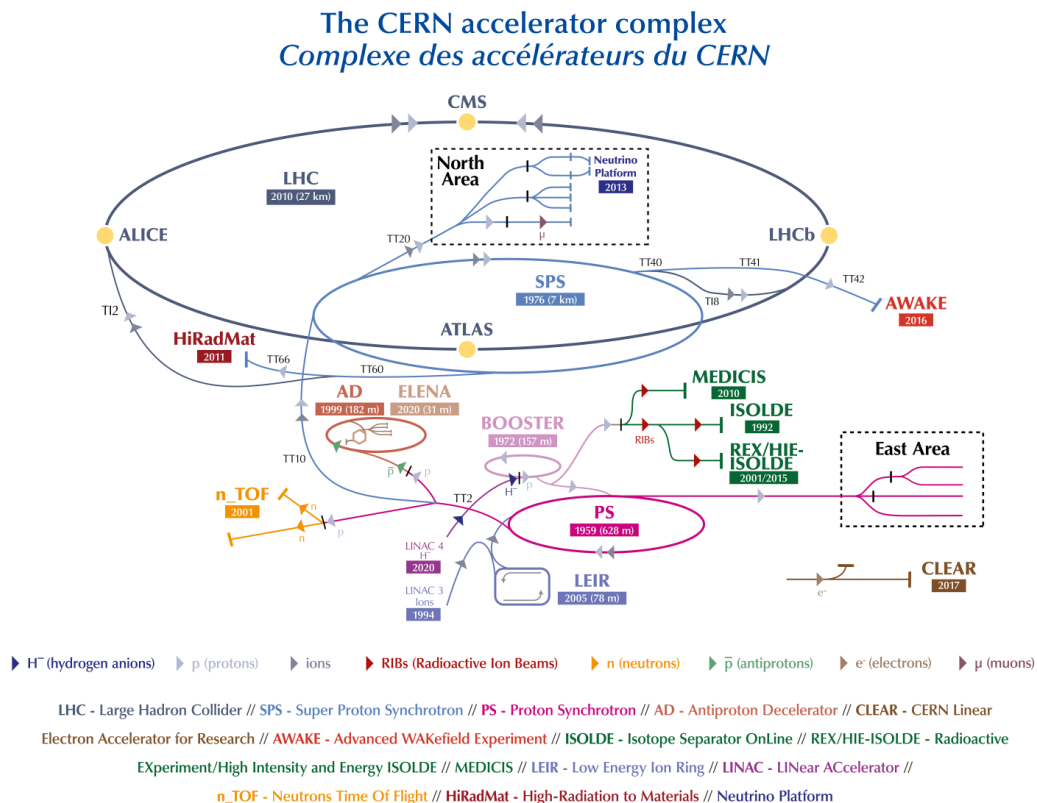


Figure 13: Schematic overview of the CERN accelerator complex and a selection of experiments using the accelerators.

4.2 Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. It began operation on 10 September 2008 and remains the latest addition to CERN's accelerator complex. The LHC consists of a 27-kilometre ring of superconducting magnets, along with several accelerating structures that boost the energy of the particles as they travel. Inside the accelerator, two high-energy particle beams travel at nearly the speed of light before being made to collide. The beams travel in opposite directions within separate beam pipes—two tubes maintained at ultrahigh vacuum. The beams can reach an unprecedented energy of 6.8 TeV, leading to a collision energy of 13.6 TeV. The beams collide at four different locations, which host the main experiments of the LHC: ATLAS, CMS, LHCb, and ALICE.

The proton beams begin their journey as negative hydrogen ions. These ions are then accelerated by the Linear Accelerator (LINAC) to 160 MeV, preparing them to enter the Proton Synchrotron Booster (PSB). During injection from LINAC into the PSB, the ions are stripped of their two electrons, leaving only protons. These protons are then accelerated to 2 GeV before passing to the PS for further acceleration to 26 GeV, and finally to the SPS for the final acceleration to 450 GeV before entering the LHC. The beams consist of 2808 bunches with 10^{11} protons per bunch, specifically chosen for an optimal combination of beam control and instantaneous luminosity delivered to the detectors.

The acceleration and control of the proton beams are managed by a state-of-the-art computing and hardware system, operated from the CERN control room. This system controls 1,232 dipole magnets, each 15 meters in length, which bend the beams, and 392 quadrupole magnets, each 5–7 meters long, which focus the beams. At the collision points, insertion magnets—made from a combination of quadrupole magnets—focus the beams into a volume of around $0.2 \mu\text{m}^2$. During the beams' revolutions, each beam gradually loses intensity due to proton collisions and radiation. When the intensity drops below a certain threshold, the beams are dumped, and new protons are injected from the SPS.

The LHC has been in operation since 2008. Each long continuous period of operation is called a run, with a short shutdown at the end of each year for detector maintenance. After each run, there is a longer shutdown to replace problematic detector parts or enhance the equipment. Currently, since 2022, the LHC is in Run 3, operating at a collision energy of 13.6 TeV. The hZa analysis uses data from Run 2, which occurred between 2015 and 2018 with a collision energy of 13 TeV, while the LbyL analysis uses data from Run 3.

The LHC does not only accelerates protons, it also facilitates Lead-proton (p-Pb) and Lead-Lead (PbPb) collisions. The PbPb collisions in Run 3 occur at a collision energy of 5.36 TeV. This is part of the heavy ion (HI) program.

4.3 Luminosity

An important quantity in high-energy particle physics is the rate of interactions at any specific point in time and space. When expressed in terms of unit area and time, in $\text{cm}^{-2}\text{s}^{-1}$, this is called instantaneous luminosity. The luminosity can be combined with the cross section measured in barns ($1\text{b} = 10^{-28} \text{m}^2$), which represents the probability of a given event occurring, to determine the total number of observed events. For Gaussian shaped beam profiles instantaneous luminosity \mathcal{L} is calculated as

$$\mathcal{L} = \frac{N}{\sigma_{tot}} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y}, \quad (39)$$

with N_1 and N_2 being the number of particles in each bunch, N_b the number of bunches, f the particle revolution frequency and σ_x, σ_y are the transversal size of the bunch at the Interaction Point, in axes x and y . For equal bunch size and circular beams the formula $L \approx fN^2/4\pi\sigma^2$ is a simplification of 39. Considering revolution frequency $f = 40^6$ Hz, particles in each bunch $N = 10^{11}$ and an interaction area $S = 4\pi\sigma^2$ with $\sigma = 16$ microns, the instantaneous luminosity of $10^{-34} \text{cm}^{-2}\text{s}^{-1}$ is derived. L is the integrated luminosity over a data period, defined as $L = \int \mathcal{L} dt$. Precise measurement of luminosity is very difficult but essential for precision physics and the observation of rare BSM processes. Constant measurements and calibrations of luminosity are carried out by the LHC and the major detectors throughout operation. For ATLAS Run 2, the current best determination of the integrated luminosity is $140.1 \pm 1.2 \text{fb}^{-1}$, an uncertainty of 0.83% [52].

A second luminosity determination is provided by leptonically decaying Z bosons, which are largely uncorrelated with the official luminosity measurements. This method involves counting the number of Z decays to leptons per operation run, specifically focusing on electron and muon final state decays.

4.4 Fiducial and total cross section

An important distinction in particle physics is between total and fiducial cross sections. Since no detector can cover the entire volume around the interaction point, it can only capture events within a subset of the phase space, known as the fiducial volume. This volume must then be extrapolated to the total volume using the property of geometrical acceptance A , defined as $N_{fid\ events}/N_{total\ events}$. Additionally, the detector cannot detect all events within this volume, leading to a detector efficiency ϵ . Combining, the general cross section for a specific detector can be defined as

$$\sigma_{fiducial} = \frac{N_{events}}{LA\epsilon} \quad (40)$$

5 ATLAS

ATLAS (A Toroidal LHC Apparatus) is a general-purpose detector located at CERN. It has a cylindrical geometry, with a height of 25 meters and a length of 44 meters. A computer-generated view of the ATLAS detector is shown in Figure 14, with a cross-section of the inner parts visible [53]. The detector consists of three main parts: the barrel and two endcaps. The barrel is aligned parallel to the beam, while the two endcaps are perpendicular, each featuring a small circular opening for the beam pipe.

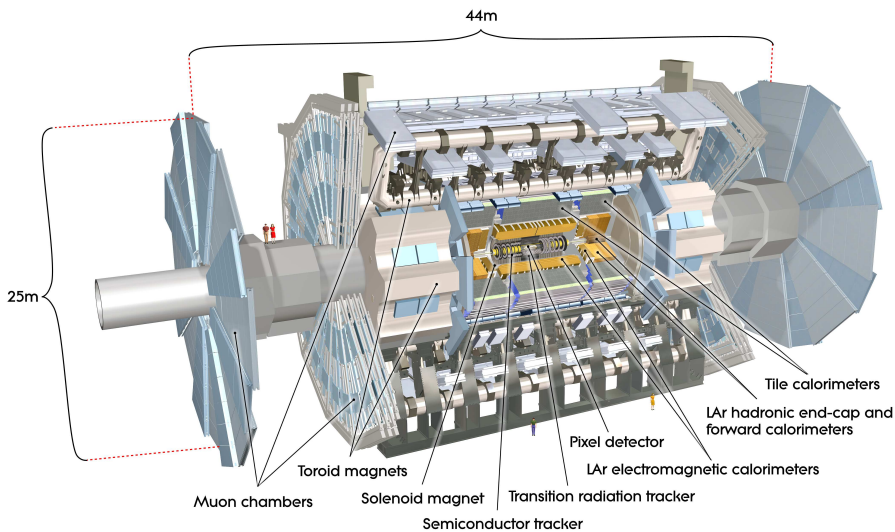


Figure 14: Overview of the ATLAS detector [53].

The origin of the ATLAS coordinate system is the Interaction Point (IP) at the center of the detector. The z-axis runs along the beam pipe, pointing to the right-hand side of the LHC's circular tangent. The x-y plane is perpendicular to the beam line and is referred to as the transverse plane. The x-axis points from the IP towards the center of the LHC ring, while the y-axis points upwards. For particle vectors, a cylindrical coordinate system is used, with r representing the radius, θ the polar angle and ϕ the azimuthal angle. For the energies of particles produced by collisions the pseudorapidity η has been used as replacement for the polar angle, defined as $\eta = -\ln(\tan\theta/2)$, as pseudorapidity differences are invariant under Lorentz transformations in the relativistic regime along the longitudinal axis. Thus, particle production is constant as a function of pseudorapidity.

The ATLAS detector consists of many different subsystems which combine their performances to reconstruct all relevant particles passing through the detector with the exception of neutrinos. A detailed schematic cross section of the particle passage is shown in Figure 15.

The first system that particles encounter is the Inner Detector (ID), which is surrounded by a solenoid magnet. The primary function of the Inner Detector is to measure the tra-

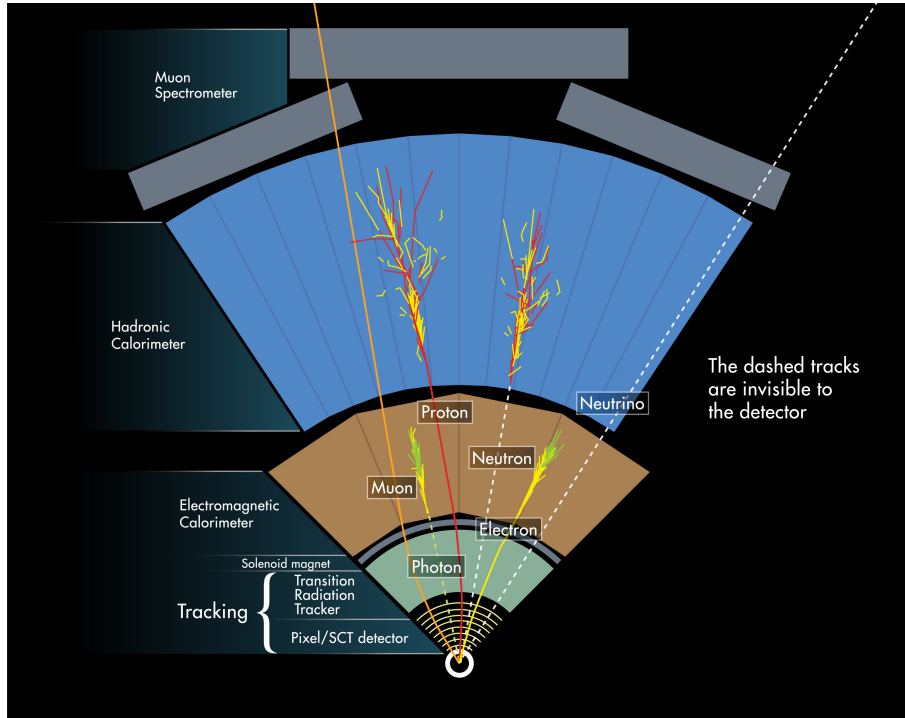


Figure 15: Illustration of passage of various particles through the detector layers: Inner Detector, Electromagnetic calorimeter, Hadronic calorimeter and Muon spectrometer [54].

jectory of charged particles. Next is the Electromagnetic Calorimeter, which is designed to stop photons and electrons, while allowing hadrons, muons, and neutrinos to pass through. Following that is the Hadronic Calorimeter, which fully stops hadrons—particles composed of quarks. Finally, the Muon Spectrometer (MS) is the last system, responsible for reconstructing muons. The MS is immersed in a toroidal magnetic field, which is used to measure the muons’ trajectory. Neutrinos are the only particles invisible to the detector, as they interact extremely weakly with matter and escape completely.

All subsystems of the detector and their basic working principles will be discussed in the following sections in more detail.

5.1 Inner detector

The ATLAS Inner Detector (ID) is designed to provide hermetic and robust pattern recognition, excellent momentum resolution, and both primary and secondary vertex measurements [53] for charged tracks. It is a cylinder 6.2 m long with a diameter of 2.1 m, immersed in a 2 T solenoidal magnetic field. A detailed view of the ID is shown in Figures 16 and 17. The detector consists of the following components, listed from innermost to outermost: IBL, Pixel Detector, SCT, and TRT.

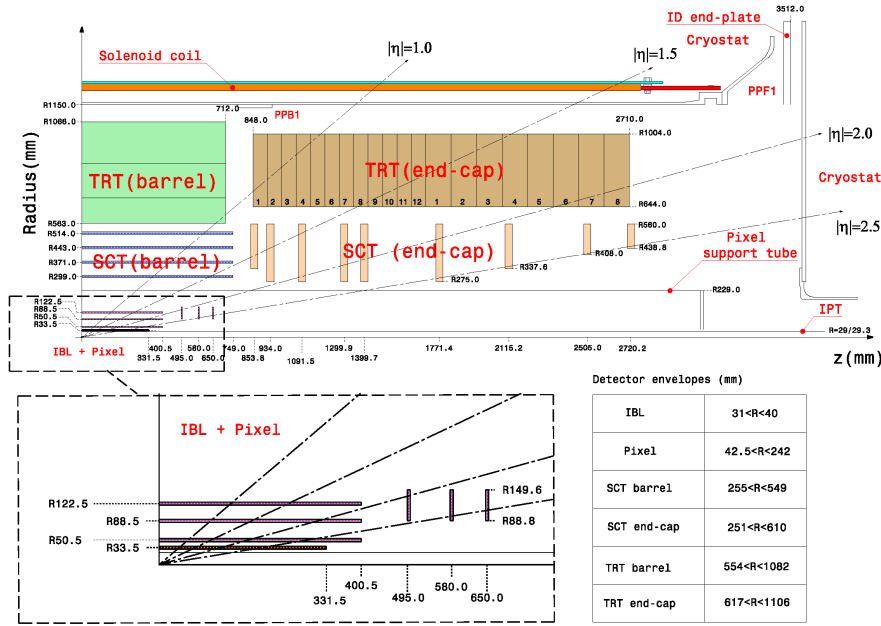


Figure 16: Plan view of a quarter-section of the ATLAS Inner Detector showing each of the major detector elements with its active dimensions and envelopes [55].

5.1.1 IBL

The first detector a particle produced by collisions encounters, outside the beam pipe, is the Insertable B Layer (IBL). It is a layer of silicon pixel detectors, with each pixel having a nominal size of $250 \times 50 \mu\text{m}^2$. With a mean sensor radius of 33.5 mm, the IBL sensors and front-end electronics are designed to cope with much higher hit rates and radiation doses [55]. In total, the IBL contains 1,736 pixel sensors, each measuring 2×2 cm, segmented in the r - ϕ and z directions, with over 12 million readout channels. The smaller layer radius and reduced pixel cell length are crucial parameters in defining the performance of the Inner Detector (ID), particularly the track-extrapolation resolution. This is achieved by providing an additional high-precision hit closer to the interaction point. Additionally, the IBL enhances the robustness of the ID when the efficiency of other detectors deteriorates due to radiation damage.

5.1.2 Pixel detector

The second detector a particle encounters, directly after the IBL, is the Pixel Detector, which consists of three concentric layers of high-precision silicon pixel detectors. The layers feature a pixel size of $50 \times 400 \mu\text{m}^2$. The purpose of the detector is the derivation of particle tracks, measuring the trajectory from the interaction region. It also aids in the determination of the interaction vertex position and the observation of secondary vertices. The Pixel Detector plays a crucial role in identifying b-quark jets, as the mass of the b

quark allows tracks to be reconstructed macroscopically in a process known as b-tagging. In total, there are 1736 pixel sensors, of size 2×6 cm, in the pixel detector, segmented in the r - ϕ and z directions, with over 80 million readout channels.

5.1.3 SCT

The Semiconductor Tracker (SCT) is the next detector a particle encounters after the Pixel Detector. In the barrel region, the SCT consists of four concentric cylindrical silicon microstrip double-layers: one layer is axial (parallel to the beam line), and another is at a stereo angle of 40 mrad (around 2°), providing measurements in two dimensions. In the endcap region, the SCT consists of nine disks with a set of radial and stereo strips. The SCT contains 2,112 modules in the barrel and 1,976 modules in the endcap, for a total of 4,088 modules (with 6.2 million readout channels). The pitch size of the strips is $80 \mu\text{m}$, providing an intrinsic space resolution of $17 \mu\text{m}$ in the r - ϕ plane and $580 \mu\text{m}$ in the z direction [53]. Typically, around eight space points per track are measured within the SCT.

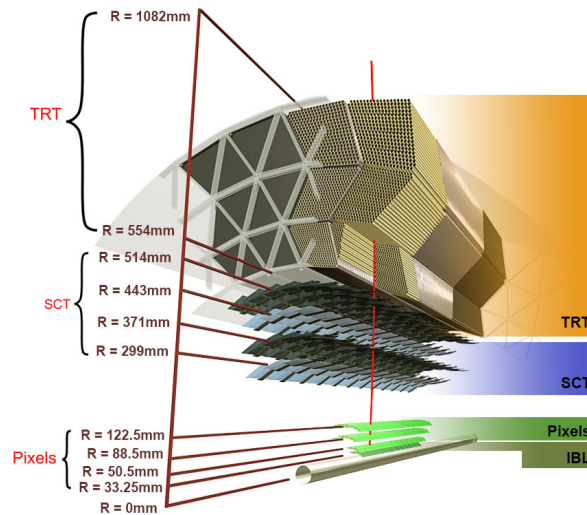


Figure 17: Drawing of the Inner Detector showing the barrel section. The beam pipe, the IBL, the Pixel layers, the four cylindrical layers of the SCT and the three layers of TRT barrel modules [56].

5.1.4 TRT

The transition radiation tracker (TRT) is the outermost part of the Inner Detector, located after the SCT. It consists of 4 mm diameter straw tubes filled with a Xe/CO₂/O₂ gas mixture (in the 70 : 27 : 3 proportion, respectively). In the barrel region, the TRT consists of up to 73 layers of straw tubes, parallel to the beam axis. In the endcap, 768

straw tubes fill 160 planes arranged radially in wheels. Charged particles traversing the straw tubes ionize the gas, creating electron-ion pairs, which generate currents that allow the determination of the minimum distance between the particle and the drift wire. Each charged particle creates signals in approximately 30 tubes, contributing significantly to the overall tracking performance. Another important function of the straw tubes is the discrimination between different charged particles. This is achieved by constructing the tube walls from materials with different refractive indices and including foam between the tubes, which emits transition radiation. The most significant amount of radiation is emitted during the passage of electrons.

Finally, the superconducting solenoid magnet system generates a magnetic field of 2 T, parallel to the beam axis. It has a diameter of 2.46 m, a length of 5.8 m, and weighs a total of 5.4 tons. Liquid helium ensures constant superconductivity by being pumped through the magnetic coils by a cryogenic system.

5.2 Calorimeters

The measurement of the energy of relativistic particles is important for any particle physics experiment. In ATLAS, for the majority of detected particles, with the exception of muons, this is done by the calorimeter system, which is responsible for stopping the particles and absorbing all their energy. Particles enter the calorimeter and initiate a particle shower, in which their energy is deposited in the calorimeter. The particle shower consists of a cascade of particles, with decreasing energy. Stopping a highly energetic particle requires a large amount of dense material. To reduce the material cost and the physical size of calorimeters, they are built as sampling calorimeters. Sampling calorimeters are constructed from alternating layers of absorbers made from dense material and active detectors used to measure the energies of particles. In ATLAS, there are two calorimeters: the electromagnetic (EM) and the hadronic. The EM calorimeter is optimized to measure photons and electrons, while the hadronic calorimeter is optimized to measure all hadrons. Both calorimeters are designed to stop particles at the expected energies produced by the collisions and aid in particle identification. A schematic drawing of the calorimeter structures is shown in Figure 18.

5.2.1 Electromagnetic calorimeter

The electromagnetic calorimeter measures the electromagnetic cascade left by photons and electrons. These two objects play a major role in the hZa and $Lb\gamma L$ analyses, in particular photons. To probe the resonances in these analyses, excellent photon energy resolution and identification are required. The EM calorimeter uses a groundbreaking accordion design for the absorbers, where the angles decrease with increasing radius, ensuring an approximately constant gap of around 4.5 mm between two absorbers. This design allows

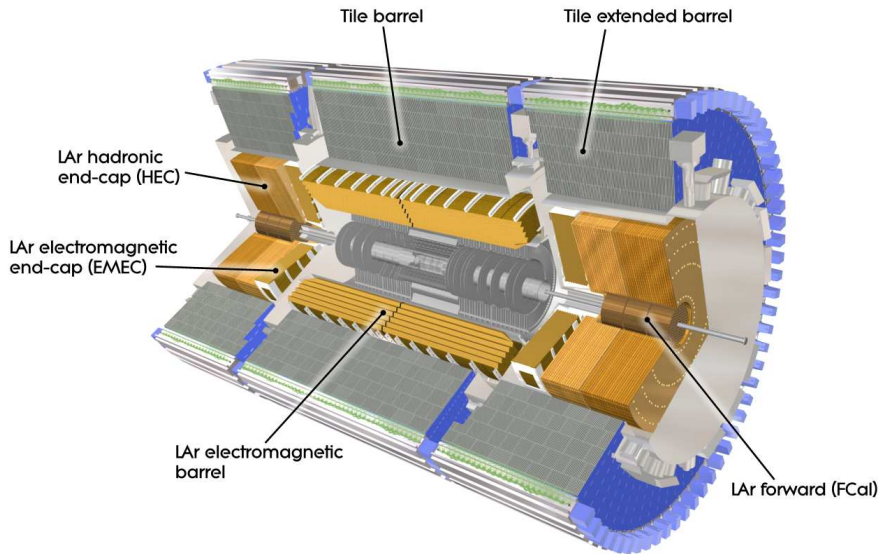


Figure 18: Cut-away view of the ATLAS calorimeter system [53].

a crossing particle to interact with the same number of absorbers regardless of its impact angle, maintaining a constant calorimeter thickness and full coverage in ϕ without any gaps [57], permitting fast detector response. The accordion design is visible in Figure 19.

The EM calorimeter uses lead as the absorber material and liquid argon as the active detector material. The readout electrodes are located in the gaps between the absorbers and consist of three conductive copper layers separated by insulating polyimide sheets. The segmentation of the calorimeter in η and in depth is obtained by etched patterns on the different layers and ϕ -segmentation is obtained by grouping together the appropriate number of electrodes. The EM calorimeter is split in two barrel and endcap regions.

In the barrel the electromagnetic calorimeter consists of two half barrels split at $\eta = 0$ with each spanning a range up to $|\eta| < 1.375$. A half-barrel is made of 1024 accordion-shaped absorbers, interleaved with readout electrodes. For ease of construction, each half-barrel has been divided into 16 modules, each covering $\Delta\eta = 22.5^\circ$. The total thickness of a module is at least 22 radiation lengths (X_0), increasing from 22 X_0 to 30 X_0 between $\eta = 0$ and $\eta = 0.8$ and from 24 X_0 to 33 X_0 between $\eta = 0.8$ and $\eta = 1.3$. The modules, shown in figure 19, have 3424 readout channels. The end-cap electromagnetic calorimeters cover a region of $1.375 < |\eta| < 3.2$ made of an inner and outer wheel. The wheels are divided into eight wedge shaped modules with 768 (256) absorbers in the outer (inner) wheel.

The electromagnetic calorimeter is segmented into four layers, longitudinal in shower depth: the presampler, layer 1, layer 2, and layer 3. The presampler (PS) layer, covering the pseudorapidity interval $|\eta| < 1.8$, is used to correct for energy loss upstream of the calorimeter. The PS consists of an active LAr layer with a thickness of 1.1 cm (0.5 cm) in

the barrel (endcap) and has a cell size of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$, with an average X_0 of 2. The first layer has a thickness of about $4.3 X_0$ at $\eta = 0$ and it is segmented into high-granularity strips in the η direction, with a typical cell size of $\Delta\eta \times \Delta\phi = 0.003 \times 0.01$ in the barrel. The granularity is sufficient to provide event-by-event discrimination between single-photon showers and two overlapping showers originating from the decays of neutral hadrons, mostly π_0 and η mesons in jets. The second layer has a thickness of about 16 X_0 at $\eta = 0$, and a cell size of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$. It collects most of the energy deposited in the calorimeter by photon and electron showers. Finally, the third layer has a cell size of $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$ and a depth of about 2 X_0 at $\eta = 0$. It is used to correct for leakage of high-energy showers into material beyond the EM calorimeter.

The interactions of photons and electrons with the ATLAS EM calorimeter produce similar electromagnetic showers, depositing a significant amount of energy in a restricted number of neighboring calorimeter cells, which can be organized as EM clusters. For the recognition and reconstruction of photons and electrons, a sliding window approach is utilized, with dimensions 3×5 in units of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$, to search for electromagnetic cluster seeds as longitudinal towers with total cluster transverse energy above 2.5 GeV. The minimum distance for the object resolution depends on the angle and the reconstructed object p_T and type. The energy resolution is measured with $Z \rightarrow ee$ decays [58] and it is found to be around $\sigma E/E = 0.0005$ or 0.05 %.

In the transition regions on both sides from the electromagnetic barrel calorimeter to the electromagnetic end-cap calorimeter there is a large amount of material in front of the first calorimeter layer. As this material can produce showers before the detector it is detrimental to the overall object quality. This region, $1.37 < |\eta| < 1.52$, is called calorimeter crack and it is omitted for standard electron and photon analysis.

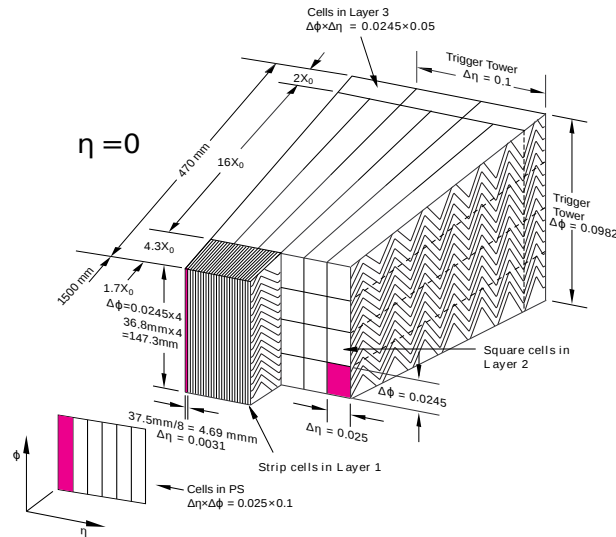


Figure 19: Longitudinal composition of a barrel module [59].

5.2.2 Hadronic calorimeter

The hadronic calorimeter surrounds the electromagnetic calorimeter and consists of the tile calorimeter, the liquid-argon hadronic end-cap calorimeter (HEC) and the liquid-argon forward calorimeter (FCAL). These detectors are responsible for the measurement of particle jets, mesons and baryons.

The tile calorimeter is a sampling calorimeter using steel as the absorber and polystyrene scintillators as the active medium [53]. It is located in the region $|\eta| < 1.7$, behind the liquid argon electromagnetic calorimeter, and is subdivided into a central barrel, 5.8 m in length, and two extended barrels, 2.6 m in length. The radial depth of the calorimeter is 7.4λ (interaction lengths), with each barrel consisting of 64 modules of size $\Delta\phi \approx 0.1$. The granularity in the tiles is coarser than in the electromagnetic calorimeter, as the showers are larger by nature.

The HEC calorimeter covers the two endcap zones of $1.5 < |\eta| < 3.2$. It consists of two cylindrical wheels and is made of copper and tungsten plates, using liquid argon as the active material. The choice of copper is made to limit the depth of the jets, due to space constraints. Each end-cap is composed of 32 wedge-shaped modules, longitudinally divided into two sections, resulting in 4 layers per endcap.

The FCAL calorimeter provides coverage in the boosted region near the beam pipe, where $3.1 < |\eta| < 4.9$. The close vicinity and coupling between these systems result in a highly hermetic design, which minimizes energy losses in cracks between the calorimeter systems and limits the backgrounds that reach the muon system. It is made of copper and tungsten, with liquid argon as the active material, similar to the HEC. The FCAL is sensitive to both hadronic and electromagnetic showers, as the electromagnetic calorimeter does not extend as close to the beam pipe.

The Zero degree calorimeter (ZDC) is important for the HI program. The primary purpose of the ZDC [60] is to detect forward neutrons with $|\eta| > 8.3$ in heavy-ion collisions. The ZDC plays a key role in determining the centrality of such collisions, which is strongly correlated to the number of very forward (spectator) neutrons. The ZDC enhances the acceptance of ATLAS central and forward detectors for diffractive processes and provides an additional minimum-bias trigger for ATLAS. There are four ZDC modules installed per arm: one electromagnetic (EM) module (about 29 radiation lengths, X_0 , deep) and 3 hadronic modules (each about 1.14 interaction lengths, λ , deep).

5.3 Muon Spectrometer

The Muon spectrometer forms the outer part of the ATLAS detector and is designed to measure the momentum of charged particles exiting the barrel and endcap calorimeters. It is immersed in a 2.5 T magnetic field produced by a toroidal magnet system, built from eight superconducting magnet coils. It covers the pseudorapidity range of $|\eta| < 2.7$ and

measures the transverse momentum of muons with resolution $\sigma_{p_T}/p_T < 0.1$, which requires a sagitta measurement of $500 \mu\text{m}$ to be made with a precision of $50 \mu\text{m}$. For this purpose, the muon spectrometer consists of a plethora of subdetectors.

The spectrometer is made up of many different types of subdetectors, azimuthally arranged in eight parts, each divided in two sectors overlapping in ϕ . The detectors are further split into chambers. The MDT (Monitored Drift Tubes) cover the region $|\eta| < 2.7$ consisting of three to eight layers of drift tubes. The tubes are filled with a gas mixture of argon and CO_2 and they have a range of shapes and sizes, due to the large number of support structures and appliances in the vicinity. In total, it consists of around 1100 chambers, which provide a spatial resolution of $\sigma(z) = 0.35 \mu\text{m}$. The CSC (Cathode-strip chambers) are used in the endcap region of $|\eta| > 2.0$, as it experiences counting rates of hits above the capability of the MDT. They combine high spatial, time, and double-track resolution with high-rate capability and low neutron sensitivity. The CSCs are chambers oriented in the radial direction, producing the position of the track by interpolation between the charges induced on neighboring cathode strips.

For the purposes of triggering by muon transverse momentum and topology, the RPC (Resistive plate chambers) and TGC (Thin gas chambers) are used. The former have lower spatial resolution compared to TGC but improved time resolution, enabling their use as fast triggers for the ATLAS detector. The RPC has a response time of 1.5 ns and the TGC has a slightly slower response time of 5 ns. They are placed in both the barrel and the endcaps.

5.4 Object reconstruction

The following section describes the interpretation of the sub-detector signals into objects meaningful for physics analysis.

5.4.1 Charged particles tracks

Charged particles tracks with transverse momentum $p_T > 0.5 \text{ GeV}$ and $|\eta| < 2.5$ are reconstructed and measured in the inner detector. A reconstructed track is required to have at least seven precision hits (from the pixel and SCT subdetectors). The track should correspond to the primary vertex along the beam pipe. For track recognition, a suite of track-fitting tools is utilized, including global χ^2 and Kalman-filter techniques [61]. Charged-particle track reconstruction in the pixel and SCT detectors begins by assembling clusters from the raw measurements and progresses to the TRT. Clusters are groups of pixels or strips with energy exceeding a threshold.

Measurement of tracks is also performed in the muon spectrometer for muons with transverse momentum ranging from 3 GeV to 3 TeV. The software used for the reconstruction of muon tracks is common to both the ID and the muon spectrometer. There are

three distinct track reconstruction strategies: stand-alone (only the muon spectrometer), combined (muon spectrometer and inner detector), and segment tag (inner detector with part of the muon spectrometer). The combination of these strategies allows for optimal acceptance and efficiency. The muon spectrometer track parameters are determined at the inner stations, which provide the first set of measurements in the muon spectrometer. The track is then propagated back to the interaction point, and the momentum is corrected for energy loss in the calorimeters. The track recognition process follows several steps, starting with the pre-processing of raw data to form drift circles in the MDT or clusters in the CSC and trigger chambers (RPCs and TGCs), followed by pattern finding and segment making, segment combining, and finally, track fitting. The efficiency of the muon performance is validated using the tag and probe method [62], using $Z \rightarrow \mu\mu$ decays.

5.4.2 Electromagnetic objects

For the standard reconstruction of electrons and photons, a seed cluster is taken from the electromagnetic calorimeter, and a loosely matching track is searched for among all reconstructed tracks [63]. A seed is a localized area in the EM calorimeter with energy deposition. Electron and photon candidates are separated by requiring electrons to have an associated track but no associated conversion [64]. Conversion is the process in an electromagnetic field where a photon converts into an electron-positron pair. In contrast, photons are defined as having no matched track or being matched to a reconstructed conversion.

The energy calibration of the detector is achieved via observation of $Z \rightarrow ee$ decays [63], which also provides an estimate of the reconstruction efficiency and systematic uncertainties. A validation of the photon calibration is also performed with $Z \rightarrow ll\gamma$ decays. The latest calibration uncertainties are around 0.05% for electrons from resonant Z-boson decays, 0.4% at $E_T \approx 10$ GeV (transverse energy) and for photons of $E_T \approx 60$ GeV they are 0.2% on average [65]. For the determination of object quality and identification, shower-shape variables (lateral and longitudinal shower profiles) are calculated using the granularity of the calorimeter. These shower-shape variables are then used to define QCD background rejection cuts, commonly known as working points (e.g. loose, tight), which define the object reconstruction efficiency and purity. This is possible due to the different shower-shapes between jets and electromagnetic objects. For the tight working point the photon identification efficiency scales from 45-60 % for $E_T = 10$ GeV to 95-98 % for $E_T > 100$ GeV [59]. Misidentified photons can be attributed to electrons or jets that did not leave a track in the ID. Finally, some form of isolation (e.g. loose, tight) is also applied to the objects, which compares the energy of the deposits within an angular cone to the energy of the object, both in tracks (p_{Tcone}) and EM calorimeter (E_{Tcone}) [59]. This mainly rejects the QCD jets, which tend to deposit energy into multiple tracks and/or

calorimeter clusters.

The definitions of the shower-shape variables used in the photon working points are as follows, shown in Figure 20, reproduced from [59]:

- f_1 - the longitudinal shower-shape discriminant, ratio of the energy in the first layer to the total energy of the EM cluster.
- $w_{\eta 2}$ - lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells.
- E_{ratio} - ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies.
- $frac_{s1}$ - shower-shape in the shower core, $(E(\pm 3) - E(\pm 1))/E(\pm 1)$, where $E(\pm n)$ is the energy in $\pm n$ strips around the strip with highest energy.
- f_{1core} - a second longitudinal shower-shape discriminant, the fraction of energy reconstructed in a ± 3 strip region (shower core) around the strip with higher energy in the first layer of the EM calorimeter (strips) with respect to the total energy of the cluster.
- R_{had} - ratio of E_T in the hadronic calorimeter to E_T of the EM cluster.
- R_ϕ - ratio of the energy in $3 \times 3 \eta \times \phi$ cells over the energy in 3×7 cells centered around the photon cluster position.
- R_{had1} - ratio of E_T in the first sampling layer of the hadronic calorimeter to E_T of the EM cluster.
- R_η - ratio of the energy in $3 \times 7 \eta \times \phi$ cells over the energy in 7×7 cells centered around the photon cluster position.
- w_{s3} - lateral shower width, $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $3 \times 2 \eta \times \phi$ calorimeter strips, and i_{max} is the index of the highest-energy strip calculated from three strips around the strip with maximum energy deposit.
- w_{stot} - total lateral width of the shower, $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $20 \times 2 \eta \times \phi$ calorimeter strips, and i_{max} is the index of the highest-energy strip measured in the strip layer.
- f_{side} - energy outside the core of the three central strips but within seven strips divided by energy within the three central strips.

- ΔE_s - difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima.

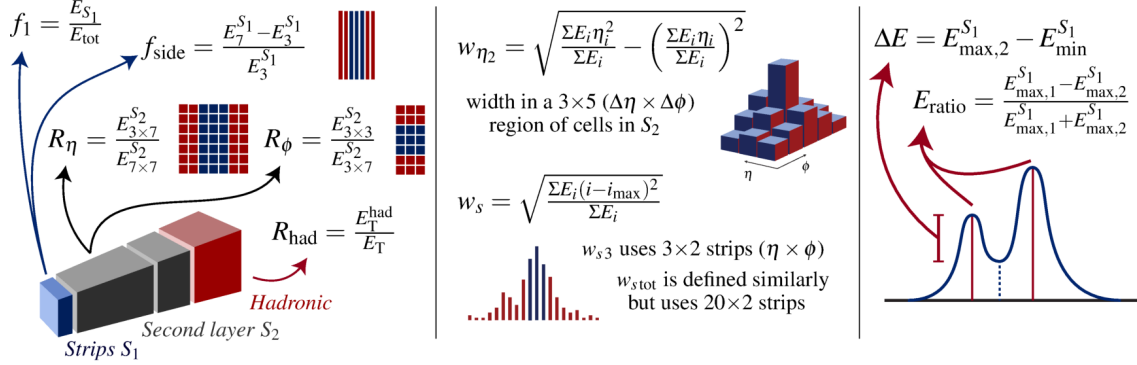


Figure 20: Schematic representation of the photon identification discriminating variables [59].

Photons are a central object in the ATLAS physics program, involved in landmark results such as the Higgs mass measurement. To achieve these high-level results, excellent signal sensitivity, energy resolution, and background rejection are required. Thus, fake photon backgrounds, such as calorimeter noise and jets reconstructed as photons, must be excluded. The ATLAS collaboration and the egamma group have created recommendations for the usage of photons with $p_T > 10$ GeV, fiducial η cuts, identification and isolation. In this thesis photons are the central objects of the hZa and LbyL analyses.

5.4.3 Jets

The quarks produced at the LHC are not asymptotically free in this energy regime, undergoing hadronization and producing jets. It is possible to treat jets as particle-like objects. They possess a cone-like structure and mainly deposit energy in the hadronic calorimeters. For jet reconstruction, clustering algorithms are utilized, two of which are the seeded cone algorithm [66] and the anti- k_t jet clustering algorithm [67]. Jet reconstruction is an involved process in ATLAS, as jets can originate from a wide variety of processes and energies. In particular, b-jets are important, as the b-mesons have a macroscopic displacement in the ID [68]. However, the exact determination of the initial jet quark or gluon remains a serious challenge [69].

5.5 Trigger system

The ATLAS detector produces an extraordinarily large amount of data, which is impossible to save and process. In addition, since most interesting processes have low cross sections,

very few of these data are actually meaningful for physics analysis. To determine useful data, a two-level trigger system is employed to record data at an average rate of 1 kHz from physics collisions, starting from an initial bunch crossing rate of 40 MHz. The system decides in real time whether to record data from a given collision [70]. The trigger system must be flexible and reliable to match the ever-changing detector conditions.

The Level-1 (L1) trigger is a hardware-based system that uses custom electronics to trigger on reduced granularity information from the calorimeter and muon detectors. The analog detector signals are digitized, calibrated by the preprocessor, and sent in parallel to the Cluster Processor (CP) and Jet/Energy-sum Processor (JEP). The CP system is able to recognize photons, electrons and τ candidates above a certain energy threshold. The L1 muon trigger utilizes the RPC detector [71]. The JEP system is able to recognize jet candidates and missing transverse energy. The L1 trigger can select events by considering event-level quantities (e.g., the total energy in the calorimeter), the multiplicity of objects above thresholds (e.g., the transverse momentum of a muon, etc.), or by applying topological requirements (such as invariant masses or angular distances) to the objects found by the electronics [70]. When an L1 trigger is observed, the rest of the detector is read out.

The second stage of the trigger, the HLT (High Level Trigger), is software-based. A list of available triggers for every run, called the menu, is essential for standard physics performance. CPUs use reconstruction algorithms to determine the final selection in a matter of up to 3 seconds. In most cases, the reconstructed objects are used to test whether the trigger is satisfied, utilizing information from the whole detector. The physics output is, on average, 1 kHz, and the permanent data storage is done in the Tier-0 computing facility located at CERN. For smooth detector operation, constant monitoring of the trigger system is required to cope with changes in detector conditions or instantaneous luminosity.

5.6 Event simulation

In physics, it is important to predict the possible results of any specific experiment. In ATLAS, there is a wide range of possible signatures due to the complexity of the SM. The complexity of the detector should also be taken into account. To successfully predict the experimental signatures of a process of interest, all elements of the process—from the proton beam to the interaction of the particles with the detector—must be considered. This should also be done for different detector conditions to minimize differences between simulation and real data. This is achieved by simulating SM or BSM processes with a dedicated generation and reconstruction framework. This framework consumes the majority of ATLAS computing resources. The flow of the simulation process is shown in Figure 21.

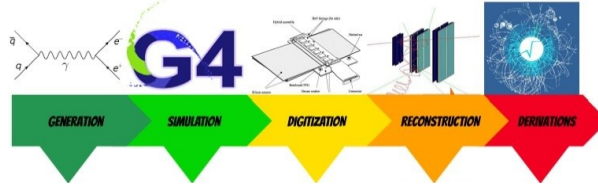


Figure 21: The event simulation production chain in ATLAS, simplified version from [72].

5.6.1 Event generation

A physics process of interest, such as Higgs production or ALP production, begins the event simulation with event generation. This is a complicated process, and there are many parameters to consider. The files that contain the simulated event information are called MC (Monte Carlo) files, as they are created from computational algorithms that rely on repeated random sampling to obtain numerical results. The initial state for a hadronic collider is proton collisions, which are simulated billions of times to produce events. The simulation is performed by standard programs that use the latest theoretical results and parton distribution functions to predict parton showers and fragmentation, leading to the observed final-state objects. Some of these programs are Pythia [73], Herwig [74], and Sherpa [75]. The event files produced are called truth files, describing the particles before any detector interaction. These files are significant for complete physics analysis, as they describe genuine signal or background events. In this thesis, the truth files are used for ALP signal validation and systematic uncertainties calculations.

5.6.2 Detector simulation

The simulation of the interactions of particles with the detector is the most computationally expensive part of event simulation. The path of 'stable' particles is simulated as they interact with the detector material and magnetic fields, produce new particles in those interactions, or potentially decay. For this, a special software called Geant4 [76] is used, which contains a realistic model of the detector at the μm level to identify every possible hit within the detector volume. Geant4 is a toolkit for simulating the passage of particles through matter. Every tiny change in the detector environment must be taken into account, as it can lead to different observations. The calorimeters are the most difficult parts to simulate due to the nature of the showers. The outcome of this step is a list of energy deposits in the various subdetectors of ATLAS.

5.6.3 Digitization

The process of digitization uses special software to transform the hits predicted by the simulation into actual digitized signals, matching those measured by the detector. These

signals are low-level information, such as charges measured by the trackers or calorimeters. This step is not as computationally expensive as the previous step, as the position of the hits leads to predictable responses. From this step onward, the events are completely indistinguishable from real events.

5.6.4 Reconstruction and Derivation

This step reconstructs the digital signals into physics objects (such as electrons and hadronic jets), producing the same kind of files as actual data that can be used for analysis. The only difference with real data is that a version of the truth record is stored, which can be useful for cross-checks in analysis. The events after reconstruction are in AOD format (Analysis Object Data). Most analyses in ATLAS utilize a subset of the recorded data in relation to objects or triggers. For example, a fully leptonic Z analysis is probably not interested in high-energy jets or photons. To limit the size of files used in physics analyses, there is a final step called derivation, which filters the useful data for specific analyses from the AOD files. Real data are also available in derivation formats. In addition, to cope with changing detector conditions, there are different MC generation campaigns corresponding to different data-taking periods. For example, mc16e corresponds to real data from 2018. Finally, only MC produced by the ATLAS central simulation process is allowed to be used in approved ATLAS results.

6 Search for Higgs boson decays into a Z boson and an ALP

6.1 Introduction

This chapter presents a reinterpretation of a Run 2 analysis conducted by the ATLAS experiment, which investigates the decays of a Higgs boson into a Z boson and an ALP (axion-like particle) that promptly decays into a pair of photons [45], process Feynman diagram is shown on Figure 22. The analysis is sensitive to ALP masses in the range of 2 GeV to 34 GeV, i.e. $m_a \leq m_h - m_Z$, with a particular emphasis on scenarios involving small ALP couplings. Such cases correspond to longer ALP lifetimes, leading to displaced photon origin vertex signatures. Additionally, for consistency with the original analysis, a dedicated search assuming promptly decaying ALPs is also performed.

In this thesis, an additional event selection proposed by the author is introduced, addressing shortcomings of the original selection and achieving competitive results. This new event selection is a significantly simplified version of the original analysis selection. The new selection utilizes the full ATLAS Run 2 dataset, along with the same background and signal Monte Carlo (MC) samples as those used in the original event selection. Furthermore, displaced vertex signatures are explored for this selection, and constraints on the ALP-photon coupling are derived across a range of masses.

This chapter is structured as follows: Section 6.2 describes the relevant MC and data samples. Section 6.3 provides a concise overview of the experimental phenomenology of displaced vertices. Section 6.4 summarizes the object selection. Section 6.5 details the reinterpretation of the ATLAS hZa analysis and presents its results. Section 6.6 introduces an optimized version of the ATLAS hZa analysis, including background estimation, systematic uncertainty treatment, and result comparisons. Finally, a summary is presented in Section 6.7.

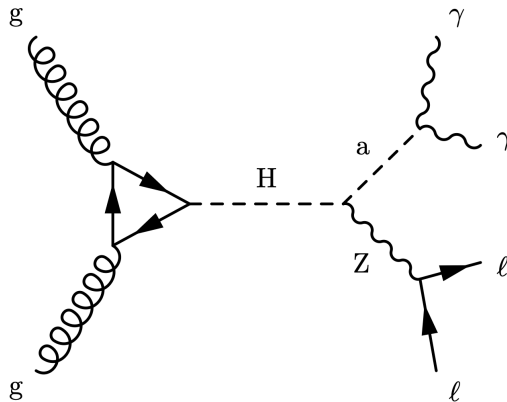


Figure 22: Feynman diagram of $Higgs \rightarrow Z a(a) \rightarrow l^+ l^- \gamma \gamma$. Higgs ggH production [77].

6.2 Dataset and MC simulated events

6.2.1 MC Signal sets

The signal samples are generated under the assumption of a SM Higgs boson produced via the gluon–gluon fusion mechanism. The gluon–gluon fusion Higgs (ggH) boson production cross section in pp collisions at $\sqrt{s} = 13$ TeV, estimated at N3LO QCD and NLO EW, is $\sigma_{SM} = 48.58$ pb for $m_h = 125.09$ GeV [49]. The SM Higgs then decays to a Z boson and a neutral pseudo-scalar a , the ALP. The a particle is relatively light (2 - 34 GeV) compared to electroweak bosons and is only allowed to decay to a pair of photons. The VBF set was found to be subdominant as it has a 12 times smaller cross section than the ggH set. There are 13 different mass points and 4 different ALP-photon couplings ($C_{\gamma\gamma} = 10^{-2}$ - prompt, 10^{-3} , 10^{-4} and 10^{-5}) simulated for the ggH set. The samples used are listed in Appendix A.1.

The ALP characteristics, like lifetime, were determined using the formulas of [26] which also include displaced signatures phenomenology. In this sample production the $h \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ was simulated where $l = (\tau, \mu, e)$ and the τ pair decays are subsequently being rejected as the final state reconstructed dilepton (μ, e) objects are outside the Z peak.

6.2.2 Monte Carlo backgrounds sets

The main source of background in this analysis comes from SM production of Z +jets and $Z\gamma$ events. Simulated samples are generated to study these backgrounds and to estimate systematic uncertainties.

The production of Z +jets was simulated with the Sherpa 2.2.1 [75] generator using next-to-leading-order (NLO) matrix elements (ME) for up to two partons, and leading-order (LO) matrix elements for up to four partons calculated with the Comix [78] and OpenLoops [79, 80, 81] libraries. They were matched with the Sherpa parton shower [82] using the MEPS@NLO prescription [83, 84, 85, 86] using the set of tuned parameters developed by the Sherpa authors. The NNPDF3.0 nnlo set of PDFs [87] was used and the samples were normalised to a next-to-next-to-leading-order (NNLO) prediction [88].

The production of $Z\gamma$ final states was simulated with the Sherpa 2.2.2 [75] generator. Matrix elements at NLO QCD accuracy for up to one additional parton and LO accuracy for up to three additional parton emissions were matched and merged with the Sherpa parton shower based on Catani–Seymour dipole factorisation [78, 82] using the MEPS@NLO prescription [83, 84, 85, 86]. The virtual QCD corrections for matrix elements at NLO accuracy were provided by the OpenLoops library [79, 80, 81]. Samples were generated using the NNPDF3.0 nnlo PDF set [87], along with the dedicated set of tuned parton-shower parameters developed by the Sherpa authors.

All generated signal MC events were processed through a full simulation of the ATLAS

detector geometry and response using the Geant4 toolkit [76]. The simulation includes multiple pp interactions per bunch crossing (pile-up), as well as the detector response to interactions in bunch crossings before and after the one producing the hard interaction. In order to model the effect of pile-up, simulated inclusive pp events are overlaid on each generated event and reweighted to match the conditions of the 2015-2018 data sample. The interactions were simulated with Pythia 8.186 using the A3 tune [73] and the NNPDF23LO 228 PDF set [89]. The samples used are listed in Appendix A.1.

6.2.3 Data

This analysis uses pp collision data collected by the ATLAS experiment from 2015 to 2018 with a centre-of-mass energy of $\sqrt{s} = 13$ TeV. After applying data quality requirements [90] the full data set corresponds to an integrated luminosity of $140.1 \pm 1.2 \text{ fb}^{-1}$ [52]. Only data periods where all the relevant detector subsystems were operational are included [91].

6.3 Displaced Vertices in experimental analysis

The structure of the ATLAS detector plays a major role in the possible experimental signatures it will be sensitive to. Decay length scales for macroscopic particles can be derived from the position of the tracker, the EM calorimeter and the hadronic calorimeter compared to the interaction point. The inner tracker or inner detector starts at around 3.3 cm from the beam pipe meaning that from that point onward one can determine the position of the displaced origin vertex for electrons, muons or jets. The EM calorimeter starts at around 1.97 meters meaning that one can be sensitive to displaced decays to photons for decay lengths up to that. Finally, for decays to muons, one can use only the muon spectrometer information and increase the distance up to 7 meters.

For $a \rightarrow \gamma\gamma$, the photons originate from the position in the detector where the ALP decays. This results in two distinct signature scenarios, as illustrated in Figure 23. For ALPs with masses below 1.5 GeV, the photons are predominantly reconstructed as a merged photon object or a photon jet in most events [92]. The likelihood of forming a merged photon object increases with the displacement of the ALP decay vertex. In such cases, it becomes impossible to calculate the ALP mass due to the absence of two distinct Lorentz vectors. Thus it can enhance the measured $hZ\gamma$ cross section. Conversely, in the exotic hZa scenario, a very low-mass ALP, characterized by a longer lifetime, has a higher probability of leaving the detector without decaying. For ALPs with mass higher than 1.5 GeV, it is possible to detect both photons and subsequently calculate the ALP mass. In this thesis, limits will be estimated for a range of ALP- γ $C_{\gamma\gamma}$ couplings focusing on scenarios where both photons are reconstructed and the ALP decays before reaching the electromagnetic calorimeter.

An especially intriguing scenario arises when both photons are successfully recon-

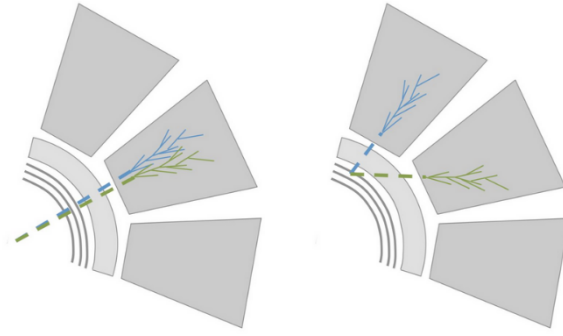


Figure 23: ALP decays to photons compared to the EM calorimeter. In the left plot is the scenario where a collimated photon jet reaches the EM calorimeter. In the right plot is the scenario where the ALP decays close to the EM calorimeter [92].

structed. There are some experimental challenges associated with ALPs with macroscopic decay lengths. The first challenge concerns the accurate reconstruction of the ALP mass. In ATLAS, the angular variables of the photon objects following default reconstruction are defined relative to the interaction point (IP). In this case, photons originate from a decay point between the IP and the EM calorimeter. This displacement causes the Lorentz vectors of the photons to shift slightly towards smaller ϕ and η angular values. As a result, the reconstructed diphoton mass exhibits smearing compared to the true ALP mass.

Furthermore, the photon ID and isolation variables can be influenced. The photon ID characterization is based on the shower-shape variables while also taking into consideration the detector structure in different pseudorapidities. The displacement of photon origin can result in the misidentification of photons if they reach the EM calorimeter at an angle that deviates from the expected trajectory for a given pseudorapidity. This introduces a systematic uncertainty, linked to the ALP decay length, which needs to be estimated. In addition, for some ALP decays which are very close to the EM calorimeter or very light ALPs, the reconstructed photons are close to each other, leading to failure of the standard isolation working points. So for most ALP analyses with photons, special care has to be given to the isolation, which is the case for the analyses described in this thesis.

6.4 Object selection

In the $h \rightarrow Za$ analysis events containing leptonically decaying Z bosons are selected. The events have to be selected by the leptonic triggers to be recorded.

6.4.1 Triggers

Data events used for this analysis are selected using an OR of single electron, single muon, di-electron and di-muon triggers with p_T thresholds and object quality criteria that depend

on the data-taking period. In more detail the triggers are shown in Tables 4 and 5.

Single Lepton Triggers		
	Electron Triggers	Muon Triggers
mc16e or data 2018	HLT_e26_lhtight_nod0_ivarlose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50
mc16d or data 2017	HLT_e26_lhtight_nod0_ivarlose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50
mc16a or data 2016	HLT_e24_lhtight_nod0_ivarlose HLT_e26_lhtight_nod0_ivarlose HLT_e60_lhmedium_nod0 HLT_e60_medium HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50
mc16a or data 2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmediumHLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50

Table 4: Single lepton triggers used by ATLAS during Run 2 for different data taking years.

Double Lepton Triggers		
	Electron Triggers	Muon Triggers
mc16e or data 2018	HLT_2e17_lhvloose_nod0_L12EM15VHI HLT_2e17_lhvloose_nod0 HLT_2e24_lhvloose_nod0	HLT_mu22_mu8noL1
mc16d or data 2017	HLT_2e17_lhvloose_nod0_L12EM15VHI HLT_2e17_lhvloose_nod0 HLT_2e24_lhvloose_nod0	HLT_mu22_mu8noL1
mc16a or data 2016	HLT_2e15_lhvloose_nod0_L12EM13VHI HLT_2e17_lhvloose_nod0	HLT_mu22_mu8noL1
mc16a or data 2015	HLT_2e12_lhvloose_L12EM10VH	HLT_mu18_mu8noL1

Table 5: Dilepton triggers used by ATLAS during Run 2 for different data taking years.

The triggers observe the deposited energy in the EM calorimeter and store the event for further processing if the energy passes the trigger threshold. The HLT triggers operate after a fast object reconstruction is finalized. Special attention should be given to the slightly different triggers between periods. In the 2015 and part of 2016 data taking, lower p_T thresholds for triggering were used which would be ineffective for handling the higher

Selection	Criteria
Leading lepton p_T	> 27 GeV
Subleading lepton p_T	> 20 GeV
Lepton η	$ \eta < 2.47$ and not $1.37 < \eta < 1.52$ (electron) $ \eta < 2.5$ (muon)
Lepton identification	Medium
Lepton isolation	FixedCutLoose (electron) / FixedCutLoose FixedRad (muon)
Photon p_T	> 10 GeV
Photon η	$ \eta < 2.47$ and not $1.37 < \eta < 1.52$
Photon identification	Loose
Photon isolation	FCLoose variation

Table 6: Summary of the object selection criteria.

instantaneous luminosity of the later years. For this reason higher p_T thresholds were implemented in subsequent years to ensure no loss of interesting events. Additionally, events passing these triggers are required to include a lepton that geometrically matches the object responsible for triggering the event.

6.4.2 Object reconstruction

The main objects used in the analysis are electrons, muons and photons. Jets can be considered secondary objects as they are largely responsible for the Z+jets background, which is the most dominant background. A summary of the object selection criteria or preselection (PS) is summarized in Table 6.

Electron objects are created from an energy deposit in the electromagnetic calorimeter matched with a track in the inner detector. They are required to be within the fiducial region of the detector, $|\eta| < 2.37$, excluding the transition region between the barrel and end-cap calorimeters ($1.37 < |\eta| < 1.52$). A likelihood based identification requirement is used to select electrons [64]. The reconstructed electrons are required to have $p_T > 20$ GeV, most energetic $p_T > 27$ GeV to match trigger requirement, and pass the “medium” likelihood identification criterion. Electron candidates must fulfill quality requirements based on the expected shower-shape and hardware condition. Additionally, a isolation criterion (“FixedCutLoose”) is applied.

Muon objects are reconstructed by matching an inner detector track with a track in the muon spectrometer [62]. Only candidates within $|\eta| < 2.5$ are considered. In addition, at this stage all muon candidates are required to have a $p_T > 20$ GeV, most energetic $p_T > 27$ GeV to match trigger requirement, and pass a “medium” muon identification requirement. In addition a isolation criterion (“FixedCutLoose FixedRad”) is applied.

Photons are reconstructed from energy clusters in the electromagnetic calorimeter. Clusters without matching inner detector tracks are classified as unconverted photon candidates while clusters matched to inner detector tracks consistent with the hypothesis of a photon conversion into e^+e^- pair are classified as converted photon candidates [93]. Reconstructed photon candidates are required to have transverse momentum $p_T > 10$ GeV, pseudorapidity $|\eta| < 2.37$, excluding the barrel/endcap calorimeter transition region $1.37 < |\eta| < 1.52$ and must pass standard quality and cleaning cuts. Furthermore, photons must pass photon identification criteria corresponding to the working point "Loose". The hZa ATLAS analysis and the optimized selection analysis use slightly different isolation criteria, which are variations of the isolation criterion working point "FCLoose".

Geometric overlap between objects passing the above selection criteria creates ambiguity in the identity of the objects. The following procedure is used to remove these ambiguities. Photons within a cone of $\Delta R = 0.3$ with electrons or muons are excluded. Electrons within a cone of $\Delta R = 0.2$ of muons are excluded.

6.4.3 Analysis tools and framework

The analyses described in this thesis utilize tools developed by the CERN collaboration and ATLAS. The sample generation, in signal and background, was performed by the ATLAS central (D)AOD production. The processing of the AODs into ntuples (simpler object centered framework) was done using the NtupleMaker based on r21 AnalysisBase, where all the CP tools are used, recommended by the physics object preparation groups of ATLAS. The tools are stored in the ATLAS tool framework called Athena with release 21 of the framework used. Finally, for the actual analysis, i.e. production of histograms and event selections our own C++ code was developed utilizing Athena EventLoop.

6.5 hZa ATLAS displaced event selection description and results

The selection criteria for the signal regions remain unchanged from the original Run 2 analysis [3]. The event selection is divided into two regions: the Z-reconstruction region and the resolved diphoton category. In general, the ATLAS prompt analysis is based on the premise that ALPs typically have low masses (ranging from 0.1 to 9 GeV) and are produced with high Lorentz boost factors. In the merged photon scenario, an excess of $hZ\gamma$ events is expected compared to the SM predictions.

6.5.1 Z boson selection

After the candidate has been successfully detected by the leptonic triggers and the object reconstruction has been completed, the Z boson has to be reconstructed. The following requirements are applied to select events with Z boson candidates:

- **2 leptons:** events with more than two leptons are rejected.
- **Leptonic angle:** the angle between the two leptons must be $\Delta R_{ll} > 0.2$ ³.
- **Opposite charge:** the two leptons should have opposite electric charge.
- **lep1,lep2 p_T :** Leading lepton $p_T > 27$ GeV and subleading lepton $p_T > 20$ GeV.
- **Same flavour:** the two leptons should have the same flavour.
- **Dilepton mass cut:** the reconstructed dilepton invariant mass must be between 81 GeV and 101 GeV.
- **Dilepton p_T cut:** the reconstructed dilepton transverse momentum is $p_T > 10$ GeV.

The first cut, after the preselection (PS) of lepton objects, where more than two leptons are rejected is designed to eliminate events involving photon pair conversions. The second cut with $\Delta R_{ll} > 0.2$, aims to suppress low-mass dilepton resonances, such as those arising from neutral pions. The cut on the dilepton mass is chosen to maximize signal significance, as studies have shown that the signal-to-background ratio is optimal within this specific mass window [3]. Finally, the dilepton p_T cut is motivated by the observation that signal distributions peak at higher transverse momentum values compared to the dominant $Z\gamma$ and Z +jets backgrounds [3]. This behaviour is driven by the recoil against the Higgs boson. Using the ALP samples of 4, 5 and 15 GeV mass the original analysis selection was reproduced. The comparison of results from the original analysis and this analysis are shown in Tables 7 and 8.

Mass	2 lep.	Opposite Lep. charge and lep. angle	lep1 ,lep2 p_T cut	lep same flavour	dilep mass cut	dilep p_T
4 GeV	100	99.56 ± 0.21	98.98 ± 0.21	98.98 ± 0.21	95.15 ± 0.21	90.23 ± 0.20
5 GeV	100	99.54 ± 0.21	98.93 ± 0.21	98.93 ± 0.21	95.06 ± 0.21	90.22 ± 0.20
15 GeV	100	99.54 ± 0.21	99.31 ± 0.21	99.31 ± 0.21	95.69 ± 0.21	90.10 ± 0.20

Table 7: Original hZa event selection efficiency. All values in % [3].

6.5.2 Diphoton selection

After Z boson candidate selection is completed and the dilepton vector is created the photon categorization starts for events with more than one photon. This selection matches the resolved selection from the original hZa analysis [3]. In every event the photons are sorted from highest to lowest p_T and then every possible diphoton vector is generated with

³The ΔR variable for a 2 particle system is defined as: $\Delta R = \sqrt{(\phi_1 - \phi_2)^2 + (\eta_1 - \eta_2)^2}$.

Mass	2 lep.	Opposite Lep. charge and lep. angle	lep1 ,lep2 p_T cut	lep same flavour	dilep mass cut	dilep p_T
4 GeV	100	99.51 ± 0.21	98.92 ± 0.21	98.92 ± 0.21	95.02 ± 0.21	90.29 ± 0.20
5 GeV	100	99.55 ± 0.21	98.93 ± 0.21	98.93 ± 0.21	94.94 ± 0.21	90.11 ± 0.20
15 GeV	100	99.52 ± 0.21	99.17 ± 0.21	99.17 ± 0.21	95.53 ± 0.21	90.02 ± 0.20

Table 8: hZa event selection efficiency from this analysis. All values in %

invariant mass $m_{\gamma\gamma}$ and transverse momentum $p_{T\gamma\gamma}$. These mass and momentum are then used to construct the X variable which originates from the kinematic relation for boosted objects $\Delta R_{\gamma\gamma} \approx \frac{2m_{\gamma\gamma}}{p_{T\gamma\gamma}}$. The definition of X is

$$X = \Delta R_{\gamma\gamma} \frac{p_{T\gamma\gamma}}{2m_{\gamma\gamma}} \quad (41)$$

Then the following cuts are applied:

- **Photon $p_T + \Delta R_{\gamma\gamma}$:** For every diphoton the photons should have $\Delta R_{\gamma\gamma} < 1.5$ and $p_T > 10$ GeV.
- **Diphoton selection:** The event includes a diphoton. If the event contains more than 2 photons the diphoton with X closest to one is selected.
- **X window:** For the selected diphoton $0.96 < X < 1.2$.
- **Four object vector:** The diphoton vector is combined with the dilepton vector to produce the dilepton+diphoton vector.
- **SR:** The dilepton+diphoton vector should have invariant mass $m_{ll\gamma\gamma}$ within the Higgs window, $110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$.
- **SR-ID:** Loose ID is applied for both photons.
- **SR-ID-ISO:** If the angular separation $\Delta R_{\gamma\gamma} > 0.22$, FCLoose isolation is applied. Otherwise no isolation is applied.

The photon selection is also described in Table 9.

The signal region (SR) in this analysis corresponds to the diphoton SR of [3] and it is the diphoton invariant mass distribution $m_{\gamma\gamma}$ within the Higgs window, $110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$, where photon ID and isolation have been applied (**SR-ID-ISO**). For the background estimation a control region (CR) is required referred as sideband (SB) region. It is the diphoton invariant mass distribution $m_{\gamma\gamma}$ in the region $80 \text{ GeV} < m_{ll\gamma\gamma} < 380 \text{ GeV}$ excluding the Higgs window of $110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$, where photon ID and isolation have been applied.

Requirements	hZa original analysis
Photon $p_T + \Delta R_{\gamma\gamma}$	$p_T > 10$ GeV and $\Delta R_{\gamma\gamma} < 1.5$, diphoton
Diphoton selection	$X = \Delta R_{\gamma\gamma} * p_{T\gamma\gamma}/2m_{\gamma\gamma}$, X closest to 1
X window	$0.96 < X < 1.20$
Four object vector	Dilepton + diphoton vector
SR	$110 < m_{ll\gamma\gamma} < 140$ GeV
SR-ID	Loose ID for both photons
SR-ID-ISO	$\Delta R_{\gamma\gamma} > 0.22$ FCLoose isolation

Table 9: Original hZa event selection.

The $m_{ll\gamma\gamma}$ window was chosen by an optimization described in [3]. The results of the diphoton selection from [3] are shown in the Table 10. The results of the diphoton selection from this analysis are shown in Table 11.

Mass	Pass preselection	photon $p_T + \Delta R_{\gamma\gamma}$ cut	X Window	SR	SR-ID	SR-ID-ISO
4 GeV	90.23 ± 0.20	14.88 ± 0.09	13.89 ± 0.08	13.38 ± 0.08	9.38 ± 0.07	5.80 ± 0.05
5 GeV	90.22 ± 0.20	14.99 ± 0.09	14.10 ± 0.08	13.58 ± 0.08	9.05 ± 0.07	6.06 ± 0.05
15 GeV	90.10 ± 0.20	14.37 ± 0.08	7.50 ± 0.06	7.03 ± 0.06	5.32 ± 0.05	3.40 ± 0.03

Table 10: Original hZa event selection efficiency- diphoton category. All values in %.

Mass	Pass preselection	photon $p_T + \Delta R_{\gamma\gamma}$ cut	X Window	SR	SR-ID	SR-ID-ISO
4 GeV	90.29 ± 0.20	14.81 ± 0.09	13.82 ± 0.08	13.44 ± 0.08	9.41 ± 0.07	5.84 ± 0.05
5 GeV	90.12 ± 0.20	15.09 ± 0.09	14.15 ± 0.08	13.67 ± 0.08	8.98 ± 0.07	6.01 ± 0.05
15 GeV	90.02 ± 0.20	14.29 ± 0.08	7.47 ± 0.06	7.08 ± 0.06	5.29 ± 0.05	3.38 ± 0.03

Table 11: Reinterpretation hZa event selection efficiency - diphoton category. All values in %.

Signal properties The diphoton invariant mass distribution $m_{\gamma\gamma}$ of the signal process $h \rightarrow Za \rightarrow ll\gamma\gamma$ is shown for various combinations of ALP mass and ALP-photon coupling $C_{\gamma\gamma}$ in Figures 24-28. This corresponds to the analysis SR and is used to illustrate the characteristics and sensitivity of the expected signal for different couplings. The results are normalized to the Higgs production cross section for the gluon-gluon fusion channel to align with the generated Monte Carlo (MC) samples. A bin width of 1 GeV is used for these plots, with all data derived from the simulated MC samples.

The simulated masses are in the range of 2 - 33 GeV (2,3,4,5,6,7,9,12,15,20 25,30,33)

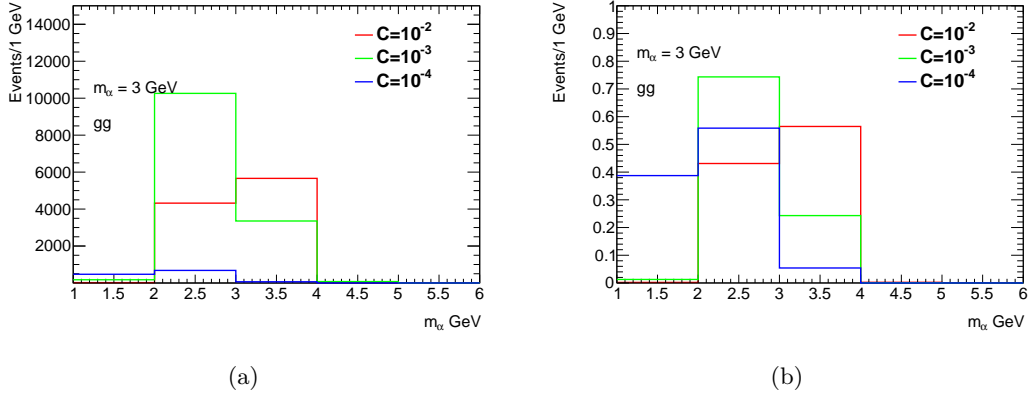


Figure 24: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 3$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to 1.

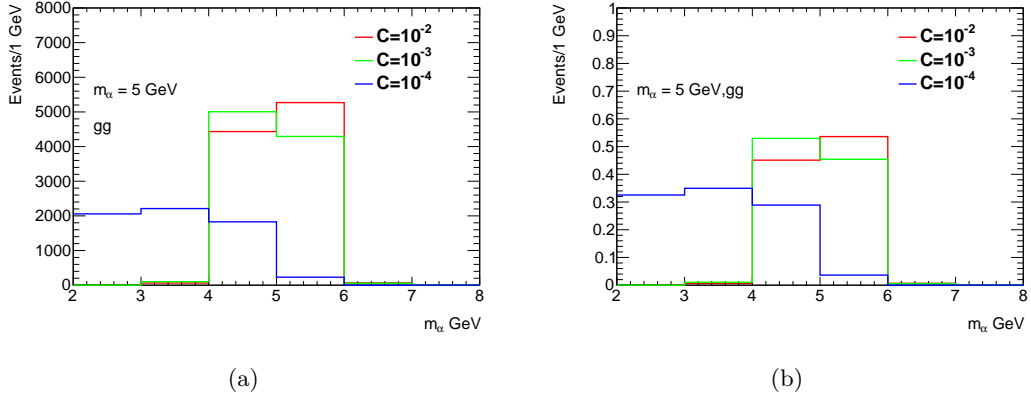


Figure 25: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 5$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to 1.

with photon-ALP couplings of $C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. This allows the limit estimation for couplings of this size for most of the parameter space of ALP masses. For low masses, $2 \leq m_{\gamma\gamma} \leq 15$ GeV, sensitivity decreases for the smaller couplings as ALPs become long-lived and decay outside the detector. For higher masses, $15 \leq m_{\gamma\gamma} \leq 33$ GeV, similar shapes to the prompt analysis were observed, within statistical uncertainty. The behaviour of $m_{\gamma\gamma}$ is consistent with expectations.

Study of broadening of displaced signatures mass peak In the low mass region, $2 \leq m_{\gamma\gamma} \leq 15$ GeV, and coupling $C_{\gamma\gamma} = 10^{-4}$ a phenomenon is observed where the ALP peak is smeared towards lower masses, as seen in Figures 24,25. A hypothesis has been

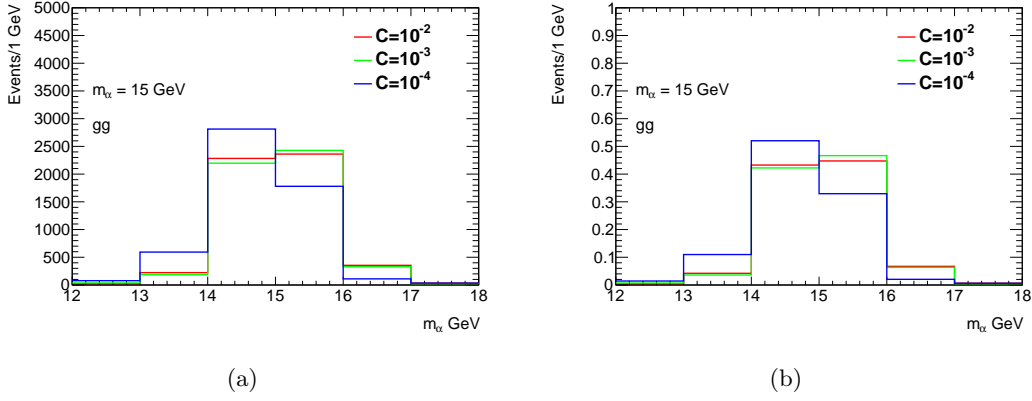


Figure 26: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 15$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to 1.

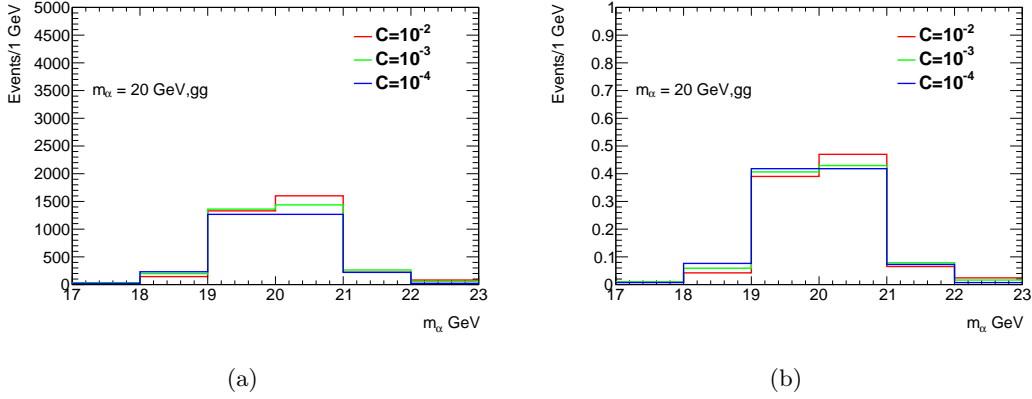


Figure 27: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 20$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to 1.

proposed that this effect is due to the shift in the decay vertex from prompt to displaced. Since all ATLAS photons are assumed to originate from the interaction point, the displaced decay vertex leads to a discrepancy in the photon momentum vectors, which is reflected in the broadening of the diphoton invariant mass. Figure 29 supports this hypothesis, as an increasing trend is observed between the mass shift and the decay length.

6.5.3 Background estimation

The main backgrounds of the hZa analysis are Z +jets and $Z\gamma$ processes. In the diphoton region the Z +jets background dominates with around 90% of the total events in the SR [3]. For Z +jets events, 80% of the reconstructed photons originate from π^0 decays, with just

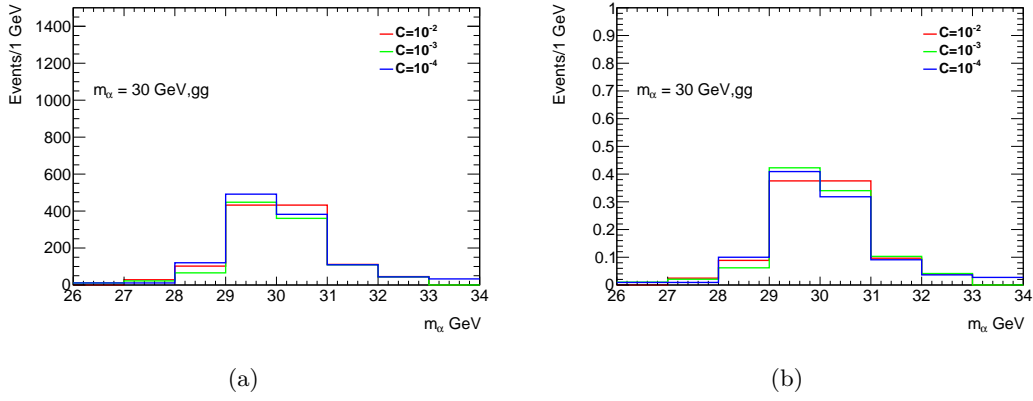


Figure 28: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 30$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. On the left plot the signal is normalized to the ggh cross section. On the right plot the signal of each coupling is normalized to 1.

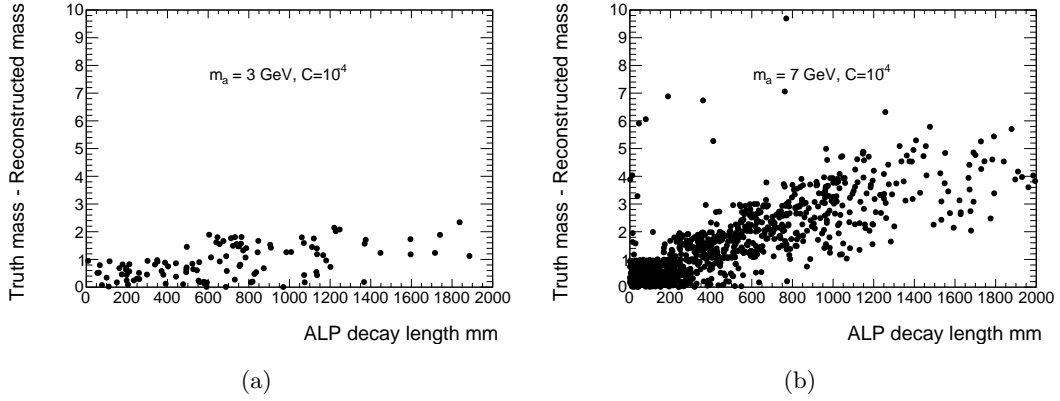


Figure 29: Plot of the absolute difference of the truth ALP mass with the reconstructed ALP mass compared to the decay length of the ALP for $m_a = 3$ GeV - $C = C_{\gamma\gamma} = 10^{-4}$ (left) and $m_a = 7$ GeV - $C = C_{\gamma\gamma} = 10^{-4}$ (right). Outliers in the right plot are attributed to ALPs where the wrong photon pair is selected.

under 20% arising from other light mesons. As a result, the majority of the background consists of jets that are misidentified as photon pairs. For $Z\gamma$ events the picture is more complex, as this background includes both of both resonant and non-resonant Z boson production plus photon radiation. In most cases, two photons have been radiated either from the initial quark interaction (ISR) or from the final state lepton pairs (FSR). FSR is strongly suppressed in the analysis due to the mass and transverse momentum cuts on the leptons, so the photons in this case are more likely to originate from different underlying charged particle interactions.

A point that should be made is that the two main backgrounds of the analysis are

not exactly orthogonal and sometimes Z+jets events may be the same as the $Z\gamma$ events leading to overlap [94]. There is a special tool in the ATLAS software network called VGammaORTool, built to discriminate between Z+jets and $Z\gamma$. This tool was used in the analysis to remove overlapping events.

Background parametrization The background parametrization used for this analysis is identical to [3], with just the necessary characteristics mentioned in this thesis. The SR is the invariant mass distribution $m_{\gamma\gamma}$ within the Higgs window, $110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$, where photon ID and isolation have been applied (**SR-ID-ISO**). A background parametrization has to be produced for this region to compare with data. The background $m_{\gamma\gamma}$ parametrization is extracted from data in the SB region, which is the region $80 \text{ GeV} < m_{ll\gamma\gamma} < 380 \text{ GeV}$ excluding the Higgs window of $110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$. For all photons in this region isolation and ID cuts have been applied [3].

The parametrization is performed on the falling part of the $m_{\gamma\gamma}$ distribution starting from 1.5 GeV, as collimated photons are dominating the lower diphoton masses. The best fit function from the SB fit was determined to be the sum of a Fermi-Dirac and an exponential, shown in Figure 30. The χ^2/ndf is 0.535 showing a good agreement between data and fit. The function is described as follows:

$$f^{bkg}(m; a, b, c, d) = \frac{a}{1 + e^{\frac{m_{\gamma\gamma} - b}{c}}} + (1 - a)e^{-\frac{m_{\gamma\gamma} - 20}{d}}, \quad (42)$$

where the first part describes the bulk of the distribution and the second part the smoothly falling tail. The function is used to predict the background on each mass bin of the SR.

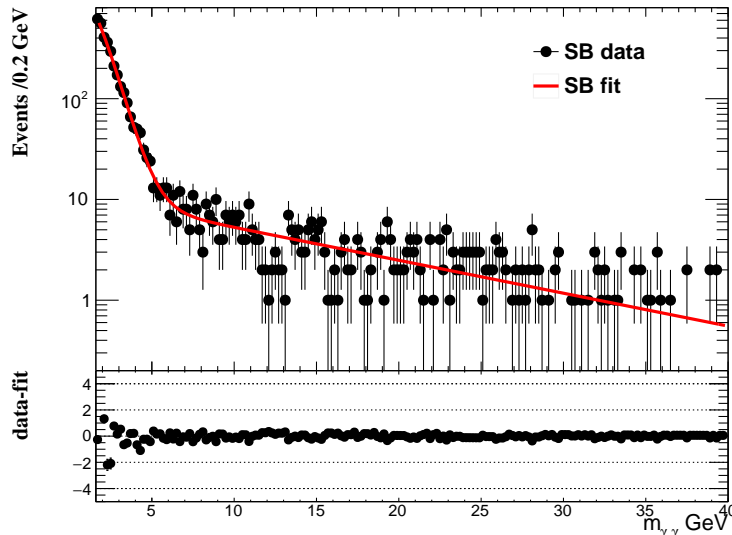


Figure 30: Data distribution $m_{\gamma\gamma}$ in the sideband region (SB), along with the fitted Fermi-Dirac function. The χ^2/ndf is 0.535 showing a good agreement between data and fit.

The final data and background $m_{\gamma\gamma}$ distributions in SR are shown in Figure 31, for 0.2 and 1 GeV bin widths respectively. The χ^2/ndf is 0.794 showing a good agreement between data and background parametrization.

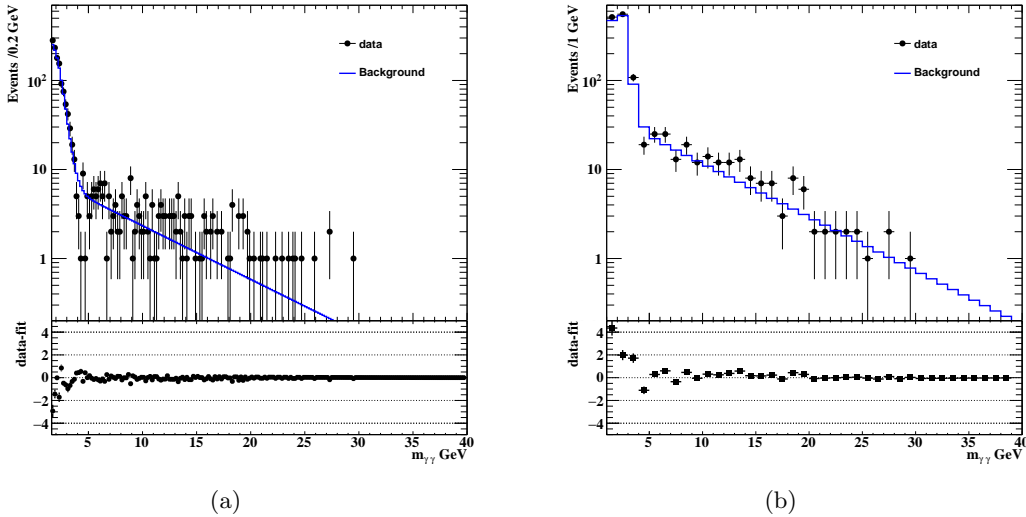


Figure 31: Background - Data $m_{\gamma\gamma}$ distributions in SR for 0.2 GeV bin width (left) and 1 GeV bin width (right). The χ^2/ndf for the Fermi-Dirac background parametrization is 0.794.

Spurious signal uncertainty The spurious signal is a non-zero signal amplitude obtained when fitting the sum of a background model plus a signal model to data that is known to be signal-free, i.e. a pure background distribution.

Following the prescription of [3], signal-free templates are obtained by fitting two alternative functions to observed data in the sideband region (SB), as shown in Figure 32. Asimov data in the signal region is obtained by scaling these templates to the corresponding number of observed events. The nominal fit procedure using the nominal background template described in Section 6.5.3 is then performed on this Asimov data to extract the signal yield for each ALP mass hypothesis. The resulting yield, called the spurious signal, N_s , is shown for the full mass and ALP-photon coupling range in Figure 33. In particular, the values for the promptly decaying samples are in very good agreement with those found in [3], and there is no significant difference as a result of the signal shape smearing arising from the increased displacement for smaller ALP-photon couplings. The same envelope function covering the results from the two alternative background functions is used for each coupling to provide a conservative estimate on the final spurious signal.

Following the description in reference [95], the spurious signal systematic is included in the likelihood fit, more details in section 6.5.5, for each mass and coupling as an additional

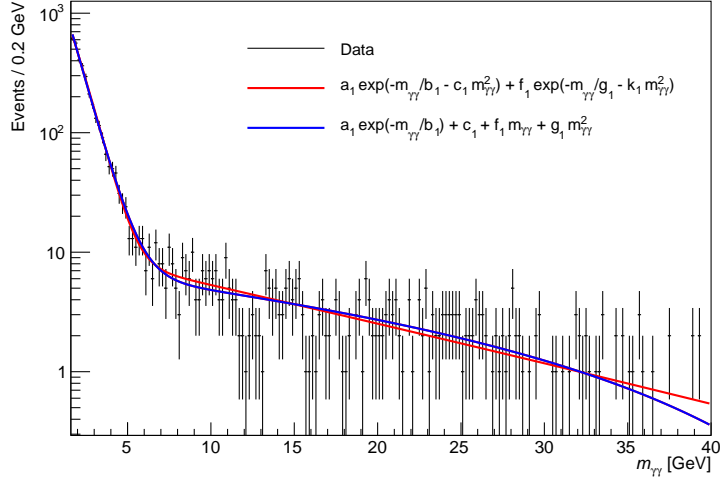


Figure 32: Data distribution in the sideband region (SB), along with fit results using two alternative background parametrizations.

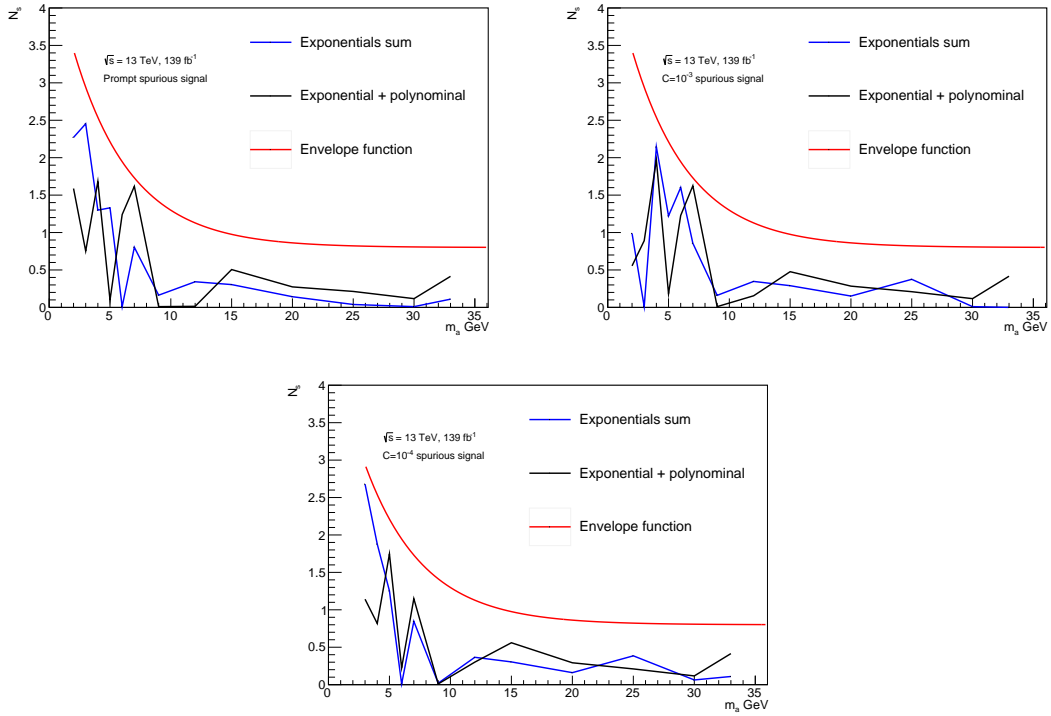


Figure 33: The spurious signal, N_s , calculated using the two alternative background functions over the ALP mass range for prompt (top left), $C = C_{\gamma\gamma} = 10^{-3}$ (top right), and $C = C_{\gamma\gamma} = 10^{-4}$ (bottom) signal samples. The red line shows the envelope used to quantify the final spurious signal uncertainty.

contribution consisting of the corresponding signal shape normalised to N_s , and controlled by a Gaussian-constrained nuisance parameter, μ_{ss} . As was the case for the prompt results in [3], the effect on the final branching ratio limits by including this systematic is found to be negligible for smaller displacements, and is on average around the percent level for larger displacements.

6.5.4 Systematic Uncertainties

Experimental particle physics involves the investigation of both statistical and systematic uncertainties. Statistical uncertainties arise from fluctuations in the data or MC samples. Systematic uncertainties, on the other hand, can originate from a variety of sources and are categorized into two types: theoretical and experimental.

Experimental systematic uncertainties stem from limited or incomplete knowledge of various aspects of the experimental setup. This includes factors such as the efficiency of the detector components, as well as energy calibration and resolution. Theoretical uncertainties, meanwhile, arise from the physics models used for the measured process, as well as potential background processes and cross sections.

To evaluate the systematic uncertainties in this analysis, the recommendations provided by the ATLAS Collaboration Combined Performance (CP) group are followed, specifically those from the egamma and muon subgroups. For all systematic uncertainties assessed using the ATLAS analysis tools, included in the Athena framework, both up and down variations are available. These variations shift the parameters used in the various reconstruction or correction tools by one standard deviation in either direction for each of the physics objects used. The use of certain working points for the analysis simplifies the process further, as weights are available for use, provided that the objects meet specific criteria. The influence of each systematic uncertainty variation is determined by calculating the difference between the weighted results (with up and down weights) and the results obtained using the nominal weights applied in the final selection cut, SR. The next section provides a comparison of each systematic uncertainty, with percentages relative to the nominal signal efficiency.

Experimental systematic uncertainties from CP recommendations

Pileup uncertainty An assumption is made a priori on the μ profile, which is the distribution of the average number of interactions per bunch crossing, in order to simulate the pileup interactions in the MC samples used in the analysis. This means that the μ distribution in the MC samples may not necessarily match the one observed in the data. A more accurate pileup description is achieved by re-weighting the MC μ distribution to match the distribution in data. Systematic uncertainty variations are provided with the re-weighting scale factors, as described in [64], and these are propagated through the analysis

resulting in a small effect on the event yield of all the signal samples, $\pm 2\text{-}4\%$, uncertainty depending on ALP mass.

Electron object uncertainties Electrons in ATLAS are subject to multiple uncertainties, which are calculated using scale factors. These scale factors correct the modeling of electrons in MC simulations relative to data, as well as the trigger efficiencies. One source of uncertainty arises from the electron energy calibration, for which scale factors are provided by the specialized tool `EgammaCalibrationAndSmearing`. Additionally, the `AsgElectronEfficiencyCorrection` tool, part of the ATLAS `egamma` suite, is used to measure electron identification efficiency in data. The simulated samples are then corrected by applying scale factors to each simulated event. These scale factors are parameterized as a function of the electron’s transverse momentum and rapidity. The uncertainties associated with these scale factors are treated as systematic uncertainties on the final signal selection efficiency. The sources of uncertainty include electron identification, electron isolation, electron reconstruction, and electron trigger efficiencies. In this specific analysis, since the electrons are produced by Z boson decays, they exhibit relatively small experimental uncertainties. The most significant electron systematic uncertainty is the electron identification efficiency, which has an uncertainty of less than $\pm 1.5\%$ across all signal samples.

Muon object uncertainties Muons in ATLAS are subject to multiple uncertainties. Since muons are much heavier than electrons and interact with different parts of the detector, their selection is performed using a specific ATLAS tool, the `MuonSelection` tool. The uncertainties in reconstruction efficiency and identification efficiency are addressed using scale factors provided by the ATLAS `MuonEfficiencyCorrections` tool. This tool also provides scale factors for the muon isolation efficiency uncertainty. The muon trigger efficiency uncertainty is handled by scale factors from the ATLAS `MuonTriggerScaleFactors` tool, which are matched to the muon trigger chain. The muon momentum scale uncertainty is addressed by scale factors from the ATLAS `MuonCalibrationAndSmearing` tool. The most significant muon systematic uncertainty is the muon reconstruction efficiency, which has an uncertainty of less than $\pm 0.5\%$ across all signal samples.

Photon identification and isolation efficiency uncertainties Photon identification and isolation efficiency systematic uncertainties are crucial for this analysis, as photon identification and isolation are key discriminators between signal and background. Uncertainties related to photon identification and isolation efficiencies are treated as systematic uncertainties, following the recommendations of the `egamma CP` group and using the ATLAS `PhotonEfficiencyCorrection` tool. This tool applies corrections to the calorimeter shower-shape variables and extracts identification scale factors as a function of the photon rapidity and transverse momentum [64]. Similarly, the isolation momenta of ID tracks

and the calorimeter isolation energy of photons in MC are adjusted to match the photon isolation efficiency measured in data [59]. These scale factors are applied to the MC to match the efficiency observed in data. The systematic variations of the scale factors are propagated through the analysis. The resulting uncertainty on the photon identification efficiency is at most $\pm 0.3\%$ across all signal samples, while the photon isolation efficiency uncertainty is at most $\pm 6.5\%$ across all signal samples.

Photon scale and resolution Uncertainties due to the photon energy scale and photon resolution are treated as systematic uncertainties, following the recommendations of the egamma CP group. The most simple decorrelation model, with one nuisance parameter each for energy scale and resolution, is used. The systematic variations of the photon energies are propagated through the analysis. It was found that the energy resolution/scale systematic uncertainty is negligible.

Luminosity The measurement of the integrated luminosity is performed by the ATLAS Luminosity Group. The uncertainty on the whole integrated luminosity for the complete ATLAS Run 2 dataset with $\mathcal{L} = 140.1 \text{ fb}^{-1}$ is 0.83%. This is derived using the ATLAS luminosity calibration procedure [52, 96].

Displaced vertices systematic uncertainty This systematic uncertainty focuses on the uncertainties of the detector response due to displaced photon origin vertices, up to 1.7 m in the x-y plane. The procedure used for the derivation of this systematic uncertainty is adopted from [92]. In the [92] $h \rightarrow aa \rightarrow 4\gamma$ analysis a novel data-driven approach was developed, to estimate systematic uncertainties of the modeling of calorimeter shower-shapes from photons with displaced production vertices. The main points of the approach are summarized in this section. For the hZa analysis the resulting systematic uncertainty histograms are used, shown in Figure 35.

The main challenge is that there is no way to identify whether photons are produced at a secondary vertex or at the primary vertex interaction point, as photons do not leave tracks in the ID. Long-lived baryons were used to quantify the difference in tracks between prompt and long-lived particles, particularly kaons, which travel a significant distance from the primary vertex before decaying into charged particles. The tracks of the decay products, originating from a displaced vertex, can be matched to reconstructed clusters in the electromagnetic calorimeter. In principle, this can be done in both data and MC, and the associated shower-shapes can then be compared. However, the correct description of the shower-shapes in MC depends not only on an accurate description of the detector response but also on the proper characterization of the particle types and energies of the tracks. To address this issue, three different regimes are tested: a near regime with tracks close to the primary vertex, a medium regime with slightly displaced tracks, and a far regime

containing all tracks up to a distance of 1.7 m from the primary vertex. The advantage of this approach is that if the near regime disagrees with the MC due to differences in the underlying nature of the charged particles, it can be used to reweight the results of the other two regimes, correcting any discrepancies.

After the weights in each regime are calculated, they are applied to the shower-shape variables. It is possible to determine both the mean and the RMS of each shower-shape variable in the medium and far regimes, for both data and MC. The mean and RMS for each regime are used to develop an interpolation for the shower-shape variables. The functions used for the interpolation are a second-order polynomial function $f(l)$ and a linear function $g(l)$, where l is the decay length. After interpolating the shower-shape variables, the identification of each photon must be recalculated. The full difference between the unmodified and modified shower-shape variables is then taken as the systematic uncertainty.

The change of the shower-shape variable, w_{tot} for Data and MC in the near, medium and far regimes is shown in Figure 34, where also the corrected MC distributions are displayed. The change of the mean-ratio as well as the RMS-ratio for the three different regimes is interpolated by the functions $f(l)$ and $g(l)$, which are shown in the lower center and lower right plots in Figure 34. The overall differences in the mean and the RMS are below 1%.

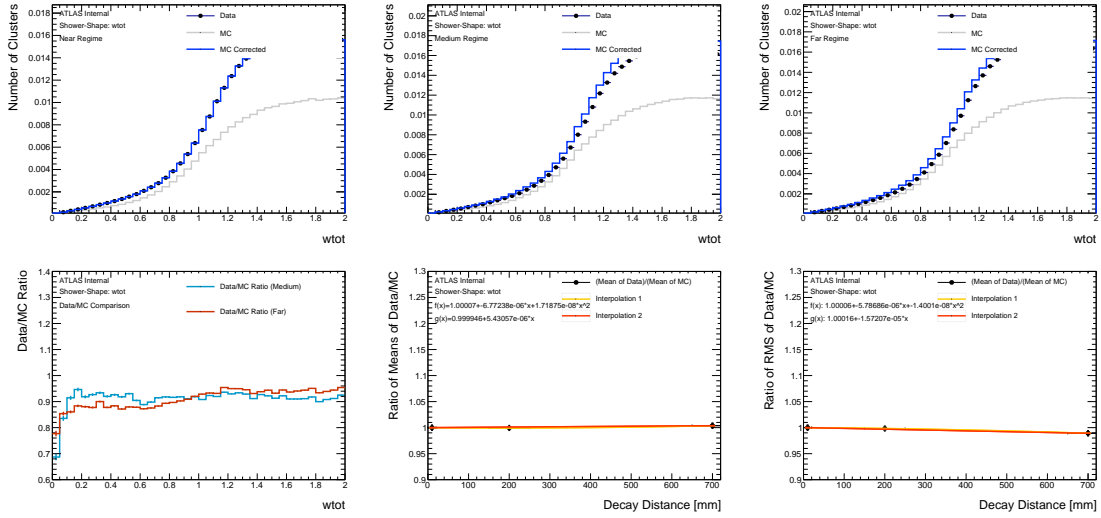


Figure 34: Shower-Shape Variable w_{tot} : Upper Row: Data and MC distribution of the shower-shape-variable for clusters which are associated to tracks originating in the near region (left), in the medium regime of the detector (middle) and the far-regime (right). Lower left: Ratio of Data to modified MC in the medium and far regime. Lower Middle: Dependence of mean of data MC vs distance of tracks; lower right: Dependence of rms of data MC vs distance of tracks [92].

Impact on the nominal ATLAS photon identification The efficiency of the photon identification at the loose working point is reevaluated w.r.t to the ALP decay radius l in the $r - \phi$ plane after fudging the shower-shapes as described below:

$$s_{new} = s_{orig} \cdot g_{mean}(l) + (s_{orig} \cdot g_{mean}(l) - \mu) \cdot slope_{RMS} \cdot l \quad (43)$$

where s_{new} and s_{orig} are the corrected and original shower-shape values, respectively. g_{mean} is the constant term of the $g(l)$ interpolation in the three regimes, $slope_{RMS}$ is the linear slope of the $g(l)$ interpolation in the three regimes, in respect to the decay length l while μ denotes the measured mean of the distribution. The observed difference between the nominal and fudged efficiencies is taken as the systematic uncertainty. The measured efficiencies and corresponding systematic uncertainty for the loose working point are shown in Fig. 35. The nominal efficiency decreases slightly with respect to the ALP decay radius, from 0.9 to 0.8. After fudging the shower-shapes, the efficiency becomes flat, remaining around 0.98. The resulting systematic uncertainty increases with the ALP decay length, rising from 0.07 to 0.23, with a maximum around $l_a = 1300$ mm.

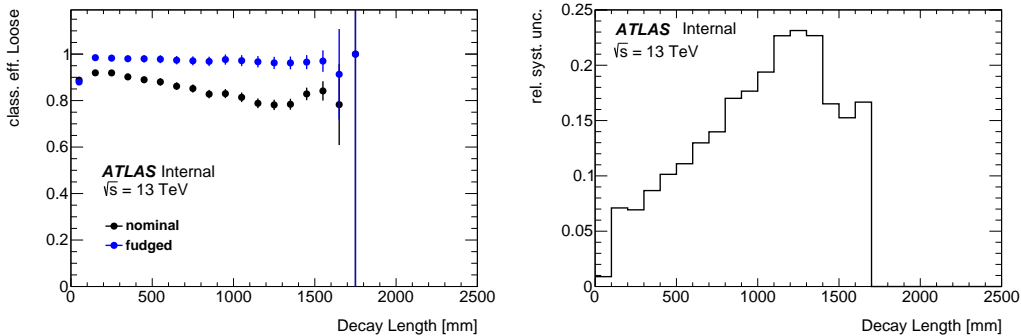


Figure 35: Systematic uncertainty of the loose photon identification working point due to the change of the photon shower-shapes with increasing ALP decay radius in the $(r - \phi)$ plane of the detector. Left: the identification efficiency before and after fudging the shower-shapes. Right: The relative difference of the efficiencies which is taken as systematic uncertainty. The difference rises with respect to the ALP decay radius and lies between 0.07 to 0.23 [92].

Then the results of the right plot of Figure 35 are used to reweight all the photons depending on the ALP decay length for all the individual ALP samples with coupling $C_{\gamma\gamma} < 10^{-2}$. The systematic uncertainty due to displaced photon origin is then calculated by comparing the diphoton invariant mass distributions, in the final analysis cut, with and without weights. This systematic uncertainty can be significant for highly displaced samples, reaching up to 33%. Two of these calculations are shown in Figure 36. Detailed Tables of the displaced systematic are shown in 12 and 13.

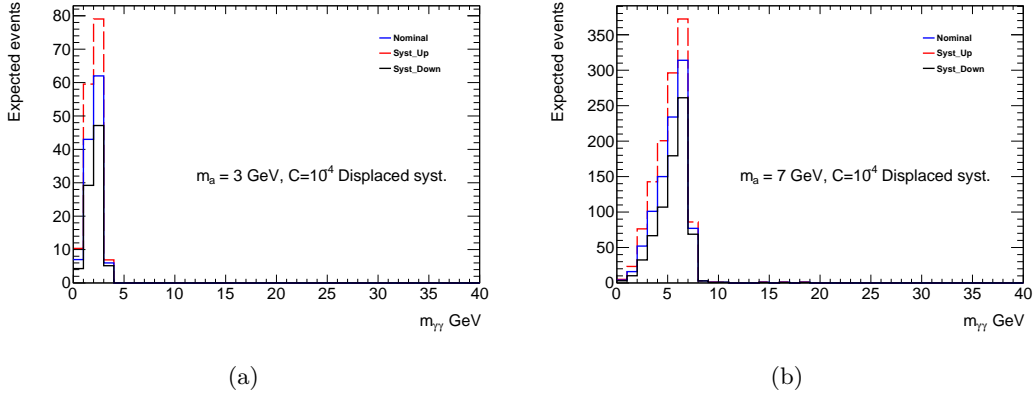


Figure 36: Systematic uncertainty due to displaced photon origin variations comparison for two different masses with $C = C_{\gamma\gamma} = 10^{-4}$, left is $m_a = 3$ GeV and right is $m_a = 7$ GeV. Nominal corresponds to no weights applied, Syst_Up to weights added and Syst_Down to weights subtracted.

Systematic uncertainty due to displaced photon origin - Up. Values in %					
$C_{\gamma\gamma}$	Signal $m_a = 2$ GeV	Signal $m_a = 3$ GeV	Signal $m_a = 4$ GeV	Signal $m_a = 5$ GeV	Signal $m_a = 7$ GeV
10^{-3}	23.71	16.02	9.23	4.44	0.33
10^{-4}	-	32.07	33.23	32.78	26.75
$C_{\gamma\gamma}$	Signal $m_a = 9$ GeV	Signal $m_a = 15$ GeV	Signal $m_a = 20$ GeV	Signal $m_a = 25$ GeV	Signal $m_a = 30$ GeV
10^{-3}	0.00	0.00	0.00	0.00	0.00
10^{-4}	19.11	6.58	1.78	0.17	0.11

Table 12: Displaced systematic uncertainty - Up variation, values in %.

Systematic uncertainty due to displaced photon origin - Down. Values in %					
$C_{\gamma\gamma}$	Signal $m_a = 2$ GeV	Signal $m_a = 3$ GeV	Signal $m_a = 4$ GeV	Signal $m_a = 5$ GeV	Signal $m_a = 7$ GeV
10^{-3}	20.79	14.51	8.48	4.10	0.30
10^{-4}	-	27.26	28.06	27.77	23.08
$C_{\gamma\gamma}$	Signal $m_a = 9$ GeV	Signal $m_a = 15$ GeV	Signal $m_a = 20$ GeV	Signal $m_a = 25$ GeV	Signal $m_a = 30$ GeV
10^{-3}	0.00	0.00	0.00	0.00	0.00
10^{-4}	16.99	6.05	1.65	0.15	0.12

Table 13: Displaced systematic uncertainty - Down variation, values in %.

6.5.5 Results

Fit strategy description To obtain an upper limit on the branching ratio for $h \rightarrow Za$, compared to the Higgs ggh production mode, a profile likelihood fit is performed on the invariant photon mass spectrum for each mass-coupling pair. The fit inputs include the signal, background estimation, real data, and systematic uncertainties. The ALP signal for each mass and coupling corresponds to the values from the respective MC in the SR and is normalized to the Higgs ggh cross section in order to produce a limit on the branching ratio. Some of the fitted ALP signal distributions are shown in Figures 24-28. The background estimation and real data correspond to the values shown in Figure 31 for each respective ALP mass. Systematic uncertainties are included in the fit for each mass-coupling pair. We distinguish systematic uncertainties by shape differences as well as normalization uncertainties. Some systematic uncertainties yield no significant shape difference but affect the signal normalization. In these cases, only the normalization uncertainty is treated in the fit, while the shape difference is neglected. The fitting procedure is performed using the Python Histfactory (pyhf) framework [97].

First, a freely floating signal parameter of interest (POI), denoted as μ , is fitted, and the most likely value for it (referred to as the best-fit) is calculated, with systematic uncertainties being constrained in the process. A pull refers to the shift of the nuisance parameter, such as the variable that describes each systematic uncertainty. The pulls of those constrained uncertainties are shown as an example for one mass-coupling point in Figure 37. To calculate upper limits on the production cross-section, the CLS method is used. The confidence level CLS is computed for several fixed values of μ . The upper limit is defined as an exclusion of signal at 95% confidence level, which is depicted in Figure 38 as an intersection with the red horizontal line at $CLS = 0.05$ for observed (solid line) values and expected from simulation (dashed line) values with $\pm 1\sigma$ (green) and $\pm 2\sigma$ (yellow). Finally, the limits are validated with a toy fit.

Limits results The fitting procedure described in the previous section yields a limit on the BR of the $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ process at 95% C.L. The following limits are produced for the three ALP-photon couplings $C_{\gamma\gamma} = 10^{-2}$ (prompt), $C_{\gamma\gamma} = 10^{-3}$ and $C_{\gamma\gamma} = 10^{-4}$, shown on Figures 39, 40 and 41 respectively.

The limit on the branching ratio can be converted into a limit on the coupling constant of the ALPs to photons, $C_{\gamma\gamma}$, following section 3.2.3. The expected branching ratio can be calculated as detailed in [26], and depends on the $C_{\gamma\gamma}$. We vary $C_{\gamma\gamma}$ in the calculation until the expected branching ratio matches the observed limit on the BR, which yields the corresponding limit on $C_{\gamma\gamma}$:

$$BR_{theo} = \Gamma_{hZa} \frac{BR_{a\gamma\gamma}}{\Gamma_h + \Gamma_{hZa}} \quad (44)$$

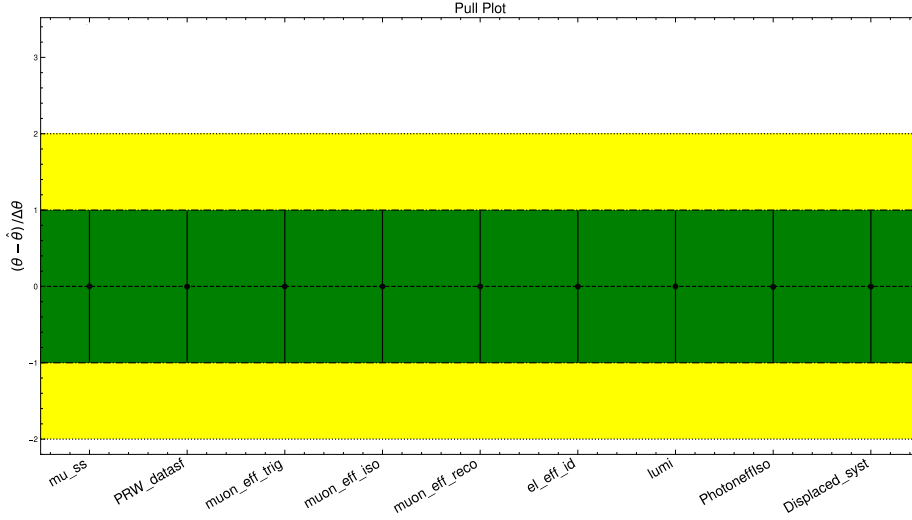


Figure 37: Pulls for the nuisance parameters in the fit for $m_a = 7$ GeV with $C_{\gamma\gamma} = 10^{-4}$.

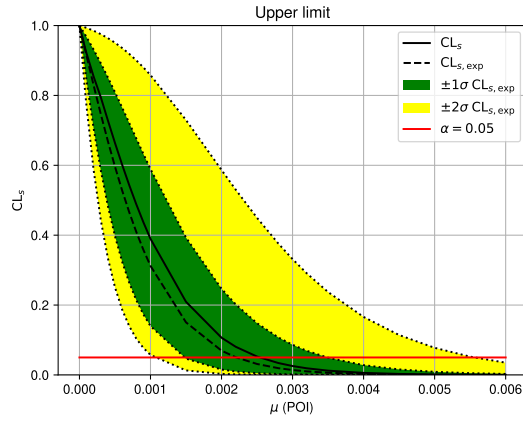


Figure 38: Observed (solid line) and expected (dashed line) CLS values for different μ with $\pm 1\sigma$ (green) and $\pm 2\sigma$ (yellow) for a fit with statistical variations and systematics applied for $m_a = 7$ GeV with $C_{\gamma\gamma} = 10^{-4}$.

$$\Gamma_{hZa} = \frac{m_h^3}{16\pi\Lambda^2} |C_{Zh}|^2 \lambda^{3/2} \left(\frac{m_Z^2}{m_h^2}, \frac{m_a^2}{m_h^2} \right) \quad (45)$$

$$\lambda(x, y) = (1 - x - y)^2 - 4xy \quad (46)$$

$$\gamma_a = \frac{m_h^2 - m_Z^2 + m_a^2}{2m_h m_a} \quad (47)$$

$$\Gamma_{a\gamma\gamma} = \frac{4\pi\alpha^2 m_a^3}{\Lambda^2} |C_{\gamma\gamma}|^2 \quad (48)$$

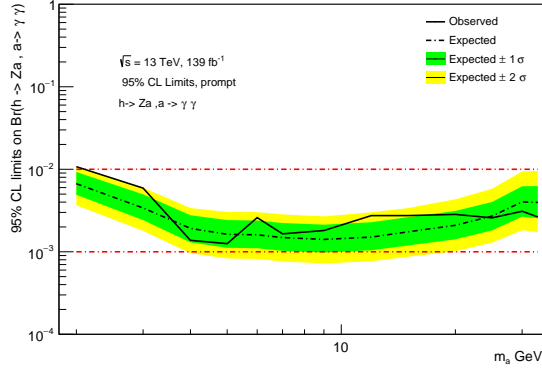


Figure 39: Upper limits, at the 95% C.L. on the branching ratio of the process $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ for $C_{\gamma\gamma} = 10^{-2}$ (prompt).

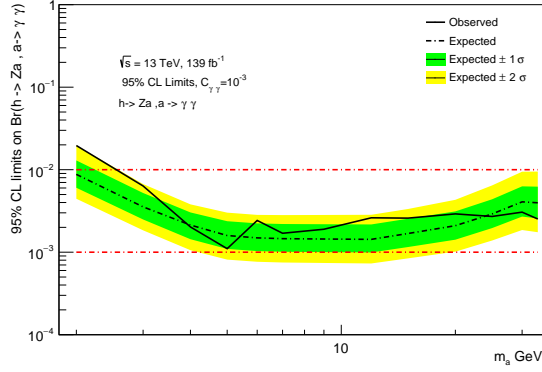


Figure 40: Upper limits, at the 95% C.L. on the branching ratio of the process $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ for $C_{\gamma\gamma} = 10^{-3}$.

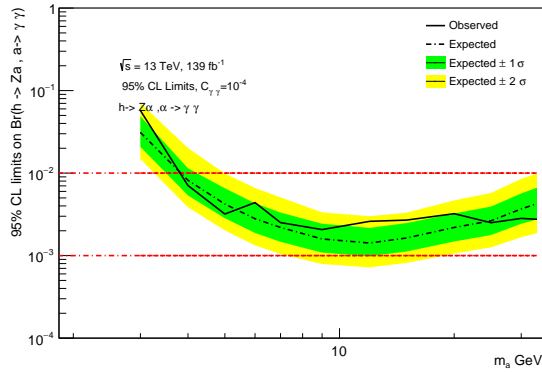


Figure 41: Upper limits, at the 95% C.L. on the branching ratio of the process $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ for $C_{\gamma\gamma} = 10^{-4}$.

The symbol definitions for equations 44 - 48 are the following. $\Gamma_h = 4.1$ MeV is the Higgs SM width. Γ_{hZa} is the measured $H \rightarrow Za \rightarrow l^+l^- \gamma\gamma$ width, derived from the branching ratio multiplied with $BR(Z \rightarrow l^+l^-) = 0.101$. $BR_{a\gamma\gamma}$ is the branching ratio of $a \rightarrow \gamma\gamma$, and it is equal to one, as all ALPs decay to photons in this hypothesis. Λ is the new physics scale, set at 1 TeV. C_{Zh} is the coupling of $H \rightarrow Za$. The masses of the particles are $m_h = 125$ GeV, $m_Z = 91.187$ GeV and the ALP mass corresponding to each branching ratio. Finally γ_a is the boost factor of the ALP and $\Gamma_{a\gamma\gamma}$ is the ALP width, which is calculated as the coupling $C_{\gamma\gamma}$ is varied, with $\alpha = 1/128$ the Electromagnetic coupling constant at the energy of the Z boson. The limits in the ALP-photon coupling with this method are shown in Figure 42.

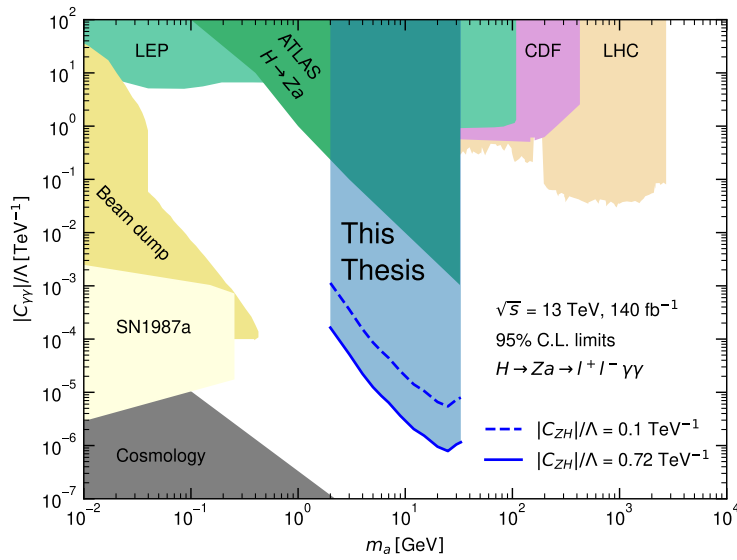


Figure 42: Limits on the ALP mass and coupling to photons at 95% C.L. The results from this analysis are shown as blue lines assuming a Higgs to ALP coupling of $C_{Zh}/\Lambda = 0.72$ TeV^{-1} (solid) and $C_{Zh}/\Lambda = 0.1$ TeV^{-1} (dashed). This plot is produced by varying the $C_{\gamma\gamma}$ coupling in the calculation until the expected branching ratio matches the observed limit on the theoretical branching ratio.

For the hZa reinterpretation analysis a second method was also used to produce limits on the ALP mass and coupling to photons. This is done for the different simulated values of the ALP coupling to photons $C_{\gamma\gamma}/\Lambda$, resulting in a grid of upper limits on C_{Zh}/Λ in the $(m_a, C_{\gamma\gamma})$ plane. Different C_{Zh} contours may then be extracted by interpolation. To reduce uncertainties on the interpolation, additional signal points are generated between the MC grid points using a lifetime reweighting approach. The resulting contour plot is shown in Figure 43. Both methods agree to within an order of magnitude. This method was ultimately used in the publication.

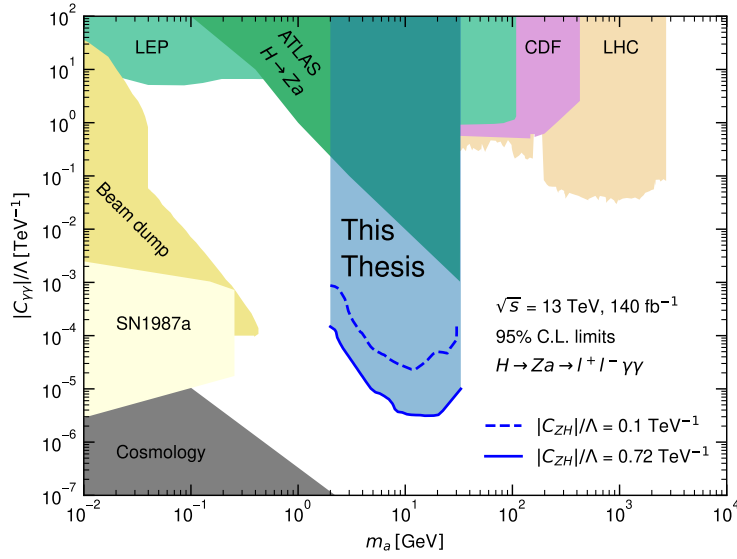


Figure 43: Limits on the ALP mass and coupling to photons at 95% C.L. The results from this analysis are shown as blue lines assuming a Higgs to ALP coupling $C_{Zh}/\Lambda = 0.72 \text{ TeV}^{-1}$ and $C_{Zh}/\Lambda = 0.1 \text{ TeV}^{-1}$ (dashed). This plot is produced using simulated values of the ALP coupling to photons produced by signal MC samples.

6.6 Optimization of the existing hZa selection

6.6.1 Introduction

As evident by the prompt limits plot, shown in Figure 39, and also the diphoton invariant mass distributions plots, shown in Figures 24-28, the hZa ATLAS analysis shows relatively low sensitivity for ALP masses larger than 15 GeV compared to the (2-15 GeV) mass range. This is also supported by a comparison with the CMS Run 2 hZa analysis [77] where the sensitivity and limits are around one order of magnitude better than ATLAS, while they should overall be comparable. This discrepancy prompted a re-evaluation of the analysis event selection to explore avenues for improved results. The following pages present the efforts to achieve results improvement using a remarkably simple selection strategy.

6.6.2 Event selection

The object selection and Z boson reconstruction follow the same procedure as in the original hZa analysis. The lepton objects are fairly standard throughout electroweak analyses with a transverse momentum (p_T) threshold of 27 GeV applied to match the trigger requirement on the leading lepton. Details on the Z selection have already been discussed on 6.5.1.

The primary difference with the hZa ATLAS analysis arises in the treatment of photons. The recommended photon objects by the egamma group are photons with $p_T > 10 \text{ GeV}$, $|\eta| < 2.37$ excluding calorimeter crack, at least passing the loose ID working point and

some form of isolation. In this case tracked based isolation ($p_{T\text{cone}}/p_T < 0.05$) is used, a variation of the FCLoose working point. These are the photon objects used in this analysis, which inherently provide substantial background rejection. Thus the photon selection is as follows:

- **2 loose photons:** 2 loose photons with $p_T > 10$ GeV, $|\eta| < 2.37$ and tracked based isolation ($p_{T\text{cone}}/p_T < 0.05$). For events with more than 2 loose photons, the two with the highest p_T are chosen, creating the diphoton object.
- **Low mass selection :** If the reconstructed invariant diphoton mass $m_{\gamma\gamma} \leq 2.5$ GeV, the event is selected. More details in section 6.6.2.
- **High mass selection:** If the reconstructed invariant diphoton mass $m_{\gamma\gamma} > 2.5$ GeV, tight photon ID is applied. If both photons pass, the event is selected. More details in section 6.6.2.
- **Higgs window :** In both cases the dilepton+diphoton vector should have invariant mass $m_{ll\gamma\gamma}$ in the Higgs window, $110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$.
- **Sideband region :** For background estimation a sideband region is made with the dilepton+diphoton vector with invariant mass $m_{ll\gamma\gamma}$ in the range $100 \text{ GeV} < m_{ll\gamma\gamma} < 110 \text{ GeV}$ and $140 \text{ GeV} < m_{ll\gamma\gamma} < 180 \text{ GeV}$.

The optimized hZa selection photon cuts are shown in Table 14.

Optimized analysis selection photon cuts	
Photon selection	2 loose photons, $p_T > 10$ GeV, $ \eta < 2.37$, track isolation
Low mass	$m_{\gamma\gamma} \leq 2.5$ GeV
High mass	$m_{\gamma\gamma} > 2.5$ GeV, 2 tight photons
Higgs window	$110 \text{ GeV} < m_{ll\gamma\gamma} < 140 \text{ GeV}$ (Signal region)

Table 14: Optimized hZa selection.

A comparison between the two selections, original and optimized, is shown in Table 15. The photon selection process in the hZa original analysis [3], as described in 6.5.2, differs significantly from the approach used in this analysis. In the original analysis, photon identification and isolation are applied later in the selection process, after the Higgs window cut, which can lead to spurious diphoton selections. Additionally, the $\Delta R_{\gamma\gamma}$ and X variable cuts, employed in the original analysis, are not used in the optimized selection. Finally, the photon isolation working points are different between the two analyses with the original hZa analysis only applying FCLoose isolation when the opening angle of the diphoton is $\Delta R_{\gamma\gamma} > 0.22$. In contrast, the optimized analysis adopts a more straightforward event selection strategy, avoiding such specific conditions.

Original hZa	Optimized hZa
$p_T + \Delta R_{\gamma\gamma}$: $p_T > 10$ GeV & $\Delta R_{\gamma\gamma} < 1.5$	Photons: 2 loose ID, $p_T > 10$ GeV, track iso.
diph.: $X = \Delta R_{\gamma\gamma} * p_{T\gamma\gamma}/2m_{\gamma\gamma}$, $X \approx 1$	Low mass : $m_{\gamma\gamma} \leq 2.5$ GeV
X window : $0.96 < X < 1.20$	High mass : $m_{\gamma\gamma} > 2.5$ GeV, tight ID
Four object vector : Dilepton + diphoton	Higgs Window : $110 < m_{l\gamma\gamma} < 140$ GeV
SR : $110 < m_{l\gamma\gamma} < 140$ GeV	
SR-ID : Loose ID for both photons	
SR-ID-ISO : $\Delta R_{\gamma\gamma} > 0.22$ FCLoose iso	

Table 15: hZa selections comparison-photon requirements.

Throughout the analysis the invariant mass distributions in three regions will be used. These are the signal region (SR), sideband region (SB) and validation region (VR). The definition of each invariant mass distribution is shown in Table 16.

Distribution	Definition
SR (Signal region)	$m_{\gamma\gamma}$ in range $m_{l\gamma\gamma}$ (110,140) GeV
SB (Sideband region)	$m_{\gamma\gamma}$ in range $m_{l\gamma\gamma}$ (100,110)&(140,180) GeV.
VR (Validation region)	$m_{l\gamma\gamma}$ in range (100,110)&(140,180) GeV.

Table 16: Definitions of distributions in optimized hZa selection.

Low mass selection motivation From the MC samples for ALP mass of 1 GeV and 2 GeV, it was observed that their behavior differs slightly from that of higher mass samples. Specifically, the opening angle between the two highest p_T photons on truth level is smaller than $\Delta R_{\gamma\gamma} < 0.2$. This leads to candidate events failing the tight photon ID criteria, as the energy clusters produced by these photons are in close proximity in the calorimeter and the shower-shapes overlap, as illustrated in Figure 44. To address this, a low-mass selection was introduced, applying only loose photon identification. The definitions of the signal region (SR) and sideband region (SB) in the diphoton + dilepton invariant mass are standard between the low mass and high mass cases, as shown in Table 16.

High mass selection motivation For MC samples of ALPs with mass greater than 2 GeV, consistent behavior was observed regarding the opening angle between photons. The angle is sufficiently large to ensure that the energy clusters produced by the photons are well-separated, preventing failure of the tight identification criteria. In the analysis, tight identification is preferred over loose identification as it improves rejection of the dominant background, Z+jets, as discussed in [59]. Consequently, the cut on the diphoton invariant mass is set at 2.5 GeV. The definitions of the signal region (SR) and sideband region (SB)

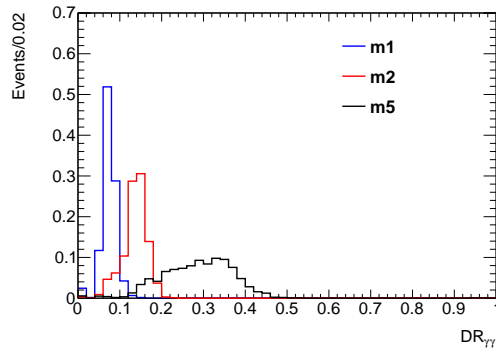


Figure 44: Opening angle $DR_{\gamma\gamma}$ of diphoton pairs for 1 GeV, 2 GeV and 5 GeV ALPs. Histograms normalized to unity.

in the diphoton + dilepton invariant mass remain standard for both low mass and high mass cases, as shown in Table 16.

MC diphoton invariant mass distributions The determination of the signal diphoton invariant mass distributions $m_{\gamma\gamma}$ is derived from MC samples and it is important for the optimization of the event selection. In particular the window in which the axion mass hypothesis is defined is set to have at least 95% of the reconstructed ALP events in the SR. The resulting ALP mass windows are shown in 17. The MC diphoton signal shapes of the

Mass (GeV)	Mass window (GeV)	Mass (GeV)	Mass window (GeV)
1	(0,2)	12	(11,13)
2	(1,3)	15	(13,17)
4	(3,5)	20	(18,22)
5	(4,6)	25	(23,27)
6	(5,7)	30	(28,32)
9	(8,10)	33	(31,35)

Table 17: ALP mass and mass window calculated from the MC of the signal sample.

invariant mass distributions for a range of masses and couplings is shown in Figures 46-50. The results are normalized to the Higgs production cross section for the gluon-gluon fusion channel, as only MC samples for this production mode were generated. A bin width of 1 GeV is used for these plots.

The diphoton signal region efficiencies of the original analysis are presented in Figures 24-28. A comparison can be made between the original ATLAS analysis and this analysis in regards to signal efficiencies, shown in Figure 45. The original ATLAS selection demonstrates higher efficiency for lower ALP masses in the range of 2–9 GeV, while the optimized selection performs better for higher ALP masses. Notably, the optimized analysis selection

efficiency increases with ALP mass, reaching a maximum of approximately 6% at $m_a = 33$ GeV, compared to only about 0.4% for the original selection. However, it is important to note that an improvement in signal selection efficiency does not necessarily translate to improved sensitivity because of the background estimation.

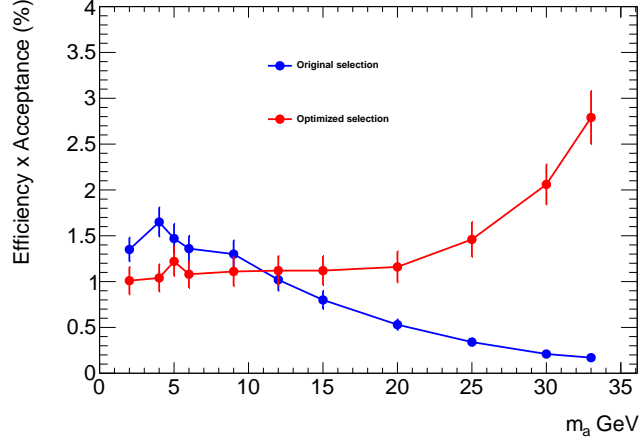


Figure 45: Selections signal efficiency \times acceptance, original (red) and optimized (blue). Acceptance ≈ 0.48 .

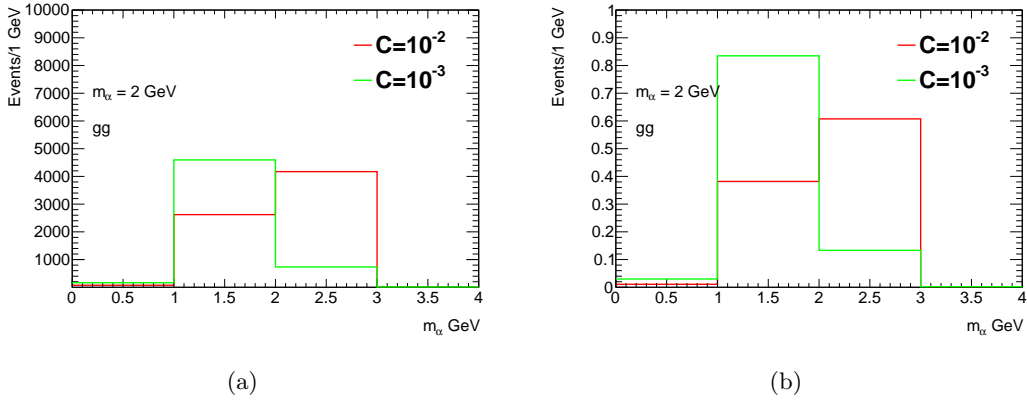


Figure 46: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 2$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}$. On the left plot the signal is normalized to the ggh cross section. On the right plot the signal of each coupling is normalized to unity.

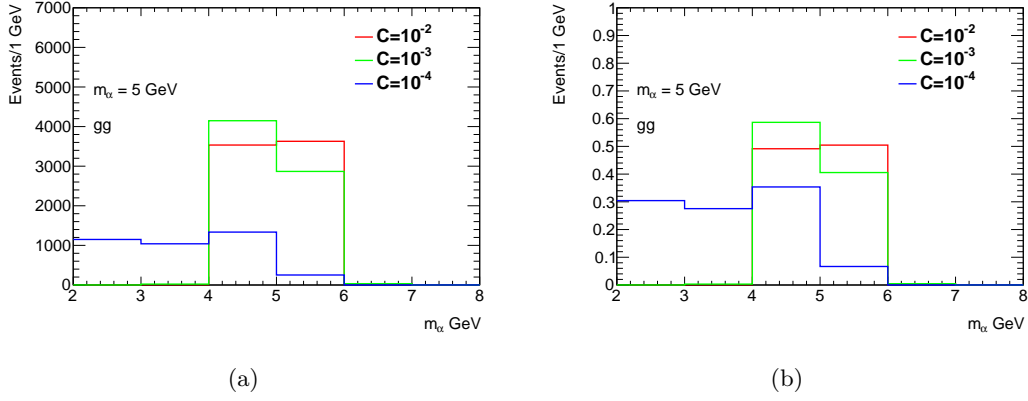


Figure 47: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 5$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to unity.

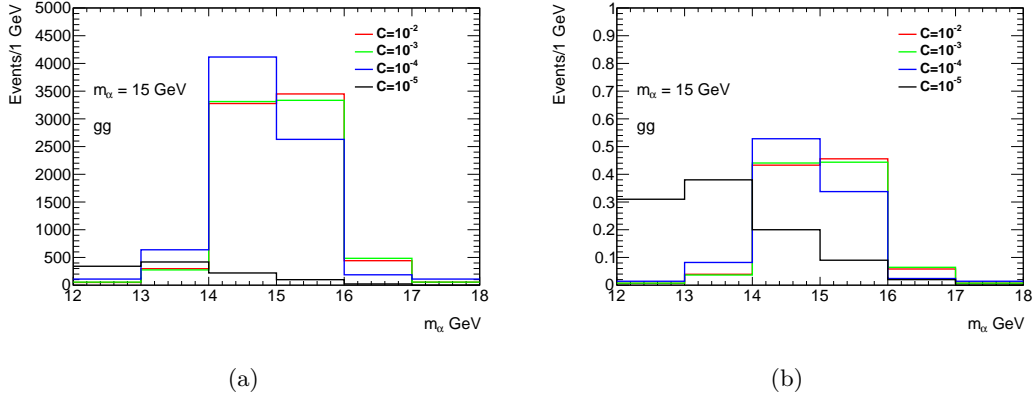


Figure 48: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 15$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to unity.

6.6.3 Background estimation

The main backgrounds of the hZa analysis are Z+jets and $Z\gamma$. The main conclusions regarding these backgrounds in the original hZa analysis remain valid. Background estimation is performed using a combination of MC simulations and sideband plots, as described in Section 6.6.2, to minimize the impact of MC uncertainties. Specifically the side band (SB) plot for the diphoton invariant mass $m_{\gamma\gamma}$ distribution, defined in Table 16, is shown in Figure 51. From this plot, the use of MC for background estimation in the low-mass region is validated after applying a correction factor of $n = 1.2120$. This factor is derived by comparing the (0–2.5) GeV mass range in the sideband (SB) histograms of MC and data,

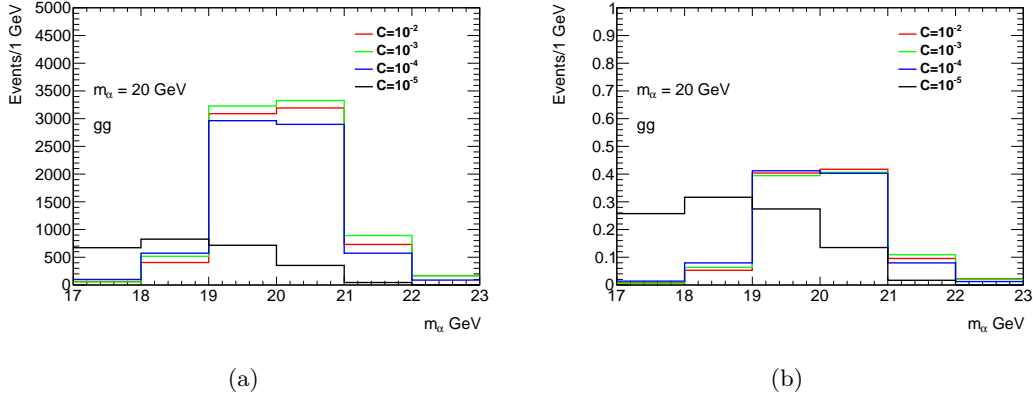


Figure 49: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 20$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to unity.

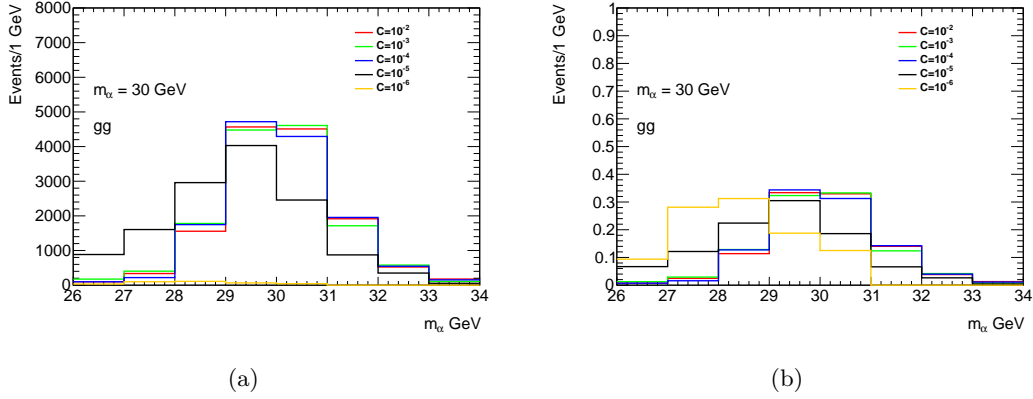


Figure 50: Signal distributions for the reconstructed $m_{\gamma\gamma}$ for $m_a = 30$ GeV for the couplings $C = C_{\gamma\gamma} = 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$. On the left plot the signal is normalized to the gg cross section. On the right plot the signal of each coupling is normalized to unity.

where the shape agreement in the SB region was found to be satisfactory. Consequently, the corrected MC background estimation for the low-mass region is shown in Figure 52. An additional background validation is performed using the invariant mass of the diphoton + dilepton in the validation region (VR), also defined in Table 16, as shown in Figure 53. This comparison demonstrates a roughly similar shape between data and MC, further supporting the reliability of the background estimation. The validation is particularly relevant because low-mass events dominate the VR, as illustrated in Figure 51.

High mass background estimation Since MC predictions are not reliable in the high-mass region, an alternative method for background estimation is necessary, leveraging the

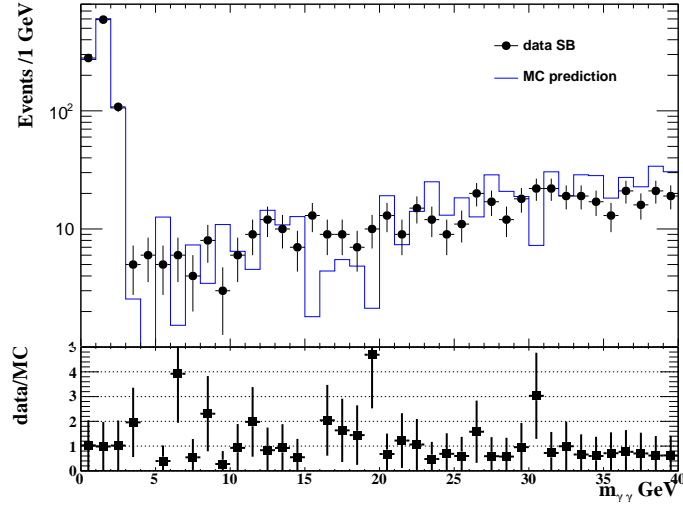


Figure 51: SB Diphoton invariant mass distribution $m_{\gamma\gamma}$ plot in region (0,40) GeV. MC is in blue and data is black.

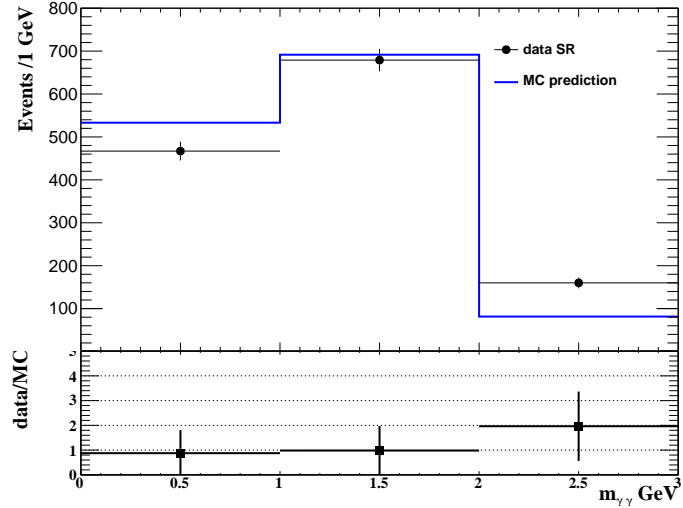


Figure 52: Diphoton invariant mass distribution $m_{\gamma\gamma}$ in region (0,3) GeV - low mass signal region showing the background estimation used. MC is in blue and data is black.

sideband (SB) and validation regions (VR). Three methods were tested for background estimation. For the first method a linear function is fitted to the SB distribution. A transfer function, derived from the VR distribution, is used to normalize the SB fit and map it into the signal region (SR), providing a background estimation. The second method involves fitting a function directly to the SR distribution while excluding the ALP mass windows (listed in Table 17). This method provides a background estimation for each mass point. Finally, the third method was the ABCD method which used both the SB and SR distributions to produce a background estimation. The key distribution plots for the

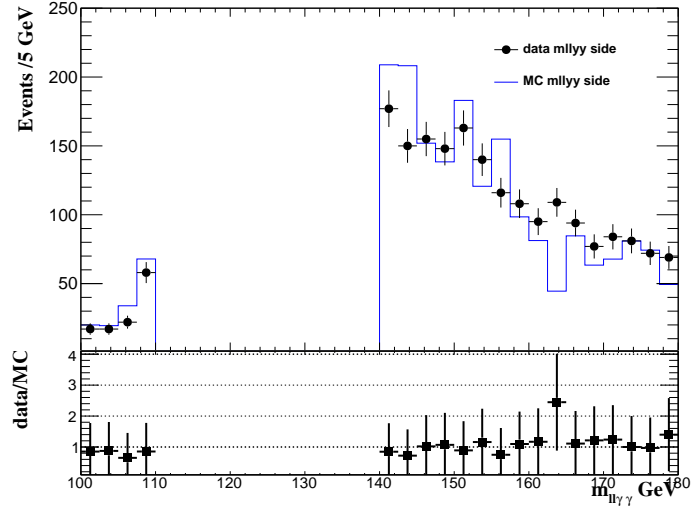


Figure 53: Diphoton+dilepton invariant mass distribution in the VR of mass range $m_{ll\gamma\gamma}$ (100,110) - (140,180) GeV. MC is in blue and data is in black.

background estimation are shown in Figures 53 and 54.

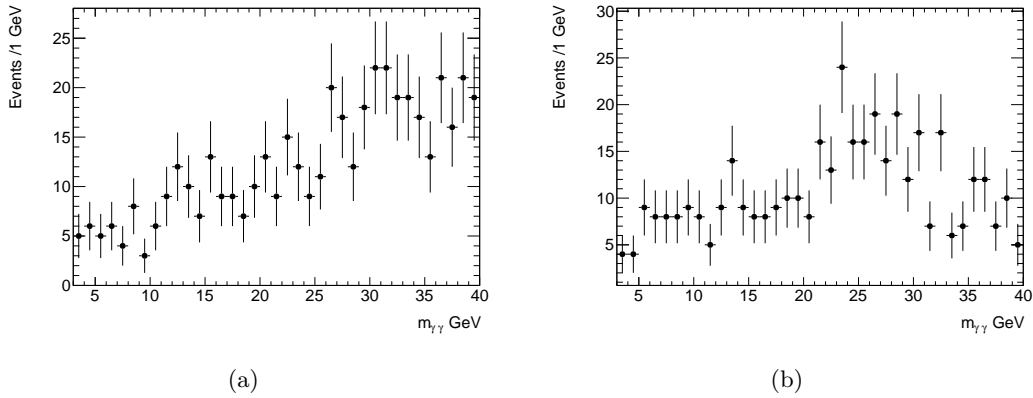


Figure 54: Diphoton invariant mass distribution in mass range (3,40) GeV $m_{\gamma\gamma}$. SB (left) and SR (right).

Method 1: Background estimation with SB distribution fit The first method for background estimation involves fitting a function to the SB distribution. This approach is based on the hypothesis that the sideband region does not contain any signal and exhibits kinematics very similar to those of the signal region. Consequently, the fitted function is expected to approximate the background distribution in the signal region, allowing any potential signal to be distinguishable. This method is commonly used in various ATLAS analyses where the MC description of the background is inadequate. Additionally, it mitigates potential background uncertainties, such as those arising from energy resolution or

normalization discrepancies, as both the SB and SR distributions are derived entirely from data. The main problem of this method is that the two regions should have a similar population of data and a transfer factor that maps the SB to the SR should be derived. The fit to the SB region is shown in Figure 55. The fitted functions used are a linear (pol1)

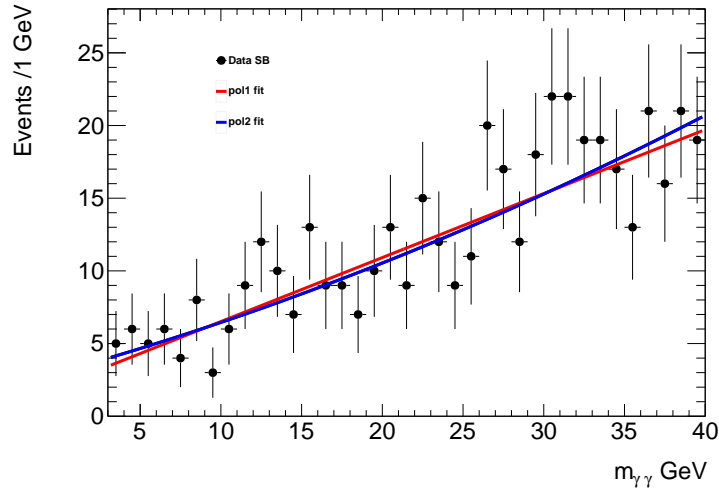


Figure 55: Diphoton invariant mass distribution in region (3,40) GeV for the SB region. The fitted pol1 and pol2 functions are shown.

and a quadratic (pol2) function, with χ^2/ndf values of 0.73 and 0.74, respectively. The two functions converge as the second-order term in the quadratic function is close to zero. Therefore, the linear function provides an adequate description of the SB distribution.

To use this function for background estimation, a transfer factor must be derived from the VR. The derivation of this factor is shown in Figure 56. In this plot, a Landau distribution is fitted to the VR data, with a χ^2/ndf of 1.18. The fitted distribution predicts 1838 events in the SR for the $m_{l\gamma\gamma}$ invariant mass range of (110, 140) GeV. The actual observed number of events in this range is 1824. The transfer factor, which maps the number of sideband events in the VR mass range (100, 110) and (140, 180) GeV to the signal region, is then calculated as: $1838/1952 = 0.9416$.

After the derivation of the transfer factor the fitted linear function of the SB can be applied to the SR as background estimation, shown in Figure 57. This is the first method for a data driven background estimation.

Method 2: Background estimation with SR fit with blinded ALP mass regions This ALP analysis uses the narrow-width ALP approximation, which is based on the assumption that ALPs only interact with photons and do not couple to other Standard Model particles, resulting in a small decay width. As a result, an ALP signal is expected to be localized around the ALP mass. From the right plot of Figure 54, it is clear that no

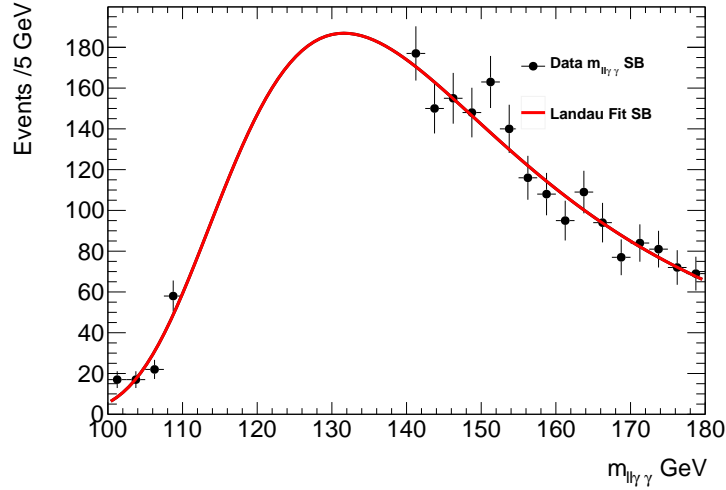


Figure 56: VR Diphoton+dilepton invariant distribution. The fitted Landau distribution is shown. The integral of the Landau distribution in the SR is 1838. The number of events in the VR are 1952.

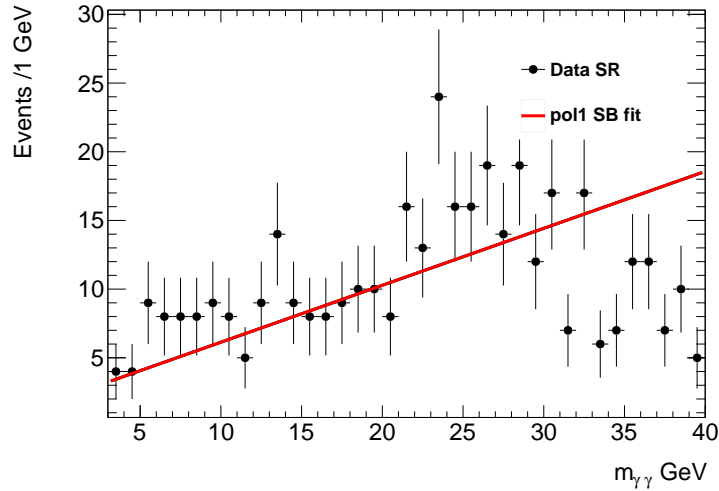


Figure 57: SR Diphoton invariant mass distribution plot in mass region (3,40) GeV. The normalized, by the transfer function from Figure 56, pol1 function is shown.

localized excess with greater than 2σ significance is observed. Still, as the null hypothesis for each ALP mass is to locate signal matching an ALP resonance the SR distributions should be blinded in the regions corresponding to Table 17, using only the data around it as fit inputs. Thus in this method the SR_method2 and SB_method2 regions are defined as shown in Table 18. A table showing the new SB regions per ALP mass is Table 19.

Three polynomial functions—pol0, pol1, and pol2—are fitted to the data, and then integrated within the blinded ALP mass region to estimate the background. The background

Distributions definitions in background estimation method 2	
Distribution	Definition
SR_method2 (Signal region)	ALP mass window in $m_{\gamma\gamma}$ distribution, table 17.
SB_method2 (Sideband region)	Data out of ALP mass window in $m_{\gamma\gamma}$ distribution, Table 19

Table 18: Definitions of distributions in method 2 of background estimation in hZa optimization analysis.

Mass (GeV)	SB_method2 (GeV)	Mass (GeV)	SB_method2 (GeV)
4	(5,11)	15	(8,13)&(17,22)
5	(3,4)&(6,11)	20	(13,18)&(22,27)
6	(3,5)&(7,11)	25	(17,23)&(27,33)
9	(4,8)&(10,14)	30	(22,28)&(32,38)
12	(6,11)&(13,18)	33	(25,31)&(35,40)

Table 19: Background estimation method 2 SB regions from SR $m_{\gamma\gamma}$ distribution.

modeling uncertainty in the blinded region is determined by subtracting the two functions that yield the smallest χ^2/ndf values. This procedure is repeated for each ALP mass. Two example plots for ALP masses of 6 and 20 GeV are shown in Figures 58 and 59.

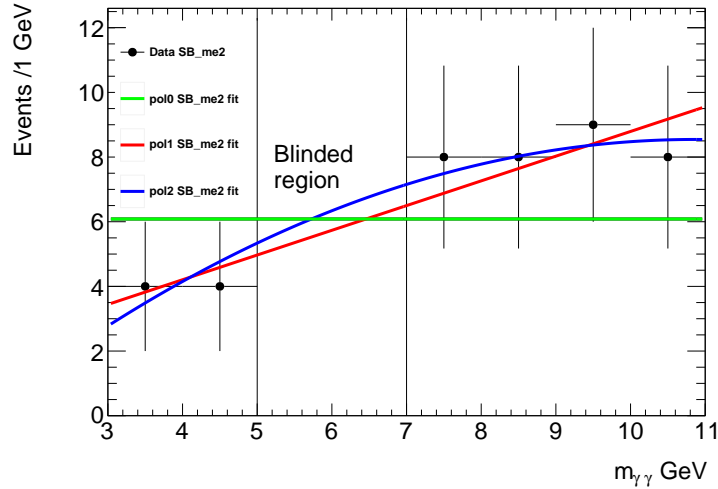


Figure 58: Background estimation fit for ALP of 6 GeV in the diphoton invariant mass distribution using the method 2 SB region defined in Table 19. The three fits are pol0, pol1 and pol2, with pol2 having the smallest χ^2/ndf value of 0.108. The subtraction of the pol1 and pol2 fits is used as background modeling uncertainty.

From the two plots it is evident that the polynomial fits are capable of describing the distributions in the SB_method2 regions and produce a background estimation in the blinded signal region SR_method2. This method is the primary method of background

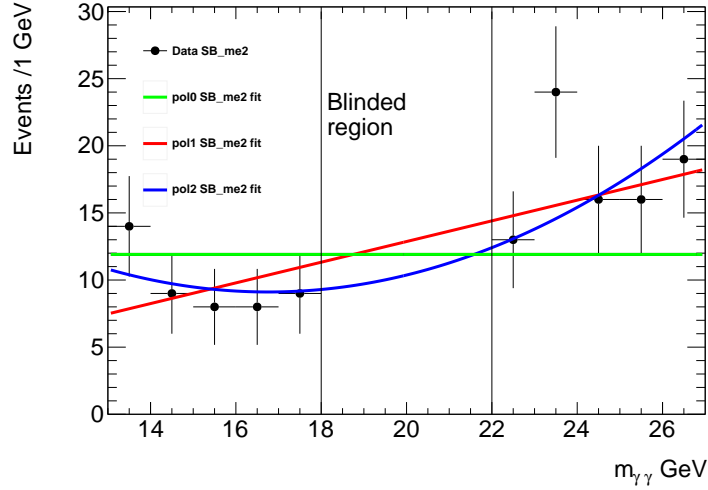


Figure 59: Background estimation fit for ALP of 20 GeV in the diphoton invariant mass distribution using the method 2 SB region defined in Table 19. The three fits are pol0, pol1 and pol2, with pol2 having the smallest χ^2/ndf value of 0.788. The subtraction of the pol1 and pol2 fits is used as background modeling uncertainty.

estimation employed in the analysis.

Method3: ABCD method The ABCD method is a background estimation method used extensively in ATLAS analyses. It is a straightforward method which requires that there are two selections that form part of the definition of the signal region, region A, which can be inverted in order to define three further regions, region B, C, and D. These control regions should ideally be rich in events produced from background processes that the method is attempting to estimate. The regions should be able split via a continuous variable. In our case the regions are as follows, defined per ALP mass point:

- Region A (Signal region) : The mass window of the ALP on the SR distribution, which is blinded according to Table 17.
- Region B : The mass window of the ALP on the SB distribution, according to table 17.
- Region C : The sideband mass range outside the ALP mass window, which is also used for method 2 shown in Table 19, on SR distribution.
- Region D: The identical mass range as region C on the SB distribution.

These regions are defined for each ALP mass point, shown in Tables 17 and 19. The goal of the ABCD method is to produce a prediction for the number of background events in the

signal region A. The assumption that underpins the ABCD method is that the following statement is true:

$$\frac{N_C^{bkg}}{N_D^{bkg}} = \frac{N_A^{bkg}}{N_B^{bkg}} \quad (49)$$

This condition will be met if the observables defining the ABCD plane are sufficiently uncorrelated for background events. The method will then calculate a transfer factor that maps the number of events in region B to region A. The regions for a 6 GeV ALP and a 20 GeV ALP are shown in Figure 60. The results for a 6 GeV ALP and a 20 GeV ALP are shown in Figures 61 and 62.

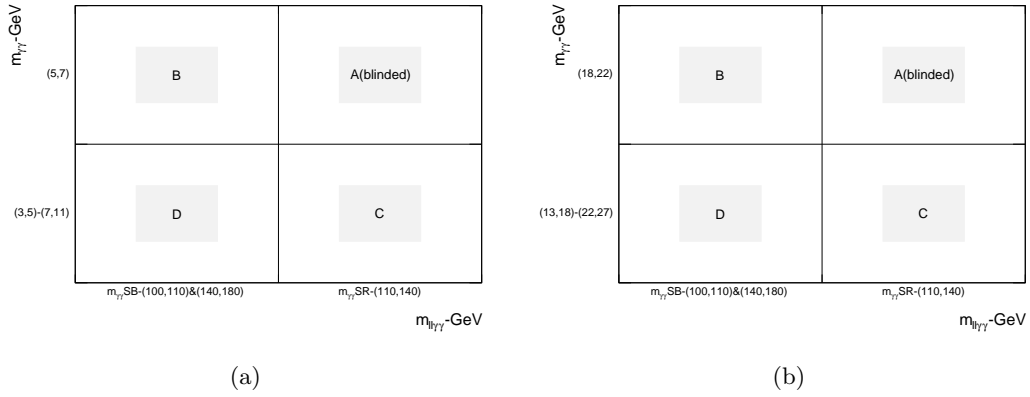


Figure 60: A representation of the ABCD regions for a 6 GeV ALP (left) and a 20 GeV ALP (right).

Overall, the background estimation is reasonable, as it agrees with the results from the other two methods. However, the fluctuations in the bins of region B carry over to the background estimation, which is undesirable. Additionally, the uncertainty of this method is higher than that of methods 1 and 2 due to the data fluctuations. This method is primarily suited for smooth backgrounds or singular binning, so this behavior is expected. Nevertheless, it remains a simple and reliable approach, serving as the secondary method for background estimation.

Methods comparison A comparison of all three methods for the 6 GeV and 20 GeV ALP is shown in Tables 20 and 21. A comparison is made with the original analysis background estimation, to illustrate the background rejection of the two event selections.

The three methods produce compatible results validating the different approaches. The results of the three methods for the high mass region are shown in Figure 63. Ultimately, method 2 was used as background estimation.

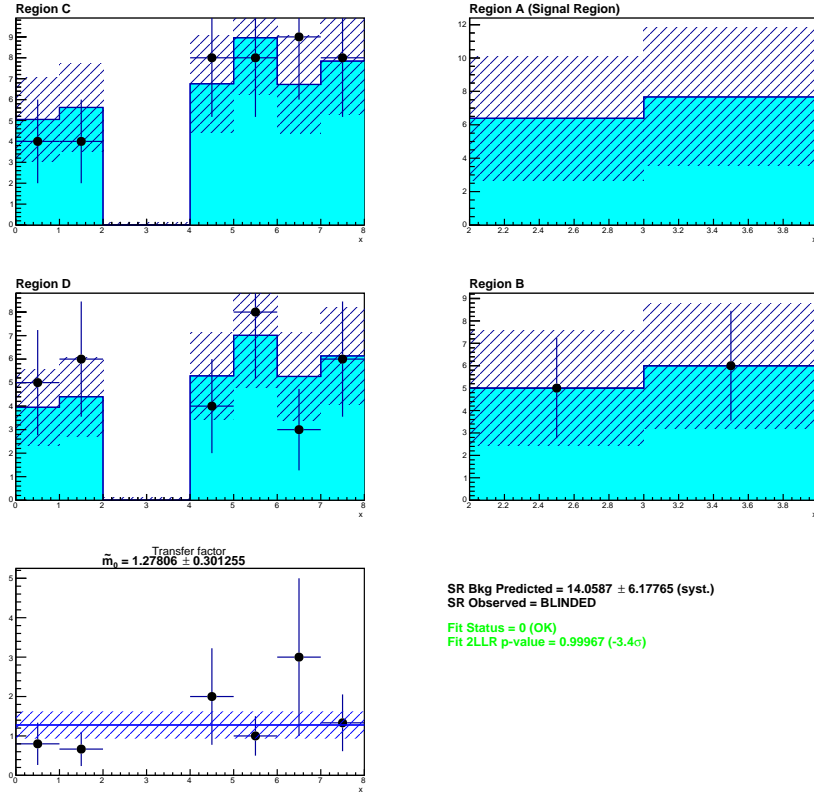


Figure 61: Background estimation for 6 GeV ALP using the ABCD method. The blue area on region A is the background prediction. A transfer function is also calculated together with a likelihood fit for the different signal region predictions.

Mass Range	Method 1	Method 2	Method 3	Orig. back. est.
(5,6)	4.49 ± 0.29	5.85 ± 0.50	6.39 ± 1.50	22.18 ± 0.94
(6,7)	4.93 ± 0.20	6.76 ± 0.75	7.67 ± 1.81	18.99 ± 0.90

Table 20: Comparison of the background estimation methods for a 6 GeV ALP. The background estimation of the original analysis is also shown.

Mass Range	Method 1	Method 2	Method 3	Orig. back. est.
(18,19)	10.17 ± 0.37	9.48 ± 0.72	9.06 ± 1.17	3.59 ± 0.36
(19,20)	10.61 ± 0.37	10.02 ± 0.80	12.95 ± 1.68	3.12 ± 0.37
(20,21)	11.05 ± 0.37	10.80 ± 0.78	16.83 ± 2.18	2.72 ± 0.33
(21,22)	11.48 ± 0.36	11.82 ± 0.65	11.65 ± 1.51	2.36 ± 0.31

Table 21: Comparison of the background estimation methods for a 20 GeV ALP. The background estimation of the original analysis is also shown.

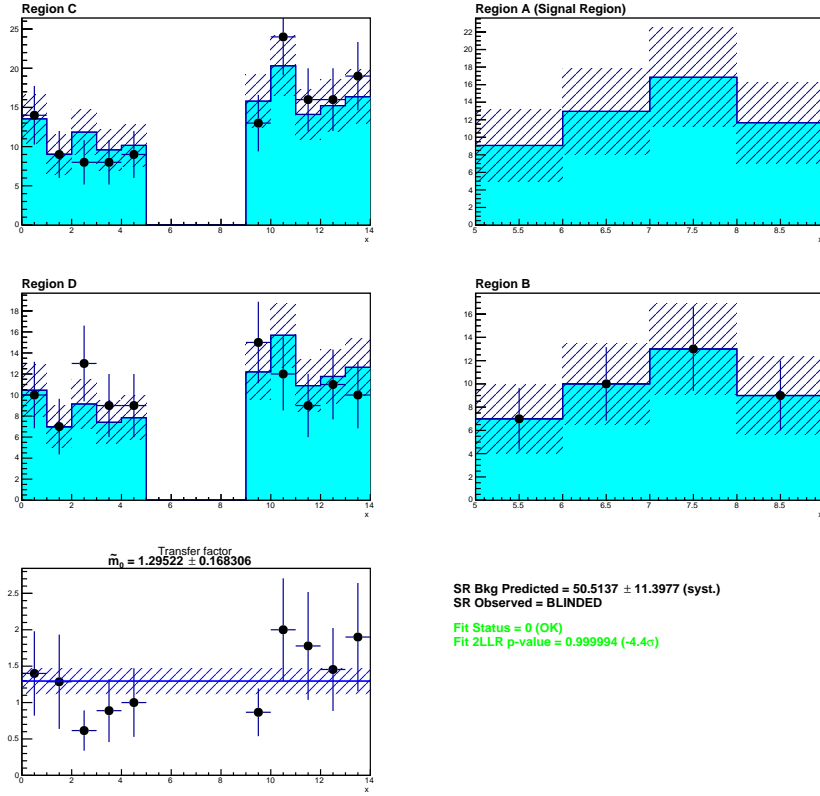


Figure 62: Background estimation for 20 GeV ALP using the ABCD method. The blue area on region A is the background prediction. A transfer function is also calculated together with a likelihood fit for the different signal region predictions.

6.6.4 Systematic uncertainties

The main systematic uncertainties are similar in both the reinterpretation and the optimized analyses. However, due to the different event selection, the systematic uncertainties have been recalculated. The average results for the prompt systematic uncertainties across the full range of ALP masses are discussed.

- Photon identification efficiency uncertainty: In the case of the optimized analysis, both tight and loose photon identification are used, compared to only loose photon identification in the original analysis, which can influence the systematic uncertainty. This systematic uncertainty is the second most important of the analysis contributing $\pm 5\text{-}6\%$ on the final limit, depending on ALP mass.
- Photon isolation efficiency uncertainty: In the case of the optimized analysis, all photons use track-based isolation, compared to a subset of photons in the original analysis that use FCLoose. This, combined with the tight photon ID, leads to the most important systematic uncertainty of the analysis contributing $\pm 8\text{-}10\%$ on the

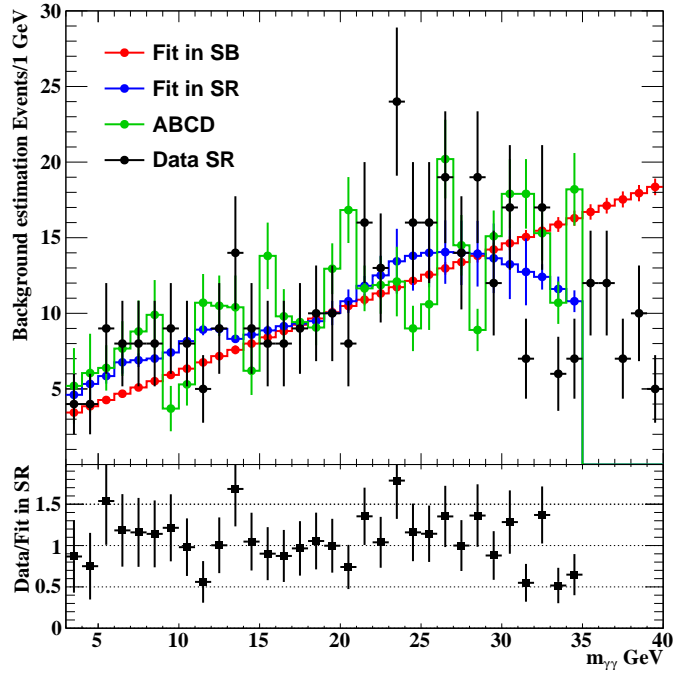


Figure 63: Background estimation results for the high mass region. Method 1 (red) - fit in SB, method 2 (blue)- fit in SR, method 3 (green)- ABCD and data in SR (black). Reasonable agreement is seen between the methods.

final limit, depending on ALP mass.

- Pile up uncertainty: In the case of the optimized analysis, slightly different signal events may be selected compared to the original analysis. Thus, the pile-up uncertainty must be taken into consideration and recalculated. This systematic uncertainty is the third most important of the analysis contributing $\pm 2-4\%$ on the final limit, depending on ALP mass.
- Electron identification efficiency uncertainty: In the case of electrons the selection in both analyses is identical. This systematic uncertainty contributes $\pm 0.5 - 0.6\%$ on the final limit, depending on ALP mass.
- Electron isolation efficiency uncertainty: This systematic uncertainty is negligible contributing $\pm 0.05\%$ on the final limit, as in the original analysis.
- Electron trigger efficiency uncertainty: This systematic uncertainty contributes $\pm 0.2 - 0.3\%$ on the final limit, depending on ALP mass.
- Muon identification/reconstruction efficiency uncertainty: In the case of muons the selection in both analyses is identical. This systematic uncertainty contributes $\pm 0.2 - 0.4\%$ on the final limit, depending on ALP mass.

- Muon isolation efficiency uncertainty: This systematic uncertainty contributes $\pm 0.2 - 0.4 \%$ on the final limit, depending on ALP mass.
- Muon trigger efficiency uncertainty: This systematic uncertainty contributes $\pm 0.9 - 1.2 \%$ on the final limit, depending on ALP mass.

The background estimations also introduce the MC uncertainty and the background modeling uncertainty. This leaves the systematic uncertainty due to displaced photon origin, which, as in the hZa original analysis, requires a more involved analysis.

Displaced photon origin systematic uncertainty The optimized hZa event selection uses slightly different ALP events than the original selection. Additionally, the photon ID working point 'Tight' is applied, which has a different classification than the 'Loose' working point, with more shower-shape variables utilized [59]. The interpolation of the shower-shape variables is identical, but afterward, the Tight ID of each photon needs to be recalculated. The measured efficiencies and corresponding systematic uncertainty for the tight working point are shown in Fig. 64, reproduced from [92]. The efficiencies decrease with respect to the ALP decay radius both before and after fudging the shower-shapes, ranging from 0.8 to 0.3, with a minimum around $l_a = 1600$ mm. The corresponding systematic uncertainty behaves similarly to the systematic uncertainty of the loose photon identification, ranging from 0.05 to 0.21.

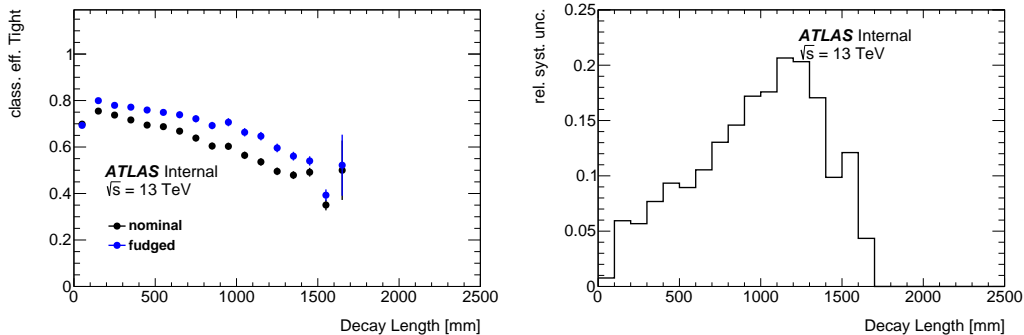


Figure 64: Systematic uncertainty of the tight photon identification working point due to the change of the photon shower-shapes with increasing ALP decay radius in the $(r - \phi)$ plane of the detector. Left: the identification efficiency before and after fudging the shower-shapes. Right: The relative difference of the efficiencies which is taken as systematic uncertainty. The difference rises with respect to the ALP decay radius and lies between 0.06 to 0.21 [92].

Subsequently, the displaced photon vertices systematic uncertainty is calculated in the same way as the original analysis, with weights applied to each photon on an event-by-event basis. For the ALP masses of 1 and 2 GeV the loose ID displaced systematic uncertainty

is used, with the loose ID weights shown in Figure 35. For the ALP masses exceeding 2.5 GeV the tight ID weights are used, shown in Figure 64. Some of these calculations are shown in Figure 65. Results are shown in Tables 22 and 23.

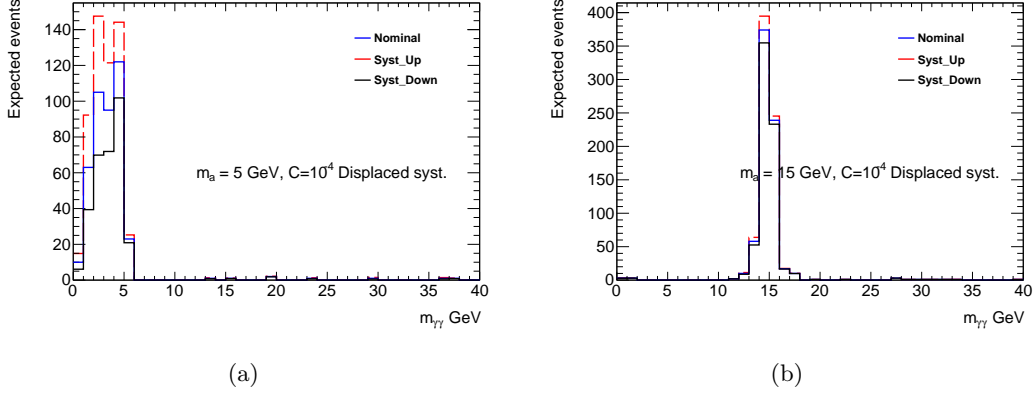


Figure 65: Systematic uncertainty due to displaced photon origin variations for two different masses with $C = C_{\gamma\gamma} = 10^{-4}$, left is 5 GeV and right is 15 GeV. Nominal corresponds to no weights applied, Syst_Up to weights added and Syst_Down to weights subtracted.

	Systematic uncertainty due to displaced photon origin - Up. Values in %							
$C_{\gamma\gamma} \setminus m_a$	4 GeV	5 GeV	7 GeV	9 GeV	15 GeV	20 GeV	25 GeV	33 GeV
10^{-3}	7.73	3.39	0.34	0	0	0	0	0
10^{-4}	33.56	30.39	23.54	16.24	4.87	0.78	0.15	0
10^{-5}	-	-	30.36	34.72	33.06	15.96	15.96	9.48

Table 22: Displaced systematic uncertainty - Up variation. Values in %

	Systematic uncertainty due to displaced photon origin - Down. Values in %							
$C_{\gamma\gamma} \setminus m_a$	4 GeV	5 GeV	7 GeV	9 GeV	15 GeV	20 GeV	25 GeV	33 GeV
10^{-3}	7.16	3.16	0.33	0	0	0	0	0
10^{-4}	28.16	25.76	20.57	14.61	4.53	0.73	0.13	0
10^{-5}	-	-	25.89	28.87	27.67	19.90	14.41	8.74

Table 23: Displaced systematic uncertainty - Down variation. Values in %.

6.6.5 Results

To obtain an upper limit on the branching ratio for $h \rightarrow Za$, compared to the Higgs gg production mode, a profile likelihood fit is performed on the invariant photon mass spectrum for each mass-coupling pair, in an identical way to the reinterpretation of the ATLAS hZa analysis. As the systematic variations are slightly different in this case the pull plot and upper limit at 95% C.L. are shown in Figures 66 and 67 respectively. The fitting procedure yields a limit on the BR of the $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ process at 95% C.L. The following limits are produced for the four couplings $C_{\gamma\gamma} = 10^{-2}$ (prompt), $C_{\gamma\gamma} = 10^{-3}$, $C_{\gamma\gamma} = 10^{-4}$ and $C_{\gamma\gamma} = 10^{-5}$, shown in Figure 68.

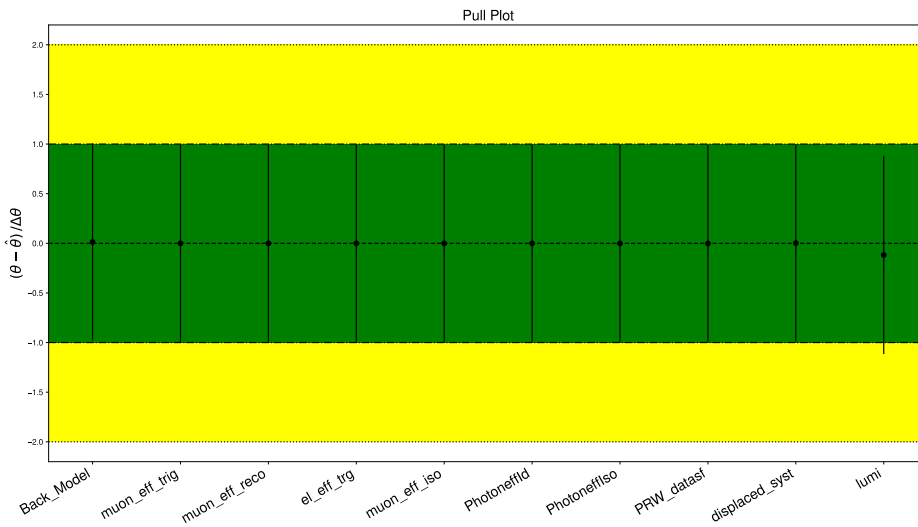


Figure 66: Pulls for the nuisance parameters in the fit for $m_a = 9$ GeV with $C_{\gamma\gamma} = 10^{-4}$.

The limit on the branching ratio can be converted into a limit on the coupling constant of the ALPs to photons $C_{\gamma\gamma}$. The formulas and methodology used is identical to the one used by the reinterpretation, in section 6.5.5, just with slightly different limits on the branching ratios. The limits in the ALP-photon coupling are shown in Figure 69.

It can be seen that the limits of the ALP coupling to photons are very similar between the two selections. This can be explained by considering that the limits depend on the relative difference in the selection for the same mass. This is expected, as the prompt selection efficiency represents the maximum number of events that can be observed, since the detector is optimized to detect prompt photons. As the displaced couplings increase, events can no longer be reconstructed, leading to smaller efficiencies. The only way to increase the coupling sensitivity is by using a larger dataset.

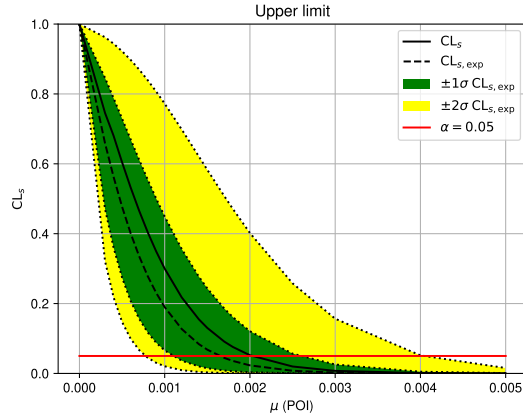


Figure 67: Observed (solid line) and expected (dashed line) CLS values for different μ with $\pm 1\sigma$ (green) and $\pm 2\sigma$ (yellow) for a fit with statistical variations and systematics applied for $m_a = 9$ GeV with $C_{\gamma\gamma} = 10^{-4}$.

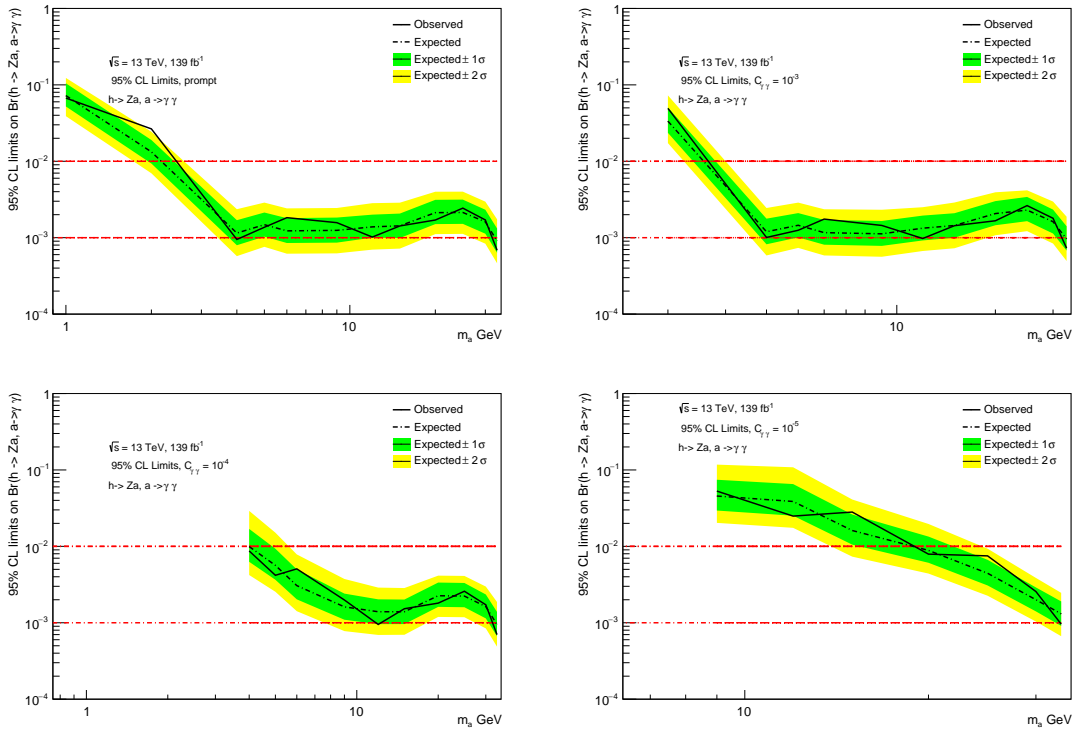


Figure 68: Upper limits, at the 95% C.L. on the branching ratio of the process $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$. Upper row left: Limits for a coupling of $C_{\gamma\gamma} = 0.01$ (prompt), right: $C_{\gamma\gamma} = 10^{-3}$. Lower row left: $C_{\gamma\gamma} = 10^{-4}$, right: $C_{\gamma\gamma} = 10^{-5}$. Also shown are the ± 1 and 2σ uncertainty bands taking into account the statistical and systematic uncertainties.

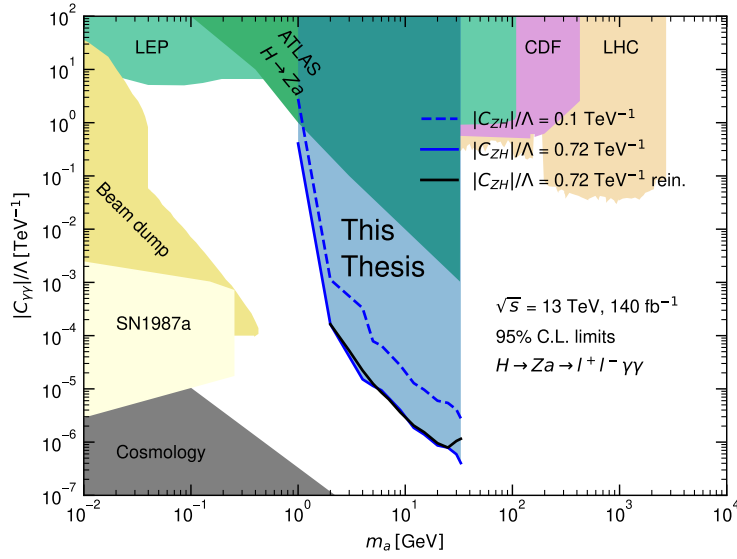


Figure 69: Limits on the ALP mass and coupling to photons at 95% C.L.. The results from this analysis are shown as blue lines assuming a Higgs to ALP coupling of $C_{Zh}/\Lambda^2 = 0.72 \text{ TeV}^{-1}$ (solid) and $C_{Zh}/\Lambda^2 = 0.1 \text{ TeV}^{-1}$ (dashed). The reinterpretation results are shown as the black line. This plot can be directly compared to 42.

6.6.6 Comparison of prompt sensitivities in the two analyses

A short comparison of the prompt sensitivities for the two analyses is shown in Figure 70. Excluding the ALP mass of 2 GeV, the limits of the two analyses are similar, with no major difference in sensitivity. For ALP masses greater than 20 GeV, the optimized analysis shows improved sensitivity due to the higher signal selection efficiency, which reaches a maximum of around 6% for $m_a = 33 \text{ GeV}$, compared to around 0.4% for the original selection. Overall, the sensitivity of the two analyses is similar and worse than the CMS analysis [77]. One possible improvement in sensitivity could be made with improved background rejection, e.g. kinematic variable signal and background classification using machine learning. Nonetheless, the hZa channel is an interesting channel, and multiple analyses will probe it in the future.

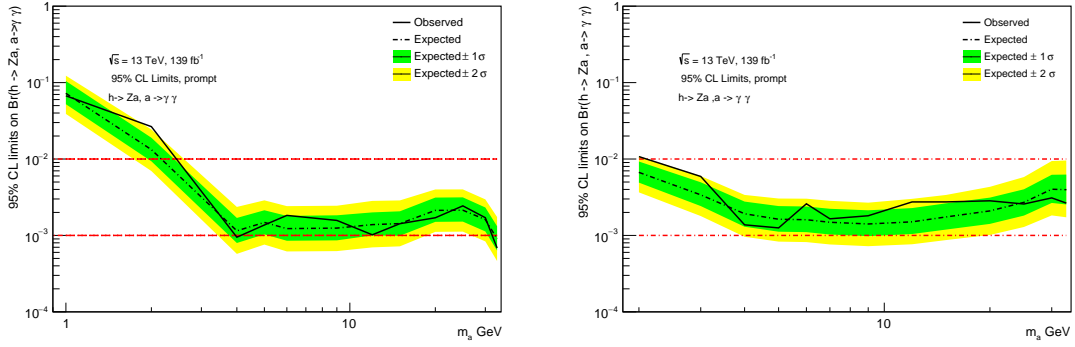


Figure 70: Upper limits, at the 95% C.L. on the branching ratio of the process $H \rightarrow Za \rightarrow l^+l^-\gamma\gamma$ for prompt coupling. Optimized analysis (left) and reinterpretation analysis (right).

6.7 Summary of $h \rightarrow Za$

Data recorded by the ATLAS experiment at the LHC, corresponding to an integrated luminosity of 139 fb^{-1} from proton-proton collisions at a center-of-mass energy 13 TeV, are used to search for a rare decay of the 125 GeV Higgs boson to a Z boson and an axion-like particle a , which is considered to have mass m_a between 2 GeV and 34 GeV. The particle a is considered to decay to a diphoton pair. Two selections were developed and compared, yielding similar results. No significant discrepancies are observed with respect to the expectation from Standard Model backgrounds and therefore upper limits are reported ranging from 0.01 to 0.001 branching ratio. These limits are the first to take displaced couplings into account in the $h \rightarrow Za$ channel.

7 Light-by-Light scattering

7.1 Introduction

Light-by-light (LbyL) scattering, $\gamma\gamma \rightarrow \gamma\gamma$, is a quantum-mechanical process that is forbidden in the classical theory of electrodynamics as electromagnetic waves in vacuum obey the principle of linear superposition [98]. In the Standard Model, which is a relativistic theory incorporating quantum mechanics, the process becomes possible at higher orders. The $\gamma\gamma \rightarrow \gamma\gamma$ reaction proceeds at lowest possible order in the fine structure constant (α_{em}) via virtual one-loop box diagrams involving charged fermions (leptons and quarks) and the W^\pm boson [47], shown in figure 71. The process involves four electromagnetic vertices, making it extremely rare compared to electromagnetic tree-level processes, such as Compton scattering. However, in various extensions of the Standard Model, additional contributions are possible, rendering the measurement of LbyL scattering sensitive to potential new physics. Relevant examples include axion-like particles (ALPs), spin-2 particles (commonly referred to as gravitons), which couple indirectly to photons [39, 99], and higher-dimension operators [100]. These contributions would produce a resonance peak in the invariant mass spectrum of the diphoton final state. These two features make the LbyL scattering a very interesting channel.

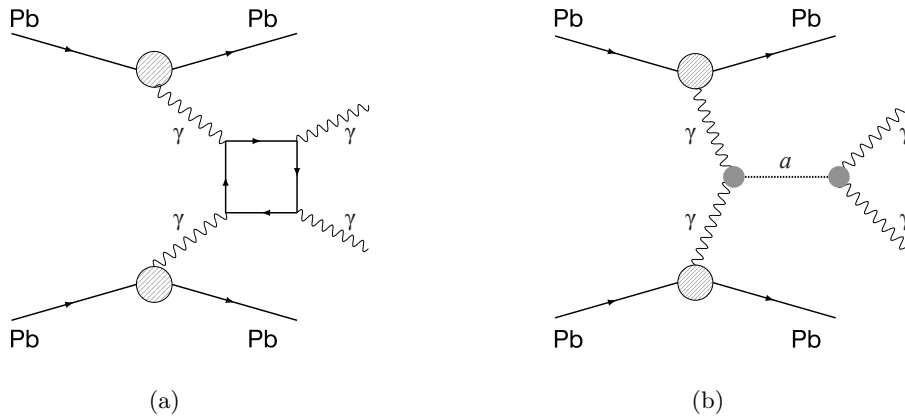


Figure 71: Schematic diagrams of (left) SM LbyL scattering and (right) ALP production in Pb+Pb [47].

7.2 LbyL scattering in ATLAS

At the LHC, proton-proton (pp) collisions are not the only type of collisions studied, as collisions involving ions of the heaviest stable element, lead (Pb), are also investigated. These are particularly interesting because many physics processes are easier to detect in p-Pb or Pb-Pb collisions. Examples include processes related to photon or gluon emissions.

In Run 3 in 2023 PbPb collisions at a center-of-mass energy of $\sqrt{s} = 5.36$ TeV were performed. Pb-Pb collisions at the LHC are typically studied in scenarios where the ions interact hadronically, producing high multiplicities and creating a volume in which the reaction participants form a domain of hot, dense quark-gluon matter. These are referred to as central collisions. Additionally, it is of significant interest to study very peripheral (or ultra-peripheral) collisions (UPCs), which are characterized by an impact parameter b greater than the sum of the radii of the ions. In such cases, the ions do not interact directly through the short-range nuclear force but instead primarily via their electromagnetic (EM) fields.

In this regime the LHC, loosely speaking, behaves like a photon-photon collider. The electric charges of the protons within the nucleus generate an electromagnetic field that can be viewed as a cloud of photons surrounding the nucleus. This model is known as the Weizsäcker-Williams method or the equivalent photon approximation (EPA) and is commonly used to simplify calculations of photon-induced interactions. The EM fields produced by the colliding Pb nuclei can thus be treated as beams of quasi-real photons with a small virtuality of $Q^2 < 1/R^2$, where R is the radius of the charge distribution leading to $Q^2 < 10^{-3}$ GeV² [101]. Indeed, ultra-relativistic ions generate strong EM fields, with magnitudes of $E \sim 10^{25}$ V/m for PbPb collisions at the LHC with a center of mass energy per nucleon pair of 5.02 TeV. In this regime, the electric and magnetic fields of each ion are perpendicular and almost transverse to the direction of propagation [47].

The process of interest for this analysis is LbyL scattering, the reaction Pb+Pb ($\gamma\gamma$) \rightarrow Pb+Pb $\gamma\gamma$. The cross section for this reaction can be calculated by convolving the respective photon flux with the elementary cross section for the process $\gamma\gamma \rightarrow \gamma\gamma$. Since the photon flux associated with each nucleus scales as Z^2 (the nuclear number), the cross section is significantly enhanced compared to proton-proton (pp) collisions, as each individual proton within a nucleus can participate in the interaction. In this measurement, the final-state signature of interest is the exclusive production of two photons. The diphoton final state is measured in the detector surrounding the Pb+Pb interaction region, with the incoming Pb ions surviving the electromagnetic interaction. Hence, one would expect the detection of two low-energy photons with no additional activity in the central detector. In particular, no reconstructed charged-particle tracks originating from the Pb+Pb interaction point are anticipated.

A measurement of the LbyL scattering cross section was performed by ATLAS using the 2015 and 2018 PbPb dataset with a centre-of-mass energy of $\sqrt{s} = 5.02$ TeV, corresponding to an integrated luminosity of 2.2 nb^{-1} [47]. This analysis follows the approach proposed in [47] and the methodology used in the previous measurements, utilizing PbPb data recorded in 2023 by the ATLAS detector at a centre-of-mass energy of $\sqrt{s} = 5.36$ TeV corresponding to a luminosity of 1.64 nb^{-1} . The main advantages over the previous analysis are lower p_T

thresholds for the photons and higher collision energy.

As this analysis is still ongoing, only general elements and results will be presented, without definitive conclusions. Key aspects of the analysis, such as event selection, triggers, background estimation, and systematic variations, will be briefly discussed. The main focus will be on the methods used for the determination of the photon identification working point, as it was the primary contribution of the author.

7.3 Data and MC samples

The data used in this measurement is from PbPb collisions with a centre-of-mass energy of $\sqrt{s} = 5.36$ TeV, recorded in the 2023 HI runs at the LHC. The full data set corresponds to an integrated luminosity of 1.64 nb^{-1} of PbPb data, with a preliminary uncertainty of 3.5%. Only high-quality data with all detectors operating normally are analyzed.

LbyL signal events are generated using SuperChic v4.2 [102]. They take into account box diagrams with leptons and quarks (such as the diagram in Figure 71), and W^\pm boson, including interference effects. The W^\pm contribution is only important for diphoton masses $m_{\gamma\gamma} > 2m_W$.

The exclusive diphoton final state can be also produced via the strong interaction through a quark loop in the exchange of two gluons in a colour-singlet state. This central exclusive production (CEP) process, $gg \rightarrow \gamma\gamma$, is modeled using SuperChic v4.2 [102]. This process has a large theoretical uncertainty, $\mathcal{O}(100\%)$, mostly related to the modeling of the rapidity gap survival probability [103]. Therefore, further constraints on the normalization of this background process are estimated directly from data, using a dedicated control region.

The $\gamma\gamma \rightarrow e^+e^-$ process is a potential background, when both leptons are reconstructed as photons, but it is also valuable for various calibration studies within the analysis. The $\gamma\gamma \rightarrow e^+e^-$ process is modeled with the Starlight + Pythia8 MC generators [104, 105], in which the cross section is computed by convolving the Pb+Pb photon flux with the leading-order formula for $\gamma\gamma \rightarrow e^+e^-$.

Finally, ALPs are also generated using SuperChic v4.2 [102]. The ALP resonances generated by the LbyL process will primarily be at rest, resulting in back-to-back energy clusters in the transverse plane of the electromagnetic calorimeter, due to the photon kinematics. This feature is also used as a selection criterion in the event selection.

All simulated events make use of a GEANT4 [76] based detector simulation [106] and are reconstructed with the standard ATLAS reconstruction software.

7.4 Event selection

7.4.1 Triggers

As with all ATLAS analyses, interesting events should be selected and stored to be analyzed, which is achieved through specified triggers. In the case of the LbyL analysis, this process is more complex, as the signal photons have much lower transverse momentum than those typically used in ATLAS, necessitating specific triggers for data collection. These triggers are designed to be sensitive to events with moderate activity in the calorimeter system and minimal signal in the rest of the detector. They are split into L1 (hardware) and HLT (software) triggers. The triggers will be briefly presented below as it is a work in progress.

The two L1 triggers used are: L1_TAU1_TE4_VTE200 and L1_2TAU1_VTE200. The TAU1 requirement means that at least one EM cluster has been registered with a minimum $p_T = 1$ GeV and similarly 2TAU1 means two EM clusters have been registered. The calorimetric TE4 requires a minimum total transverse energy of 4 GeV and VTE200 vetoes events with total transverse energy exceeding 200 GeV. The L1 trigger efficiency is calculated as a ratio of events passing the logical OR of L1_TAU1_TE4_VTE200 and L1_2TAU1_VTE200 triggers to all UPC events. Trigger efficiency is calculated as a function of the sum of the transverse energy of two EM clusters and parametrized using an error function:

$$\epsilon_{L1} = 0.5(\text{erf}((E_T^{\text{cluster1}} + E_T^{\text{cluster2}} - p_1)/p_2) + 1), \quad (50)$$

where the obtained values of parameters are: $p_1 = 5.71914$ and $p_2 = 2.57668$. Results are compared with the fit to the 2018 data trigger efficiency [107], shown in figure 72. During Run 2 data-taking a logical OR of two triggers L1_TAU1_TE4_VTE200 and L1_2TAU1_VTE50 was used. The efficiency value is about 50% for events with $E_T^{\text{cluster1}} + E_T^{\text{cluster2}} = 5$ GeV and about 95% for $E_T^{\text{cluster1}} + E_T^{\text{cluster2}} = 10$ GeV. The distribution reaches a plateau for events where the sum of the transverse energy is equal to 15 GeV. The trigger efficiency was found to have a higher energy threshold compared to the Run 2 reference, due to the increased noise level.

In Run 3 there are also HLT triggers which were specifically created for the LbyL analysis. These are:

- HLT_mb_sp_vpix30_hi_FgapAC5_L1TAU1_TE4_VTE200
- HLT_mb_sp_vpix30_hi_FgapAC5_2g0_etcut_L12TAU1_VTE200
- HLT_mb_sp_vpix30_hi_FgapAC5_L12TAU1_VTE200

They include three sets of requirements:

- L1* e.g. L1TAU1_VTE200 refers to the level one seed used, described in L1 above.

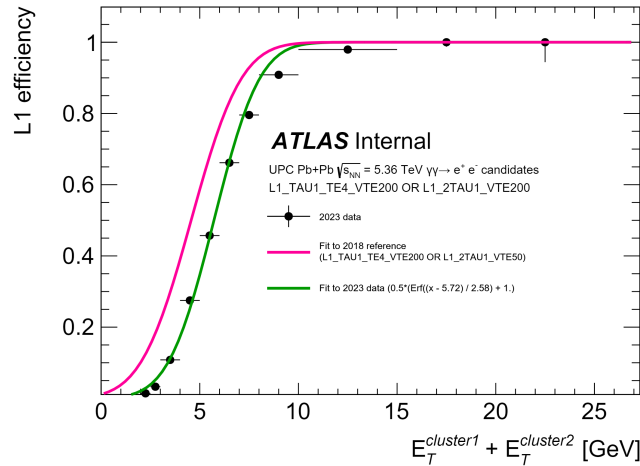


Figure 72: L1 trigger efficiency as a function of the sum of the transverse energy of two EM clusters with statistic uncertainties, error function fitted to data, and a fit to 2018 data.

- FCal veto, i.e. rejection of events with $\sum E_T^{FCal} > 5$ GeV on any side of FCal (imposed in hi_FgapAC5).
- Low activity in the ID, defined by a maximum number of 30 hits in the Pixel Detector (imposed in mb_sp_vpix30).

For the calculation of the trigger systematic uncertainties various supporting triggers are targeting the electron-positron channel. Particularly interesting are the following sources of uncertainties:

- Track uncertainties - uncertainty on the hits used for the charged particle tracks construction.
- Track acoplanarity - uncertainty introduced by the acoplanality of tracks compared to reconstructed particle objects.
- Diphoton angle - uncertainty on the kinematics due to low photon p_T on EM calorimeter clusters.
- ZDC calorimeter uncertainties - the ZDC detectors are used for the measurement of the event hadronic activity.

7.4.2 Object reconstruction

Photons are reconstructed from EM clusters in the calorimeter and tracking information provided by the ID, which allows for the identification of photon conversions [59]. Selection requirements are applied to remove EM clusters with significant energy contributions from

poorly functioning calorimeter cells, and a timing requirement is imposed to reject out-of-time candidates. An energy calibration specifically optimized for photons [63] is applied to the candidates, accounting for upstream energy loss and both lateral and longitudinal shower leakage. The calibration is derived for nominal pp collisions, with dedicated factors applied to account for the negligible contribution from multiple Pb+Pb collisions at the same bunch crossing. The AOD format is used to store the selected events. The processing of the AODs into ntuples (simpler object centered framework) was done using the NtupleMaker based on r24 AnalysisBase, where all the CP tools are used, recommended by the physics object preparation groups of ATLAS. The tools are stored in the ATLAS tool framework called Athena with release 24 of the framework used, for Run 3 analyses.

7.4.3 Analysis selection

The event selection is relatively simple as the LbyL scattering process is rare and clearly defined. The main objective is the limitation of the two irreducible backgrounds $\gamma\gamma \rightarrow e^+e^-$ (electron-positron) and central exclusive process $gg \rightarrow \gamma\gamma$ (CEP), which persist in a standard two photon selection. The required cuts are as follows:

- Exactly 2 photons with $p_T > 2$ GeV and $|\eta| < 2.37$, excluding crack region.
- Diphoton invariant mass $m_{\gamma\gamma} > 4$ GeV.
- Diphoton transverse momentum $p_{T\gamma\gamma} < 2$ GeV.
- Diphoton Acoplanarity $A < 0.01$, described by equation 51.
- No standard tracks.
- No pixel tracks.
- Photon Identification, Run 3 working point.
- Any of the specified LbyL triggers detailed in 7.4.1.

The selection of these specific cuts is straightforward. The p_T of the photons is the lowest currently considered safe, before photon reconstruction uncertainties become dominant. The diphoton kinematic cuts are motivated by the diphoton signal kinematics, aiming to suppress irreducible backgrounds and fake photon sources. The photons are required to be emitted in a back-to-back topology, which can be quantified using the reduced acoplanarity A , a measure of the deviation of the two particles from being coplanar, defined as:

$$A = 1 - \frac{|\Delta\phi_{\gamma\gamma}|}{\pi} \quad (51)$$

where $\Delta\phi_{\gamma\gamma}$ is the difference in the azimuthal angle of the two photons. It has been found to be strongly correlated with the CEP and electron-positron backgrounds, which

primarily occur at values $A > 0.01$. The two 'no tracks' cuts are applied to suppress the electron-positron and fake photon backgrounds. Photon identification, with a working point developed by the author, is used to further reduce the fake photon background.

7.5 Backgrounds

Multiple backgrounds exist that produce a signature identical to LbyL scattering. Therefore, the event number and characteristics of these backgrounds must be identified within the signal region. The largest contribution arises from the production of an electron-positron pair ($\gamma\gamma \rightarrow e^+e^-$). These events can fake signal signatures if the electrons are misidentified as photons or in case a hard-bremsstrahlung photon is emitted. Another contribution comes from the central exclusive process (CEP) $gg \rightarrow \gamma\gamma$ which produces an identical final state to the signal. Finally, fake photons can be a source of background mainly caused by calorimeter noise. The relevant backgrounds are shown in figure 73.

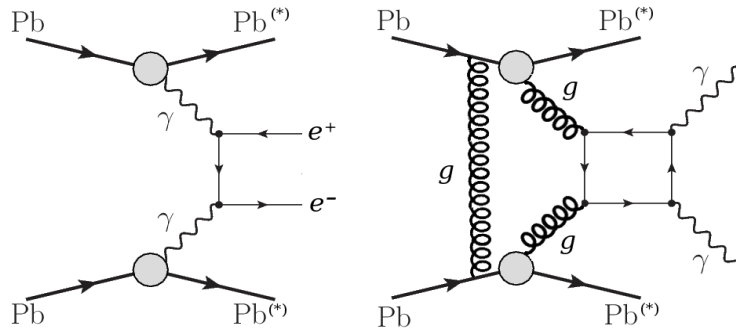


Figure 73: Feynman diagrams of the $\gamma\gamma \rightarrow e^+e^-$ (left) and CEP background (right) [47].

7.5.1 Electron-Positron background

The electron-positron process $\gamma\gamma \rightarrow e^+e^-$ is an important source of background, as it is a tree level process. In such cases, if the charged particle tracks are not reconstructed, these electrons may be misidentified as photons. Suppressing this background is crucial for ensuring the overall quality of the LbyL events.

The main way of suppression is via a tracks veto on two differently defined tracks, pixel and charged particle. Charged particle tracks are defined as tracks with $p_T > 100$ MeV, $|\eta| < 2.5$, at least one hit in the Pixel detector and at least six hits in the Pixel and SCT detectors in total. Pixel tracks are reconstructed using only the information from the Pixel detector, and are required to have $p_T > 50$ MeV, $|\eta| < 2.5$ and at least three hits in the Pixel detector. In order to suppress fake pixel tracks due to noise in the Pixel detector, only pixel tracks with $\Delta\eta < 0.5$ from the photons are considered. Pixel tracks, in particular, provide a more stringent cut, significantly reducing the number of background

events. The contribution of this background to the signal selection is calculated using data and simulated $\gamma\gamma \rightarrow e^+e^-$ events. A control region $C(N_{\text{Trk}} = 2)$ containing exactly two charged particles tracks is defined which is expected to be dominated by this background as the two electrons are expected to be reconstructed as two charged particle tracks. Each electron can produce both a pixel track and a charged particle track. Mistag events are those that have only a single reconstructed standard track. $N_{\text{PixTrk}=n}$ corresponds to n pixel tracks reconstructed in mistag events. The mistag probability is obtained by taking the fraction of events in $\text{CR}(N_{\text{Trk}} = 1)$, where only one charged track is reconstructed.

$$p_{\text{mistag}} = \frac{N_{\text{PixTrk}=1}}{N_{\text{PixTrk}=1} + N_{\text{PixTrk}=2}} \quad (52)$$

The conditional probability of reconstructing both pixel tracks when only one standard track is reconstructed can be written as $1 - p_{\text{mistag}}$. A binomial distribution is then constructed to determine the probability of different misidentification scenarios.

$$p_{N_{\text{PixTrk}}=2} = (1 - p_{\text{mistag}})^2 \quad (53)$$

$$p_{N_{\text{PixTrk}}=1} = 2 \cdot p_{\text{mistag}} \cdot (1 - p_{\text{mistag}}) \quad (54)$$

$$p_{N_{\text{PixTrk}}=0} = p_{\text{mistag}}^2 \quad (55)$$

The derived conditional probabilities are then combined with two CR regions, where the number of charged particle tracks is zero. These are $N_{\text{CR}(\text{PixTrk} = 1)}$, $N_{\text{CR}(\text{PixTrk} = 2)}$ where 1,2 are the numbers of reconstructed pixel tracks respectively. The expected contribution of $\gamma\gamma \rightarrow e^+e^-$ events in the signal regions is given by

$$\text{SR}_{\text{expected}} = (N_{\text{CR}(\text{PixTrk} = 1)} + N_{\text{CR}(\text{PixTrk} = 2)}) \cdot \frac{p_{N_{\text{PixTrk}}=0}}{1 - p_{N_{\text{PixTrk}}=0}} \quad (56)$$

As the analysis is not yet completed and numerical results for the background estimation are not available. However, the method is identical to the one used in the Run 2 analysis leading to results shown in [47]. The number of $\gamma\gamma \rightarrow e^+e^-$ events in the signal region of the Run 2 analysis was estimated to be $N_{\gamma\gamma \rightarrow e^+e^-} = 15 \pm 7$, where the uncertainty accounts for the p_{mistag} uncertainty and the limited event yield in $N_{\text{CR}(\text{PixTrk} = 1,2)}$.

7.5.2 Central exclusive production

The Central exclusive production (CEP) process $gg \rightarrow \gamma\gamma$ is an important background of the analysis. The kinematics of the background are slightly different than the LbyL process due to the emission of ISR gluons. This emission influences the acoplanality, described in equation 51, of the diphoton events and provides a very powerful cut on the irreducible backgrounds as the LbyL photons cannot emit ISR. This behaviour has been observed in both data and MC as shown in figure 74, which presents the relevant plot for the Run 2 LbyL analysis.

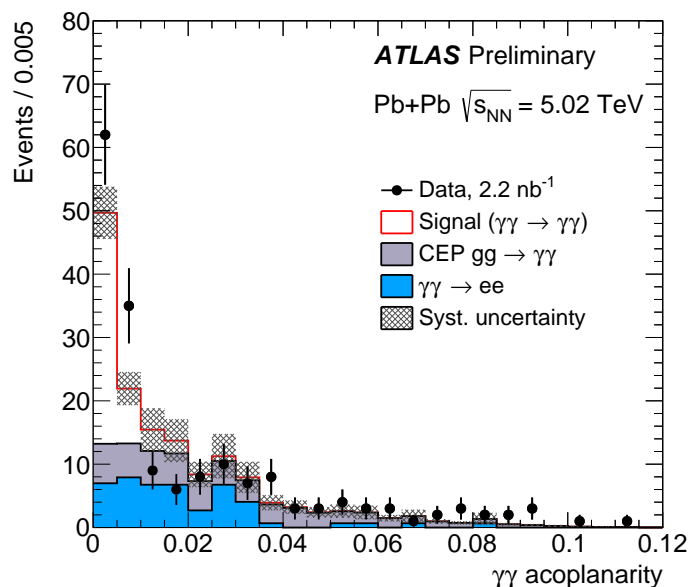


Figure 74: The diphoton acoplanarity distribution for events satisfying signal region selection, but before applying $A < 0.01$ requirement. The CEP background is normalized in the $A > 0.01$ control region. Data are shown as points with statistical error bars, while the histograms represent the expected signal and background levels. The shaded band represents the uncertainties on signal and background predictions, excluding the uncertainty on the luminosity [47].

The background estimation of CEP is done with MC with a validation of the normalization in the $A > 0.01$ control region, which is then used in the SR. The normalization is constrained using the condition:

$$N_{data}(A > 0.01) = n(N_{gg \rightarrow \gamma\gamma}(A > 0.01) + N_{sig}(A > 0.01) + N_{\gamma\gamma \rightarrow e^+e^-}(A > 0.01)) \quad (57)$$

where n is the normalization factor, N_{data} denotes the number of observed events, $N_{gg \rightarrow \gamma\gamma}$ is the CEP expected events from MC simulation, N_{sig} is the number of signal events from MC simulation and $N_{\gamma\gamma \rightarrow e^+e^-}$ is the electron-positron background yield. The electron-positron background is derived by the method shown in 7.5.1.

As the analysis is not yet completed and numerical results for the CEP background estimation are not available. However, the method is identical to the one used in the Run 2 analysis leading to results shown in [47]. The number of $gg \rightarrow \gamma\gamma$ events in the signal region of the Run 2 analysis was estimated to be $N_{gg \rightarrow \gamma\gamma} = 12 \pm 3$, where the uncertainty accounts for shape uncertainties in the acoplanarity region, exact simulation of the ISR gluon radiation and normalization uncertainty.

7.5.3 Fake photons background

As the photons used in the analysis have a low transverse momentum threshold of $p_T > 2$ GeV fake photon signatures can be produced by cosmic muons or noise in the EM calorimeter. In the case of cosmic muons, if they do not cross the pixel detector, no charged track is reconstructed. The background can be estimated by utilizing tracks reconstructed using only the muon spectrometer. In the case of calorimeter noise, different shower-shapes than those of real photons are expected. Thus, the shower-shapes can be used to distinguish between real and fake photons. The methodology and framework used for this classification will be discussed extensively in the next section.

7.6 Photon identification

7.6.1 Motivation

The main working points used by ATLAS for photon identification are called 'loose' and 'tight'. These working points utilize EM calorimeter variables called shower-shapes to produce a classification which discriminates between good quality photons and fake photons [59]. As these working points are created for photons that have high enough p_T to be observed by a photon trigger they are optimized for photons with $p_T > 20$ GeV. The two ATLAS working points were tested for the Run 3 analysis using the LbyL signal MC sample, shown in figure 75. The identifications were imposed on signal photons with $p_T > 2$ GeV and $|\eta| < 2.37$ (excluding calorimeter crack). The efficiency is given by the ratio of identified photons to the total number of reconstructed photons. The tight working point includes more shower-shape variables than the loose working point, resulting in a stricter selection and improved background rejection [59]. In the photon p_T region of 2-5 GeV, where the majority of the LbyL signal is expected according to the MC, the tight working point achieves an efficiency of 18%, while the loose working point has an efficiency of 50%, leading to a significant loss of signal.

Another issue with the default working points is their background rejection. The working points are optimized to reject jets reconstructed as photon backgrounds, where jets deposit part of their energy in the EM calorimeter without associated charged particle tracks. This differs slightly from the background in the LbyL selection, which is mainly caused by calorimeter noise. Therefore, the option to tune the photon identification for this new background should be considered. These issues were identical in the Run 2 analysis. However, the very loose photon working point developed then does not exhibit good behavior for the Run 3 analysis photon p_T threshold of 2 GeV, compared to 2.5 GeV for Run 2.

To summarize a method of photon identification should be developed that optimizes the efficiency for the LbyL $\gamma\gamma \rightarrow \gamma\gamma$ signal and achieves good background rejection. The

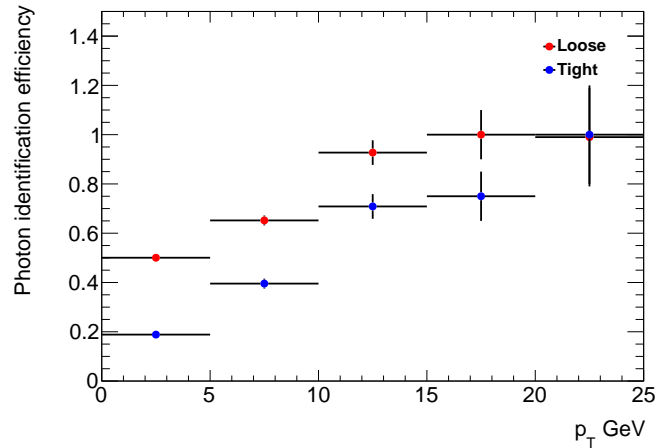


Figure 75: Efficiency of the ATLAS photon identification working points loose and tight for the LbyL signal in photon p_T , with $p_T > 2$ GeV.

method uses machine learning methods to train classifiers that discriminate between photon signal and background.

7.6.2 Signal and background photons

Signal photons are defined based on good-quality photons, according to ATLAS CP recommendations, from LbyL MC simulation. One photon per event is selected, and they are required to pass photon $p_T > 2$ GeV and $|\eta| < 2.37$, excluding the crack region, criteria.

Background photons are selected from a “fake-photon control region” in data. The fake-photon control region is defined by imposing the following criteria:

- LbyL trigger requirement.
- Only one photon candidate per event.
- Photon $p_T > 2.0$ GeV and $|\eta| < 2.37$ excluding the crack region.
- No tracks originating from a primary vertex in the event.

This region is produced using data from the 2023 PbPb UPC data period. Possible LbyL candidate events with two photons are excluded. Events with one photon are selected to compare the photon shower-shape variables between the background and signal. No cut on acoplanarity is imposed on the background photons. This region is dominated by fake-photon events, mostly originating from muons crossing the EM calorimeter and from calorimetric noise.

The photon shower-shapes are variables calculated from measurements in the calorimeters that characterize the shape of the clusters and the detector response to incoming particles [59]. The discriminating information of the classifiers comes from the following

calorimeter shower-shape variables : f_1 , $w_{\eta 2}$, E_{ratio} , $frac_{s1}$, f_{1core} , R_ϕ and R_{had} . Their definitions are given in section 5.4.2. Figures 76–79 show a comparison of the shower-shapes for signal (blue) and background photons (red).

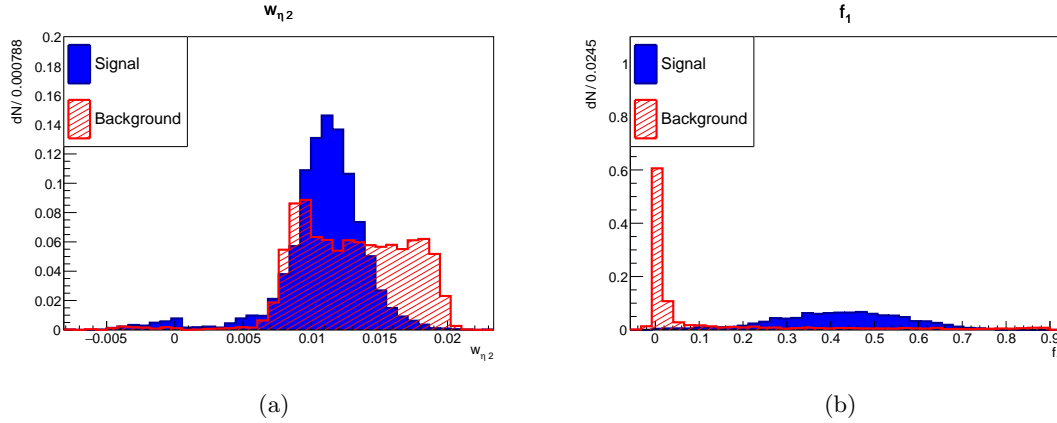


Figure 76: Photon shower-shape variables from background (red) and signal (blue). Shower-shapes $w_{\eta 2}$ (left) and f_1 (right).

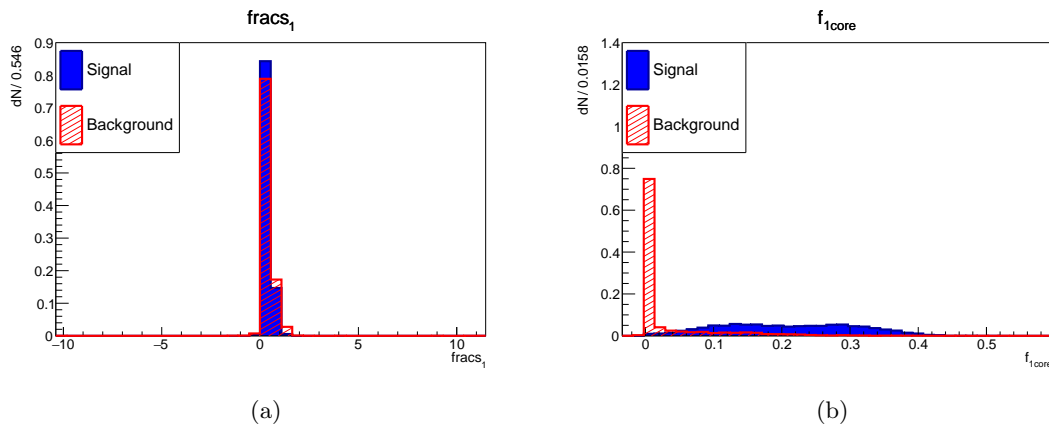


Figure 77: Photon shower-shape variables from background (red) and signal (blue). Shower-shapes $frac_{s1}$ (left) and f_{1core} (right).

The variables that according to the Figures 76-79 have the best discriminating value are E_{ratio} , f_1 and $w_{\eta 2}$. However, correlations in the variables can create a different image during the classification. The discriminating power of f_1 variable is a property of low-energy photons. Low- E_T photons deposit most of their energy in the strips and cannot penetrate very far into the calorimeter. On the other hand, calorimeter noise tends to concentrate at $f_1 \approx 0$. Moreover, particles such as muons lose energy uniformly while propagating through the calorimeter, producing clusters with a small energy fraction in the strip section of the

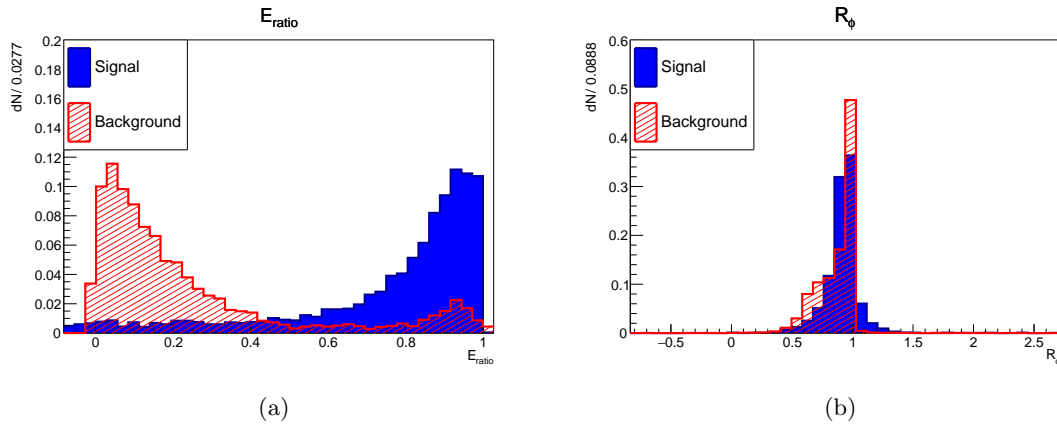


Figure 78: Photon shower-shape variables from background (red) and signal (blue). Shower-shapes E_{ratio} (left) and R_ϕ (right).

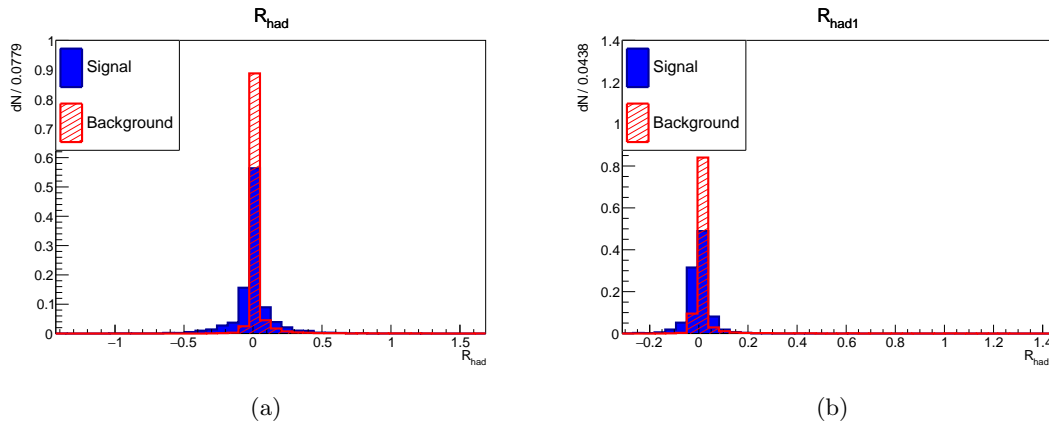


Figure 79: Photon shower-shape variables from background (red) and signal (blue). Shower-shapes R_{had} (left) and R_{had1} (right).

calorimeter [108].

In contrast to common practices for classification using shower-shape variables, p_T reweighting was not applied. This decision was based on the distinct p_T distributions of the signal and background: the signal exhibited an exponential distribution, whereas the background showed an approximately uniform distribution, as illustrated in Figure 80. Applying p_T reweighting would have required assigning large weights to the signal, which was deemed impractical. Nevertheless, steps were taken to ensure that the classifiers did not inadvertently learn to discriminate based on photon p_T , achieved by using a subset of the shower-shape variables. This was verified through checks using both MC signal and data.

In the Run 3 LbyL analysis, two different classifiers were trained for photon identifica-

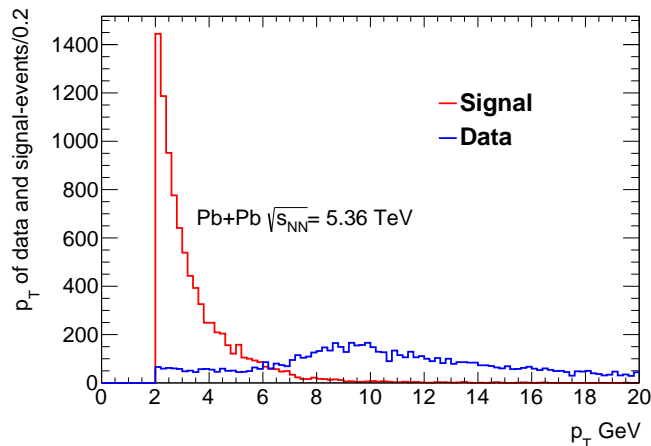


Figure 80: Photon p_T distributions in background (data) and signal photons.

tion (PID): a Boosted Decision Trees (BDT) classifier, described in Section A.5.3, and a Neural Network (NN) classifier, as detailed in Section A.5.2, referred to as NN3 to distinguish it from the Run 2 NN classifier. The primary motivation for training two classifiers was to facilitate comparisons between different classification methods and regions, ultimately selecting the most effective classifier for the analysis. The selection of the most important variables was carried out in two steps for each method. First, the classifier for each eta region, as defined in Table 24, was trained and optimized using all available shower-shape variables. During training, a ranking of the variables was generated based on their separation power, which quantifies the classifier’s ability to discriminate using only that variable, as described in [109]. By comparing the rankings across eta regions, seven variables were identified as the most significant for both classification methods: f_1 , $w_{\eta 2}$, E_{ratio} , $frac{s_1$, f_{1core} , R_ϕ , and R_{had} . These variables were subsequently used to train the classifiers. The performance of the classifiers trained with these seven variables was compared to those trained with the full set of variables, and only a slight decrease in performance was observed when using the reduced set.

The first method employs a BDT classifier, using all seven variables as initial inputs for training. The second method, NN3, utilizes a neural network classifier trained on the same input variables. Since the shower-shapes exhibit a dependence on the pseudorapidity η , both methods are optimized within four η bins, motivated by the structure of the electromagnetic calorimeter. These bins are defined such that two cover the barrel calorimeter and two cover the end-cap region, as shown in Table 24. Both methods are optimized by maximizing the significance, defined as $\sigma = S/\sqrt{S+B}$. The signal efficiencies are around 93 % for both methods, when combining all eta regions. The event population in the eta1 and eta2 regions is significantly larger than that in the eta3 and eta4 regions, consistent with the expected photon pseudorapidity distributions. The original Run 2 NN is also

utilized as an additional working point for validation purposes, with a signal efficiency of approximately 88 %.

Region name	Pseudorapidity in region
eta1	$ \eta < 0.6$
eta2	$0.6 \leq \eta < 1.37$
eta3	$1.52 \leq \eta < 1.81$
eta4	$1.81 \leq \eta < 2.37$

Table 24: Eta regions

7.6.3 Training MC and data

To compare the two methods after optimization, the full dataset was divided into two subsets: the training set and the testing set. The separate testing set is crucial for classifier training, as it ensures the model is not overfitting—meaning it learns only from the training data and fails to generalize well to new data. The two datasets are shown in Table 25.

Dataset	Data	MC
Training	9000 events of UPC - HI Run 2023	9000 events Superchic 3.0 2024 LbyL signal
Testing	2000 events of UPC - HI Run 2023	2000 events Superchic 3.0 2024 LbyL signal

Table 25: Definition of the training and testing dataset

7.6.4 BDT PID

For training, the TMVA framework was used [109]. The inputs were derived from ntuples of data and MC. Multiple architectures with varying numbers of trees, cuts, and bagging were tested. The number of trees was set to 2000 for all regions. Bagging was set at 0.6, representing the relative size of the bagged event sample compared to the original data sample. Bagging is used to reduce statistical fluctuations. Finally, the number of cuts, which determines the number of segments in each variable, was set to 50.

The output value is in between -1 and 1. When the output value is larger than a given threshold a ($y > a$) the input is classified as a photon while for ($y < a$) it is considered background. The single output parameter allows for selecting the desired signal efficiency by adjusting the threshold above which an object is classified as a signal photon, while also providing simple tuning. The threshold is optimized by maximizing $\sigma = S/\sqrt{S+B}$.

For the BDT network the input variables are directly taken from the shower-shape variables, as no further preparation is needed for a BDT. Additionally, every set of input variables is labeled with 1 for a signal photon and 0 for background.

After the training the BDT is evaluated on the testing dataset to ensure that it is neither over-fitted nor under-fitted. The results for the BDT for different eta regions are shown on Table 26. There are only minor differences in the background rejection and

Dataset	eta1	eta2	eta3	eta4
Training background rejection	$97.8 \pm 0.2\%$	$98.0 \pm 0.2\%$	$97.7 \pm 0.5\%$	$97.8 \pm 0.6\%$
Testing background rejection	$96.9 \pm 0.6\%$	$97.3 \pm 0.6\%$	$96.6 \pm 1.4\%$	$98.1 \pm 1.3\%$
Training signal efficiency	$93.8 \pm 0.4 \%$	$92.3 \pm 0.5\%$	$94.0 \pm 0.8\%$	$94.7 \pm 0.7\%$
Testing signal efficiency	$92.5 \pm 0.9\%$	$91.4 \pm 1.0\%$	$91.5 \pm 2.2\%$	$91.7 \pm 2.7\%$
Classifier cut off	-0.4	-0.9	-0.7	-0.2

Table 26: BDT classifier background rejections and signal efficiencies for the different selection methods and datasets.

signal efficiency between training and testing samples, which shows that the BDT is not over-fitted. The differences are within the statistical uncertainties of the testing samples.

7.6.5 NN3 PID

For the DNN network, which is also trained on the TMVA framework, the input variables are normalized using Gaussian-parameter normalization. For every input variable the mean μ and the standard deviation σ are determined such that the mean of all transformed training samples equals to 0 and the standard deviation is 1. Therefore, every input variable has specific parameters for the transformation. Additionally, each set of input variables is labeled with 1 for a signal photon and 0 for background.

A DNN, similar to the one used in Run 2, was trained using the same variables as the BDT. Multiple network architectures were tested, and it was observed that simpler architectures yielded better results. The selected network consists of three hidden layers with 21, 14, and 7 neurons, respectively, and employs rectified linear unit (ReLU) activation functions. The output layer contains a single node, with its value ranging between 0 and 1, set by a sigmoid function. This single output parameter enables the selection of the desired signal efficiency by adjusting the classification threshold for photons to be identified as signal. The threshold is determined by maximizing the significance. To prevent overfitting, dropout layers with a rate of 0.05 were included between the hidden layers during training.

For the training, a common configuration was chosen for all regions. The training dataset is split into two subsets: 90% for the actual training and 10% for validating the network's performance after each iteration. The ADAM optimizer was used, and the mean squared error was chosen as the loss function. The network weights were initialized using the Xavier method, described in section A.5.2. Training was conducted over 200 epochs, with validation accuracy evaluated after each epoch to identify the best configuration and prevent overtraining. An early stopping mechanism was implemented, allowing training

to terminate before completing all 200 epochs if no significant improvement in validation accuracy was observed within 15 consecutive epochs. In most cases, training concluded after approximately 150 epochs. The network optimization was further stabilized by the inclusion of weight decay, which helped prevent erratic changes in the network weights.

After the training the network is evaluated on the testing dataset to ensure that it is neither over-fitted nor under-fitted. The results for the NN3 network for different eta regions are shown in Table 27. There are only minor differences in the background rejection

Dataset	eta1	eta2	eta3	eta4
Training background rejection	$95.5 \pm 0.3\%$	$94.1 \pm 0.4\%$	$87.8 \pm 1.1\%$	$91.0 \pm 1.2\%$
Testing background rejection	$94.2 \pm 0.8\%$	$94.9 \pm 0.7\%$	$90.2 \pm 2.2\%$	$90.6 \pm 2.8\%$
Training signal efficiency	$93.9 \pm 0.4\%$	$92.3 \pm 0.4\%$	$93.0 \pm 0.8\%$	$94.1 \pm 1.0\%$
Testing signal efficiency	$92.6 \pm 0.9\%$	$91.9 \pm 0.9\%$	$94.3 \pm 1.8\%$	$91.6 \pm 2.7\%$
Classifier cut off	0.75	0.75	0.55	0.50

Table 27: NN3 classifier background rejections and signal efficiencies for the different selection methods and datasets.

and signal efficiency between training and testing samples, which shows that the network is not over-fitted. The differences are within the statistical uncertainties of the testing samples.

7.6.6 Classifiers comparison

The comparison of the three classifiers is useful for validation purposes. To facilitate this comparison, the background rejection versus signal efficiency is plotted, commonly referred to as a ROC curve. For each classifier, the efficiencies are derived from the training sample. The results are shown in Figures 81 and 82 for the different eta regions, with the values of the ROC curve intervals provided in Table 28. The results show that the BDT provides the best background rejection and signal efficiency. The NN3 network achieves the second-best performance on this dataset, while the NN2 network ranks third. However, the NN2 network demonstrates comparable performance when the photon transverse momentum cut is adjusted to 2.5 GeV, as used in the Run 2 analysis. Thus, the BDT classifier will be used for PID in the Run 3 LbyL analysis.

In most cases involving shower-shape variables, NN classifiers are considered optimal for classification. However, in this instance, the BDT classifier outperforms the NN. The exact reason for this is not entirely understood, but it may stem from a combination of factors. Firstly, the size of the training samples may be insufficient for the NN to be fully optimized, which is supported by the substantial improvement of the BDT over the NN in the eta3 and eta4 regions, which have the lowest number of events. A second possible cause could be the behavior of certain variables in the background and signal samples,

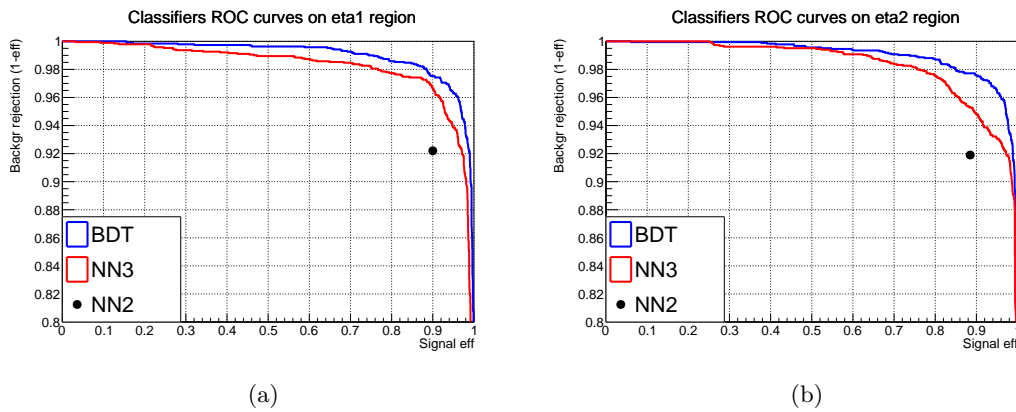


Figure 81: Background rejection vs signal efficiency of the different photon identification methods, for pseudorapidity regions η_1 and η_2 . The results of the BDT network trained for this analysis is shown in blue, while the red curve describes the results of the Run 3 neural network (NN3) and the black dot shows the result of using the Run 2 neural network (NN2) on the new datasets.

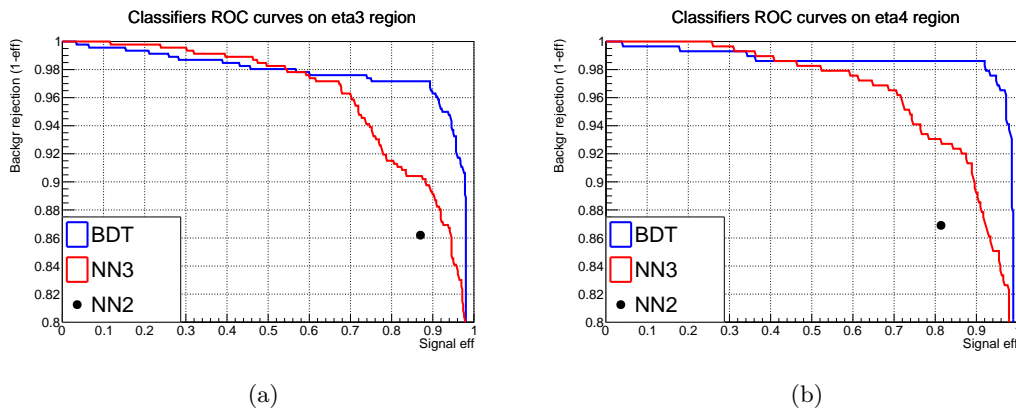


Figure 82: Background rejection vs signal efficiency of the different photon identification methods, for pseudorapidity regions η_3 and η_4 . The results of the BDT network trained for this analysis is shown in blue, while the red curve describes the results of the Run 3 neural network (NN3) and the black dot shows the result of using the Run 2 neural network (NN2) on the new datasets.

specifically the f_1 , f_{1core} , and E_{ratio} variables. For these variables, simple orthogonal cuts can effectively discriminate between signal and background (e.g., $f_1 < 0.1$ exclusively identifies background). The BDT classifier is able to identify these cuts quickly and then optimize on the remaining events. In contrast, the NN classifier may struggle to find these simple cuts, resulting in events being treated as noise, which negatively impacts the optimization. A potential solution would be to apply the cuts before training begins, but

ROC intervals/Eta region	BDT	NN3
eta1	0.985	0.983
eta2	0.988	0.984
eta3	0.976	0.961
eta4	0.977	0.963

Table 28: ROC curves intervals for the different eta regions and classifiers.

this would lead to prohibitively low statistics.

Another important factor to consider is the nature of the problem at hand. The goal of the LbyL analysis is to measure the cross section, which requires a certain level of precision. To achieve this precision, systematic uncertainties in the signal and background estimations must be kept to a minimum. Given that the fake photons background is significant in this analysis, improving its rejection is crucial for the accurate measurement and validation of the analysis. While a slightly lower signal efficiency may introduce some uncertainty in the calculation of the total cross section, this is outweighed by the effect of reducing background uncertainties.

7.6.7 Signal efficiency and background rejection studies

To validate the photon PID efficiencies extracted from simulated samples, the efficiency is also measured using data. This is done with two different methods: the first method evaluates the signal PID efficiency, while the second method assesses the background rejection.

Final-state radiation photons For the first method $\gamma\gamma \rightarrow e^+e^-$ events where one of the electrons is emitting an Final-State Radiation (FSR) photon are used. The selection criteria for the FSR events were determined in the previous analysis [47]. The following selection criteria are imposed to select such events with an FSR photon:

- Any of the specialized triggers mentioned in Section 7.4.1.
- Two-opposite charge tracks with $p_T > 0.5$ GeV.
- Invariant mass of the two-track system with $M_{2tracks} > 1.5$ GeV.
- A photon candidate with a minimum energy of $E_T > 2.0$ GeV and $|\eta| < 2.37$ excluding the crack region is required.

The photon candidate must be separated from both tracks with $\Delta R > 0.30$ to ensure the calorimeter clusters of the electron and the photon candidate are isolated, preventing any energy leakage between them. From simulation, it was determined that a further cut on the transverse momentum of the two track plus photon system $p_{T-2tracks\gamma} < 1\text{GeV}$ is required to suppress contamination from other processes.

This selection yields 2636 photon candidates using the BDT classifier. In Figure 83 the resulting single photon PID efficiencies in dependence of the photons transverse momentum p_T and pseudorapidity η are shown. A good agreement is observed for all the p_T and η regions. In addition, a validation is performed on the classifier signal efficiency.

For the BDT classifier, good agreement between data and the LbyL simulated sample is observed at all p_T regions. In regions with $p_T > 10$ GeV, the statistical uncertainty dominates. The efficiencies as a function of η show deviations of up to 4% in the very forward region, but this is also not considered significant.

The ratio between data and simulation is used as MC-to-data scale factors (SFs) to correct the simulation. For the nominal analysis the SFs parametrized by p_T^γ are used and the SFs parametrized by η are taken as a systematic variation to derive uncertainties.

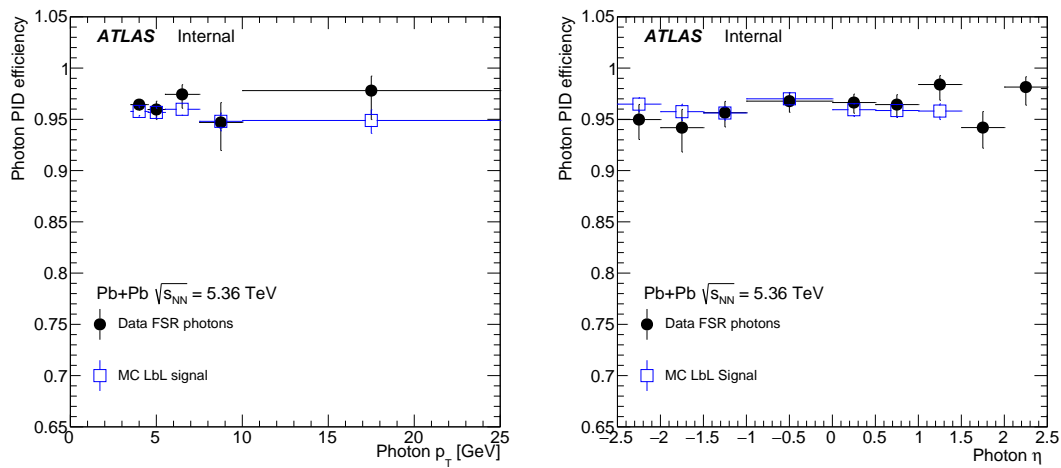


Figure 83: Run3 BDT efficiency as a function of photon p_T (left) and η (right) extracted from FSR event candidates in 2023 PbPb data (Black full markers) and MC simulation (Blue markers).

Background rejection As a second validation of the BDT working point, another data-driven method was employed. Events that were detected with strictly one photon, with $p_T > 2$ GeV and $|\eta| < 2.37$ excluding calorimeter crack, and no quality cuts were used to create a low quality dataset. This will help determine if there is any bias in the classifiers towards certain p_T or η values. If such bias is detected, it could impact the analysis results, particularly the unfolded cross-section measurement. The measured property in question is background rejection—namely, the fraction of events that pass the classifier compared to total events ($1 - N_{passed}/N_{total}$). A comparison is made between the training dataset and the low quality dataset. The results are shown in figure 84.

In Figure 84, no major bias is observed in the photon transverse momentum with respect

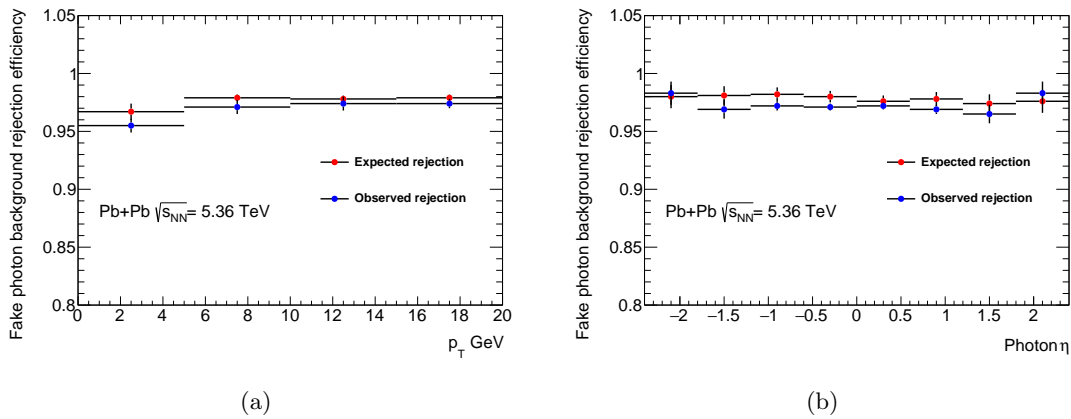


Figure 84: Background rejection expressed in fractional form for the BDT. The expected rejection is calculated using the training dataset, while the observed rejection is calculated using the low quality dataset. The background rejection is shown in correlation with p_T and η .

to background rejection. The average rejection rate is around 97% for the BDT classifier. The slightly reduced rejection in the p_T range of 2–5 GeV is attributed to the detector’s smaller response at these lower energies, which is consistent with expectations based on the calorimeter construction. This smaller response leads to more similar shower-shapes, making background discrimination more challenging.

Similarly, in Figure 84, no significant bias is observed in photon pseudorapidity (η) with respect to background rejection. A smoother trend is seen, largely due to the division into the eta regions, which is the recommended practice by ATLAS for optimizing performance. The background rejection in η is an average of around 97 % in all regions for BDT.

7.6.8 Conclusions

The development and validation of the Run 3 PID classification have been overall successful. The BDT classifier achieves a signal efficiency of approximately 94%, which is comparable to the 95% efficiency of the Run 2 analysis classifier. The background rejection of the BDT classifier is around 97%, slightly lower than the 98% achieved by the Run 2 classifier. These small differences can be attributed to varying detector conditions between Run 2 and Run 3. The conditions for Run 3 are not yet fully understood, leading to reduced photon object quality and less precise luminosity measurements. Additionally, the inclusion of lower photon p_T thresholds introduced more electromagnetic calorimeter noise, which affected classifier performance.

The decision to lower the photon p_T cut is well justified. A comparison of diphoton events in the signal MC indicates that this adjustment increases the number of events by

a factor of 1.5, which compensates for the noise increase, compared to the Run 2 analysis, $p_T > 2.5$ GeV cut. Thus if the experimental challenges are surpassed the integrated fiducial cross section of the Run 3 analysis is expected to be higher than the Run 2 analysis.

Photon identification plays a crucial role in the LbyL analysis, as fake photons contribute significantly to the background estimations for both electron-positron and CEP irreducible backgrounds. Consequently, the photon identification working point must be developed carefully and in advance of the background estimation. In comparison to Run 2, the Run 3 photon classification has been optimized to better reject fake photons in control background regions, resulting in smoother and more reliable background performance. Additionally, the Run 3 classification demonstrates improved consistency when compared to FSR photons, as shown in figure 83. This enhancement reduces the systematic uncertainty associated with photon identification, even in the presence of increased calorimeter noise. The implementation of the classification network using TMVA in Run 3 is also significantly simpler and more intuitive than the network implementation used in Run 2. Overall, the Run 3 photon classification is not only more robust but also more streamlined, effectively addressing several shortcomings observed in the Run 2 classification.

7.7 Results

The Run 3 analysis of the LbyL signal will produce similar results to the Run 2 analysis [47]. The major differences are the lower photon p_T threshold used in Run 3 and also the higher energy of collisions in Run 3. The results will be given in plots of kinematic distributions and integrated fiducial and differential cross section. The integrated fiducial cross section of the Run 2 analysis was measured at $\sigma_{fid} = 120 \pm 17(\text{stat.}) \pm 13(\text{syst.}) \pm 4(\text{lumi.})$ nb compared to the theoretical 80 ± 8 nb [110]. The Run 3 cross section is expected to be higher due to the lower p_T threshold and the higher collision energy. Regarding ALPs in the Run 2 analysis, integrated cross sections above 2 to 70 nb were excluded at the 95 % CL, depending on the diphoton invariant mass in the range 6–100 GeV. Slightly improved performance depending on the ALP selection efficiency is expected in the Run 3 analysis.

8 Conclusion

The $h \rightarrow Za$ analysis uses ATLAS Run 2 pp collision data to search for the rare decay of the 125 GeV Higgs boson into a Z boson and an axion-like particle (ALP- a). The mass of ALP is considered to be between 2 GeV and 33 GeV, with the ALP decaying into a diphoton pair. This analysis reinterprets a previous $h \rightarrow Za$ analysis [3], extending its sensitivity to ALPs with macroscopic lifetimes of up to 1.97 m by searching for displaced photon origin vertices. Additionally, an optimized $h \rightarrow Za$ selection is introduced, providing higher signal efficiency, and its performance is compared to the reinterpretation of the original ATLAS $h \rightarrow Za$ analysis.

No significant discrepancies are observed relative to the expectations from Standard Model backgrounds. Consequently, upper limits on the branching ratio are reported ranging from 0.1 % to 10^{-3} %. These are used to produce exclusion contours in the m_a - $|C_{\gamma\gamma}|/\Lambda$ plane. The two $h \rightarrow Za$ analyses provide constraints across a range of ALP masses and couplings, demonstrating good agreement overall.

The LbyL analysis concerns measurements of light-by-light scattering in lead-lead ultra-peripheral collision events from the 2023 data-taking campaign. Although still ongoing, the analysis aims to become the first to present a direct observation of light-by-light scattering in Run 3 data upon publication. A key component of the study is the development of a dedicated photon identification strategy utilizing advanced classification methods based on machine learning. Thus, signal photons are selected with transverse momentum $p_T > 2$ GeV, achieving high signal efficiency and background rejection.

A Appendix

A.1 Filenames of MC samples

The signal samples used in this analysis are listed for a single campaign (mc16e) in Table 30.

The background samples are listed in table 29.

Full sample name
mc16_13TeV.364502.Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70.deriv.DAOD_STDM3.e5928_s3126_r10724_p3970
mc16_13TeV.364500.Sherpa_222_NNPDF30NNLO_eegamma_pty_7_15.deriv.DAOD_STDM3.e5928_s3126_r10724_p397
mc16_13TeV.364501.Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r10724_p3970
mc16_13TeV.364505.Sherpa_222_NNPDF30NNLO_mumugamma_pty_7_15.deriv.DAOD_STDM3.e5988_s312_r10724_p3970
mc16_13TeV.364506.Sherpa_222_NNPDF30NNLO_mumugamma_pty_15_35.deriv.DAOD_STDM3.e5988_s312_r10724_p3970
mc16_13TeV.364507.Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70.deriv.DAOD_STDM3.e5988_s312_r10724_p3970
mc16_13TeV.364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_STDM3.e5299_s3126_r10724_p3970
mc16_13TeV.364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_STDM3.e5299_s3126_r10724_p3970
mc16_13TeV.364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_STDM3.e5299_s3126_r10724_p3970
mc16_13TeV.364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_STDM3.e5271_s3126_r10724_p3970
mc16_13TeV.364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_STDM3.e5271_s3126_r10724_p3970
mc16_13TeV.364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_STDM3.e5271_s3126_r10724_p3970

Table 29: Overview of generated signal samples with full names of the MC background sets.

Full sample name
mc16_13TeV.603899.PhPy8EG_AZNLO_ggH125_mA2p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603842.PhPy8EG_AZNLO_ggH125_mA3p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603843.PhPy8EG_AZNLO_ggH125_mA4p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603844.PhPy8EG_AZNLO_ggH125_mA5p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603845.PhPy8EG_AZNLO_ggH125_mA6p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603846.PhPy8EG_AZNLO_ggH125_mA7p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603847.PhPy8EG_AZNLO_ggH125_mA9p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603848.PhPy8EG_AZNLO_ggH125_mA12p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603849.PhPy8EG_AZNLO_ggH125_mA15p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603850.PhPy8EG_AZNLO_ggH125_mA20p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603909.PhPy8EG_AZNLO_ggH125_mA25p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603910.PhPy8EG_AZNLO_ggH125_mA30p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603911.PhPy8EG_AZNLO_ggH125_mA33p0_Cyy0p001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603912.PhPy8EG_AZNLO_ggH125_mA3p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603913.PhPy8EG_AZNLO_ggH125_mA4p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603914.PhPy8EG_AZNLO_ggH125_mA5p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603915.PhPy8EG_AZNLO_ggH125_mA6p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603916.PhPy8EG_AZNLO_ggH125_mA7p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603917.PhPy8EG_AZNLO_ggH125_mA9p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603918.PhPy8EG_AZNLO_ggH125_mA12p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603919.PhPy8EG_AZNLO_ggH125_mA15p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603920.PhPy8EG_AZNLO_ggH125_mA20p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603921.PhPy8EG_AZNLO_ggH125_mA25p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603922.PhPy8EG_AZNLO_ggH125_mA30p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603923.PhPy8EG_AZNLO_ggH125_mA33p0_Cyy0p0001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603924.PhPy8EG_AZNLO_ggH125_mA7p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603925.PhPy8EG_AZNLO_ggH125_mA9p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603926.PhPy8EG_AZNLO_ggH125_mA12p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603927.PhPy8EG_AZNLO_ggH125_mA15p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603928.PhPy8EG_AZNLO_ggH125_mA20p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603929.PhPy8EG_AZNLO_ggH125_mA25p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603930.PhPy8EG_AZNLO_ggH125_mA30p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603931.PhPy8EG_AZNLO_ggH125_mA33p0_Cyy0p00001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603932.PhPy8EG_AZNLO_ggH125_mA30p0_Cyy0p000001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726
mc16_13TeV.603933.PhPy8EG_AZNLO_ggH125_mA33p0_Cyy0p000001_Czh1p0.merge.AOD.e8324_e8455_s3889_r10724_r10726

Table 30: Overview of generated signal samples with full names of the MC signal sets.

A.2 Higgs boson theoretical elements

A.2.1 Higgs symmetry breaking mechanism

The Higgs mechanism has an analogy in solid state physics and specifically the explanation of the properties of superconductors. The Meissner effect is the observed exclusion of the Electromagnetic fields from a superconductor. To explain this phenomenon Landau and Ginzburg [111] used an external complex scalar field which is invariant under the abelian U(1) symmetry. For a specific potential the field ϕ would acquire a Vacuum Expectation Value (VEV) of $\langle\phi\rangle = \phi_0$ and break the U(1) symmetry. Then the Lagrangian would acquire a term of the form:

$$\Delta\mathcal{L} = e^2\phi_0^2 A_\mu A^\mu \quad (58)$$

which is a photon mass term, $m_A^2 = 2e^2\phi_0^2$. Then the Electromagnetic field can penetrate the superconductor only to the depth m_A^{-1} .

In the 1960s this mechanism, by which spontaneous symmetry breaking generates a mass for a gauge boson, was explored and generalized to the non-Abelian case by Higgs, Kibble, Guralnik, Hagen, Brout, and Englert, and is now known as the Higgs mechanism [11, 12, 13]. The Higgs mechanism extends straightforwardly to systems with non-Abelian gauge symmetry, and it is described in the following.

Consider a system of scalar fields ϕ_i that appear in a Lagrangian which is invariant under a symmetry group G, which is represented by the following transformation:

$$\phi_i = (1 + i\alpha^\alpha \tau^\alpha)_{ij} \phi_j \quad (59)$$

It is convenient to write the ϕ_i as real-valued fields, for example, writing n complex fields as $2n$ real fields. Then the group representation matrices τ^α , with structure constants α^α , must be pure imaginary and, since they are Hermitian, antisymmetric. So then one can write $\tau_{ij}^\alpha = iT_{ij}^\alpha$ so that the T^α group representation matrices are real and antisymmetric.

Promoting the symmetry group G to a local gauge symmetry, the covariant derivative of ϕ_i is:

$$D_\mu\phi = (\partial_\mu - ig\tau^\alpha A_\mu^\alpha)\phi = (\partial_\mu + gT^\alpha A_\mu^\alpha)\phi \quad (60)$$

Then the kinetic energy of ϕ is:

$$\frac{1}{2}(D_\mu\phi_i)^2 = \frac{1}{2}(\partial_\mu\phi_i)^2 + gA_\mu^\alpha(\partial_\mu\phi_i T_{ij}^\alpha\phi_j) + \frac{1}{2}g^2 A_\mu^\alpha A^{b\mu}(T^\alpha\phi)_i(T^b\phi)_i \quad (61)$$

Now let ϕ_i acquire vacuum expectation values

$$\langle\phi_i\rangle = (\phi_0)_i \quad (62)$$

Then it is possible to expand ϕ_i around these values. The last term of the Lagrangian then contains a form with the structure of a gauge boson mass:

$$\Delta\mathcal{L} = \frac{1}{2}m_{ab}^2 A_\mu^a A^{b\mu} \quad (63)$$

with the mass matrix:

$$m_{ab}^2 = g^2 (T^a \phi_0)_i (T^b \phi_0)_i \quad (64)$$

This matrix is positive semi-definite since any diagonal term, in any basis, acquires a positive mass:

$$m_{aa}^2 = g^2 (T^a \phi_0)^2 \geq 0 \quad (65)$$

There might be some generator T^α which leaves the vacuum G invariant:

$$T^a \phi_0 = 0 \quad (66)$$

In that case there is no contribution to the mass matrix and the gauge boson remains massless.

The generators which accompany spontaneous symmetry breaking are called Goldstone bosons. They are spin zero and they share the quantum numbers of the spontaneously broken internal symmetry generators. The Goldstone theorem describes the phenomenon more thoroughly, which is given as follows: Let a theory L be invariant under transformations of a continuous group G with n generators. If there is a spontaneous symmetry breaking such that the vacuum remains invariant under the action of a subgroup of $G' \subset G$ with $m < n$ generators, then, massless spin-0 particles will emerge in equal number to the generators of G that do not leave the vacuum invariant [112].

A.2.2 Weak bosons and fermion masses

The concept of spontaneous symmetry breaking was introduced in order to solve the problem of the fermion and weak gauge boson masses. The starting point is a $SU(2) \times U(1)$ symmetry with four massless generators which is invariant for the Lagrangian at high energies [9]. The theory at low energies should be a $U(1)_{EM}$ plus massive weak gauge bosons. The way to achieve that is by introducing an additional doublet of scalar fields ϕ with hypercharge $Y = +1$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (67)$$

If ϕ acquires a vacuum expectation value v , $v = 246$ GeV [6], calculated from muon decay, then without loss of generality it can be considered real and in the lower component:

$$\phi = -\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (68)$$

As the field possesses weak isospin and hypercharge its kinetic term can be given by the Lagrangian:

$$\Delta\mathcal{L} = \frac{1}{2} (0 \quad v) (gA_\mu^\alpha \tau^a + \frac{1}{2} g' B_\mu) (gA_\mu^\alpha \tau^a + \frac{1}{2} g' B_\mu) \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (69)$$

The evaluation of the matrix product, using the Pauli matrices properties, leads to:

$$\Delta\mathcal{L} = \frac{1}{2} \frac{v^2}{4} [g^2(A_\mu^1)^2 + g^2(A_\mu^2)^2 + (-gA_\mu^3 + g'B_\mu)^2] \quad (70)$$

As the terms have the form $mA_\mu A^\mu$ there are 3 massive bosons which can be written as:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(A_\mu^1 \pm iA_\mu^2) \quad (71)$$

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gA_\mu^3 - g'B_\mu) \quad (72)$$

The fourth vector field, orthogonal to Z_μ remains massless and is identified as the Electromagnetic vector potential:

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gA_\mu^3 + g'B_\mu) \quad (73)$$

$$m_W = \frac{vg}{2}, \quad m_Z = \sqrt{g^2 + g'^2} \frac{v}{2}, \quad m_A = 0 \quad (74)$$

The masses of the gauge bosons are given by the equations above on tree level, as loop corrections can change the W^\pm and Z masses slightly as they are not protected by any symmetry. The photon remains massless as its mass is protected by the gauge invariance of the $U(1)_{EM}$ group. This leads to a new covariant derivative for the broken symmetry as the covariant derivative of the unbroken symmetry can be expressed with the mass eigenstates of the weak bosons and photon:

$$D_\mu \equiv \partial_\mu - g \frac{1}{\sqrt{2}}(W_\mu^+ T^+ + W_\mu^- T^-) - i \frac{1}{\sqrt{g^2 + g'^2}} Z_\mu (g^2 T^3 - g'^2 Y) - i \frac{gg'}{\sqrt{g^2 + g'^2}} A_\mu (T^3 + Y) \quad (75)$$

Here $T^\pm = (T^1 \pm iT^2)$ so that in the spinor representation of $SU(2)$, $T^\pm = \sigma^\pm$, where σ^\pm are the Pauli matrices for ladder operators. The last term of the covariant derivative shows that the field A_μ couples to a linear combination of the gauge generators of the original $SU(2) \times U(1)$ theory, so it remains massless. The electric charge operator can be identified as the sum of the weak isospin and the hypercharge operators:

$$Q = (T^3 + Y) \quad (76)$$

Also the electric charge e of the electron can be identified as:

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \quad (77)$$

To further simplify the covariant derivative the weak mixing angle θ_w is defined as the angle appearing in the change of basis from (A^3, B) to (Z, A) :

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} \cos \theta_w & -\sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} A^3 \\ B \end{pmatrix} \quad (78)$$

where :

$$\sin \theta_w = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \cos \theta_w = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (79)$$

Hence the covariant derivative for massive gauge bosons is:

$$D_\mu \equiv \partial_\mu - g \frac{1}{\sqrt{2}} (W_\mu^+ T^+ + W_\mu^- T^-) - i \frac{g}{\cos \theta_w} Z_\mu (T^3 - \sin^2 \theta_w) - ie A_\mu Q \quad (80)$$

where $g = e / \sin \theta_w$. This equation shows that the couplings of all of the weak bosons are described by two parameters: the well-measured electron charge e and a new parameter θ_w .

Finally the weak bosons masses are related, on tree level, as:

$$m_W = m_Z \cos \theta_w \quad (81)$$

The phenomenon of ElectroWeak Symmetry Breaking (EWSB) can also be described in the following manner. In the original unbroken $SU(2)$ symmetry the W^\pm and Z bosons are massless. After the symmetry breaking the would be Goldstone bosons get absorbed by the three gauge bosons of the broken symmetry generators. This gives the bosons mass and a associated necessary third polarization degree of freedom. So initially the theory had four massless bosons and three scalars so a total of $4 * 2 + 3 * 1 = 11$ degrees of freedom. After symmetry breaking there are 3 massive bosons and 1 massless boson for a total of $3 * 3 + 1 * 2 = 11$ degrees of freedom, meaning that the theory maintains the same number of degrees of freedom.

When the Higgs mechanism is active the fermions can acquire mass by interacting with the Higgs field. In the Standard Model the left handed leptons $(e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau)$ and the left handed quarks (u, d, s, c, b, t) pair up to transform in the $SU(2)$ representation. There are three generations of doublet pairs of quarks and leptons:

$$Q_L^i = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad (82)$$

$$L_L^i = \begin{pmatrix} e_L \\ \nu_{eL} \end{pmatrix}, \begin{pmatrix} \mu_L \\ \nu_{\mu L} \end{pmatrix}, \begin{pmatrix} \tau_L \\ \nu_{\tau L} \end{pmatrix} \quad (83)$$

Here $i = 1, 2, 3$ corresponding to generations and each element transforms as a left handed Weyl-spinor. The right handed quarks and leptons are organized as singlets:

$$e_R^i = (e_R, \mu_R, \tau_R), \quad \nu_R^i = (\nu_{eR}, \nu_{\mu R}, \nu_{\tau R}), \quad u_R^i = (u_R, c_R, t_R), \quad d_R^i = (d_R, s_R, b_R) \quad (84)$$

Here the right handed neutrino particles are included even though they have never been observed and they are a singlet under the Standard Model symmetries. The charges of the elementary particles in the electroweak symmetry are given in the table 31.

Using 67 for the complex Higgs doublet field ϕ it is possible to write the quark masses as follows, with matrix elements Y :

$$\mathcal{L}_{mass} = -Y_{ij}^d \bar{Q}^i \phi d_R^j - Y_{ij}^u \bar{Q}^i \tilde{\phi} u_R^j + h.c \quad (85)$$

Here i, j are indexes from 1 to 3 for the generations and $\tilde{\phi} \equiv i\sigma^2 \phi$ using the $SU(2)$ group properties. Each term is $SU(3) \times SU(2) \times U(1)$ invariant. This representation of quarks is called the flavour basis. When the Higgs ϕ field acquires a VEV it is possible to rewrite the Lagrangian as:

$$\mathcal{L}_{mass} = -\frac{u}{\sqrt{2}} Y_{ij}^d d_L^i d_R^j - \frac{u}{\sqrt{2}} Y_{ij}^u u_L^i u_R^j + h.c \quad (86)$$

Fermion Family	Left chiral fermions				Right chiral fermions			
		Electric Charge Q	Weak Isospin T3	Weak Hypercharge Y		Electric Charge Q	Weak Isospin T3	Weak Hypercharge Y
Leptons	ν_e, ν_μ, ν_τ	0	1/2	-1	ν_R	0	0	0
	e_L, μ_L, τ_L	-1	-1/2	-1	e_R, μ_R, τ_R	-1	0	-2
Quarks	u_L, c_L, t_L	2/3	1/2	1/3	u_R, c_R, t_R	2/3	0	4/3
	d_L, s_L, b_L	-1/3	-1/2	1/3	d_R, s_R, b_R	-1/3	0	-2/3

Table 31: Electroweak charges of elementary particles

Using the $U(1)_6$ global symmetry of the Standard Model we can freely change basis and replace the Yukawa terms with masses after diagonalizing the Yukawa matrices. This changes the flavour basis to the mass basis m_i . Then the Lagrangian can be written as

$$\mathcal{L}_{mass} = -m_j^d d_L^j d_R^j - m_j^u u_L^j u_R^j + h.c \quad (87)$$

Then the Yukawa coupling, the interaction of form $L \approx g\phi\bar{\psi}\psi$ stemming from a field potential, of the interaction of each quark with the Higgs can be written as $y = \sqrt{2}\frac{m_f}{u}$, which scales linearly with mass. This highlights that the Higgs field interacts more strongly with heavier quarks. Particularly the Yukawa coupling between the top quark and the Higgs is close to unity.

Another consequence of shifting from the flavour matrix to the mass matrix is a quark mixing caused in the kinetic terms. This doesn't influence the B_μ and A_μ^3 fields as the flavour terms do not mix up-and down-type quarks meaning that terms of the form $\bar{u}_R y u_L$

remain unchanged. The only things that are sensitive to the flavor rotations are the W_μ^\pm couplings. Thus the Lagrangian can be expressed as follows:

$$\begin{aligned} \mathcal{L}_{mass-basis} = & \frac{e}{\sin\theta_w} Z_\mu J_\mu^Z + e A_\mu J_\mu^{EM} - m_j^d (\bar{d}_L^j d_R^j + \bar{d}_R^j d_L^j) - m_j^u (\bar{u}_L^j u_R^j + \bar{u}_R^j u_L^j) \\ & + \frac{e}{\sqrt{2}\sin\theta_w} [W_\mu^+ \bar{u}_L^i \gamma^\mu (V)^{ij} d_L^j + W_\mu^- \bar{d}_L^i \gamma^\mu (V^\dagger)^{ij} u_L^j] \end{aligned} \quad (88)$$

where J^Z and J^{EM} are the neutral and electromagnetic currents respectively and where V is the CKM (Cabibbo-Kobayashi-Maskawa) matrix. In the CKM matrix [6] the objects coupling to the up-type quarks via charged-current interactions, are a mixture of various physical flavours of quarks, which can be written in matrix form.

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97373 \pm 0.00031 & 0.2243 \pm 0.0008 & 0.00382 \pm 0.00020 \\ 0.221 \pm 0.004 & 0.975 \pm 0.006 & 0.0408 \pm 0.0014 \\ 0.0086 \pm 0.0002 & 0.0415 \pm 0.0009 & 1.014 \pm 0.029 \end{pmatrix} \quad (89)$$

The CKM matrix is the main observed source of CP violation in the SM. As the matrix is nearly diagonal fewer CP violation than expected is found. The CP violation has a pivotal role in baryogenesis, a physical process that is hypothesized to have taken place during the early universe to produce baryonic asymmetry, the imbalance of matter (baryons) and antimatter (antibaryons). Currently an overabundance of matter is observed in the universe, compared to the SM CP violation.

A similar treatment is possible with the leptons, albeit with only the $\tilde{\phi}$ as the neutrino mass mechanism is currently not explained with the Standard Model. This creates terms of the following form in the Lagrangian:

$$\mathcal{L}_{mass} = -\frac{u}{\sqrt{2}} \lambda_{ij}^e e_L^i e_R^j - \frac{u}{\sqrt{2}} \lambda_{ij}^e e_R^i e_L^j \quad (90)$$

The charged leptons (i.e. the electron, muon, and tau) obtain an effective mass through interaction with the Higgs field, but the neutrinos remain massless. The zero mass of neutrino is in close agreement with current direct experimental observations of the mass which is around 2 eV [6]. Ultimately, the Yukawa couplings of the leptons with the Higgs field are given by $y = \sqrt{2} \frac{m_f}{u}$. This extreme difference between the Yukawa coupling of an electron and the coupling of a top quark is an example of fine tuning and it is one of the biggest mysteries in particle physics.

A.2.3 Unitarity arguments for Higgs boson

Using the problem of weak gauge boson scattering an argument can be established for the usefulness of the Higgs boson in the Standard Model as well as a limit on its mass. Beginning from the Goldstone bosons equivalence theorem [113] which states that for sufficiently

high energies the longitudinal polarization of the weak vector bosons dominates the transverse and the vector boson behaves like the Goldstone boson it absorbed. Considering the scattering of longitudinal W bosons off longitudinal Z bosons leads to the evaluation of:

$$\sigma(W_L^+(p_1)Z_L(p_2) \rightarrow W_L^+(p_3)Z_L(p_4)) \quad (91)$$

The following Feynman diagrams contribute to this process:

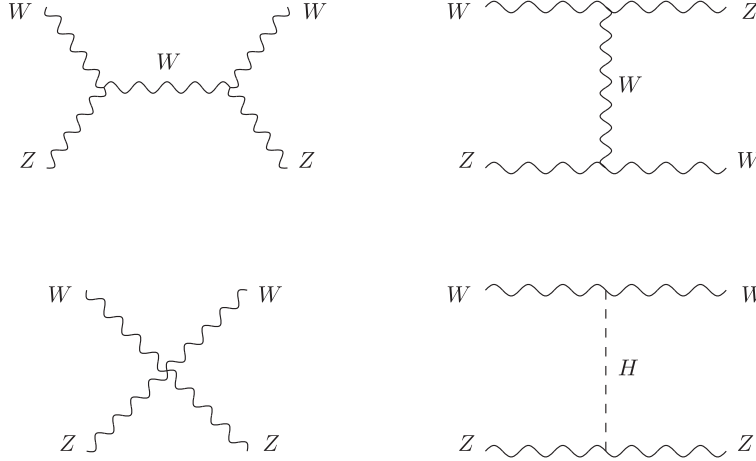


Figure 85: The following four diagrams contribute to the gauge boson scattering at tree level. Top left: s channel W exchange, Top right: u channel W exchange, Bottom left: four point vertex, Bottom right: t channel Higgs exchange [114].

Taking only the diagrams without taking the Higgs exchange into consideration, one can reach the following expression for the matrix element:

$$\mathcal{M}_{tot} = -\frac{m_Z^2}{4m_W^4} e^2 \cot^2 \theta_w (s + u) + O(1) = \frac{t}{v^2} + O(1) \quad (92)$$

where v is the value of the Higgs potential VEV as u is the Mandelstam variable. The cross section is proportional to t so it diverges for high energies. Hence the electroweak force becomes non-perturbative for the TeV energy regime and the force would exhibit similar features like QCD for low energies. The way to regulate this behaviour is by considering the Higgs diagram which contributes in the matrix element:

$$\mathcal{M}_h = -\frac{e^2}{4m_Z^2 \sin^2 \theta_W \cos^2 \theta_W} \frac{t^2 (t - 4m_W^2)(t - 4m_Z^2)}{(t - m_h^2)(t - 2m_W^2)(t - 2m_Z^2)} = -\frac{t^2}{v^2(t - m_h^2)} + O(1) = -\frac{t}{v^2} + O(1) \quad (93)$$

So the high energy behaviour of \mathcal{M}_{tot} is completely canceled by \mathcal{M}_h leading to good ultra-violet behaviour of the theory. However if m_h is too large, the theory is non perturbative until the Higgs diagram contribution starts to occur. Using partial wave unitarity a bound on the Higgs mass is found [115]:

$$m_h \leq \sqrt{\frac{16\pi}{3}} \frac{1}{v} \approx 1 \text{ TeV} \quad (94)$$

This is an interesting result as it motivates that the Higgs boson or something else that serves its function will be discovered at less than the TeV scale, which is accessible by the Large Hadron Collider.

A.3 Standard Model Lagrangian

The combination of the three fundamental forces and the Higgs field is called the Standard model of particle physics. It is a theory of $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry. A way to formulate an invariant Lagrangian of the standard model is the following, reproduced from [111]. Here the hypercharge field is written as B , the weak isospin field is written as W and the chromodynamic field is written as G :

$$\begin{aligned}
\mathcal{L} = & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}G_{\mu\nu}G^{\mu\nu} \\
& \text{(gauge U(1), SU(2), SU(3) self-interaction terms)} \\
& + \bar{L}_L^i \sigma^\mu i D_\mu L_L^i + \bar{e}_R \sigma^\mu i D_\mu e_R + (h.c) \\
& \text{(Lepton Dynamical terms)} \\
& + \bar{Q}_L^i \sigma^\mu i D_\mu Q_L^i + \bar{d}_R \sigma^\mu i D_\mu d_R + \bar{u}_R \sigma^\mu i D_\mu u_R + (h.c) \\
& \text{(Quark dynamical term)} \\
& + |D_\mu \phi|^2 + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \\
& \text{(Higgs dynamical and mass term)} \\
& - Y_{ij}^d \bar{Q}_L^i \phi d_R^j - Y_{ij}^d Q_L^i \tilde{\phi} d_R^j + (h.c) \\
& \text{(Quarks mass terms)} \\
& - Y_{ij}^d \bar{L}_L^i \phi d_R^j - Y_{ij}^d L_L^i \tilde{\phi} d_R^j + (h.c) \\
& \text{(Leptons mass terms)}
\end{aligned} \tag{95}$$

where h.c means hermitian conjugate of the preceding terms. In addition, the σ symbolizes the Pauli matrices and it is the Weyl part of the spinor. The derivative operators D_μ describe the interactions of the fermions and the bosons.

$$D_\mu L_L^i = [\partial_\mu - i\frac{g'}{2}B_\mu + i\frac{g}{2}W_\mu]L_L^i \tag{96}$$

$$D_\mu Q_L^i = [\partial_\mu + i\frac{g'}{6}B_\mu + i\frac{g}{2}W_\mu + ig_s G_\mu]Q_L^i \tag{97}$$

$$D_\mu \nu_R = \partial_\mu \nu_R \tag{98}$$

$$D_\mu e_R = [\partial_\mu - ig' B_\mu]e_R \tag{99}$$

$$D_\mu u_R = [\partial_\mu + i\frac{2g'}{3}B_\mu + ig_s G_\mu]u_R \tag{100}$$

$$D_\mu d_R = [\partial_\mu - i\frac{g'}{3}B_\mu + ig_s G_\mu]d_R \tag{101}$$

$$D_\mu \phi = [\partial_\mu + i\frac{g'}{2}B_\mu + i\frac{g}{2}W_\mu]\phi \tag{102}$$

The quarks form a triplet representation under the $SU(3)$ group which consists of the three colours red, green and blue. The Lagrangian shows a complete picture of the standard

model interactions which is necessary for calculations as the couplings defer in magnitude and a vast extent of possible interactions can originate from relatively simple terms. A map of the possible Standard model interactions is shown below:

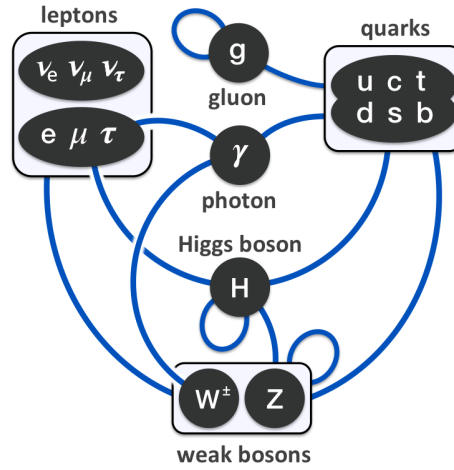


Figure 86: The interactions of the standard model

A.4 Beyond the Standard Model theories

A.4.1 EFT essentials

To understand and explain the phenomenology and properties of extending the SM particle content, a useful tool are Effective Field Theories (EFT). The basic idea underlying the construction of an EFT is that, in a situation where one is faced with a quantum field theory with two (or more) very different energy or length scales, one can construct a simpler theory by performing a systematic expansion in the ratio of these scales [116]. In view of the fact that the Standard Model leaves many questions unanswered, it is plausible that there should exist some “physics beyond the Standard Model” involving new heavy particles with masses M much above the scale of electroweak symmetry breaking or light particles with masses below the electroweak scale that originate from an underlying BSM theory that starts operating at a higher energy Λ .

This can be realized by extending the familiar Standard Model Lagrangian with higher-dimensional local operators built out of Standard Model fields. It is also possible to introduce an entire new field α which can be a scalar, a boson or a fermion and build all the possible operators that include Standard Model fields and this new field. The new operators $O(n)_i$ with mass dimension $D = 4 + n$, where n is positive, must respect the symmetries of the Standard Model, such as Lorentz invariance and gauge invariance. There is of course an infinite set of such operators, but importantly there exists only a finite set of operators for each dimension D , and the contributions of these operators to any given observable are suppressed by powers of $1/\Lambda^{D-4}$ relative to the contributions of the operators of the Standard Model. The EFTs are full-fledged quantum field theories, and one can compute measurable quantities such as cross sections without any reference or input from an underlying UV theory.

Fermi theory of Weak Interactions The classic example of an EFT is the Fermi theory of low-energy weak interactions [117]. The full (UV) theory is the SM, and we can match on to the EFT by transitioning to a theory valid at momenta small compared to $M_{W,Z}$. Since the weak interactions are perturbative, the matching can be done order by order in perturbation theory.

The W boson interacts with leptons with the weak current:

$$j_W^\mu = \nu_l \gamma^\mu P_L l \quad (103)$$

where $P_L = (1 - \gamma^5)/2$ is the left handed projection operator. The tree level amplitude for the muon decay $\mu^- \rightarrow \bar{\nu}_e \nu_\mu e^-$ is

$$\mathcal{A} = \left(\frac{-ig}{\sqrt{2}} \right)^2 (\bar{e} \gamma^\mu P_L \nu_e) (\bar{\nu}_\mu \gamma^\nu P_L \mu) \frac{-ig_{\mu\nu}}{p^2 - M_W^2} \quad (104)$$

where $g/\sqrt{2}$ is the W coupling constant. For low momentum transfer $p \ll M_W$ the propagator can be expanded

$$\frac{1}{p^2 - M_W^2} = -\frac{1}{M_W^2} \left(1 + \frac{p^2}{M_W^2} + \frac{p^4}{M_W^4} + \dots \right) \quad (105)$$

giving different orders in the EFT parameter p/M_W . Retaining only the first term gives

$$\mathcal{A} = \frac{1}{M_W^2} \left(\frac{-ig}{\sqrt{2}} \right)^2 (\bar{e}\gamma^\mu P_L \nu_e)(\bar{\nu}_\mu \gamma^\nu P_L \mu) + O\left(\frac{1}{M_W^4}\right) \quad (106)$$

which is the same amplitude as that produced by the lowest order Lagrangian for muon decay in the EFT. In terms of Feynman diagrams the process is shown in Figure 87. Finally,

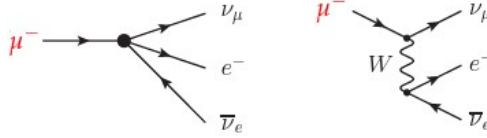


Figure 87: Leading order muon decay in the Fermi theory (left) and the SM (right) [118].

the constant G_F is defined as

$$\frac{G_F}{\sqrt{2}} \equiv \frac{g^2}{8M_W^2} = \frac{1}{2v^2} \quad (107)$$

where v is the scale of electroweak symmetry breaking. So the muon decay Lagrangian is

$$\mathcal{L} = -\frac{4G_F}{\sqrt{2}} (\bar{e}\gamma^\mu P_L \nu_e)(\bar{\nu}_\mu \gamma^\nu P_L \mu) \quad (108)$$

The EFT Lagrangian equation is the low-energy limit of the SM. The EFT no longer has dynamical W bosons, and the effect of W exchange in the SM has been included via dimension-six four-fermion operators. The procedure used here is referred to as “integrating out” a heavy particle, the W boson. The mass of the W boson is treated as large compared to the other scales of the problem. The Lagrangian can be used to calculate the muon decay rate

$$\Gamma(\mu^- \rightarrow \bar{\nu}_e \nu_\mu e^-) = \frac{G_F^2 m_\mu^5}{192\pi^2}. \quad (109)$$

Using the experimental value of the muon lifetime 2.197×10^{-6} s, $G_F \approx 1.16 \times 10^5$ GeV². Using $G_F \approx 1/\Lambda^2$ gives $\Lambda \approx 300$ GeV. This indicates that the EFT has a scale of order Λ . Historically, when the SM was developed, G_F was fixed from muon decay, but the values of M_W and M_Z were not known. Their values were not needed to apply the Fermi theory to low-energy weak interactions or link it to pion decay. Thus an EFT expansion of the SM doesn't need to know the full UV theory to make accurate predictions.

A.4.2 QCD Axion

The strong CP problem, discussed in 2.2.2, can be solved by introducing an additional chiral symmetry which leads to a very natural solution, as one can effectively rotate the θ QCD vacuum configuration away [2] and restore the axial symmetry. Two suggestions were made for the new chiral symmetry, the u-quark could be massless or the Standard Model has an additional global $U(1)$ chiral symmetry.

The suggestion involving the u-quark is fairly simple as the zero mass will allow a rotation of the quark basis with any phase θ which can cancel out the angle, making an exact chiral symmetry. However this has been disproved as the u-quark is confirmed to have a finite mass.

The next suggestion is more promising as it involves introducing a new $U(1)_{PQ}$ chiral symmetry. This was first suggested by Peccei and Quinn in 1977 [119] and it is an elegant solution of the strong CP problem. The introduced $U(1)_{PQ}$ symmetry spontaneously breaks and effectively replaces the static CP-violating angle θ with a dynamical CP-conserving field called the axion. The axion is the Nambu-Goldstone boson of the broken $U(1)_{PQ}$ symmetry [1]. Under a $U(1)_{PQ}$ transformation, the axion field $a(\chi)$ translates as

$$a(\chi) \rightarrow a(\chi) + \alpha f_a \quad (110)$$

where f_a is the energy scale of the symmetry breaking of $U(1)_{PQ}$. The axion is a pseudoscalar particle. To make the Lagrangian invariant the following axion interaction terms have to be included:

$$\mathcal{L} = L_{SM} + \theta \frac{g^2}{32\pi^2} F_{\mu\nu}^\alpha \tilde{F}^{\mu\nu\alpha} - \frac{1}{2} \partial_\mu a \partial^\mu a + L_{int}[\partial^\mu a; \Psi] + \xi \frac{a}{f_a} \frac{g^2}{32\pi^2} F_{\mu\nu}^\alpha \tilde{F}^{\mu\nu\alpha}. \quad (111)$$

The last term ensures that the $U(1)_{PQ}$ current has a chiral anomaly and represents the axion potential [119]. Its minimum occurs at $\langle a \rangle = f_a \theta / \xi$. Expanding around the minimum gives the axion a mass

$$m_a^2 = \left\langle \frac{\partial^2 V_{eff}}{\partial a^2} \right\rangle = -\frac{\xi}{f_a} \frac{g^2}{32\pi^2} \frac{\partial}{\partial a} F_{\mu\nu}^\alpha \tilde{F}^{\mu\nu\alpha} \Big|_{\langle a \rangle = f_a \frac{\theta}{\xi}}. \quad (112)$$

In the original PQ theory the value of f_a coincided with the value of the electroweak symmetry breaking. However due to observed differences in the $K \rightarrow \pi + E_T$ branching ratio found by the KEK experiment the model has been ruled out [119]. However, models with $f_a \gg v$, called invisible axion models, are still viable. Two kinds of models have been proposed the KSVZ and the DFSZ model. The KSVZ model due to Kim, Shifman, Vainshtein and Zakharov [120] introduces a scalar field σ with $f_a = \langle \sigma \rangle \gg v$ and a super-heavy quark Q with $M_Q \approx f_a$ as the only fields carrying PQ charge. Thus, this model doesn't interact with leptons and interacts with the ordinary quarks only on higher orders via the QCD and QED anomalies. The DFSZ model due to Dine, Fischler, Srednicki and

Zhitnisky [121] adds to the original PQ model a scalar field ϕ which carries PQ charge and $f_a = \langle \phi \rangle \gg v$. In this case the quarks and leptons may have tree-level axion couplings. All of these models contain at least one electroweak singlet scalar that acquires a vacuum expectation value and thereby breaks the PQ symmetry [6]. The KSVZ and DFSZ models are frequently used as benchmark examples, but other models exist where both heavy quarks and Higgs doublets carry PQ charges (see [122] for an extensive review).

In the majority of the axion models the axion mass is given by the following formula which includes state of the art calculations in QED and chiral perturbation theory.

$$m_a = 5.691(51) \frac{10^9 \text{GeV}}{f_a} \text{meV}. \quad (113)$$

This value is also confirmed by QCD lattice simulations [6]. The axion mass in most models is below the eV scale.

The axion's two-photon interaction plays a key role for many searches

$$\mathcal{L}_{a\gamma\gamma} = -\frac{g_{a\gamma\gamma}}{4} a F_{\mu\nu} \tilde{F}^{\mu\nu} \quad (114)$$

where \tilde{F} is the dual Electromagnetic tensor. The coupling constant is given by [123]

$$g_{a\gamma\gamma} = \frac{a}{2\pi f_a} \left(\frac{E}{N} - 1.92(4) \right) = \left(0.203(3) \frac{E}{N} - 0.39(1) \right) \frac{m_a}{\text{GeV}^2} \quad (115)$$

where E and N are the electromagnetic and color anomalies of the axial current associated with the axion. In general, a broad range of axions masses and couplings is possible motivating axion searches with multiple distinct experimental methods. The current axion-photon coupling limits per mass are shown in 88.

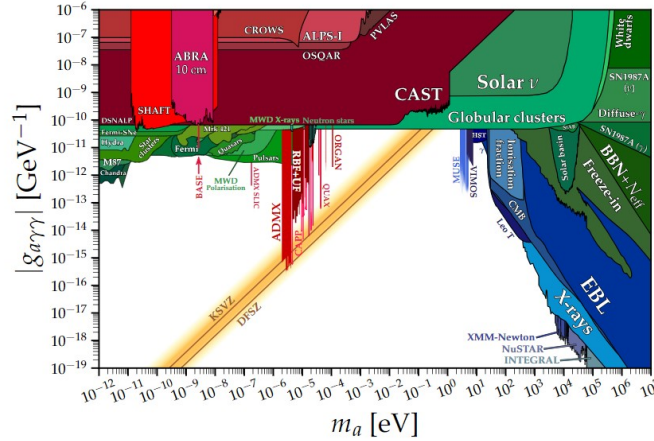


Figure 88: Exclusion plot for the Axion-photon coupling [124].

A.4.3 ALPs

Effective Lagrangian for ALPs The existence of a new spin-0 resonance a is assumed, which is a gauge-singlet under the SM gauge group. Its mass m_a is assumed to be smaller

than the electroweak scale v . A natural way to get such a light particle is by imposing a shift symmetry, $a \rightarrow a + c$, where c is a constant. The UV theory is assumed to be CP invariant, and that CP is broken only by the SM Yukawa interactions. The reason for that is that CP violation in the UV theory can lead to numerous flavour changing currents in the SM [26]. The particle a is supposed to be odd under CP, a pseudoscalar. The most general effective theory Lagrangian with operators up to dimension 5, in the unbroken phase of the electroweak theory, reads [125]

$$\begin{aligned} \mathcal{L}^{D \leq 5} = & \frac{1}{2} \partial_\mu a \partial^\mu a - \frac{m_{a,0}^2}{2} a^2 + \frac{\partial^\mu a}{\Lambda} \sum_F \overline{\psi}_F C_F \gamma_\mu \psi_F \\ & + g_s^2 C_{GG} \frac{a}{\Lambda} G_{\mu\nu}^\alpha \tilde{G}^{\mu\nu\alpha} + g^2 C_{WW} \frac{a}{\Lambda} W_{\mu\nu}^\alpha \tilde{W}^{\mu\nu\alpha} + g_s^2 C_{BB} \frac{a}{\Lambda} B_{\mu\nu} \tilde{B}^{\mu\nu} \end{aligned} \quad (116)$$

where the explicit shift-symmetry breaking mass term $m_{a,0}$ is included. $G_{\mu\nu}^\alpha$, $W_{\mu\nu}^\alpha$ and $B_{\mu\nu}$ are the field strength tensors of $SU(3)_C$, $SU(2)_L$ and $U(1)_Y$, and g_s , g and g' denote the corresponding coupling constants. The dual field strength tensors are defined as $\tilde{B}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} B_{\alpha\beta}$ etc. (with $\epsilon^{0123} = 1$). The sum in the first line extends over the chiral fermion multiplets F of the SM. The quantities C_F are hermitian matrices in generation space. The suppression of the dimension-5 operators is indicated with a new-physics scale Λ , which is the characteristic scale of global symmetry breaking and the mass generation of the ALP, assumed to be above the weak scale. In the literature on axion phenomenology one often eliminates Λ in favor of the "axion decay constant" f_a , defined here such that $\Lambda/|C_{GG}| = 32\pi^2 f_a$ [26], in an analogy with pion physics.

The ALP can receive a mass by means of either an explicit soft breaking of the shift symmetry or through non-perturbative dynamics, like in the case of the QCD axion [1, 26]. In the case of explicit symmetry breaking QCD dynamics provide a mass term given by

$$m_{a,dyn} = 5.691(51) \frac{10^9 \text{ GeV}}{f_a} \text{ meV} \approx 1.8 \text{ MeV} |C_{GG}| \left[\frac{1 \text{ TeV}}{\Lambda} \right]. \quad (117)$$

When an explicit symmetry-breaking mass term $m_{a,0}$ is included in the effective Lagrangian, the resulting mass squared $m_a^2 = m_{a,0}^2 + m_{a,dyn}^2$ becomes a free parameter, with $m_a \ll u$.

In the dimension-5 effective Lagrangian there are no ALP couplings to the Higgs doublet ϕ . The only candidate for such an interaction is

$$O_{Zh} = \frac{\partial^\mu a}{\Lambda} (\phi^\dagger i D_\mu \phi + h.c.) = -\frac{g}{2c_w} \frac{\partial^\mu a}{\Lambda} Z_\mu (u + h)^2 \quad (118)$$

where $c_w \equiv \cos\theta_w$ denotes the cosine of the weak mixing angle, and the last expression holds in unitary gauge. Despite appearance, this operator does not give rise to a tree-level $h \rightarrow Za$ matrix element as the resulting tree-level graphs precisely cancel each other [126].

In addition, the equations of motion of the Higgs doublet and the SM fermions can be used to express the operator as a fermionic operator [126], canceling it from the EFT basis via field redefinitions.

At dimension 6 and higher several additional operators with higgs-ALP coupling can arise. They are [26]

$$\mathcal{L}^{D \geq 6} = \frac{C_{ah}}{\Lambda^2} (\partial^\mu a) (\partial_\mu a) \phi^\dagger \phi + \frac{C'_{ah}}{\Lambda^2} m_{a,0}^2 a^2 \phi^\dagger \phi + \frac{C_{Zh}^{(7)}}{\Lambda^3} (\partial^\mu a) (\phi^\dagger D_\mu \phi + h.c.) + \dots \quad (119)$$

The first two terms are the leading Higgs portal interactions, which give rise to the decay $h \rightarrow aa$. The second term is only allowed if the Lagrangian contains an explicit ALP mass term, as it otherwise violates the shift symmetry. This explains its relative suppression by $m_{a,0}^2/m_h^2$. The third term is the leading operator mediating the decay $h \rightarrow Za$ at tree level [126]. These operators produce the tree level Higgs-ALP interactions at colliders.

After electroweak symmetry breaking (EWSB), the effective Lagrangian contains couplings of the pseudoscalar a to $\gamma\gamma$, $Z\gamma$ and ZZ . Also, the flavour-diagonal currents can couple to a . The relevant terms for ALP decays are shown below:

$$\mathcal{L}_{eff}^{D \leq 5} \in e^2 C_{\gamma\gamma} \frac{a}{\Lambda} F_{\mu\nu} \tilde{F}^{\mu\nu} + \frac{2e^2}{s_w c_w} C_{\gamma Z} \frac{a}{\Lambda} F_{\mu\nu} \tilde{Z}^{\mu\nu} + \frac{e^2}{s_w^2 c_w^2} C_{ZZ} \frac{a}{\Lambda} Z_{\mu\nu} \tilde{Z}^{\mu\nu} + \sum_F \frac{C_{ff}}{2} \frac{\partial^\mu a}{\Lambda} \bar{f} \gamma_\mu \gamma_5 f, \quad (120)$$

where $c_w = \cos \theta_w$ and $s_w = \sin \theta_w$ and

$$C_{\gamma\gamma} = C_{WW} + C_{BB}, C_{\gamma Z} = c_w^2 C_{WW} - s_w^2 C_{BB}, C_{ZZ} = c_w^4 C_{WW} + s_w^4 C_{BB}. \quad (121)$$

All the relevant ALP - SM particle interactions are described by the Feynman rules created by the effective Lagrangian. In addition, it allows the calculation of loop-order effects that lead to subtle ALP phenomena. This creates a versatile and general framework that can describe the behaviour of ALPs over a wide mass range.

Anomalous magnetic moment of the muon & ALPs The persistent deviation of the measured value of the muon anomalous magnetic moment $\alpha_\mu = (g - 2)_\mu/2$ [29] from its SM value provides one of the most compelling hints for new physics. The difference $\alpha_\mu^{exp} - \alpha_\mu^{SM} = (29.3 \pm 7.6) \times 10^{-10}$ which is a 4 sigma difference from zero ⁴. It has been emphasized recently that this discrepancy can be accounted for by an ALP with an enhanced coupling to photons [128]. At one-loop order, the effective Lagrangian gives rise to the contributions to α_μ shown in Figure 89. The first diagram in which a ALP couples to a muon gives the wrong sign [128], but this effect can be overcome by the contribution of the second diagram which involves the photon-ALP or γ -Z coupling, if $C_{\gamma\gamma}$ is sufficiently large. To briefly elaborate according to reference [26], if the Wilson coefficients $c_{\mu\mu}$ and $C_{\gamma\gamma}$ are of similar magnitude, then a logarithmically enhanced contribution is the parametrically

⁴By the time of the thesis publication on new theoretical result [127] was available which greatly reduced the tension between theory and experiment, validating SM.

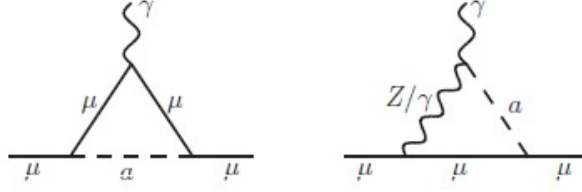


Figure 89: One-loop diagrams contributing to the anomalous magnetic moment of the muon [26].

largest one-loop correction. It gives a positive shift of α_μ provided the product $c_{\mu\mu}C_{\gamma\gamma}$ is negative. The correction proportional to $C_{\gamma Z}$ is suppressed by $(1 - 4s_w^2)$ and hence is numerically subdominant. Finally the $(c_{\mu\mu})^2$ term is suppressed when $m_a^2 \gg m_\mu^2$ while the remaining terms remain unsuppressed.

Figure 90 shows the regions in the parameter space of the couplings $c_{\mu\mu}$ and $C_{\gamma\gamma}$ in which the experimental value of the muon anomalous magnetic moment can be explained in terms of the ALP-induced loop corrections of the Lagrangian. The allowed parameter space increases as the ALP mass increases compared to the muon mass. The explanation of the anomaly is possible without fine tuning as long as the two couplings have the same or similar magnitude.

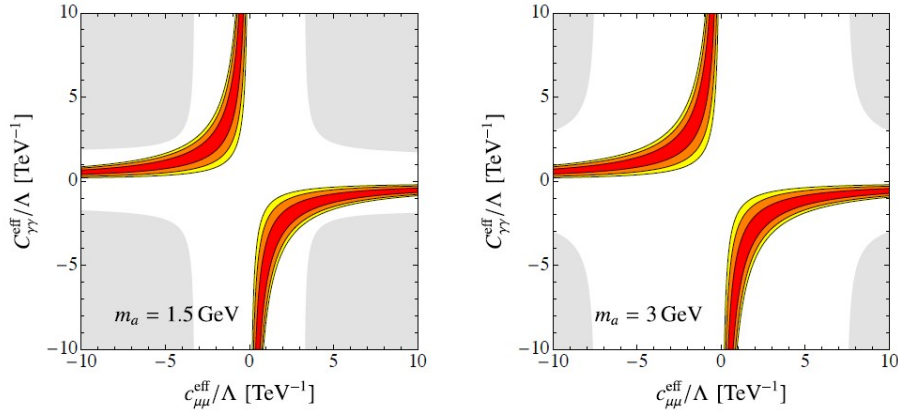


Figure 90: Regions in ALP parameter space where the g-2 anomaly is reproduced at 68/ (red) , 95 / (orange) and 99 / (yellow) CLs for $m_a = 1.5$ and 3 GeV. The grey regions are excluded by Babar dark-photon searches [26].

A.5 Machine Learning

In this thesis Machine learning methods were used by the author to develop a classification for photon identification in the LbyL analysis of chapter 7. The classifiers developed were a Deep Neural Network and a Boosted Decision Tree. A brief description of their general characteristics will be made.

A.5.1 Introduction

Machine learning (ML) is an area of study in computer science concerned with the development and study of statistical algorithms that can learn from data and generalize to unseen data and perform tasks without explicit instructions. Following the explosion of computer performance and availability this area has seen a stupendous advancement in interest and performance. The main advantages of current ML include adaptability to different conditions, capacity of incorporating immense amounts of data and credibility. In modern day particle physics there is a plethora of available data making ML attractive for background estimations and classifications.

In this thesis two different schools of ML will be discussed, with major differences in their philosophy and algorithms: Deep Neural Networks (DNN) and Boosted Decision Trees (BDT). DNNs consist of a network of single computational units and originally were inspired by the neurons and synapses in the human brain, with modern computing power allowing to reach higher levels of complexity compared to older ML methods. BDT is a collection of decision trees, a hierarchical model that uses a tree-like model of cuts or decisions to discriminate between possible outcomes, with boosting being a method of improving tree construction and performance. The architectures and algorithms of these models are extremely different, nonetheless they can be used for the same problem.

In this thesis the ML problem at hand can be described as supervised learning. Supervised means that for the creation of the classifiers datasets with labels were used (0,1). The outcome is the classification of events between these two datasets, creating a classifier.

A.5.2 Deep neural networks

A Deep Neural Network (DNN) is characterized by an architecture of interconnected neurons. The main algorithms that characterize the network is the description of neurons and connections, leading to each neuron producing a certain response at a given set of input signals. A neuron is a mathematical function conceived as a model of a biological neuron and it is the elementary unit of a DNN. These neurons are stacked into layers, a structure in the model's architecture, which takes information from the previous layers and then passes it to the next layer. The model architecture is the structure of the model and its features. The layer where the inputs are introduced is called input layer, the information output

is called output layer and the layers in between are called hidden layers. By applying an external signal to some input neurons the network is put into a defined state that can be measured from the response of one or several output neurons. These networks are capable of learning complex, non-linear relations when trained on a sufficiently large amount of training data [129]. An example of a DNN is shown in figure 91.

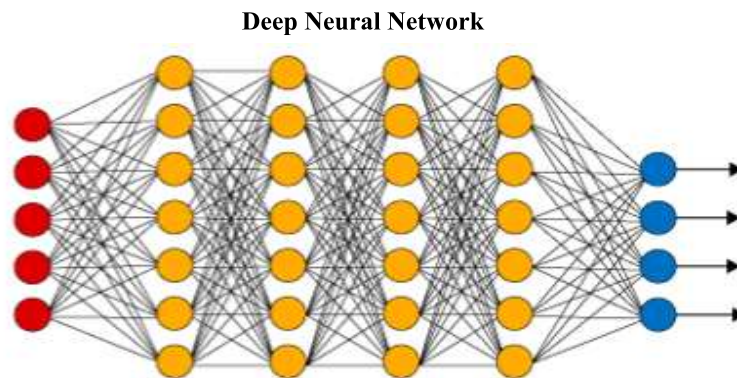


Figure 91: Example of a DNN. Inputs (red) are linked via connections (black lines) through the hidden layers (yellow) to the outputs (blue) [130].

Layers The number of layers is a key property of DNN architecture. The choice of layer types depends on the problem for which the network is used. The simplest layer type is the dense layer. It connects each output of the previous layer to each input of the dense layer. Each neuron to neuron link is assigned a weight and values are passed along the neural links that multiply each input x_k by the weights associated with the corresponding links. The mathematical relation of the weighted sum is given by $\sum^k (w_i x_i) + b$ which is then passed to a activation function. The factor b is the bias and can be used in conjunction with optimization methods. In the beginning all neurons are connected but throughout the training the weights are updated and the weights of connections with zero sway to the network are removed. The main disadvantages is that for each hidden layer added the trainable weight parameters increase, requesting more computing resources and making training harder. Other kinds of layers are also available but this thesis only deals with the dense structure.

Dropout If the network has a sufficient size it starts learning the dataset and not generalizing. This is known as overtraining. To combat this an option is added to remove a percentage of neurons randomly in each iteration during training, called dropout. Dropout introduces noise into the network architecture during training, improving the performance. The training becomes slower but the risk of overfitting is reduced substantially. Depending

on the dataset size and complexity the dropout rate could be set between 0.05 and 0.3.

Activation functions An activation function is the function which is applied to the weighted sum per neuron as shown in Figure 92. Functions like RELU (Rectified Linear Unit) are commonly used, particularly for the hidden layers, given by

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (122)$$

This simple function is easy to implement and interpret algorithmically and the calculation of outputs is straightforward. Functions like this are easily differentiable, allowing for error propagation. Other common activation functions include identity, hyperbolic tangent and Gaussian. When an output between 0 and 1 is preferred, a sigmoid function can be utilized.

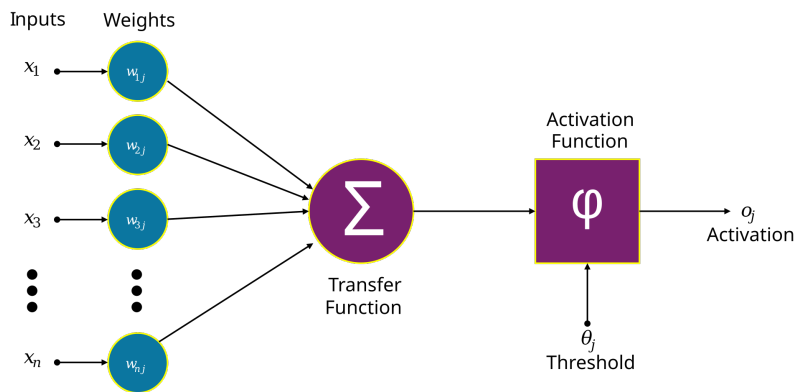


Figure 92: Example of neuron. The inputs x_i are weighted and then combined. A bias b can also be included. Then the sum is applied to the activation function to produce the output.

Training of neural networks The training of the neural network is the optimization of weights and biases of the network so that it describes the desired probability function. To train a network four ingredients are needed: a defined model architecture, a dataset which includes the expected outcomes, an optimizer and a way to evaluate the performance.

Training data The most involved step during the network training is the preparation and optimization of the dataset used. The dataset should be diverse so that the network can observe different patterns and conditions and generalize. It should be unbiased so that it can be flexible but also maintain a degree of simplicity. The dataset should be uniform in the parameter space. This can be implemented mathematically via a normalization or functional method. Non-uniformities in the data tend to bias the network towards thinking that certain parameter regions are more important than others. For variables that may

not be by definition uniform the implementation of events weights can correct the relative importance. Finally, the dataset should be large enough to allow proper training. A general rule of thumb is that the number of the training examples should always significantly exceed the number of trainable parameters, weights and biases. There is no applicable method to predict the necessary size of the dataset to achieve training. Many times in classification problems the size is limited by the number of events in one category to avoid duplicates.

Taking all these considerations into account a training dataset can be constructed. Then to optimize and quantify the network performance the dataset is split into at least three distinct parts. These are:

- Training dataset
- Validation dataset
- Test dataset

The largest fraction of the available data is used for the training dataset. It is used for optimizing the networks parameters during training. The training is usually performed on this dataset multiple times, to achieve gradual improvement. Each iteration over the full set is called an epoch.

The validation dataset is a smaller dataset used to evaluate the network performance after each training epoch. An epoch refers to one full passing of the training dataset through the algorithm. The datasets are independent leading to an unbiased performance measurement. The performances of the training and validation datasets should agree as when the network becomes overtrained the performance of the training dataset tends to be higher than the validation dataset.

Finally, after the network has been optimized and concluded its training, the performance should be evaluated on a testing dataset. This is performed completely by the user and special care should be taken to understand how the network may misbehave. For example the network may use a biased variable during training and one way to observe that is by implementing cuts in this variable and checking if the different segments have similar behavior. The main point to keep in mind is that the testing dataset should be used by the user to ask questions to the network and not as a flat performance metric.

In addition, the receiver operating characteristic (ROC) curve is a representation of the capacity of a binary classifier to separate the two classes, as its discrimination threshold is varied. It is plotting the true positive rate (or a measure of the proportion of actual positives that are correctly identified as such) against the false positive rate (or fall-out, actual negatives improperly identified as positive), obtained when scanning the classifier output. In the context of signal and background, it shows signal efficiency versus background efficiency (or background rejection, defined as $(1 - \text{efficiency})$).

Performance evaluation For the assessment of the network performance some form of metric is needed. This performance measurement is called loss function and it is used during the training. For the classification networks developed in this thesis there are only two possible outcomes, 0 (background) and 1 (signal). A probability p can be applied to the 1 outcome, with $1-p$ corresponding to 0. The binary cross entropy function can be used as loss function, defined as:

$$L(y_{true}, y_{pred}) = \frac{1}{N} \sum_i^N -(y_{i,true} \log(y_{i,pred}) + (1-y_{i,true}) \log(1-y_{i,pred})) \quad (123)$$

where N is the total number of inputs, y_{true} is the real classification label of the event and y_{pred} is the network prediction. This function measures the distance between the correct value and the predicted value and the closer to zero it is, the better the network predictions.

Optimizer Optimizers are algorithms or methods used to minimize the loss function. Optimizers are dependent on model's learnable parameters i.e weights & biases. Optimizers help to know how to change weights and the learning rate of the neural network to reduce losses. One optimizer commonly used is ADAM which will be described in the next section.

Training procedure After the preparation of the four previous steps, the network training can begin. The information is propagated through the network from the input to the output.

The initialization of the weights can be done with multiple methods, with the Xavier method [131] commonly being used where every weight is drawn by a Gaussian distribution whose mean is zero and the standard deviation σ is given by:

$$\sigma = \sqrt{\frac{2}{N_{in} + N_{out}}} \quad (124)$$

where N_{in} is number of each neuron inputs and N_{out} is number of each neuron output.

The data are funneled through the network in batches with an output prediction y_{pred} being produced after each batch. The optimal batch size depends on the problem at hand but normally it is between 50 and 300. The batch size can also affect the convergence of the model, meaning that it can influence the optimization process and the speed at which the model learns. The loss score is calculated by the loss function and then given to the optimizer which updates the weights, to minimize the loss function and the next batch is processed. This process is repeated for the full dataset multiple times with the dataset being shuffled after every epoch to prevent bias of the ordering on the update of the weights. The flow of information is shown in figure 93.

There are many different optimizers available, to find the optimal weights and biases which minimize the loss function. One of the most simple is gradient descent. Gradient descent is based on a convex function and tweaks its parameters iteratively to minimize a

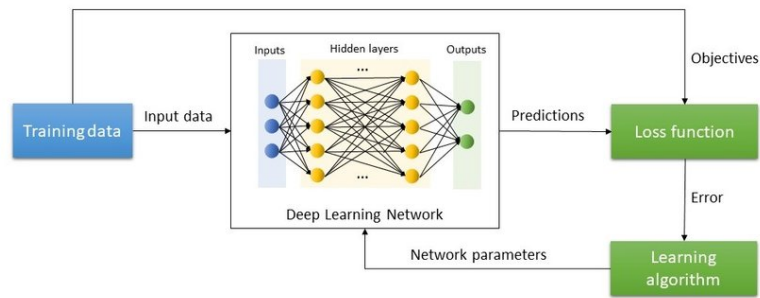


Figure 93: Schematic drawing of the training process in DNN. The flow of information through the network is denoted with arrows [132].

given function to its local minimum. Gradient descent iteratively reduces a loss function by moving in the direction opposite to that of steepest ascent. It is dependent on the derivatives of the loss function for finding minima. It can be written, in one dimension, as:

$$W_{i,j}^{new} = W_{i,j} - a \frac{\partial L(x_b, y_b)}{\partial W_{i,j}} \quad (125)$$

Here $L(x_b, y_b)$ is the value of the loss function corresponding to the randomly chosen input batch x_b and expected output y_b . The indices i and j denote the neurons in the corresponding set of layers, describing the weights W . The parameter a is called the learning rate. It determines the magnitude in which the weights are updated after each iteration. The effects of too small or too large learning rate are shown in figure 94. As the gradient descent method may get stuck in local minima, a parameter can be added in the algorithm called momentum. Momentum simulates the inertia of an object when it is moving, that is, the direction of the previous update is retained to a certain extent during the update, while the current update gradient is used to fine-tune the final update direction. In this way, the network stability can be increased to a certain extent, so that it learns faster.

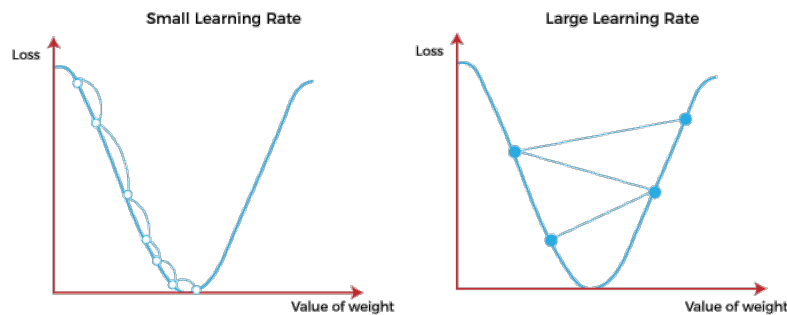


Figure 94: Gradient descent with a small learning rate (left) and large learning rate (right).

The current state of the art optimizer is called adaptive moment estimation (Adam) [133]. It is a method that computes adaptive learning rates for each parameter. It stores

both the decaying average of the past gradients, similar to momentum and also the decaying average of the past squared gradients. This can change the learning rates automatically for each neuron and make the training process faster.

To finish the training a convergence has to be reached when the agreement between the network and the true input can no longer be improved. This is called convergence steps and it is the number of training epochs without improvement in test error to be performed before ending the training phase. With all these customized parameters and algorithms DNN frameworks are very adaptive and a keystone of modern particle physics research.

A.5.3 Boosted decision trees

Decision trees were first developed in 1984 by Breiman et al [134], who proposed the CART algorithm (Classification And Regression Trees) with a complete and functional implementation of decision trees. The concept of a decision tree is to use a sequence of binary decisions to classify events. It can categorize events to many different classes but in this thesis only binary trees will be considered with output 1 (signal) and 0 (background).

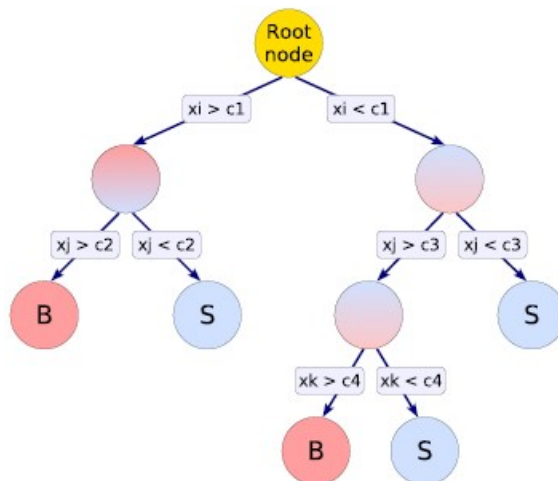


Figure 95: Schematic of a decision tree which discriminates between signal(S) and background(B) [109].

Algorithm Mathematically, decision trees (DT) are rooted binary trees as shown in figure 95. The tree starts from a root node [135]. Each node can be recursively split into two branches, until some stopping condition is reached. The stopping condition where the tree is exited is called leaf. If there was only one variable the goal of the tree would be to find the optimal cut to discriminate between signal and background. It can also find segments inside the variable with sufficient depth. For two variables and more this can be simply extended to a more complicated tree. The optimal cut can be found by

utilizing the cross entropy formula, identically to the DNN, intending to maximize the information gain after the cut is applied. The information concerning cuts is stored as a tree hyperparameter, variables which describe the decision tree implementation.

This is a greedy algorithm, not guaranteed to find the optimal solution. At each node, all variables can be considered, even if they have been used in a previous iteration: this allows to find intervals of interest in a particular variable, instead of being limited to using each variable only once. An advantage of DT over DNN is that they are humanly readable and their cuts create a selection, similar to one a physicist would use. This makes them intuitively much easier to understand and explain.

The tree performance can be very easily quantified. One way is to use the purity, $s/(s+b)$, where s (b) is the sum of weights of signal (background) events that ended up in this leaf during training. Another similar way is via significance, $s/\sqrt{s+b}$, very common for physics processes. Finally, for binary classification purposes, signal or background (mathematically typically +1 for signal and 0 or -1 for background) is assigned depending on whether the purity is above or below a specified critical value.

Tree hyperparameters The number of hyperparameters of a DT is relatively limited and they are mainly concerning the optimizations of the splits used by the DT. The choice of discriminating variables for optimal separation is essential for the construction of a good DT. Another point is the termination of the tree growing which can be done with many methods. For the problems of particle physics discussed in this thesis the main termination criteria are low statistics for further optimization of the node or an insufficient improvement with a larger tree.

The core of a decision tree algorithm resides in how a node is split into two. For this an impurity function $i(t)$ should be developed for every node, which describes the extent of the signal-background mix. It should have a maximum for no separation, equal mix of signal and background and a minimum for perfect separation. In addition, it should optimize both signal and background isolation while rewarding purer nodes. A figure of merit can be constructed with this impurity function, as the decrease of impurity for a split S of node t into two children t_P (pass) and t_F (fail):

$$\Delta i(S, t) = i(t) - p_P i(t_P) - p_F i(t_F) \quad (126)$$

where p_P , p_F are probabilities of passing and failing the split. The goal of the split is to maximize this function. The similarity to the cross entropy function is not accidental as it is one of the widely used functions together with the error miscalculation function. The Gini diversity index is currently the most popular in decision tree implementations [135]. Gini impurity measures how often a randomly chosen element of a set would be incorrectly labeled if it were labeled randomly and independently according to the distribution of labels in the set.

Variable selection Overall, decision trees are very resilient to most factors affecting variables. For decision trees the CPU consumption scales as $nN\log N$ with n variables and N training events. In physics most DTs use between 10-15 variables with as little noise and systematic uncertainties as possible. Unlike DNNs, DTs don't require any variable transformation or normalization. An interesting problem is the selection of a subset of variables to train the decision tree as variables are often correlated, they come in large numbers, and can be more or less discriminating in various regions of the input-feature phase space. A solution is starting with all the given variables and removing sets of variables and comparing the network performance, which will usually be best with all the possible variables but the DT could be overtrained. The user should try to find a set of variables that creates a simple but efficient tree, taking into consideration the problem at hand and technical features the tree should possess.

Pruning Pruning is the process of cutting back a tree from the bottom up after it has been built to its maximum size. Its purpose is to remove statistically insignificant nodes and thus reduce the overtraining of the tree. For simple decision trees it has been found to be beneficial to first grow the tree to its maximum size and then cut back, rather than interrupting the node splitting at an earlier stage. This is because apparently insignificant splits can nevertheless lead to good splits further down the tree. Pruning tends to decrease the performance but increase generalization. For boosted decision trees pruning is not so essential due to the method of boosting.

Boosted decision trees A disadvantage of decision trees is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. This can cause the wrong selection of variables to be cut as the statistical fluctuations make one variable look better than the one that should be used, altering the tree structure drastically. This problem is overcome by training a forest of decision trees and classifying the events on a majority vote of the classifications done by each tree in the forest. The trees are derived from the same training sample with the events subjected to a process known as boosting which assigns weights to them. Boosting increases the statistical stability of the classifier and is able to drastically improve the separation performance compared to a single decision tree.

This process prevents the straightforward interpretation of individual trees in the classifier but creates a stronger classification. The key to this method are the weak classifiers, which are small trees with small depth and taken individually low classification power. By limiting the tree depth during the tree building process (training), the tendency of overtraining for simple decision trees which are typically grown to a large depth and then pruned, is almost completely eliminated.

There are several different boosting algorithms with the most popular for BDTs being

Adaboost and Gradient boost [109]. The Adaboost algorithm is relatively simple. Firstly, one low depth tree is built which misclassifies some events. These events are assigned higher weights than the events that were properly classified, with the weights been calculated from the misclassification rate of the previous tree. Then events are normalized to have the same number of total events. Following this, after a number of trees is trained, the result of the classifier comes from the majority vote of the trees given by :

$$y(x) = \frac{1}{N} \sum_i^N \ln(a_i) h(x) \quad (127)$$

where x is an event input, $h(x)$ the output of a single classifier, N is the number of classifiers and a_i the weight of each classifier. Thus, the algorithm performance is augmented significantly. The algorithm can also utilize adaptable learning rate to alter the weights.

The gradient boost algorithm works by minimizing the difference true value y and model response $F(x)$ obtained from the training sample. A slightly different loss function is used and then the weights are calculated using the current gradient of the loss function and then growing a classification tree whose leaf values are adjusted to match the mean value of the gradient [109]. The learning rate can also be reduced by using a shrinkage parameter, which controls the weight of the individual trees. A small shrinkage (0.1-0.3) demands more trees to be grown but can significantly improve the accuracy of the prediction in difficult settings making more robust trees.

BDT conclusions Overall the BDT classification is extremely powerful for physics problems as it can automatically cope for some of the shortcomings the data have. The main problem it faces is that it is naturally limited by the amount of available data, while also being sensitive to micro-differences in the data quality. Nonetheless, it is a very powerful method and a cornerstone of modern particle physics.

A.5.4 DNNs and BDTs in ATLAS

ATLAS has achieved some remarkable results in the last 15 years and a major reason is the implementation of machine learning in many ATLAS analyses. In particular, DNN and BDT classifiers are widely used for various problems that arise during the analysis that require data driven approaches to be resolved. In this thesis ML is used to develop a photon identification classification for the LbyL analysis of chapter 7. A DNN and a BDT classification was developed. Other networks tested, like MLP, had suboptimal performance. This methodology was used to combat uncertainties from the classification performance. In this analysis for the creation of the classifiers the TMVA framework was used [109]. A more thorough discussion of the classifiers characteristics, training and validation will be made in chapter 7.

Bibliography

- [1] Steven Weinberg. “A New Light Boson?” In: *Phys. Rev. Lett.* 40 (4 Jan. 1978), pp. 223–226. DOI: [10.1103/PhysRevLett.40.223](https://doi.org/10.1103/PhysRevLett.40.223). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.40.223>.
- [2] Roberto D. Peccei. “The Strong CP Problem and Axions”. In: *Axions*. Springer Berlin Heidelberg, 2008, pp. 3–17. ISBN: 9783540735182. DOI: [10.1007/978-3-540-73518-2_1](https://doi.org/10.1007/978-3-540-73518-2_1). URL: http://dx.doi.org/10.1007/978-3-540-73518-2_1.
- [3] ATLAS Collaboration. *Search for the decay of the Higgs boson to a Z boson and a light pseudoscalar particle decaying to two photons*. 2023. arXiv: [2312.01942](https://arxiv.org/abs/2312.01942) [[hep-ex](#)].
- [4] X. Fan et al. “Measurement of the Electron Magnetic Moment”. In: *Physical Review Letters* 130.7 (Feb. 2023). ISSN: 1079-7114. DOI: [10.1103/physrevlett.130.071801](https://doi.org/10.1103/physrevlett.130.071801). URL: <http://dx.doi.org/10.1103/PhysRevLett.130.071801>.
- [5] [File:StandardModelofElementaryParticles.svg-WikimediaCommons.\(n.d.\).WikimediaCommons.https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg](#). [Accessed 13-11-2024].
- [6] R. L. Workman et al. “Review of Particle Physics”. In: *PTEP* 2022 (2022), p. 083C01. DOI: [10.1093/ptep/ptac097](https://doi.org/10.1093/ptep/ptac097).
- [7] ATLAS. “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 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020). URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
- [8] 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. B* 716 (2012), pp. 30–61. DOI: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021). arXiv: [1207.7235](https://arxiv.org/abs/1207.7235) [[hep-ex](#)].
- [9] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley (1995) 842 p. Westview Press, 1995.

- [10] C. Abel et al. “Measurement of the Permanent Electric Dipole Moment of the Neutron”. In: *Phys. Rev. Lett.* 124 (8 Feb. 2020), p. 081803. DOI: [10.1103/PhysRevLett.124.081803](https://doi.org/10.1103/PhysRevLett.124.081803). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.124.081803>.
- [11] 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](https://doi.org/10.1103/PhysRevLett.13.321). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [12] 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](https://doi.org/10.1103/PhysRevLett.13.508). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: [10.1103/PhysRevLett.13.585](https://doi.org/10.1103/PhysRevLett.13.585). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.585>.
- [14] John Ellis. *Higgs Physics*. 2013. arXiv: [1312.5672](https://arxiv.org/abs/1312.5672) [hep-ph].
- [15] Brian Batell, Akshay Ghalsasi, and Mudit Rai. “Dynamics of dark matter misalignment through the Higgs portal”. In: *JHEP* 01 (2024), p. 038. DOI: [10.1007/JHEP01\(2024\)038](https://doi.org/10.1007/JHEP01(2024)038). arXiv: [2211.09132](https://arxiv.org/abs/2211.09132) [hep-ph].
- [16] Javier Rubio. “Higgs Inflation”. In: *Frontiers in Astronomy and Space Sciences* 5 (Jan. 2019). ISSN: 2296-987X. DOI: [10.3389/fspas.2018.00050](https://doi.org/10.3389/fspas.2018.00050). URL: <http://dx.doi.org/10.3389/fspas.2018.00050>.
- [17] Ichiro Oda. “Higgs Potential from Weyl Conformal Gravity”. In: *Mod. Phys. Lett. A* 35.37 (2020), p. 2050304. DOI: [10.1142/S0217732320503046](https://doi.org/10.1142/S0217732320503046). arXiv: [2006.10867](https://arxiv.org/abs/2006.10867) [hep-th].
- [18] A. D. Martin et al. “Parton distributions for the LHC”. In: *The European Physical Journal C* 63.2 (July 2009), pp. 189–285. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-009-1072-5](https://doi.org/10.1140/epjc/s10052-009-1072-5). URL: <http://dx.doi.org/10.1140/epjc/s10052-009-1072-5>.
- [19] Georges Aad et al. “A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery”. In: *Nature* 607.7917 (2022). [Erratum: *Nature* 612, E24 (2022)], pp. 52–59. DOI: [10.1038/s41586-022-04893-w](https://doi.org/10.1038/s41586-022-04893-w). arXiv: [2207.00092](https://arxiv.org/abs/2207.00092) [hep-ex].
- [20] S Dawson et al. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. en. 2011. DOI: [10.5170/CERN-2011-002](https://doi.org/10.5170/CERN-2011-002). URL: <http://cds.cern.ch/record/1318996>.
- [21] Bruce Mellado Garcia et al. “CERN Report 4: Part I Standard Model Predictions”. In: (2016). URL: <https://cds.cern.ch/record/2150771>.

- [22] S Dawson et al. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. en. 2011. DOI: [10.5170/CERN-2011-002](https://doi.org/10.5170/CERN-2011-002). URL: <http://cds.cern.ch/record/1318996>.
- [23] Georges Aad et al. “Evidence of off-shell Higgs boson production from ZZ leptonic decay channels and constraints on its total width with the ATLAS detector”. In: *Phys. Lett. B* 846 (2023), p. 138223. DOI: [10.1016/j.physletb.2023.138223](https://doi.org/10.1016/j.physletb.2023.138223). arXiv: [2304.01532](https://arxiv.org/abs/2304.01532) [hep-ex].
- [24] Aleksandr Azatov et al. “New Prospects for Higgs Compositeness in $h \rightarrow Z\gamma$ ”. In: *Phys. Rev. D* 88.7 (2013), p. 075019. DOI: [10.1103/PhysRevD.88.075019](https://doi.org/10.1103/PhysRevD.88.075019). arXiv: [1308.2676](https://arxiv.org/abs/1308.2676) [hep-ph].
- [25] Qing-Hong Cao et al. “Signature of pseudo Nambu–Goldstone Higgs boson in its decay”. In: *Physics Letters B* 789 (Feb. 2019), pp. 233–237. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2018.12.040](https://doi.org/10.1016/j.physletb.2018.12.040). URL: <http://dx.doi.org/10.1016/j.physletb.2018.12.040>.
- [26] Martin Bauer, Matthias Neubert, and Andrea Thamm. “Collider probes of axion-like particles”. In: *Journal of High Energy Physics* 2017.12 (Dec. 2017). ISSN: 1029-8479. DOI: [10.1007/jhep12\(2017\)044](https://doi.org/10.1007/jhep12(2017)044). URL: [http://dx.doi.org/10.1007/JHEP12\(2017\)044](http://dx.doi.org/10.1007/JHEP12(2017)044).
- [27] ATLAS and CMS Collaborations. *Evidence for the Higgs boson decay to a Z boson and a photon at the LHC*. 2023. arXiv: [2309.03501](https://arxiv.org/abs/2309.03501) [hep-ex].
- [28] Hans Peter Nilles et al. “CP violation from string theory”. In: *Physics Letters B* 786 (2018), pp. 283–287. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2018.09.053>. URL: <https://www.sciencedirect.com/science/article/pii/S0370269318307548>.
- [29] D. P. Aguillard et al. “Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm”. In: *Phys. Rev. Lett.* 131 (16 Oct. 2023), p. 161802. DOI: [10.1103/PhysRevLett.131.161802](https://doi.org/10.1103/PhysRevLett.131.161802). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.131.161802>.
- [30] Ulrich Ellwanger and Stefano Moretti. *Possible Explanation of the Electron Positron Anomaly at 17 MeV in ^8Be Transitions Through a Light Pseudoscalar*. 2016. arXiv: [1609.01669](https://arxiv.org/abs/1609.01669) [hep-ph]. URL: <https://arxiv.org/abs/1609.01669>.
- [31] S. Ambrosanio et al. *SUSY Long-Lived Massive Particles: Detection and Physics at the LHC*. 2000. arXiv: [hep-ph/0012192](https://arxiv.org/abs/hep-ph/0012192) [hep-ph].
- [32] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. *The Physics of the Dark Photon: A Primer*. Springer International Publishing, 2021. ISBN: 9783030625191. DOI: [10.1007/978-3-030-62519-1](https://doi.org/10.1007/978-3-030-62519-1). URL: <http://dx.doi.org/10.1007/978-3-030-62519-1>.

- [33] Xiao-Rui Wong and Ke-Pan Xie. “Freeze-in of WIMP dark matter”. In: *Phys. Rev. D* 108.5 (2023), p. 055035. DOI: [10.1103/PhysRevD.108.055035](https://doi.org/10.1103/PhysRevD.108.055035). arXiv: [2304.00908](https://arxiv.org/abs/2304.00908) [[hep-ph](#)].
- [34] Quentin Bonnefoy, Christophe Grojean, and Jonathan Kley. “Shift-Invariant Orders of an Axionlike Particle”. In: *Physical Review Letters* 130.11 (Mar. 2023). ISSN: 1079-7114. DOI: [10.1103/physrevlett.130.111803](https://doi.org/10.1103/physrevlett.130.111803). URL: <http://dx.doi.org/10.1103/PhysRevLett.130.111803>.
- [35] Antonio Pich. *Effective Field Theory with Nambu-Goldstone Modes*. 2020. arXiv: [1804.05664](https://arxiv.org/abs/1804.05664) [[hep-ph](#)].
- [36] Matthew Kleban and Raul Rabadan. *Collider Bounds on Pseudoscalars Coupling to Gauge Bosons*. 2005. arXiv: [hep-ph/0510183](https://arxiv.org/abs/hep-ph/0510183) [[hep-ph](#)].
- [37] I. Brivio et al. “ALPs effective field theory and collider signatures”. In: *The European Physical Journal C* 77.8 (Aug. 2017). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-017-5111-3](https://doi.org/10.1140/epjc/s10052-017-5111-3). URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5111-3>.
- [38] Ken Mimasu and Verónica Sanz. *ALPs at Colliders*. 2015. arXiv: [1409.4792](https://arxiv.org/abs/1409.4792) [[hep-ph](#)].
- [39] Simon Knapen et al. “Searching for Axionlike Particles with Ultraperipheral Heavy-Ion Collisions”. In: *Physical Review Letters* 118.17 (Apr. 2017). ISSN: 1079-7114. DOI: [10.1103/physrevlett.118.171801](https://doi.org/10.1103/physrevlett.118.171801). URL: <http://dx.doi.org/10.1103/PhysRevLett.118.171801>.
- [40] Gautam Rupak and Elizabeth H. Simmons. “Limits on pseudoscalar bosons from rare Z decays at LEP”. In: *Physics Letters B* 362.1–4 (Nov. 1995), pp. 155–163. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(95\)01152-g](https://doi.org/10.1016/0370-2693(95)01152-g). URL: [http://dx.doi.org/10.1016/0370-2693\(95\)01152-G](http://dx.doi.org/10.1016/0370-2693(95)01152-G).
- [41] David Curtin et al. “Exotic decays of the 125 GeV Higgs boson”. In: *Physical Review D* 90.7 (Oct. 2014). ISSN: 1550-2368. DOI: [10.1103/physrevd.90.075004](https://doi.org/10.1103/physrevd.90.075004). URL: <http://dx.doi.org/10.1103/PhysRevD.90.075004>.
- [42] “Search for new phenomena in events with at least three photons collected in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *The European Physical Journal C* 76.4 (Apr. 2016). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-016-4034-8](https://doi.org/10.1140/epjc/s10052-016-4034-8). URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4034-8>.
- [43] ATLAS Collaboration. *Search for short- and long-lived axion-like particles in $H \rightarrow aa \rightarrow 4\gamma$ decays with the ATLAS experiment at the LHC*. 2023. arXiv: [2312.03306](https://arxiv.org/abs/2312.03306) [[hep-ex](#)].

- [44] “Search for light bosons in decays of the 125 GeV Higgs boson in proton-proton collisions at $\sqrt{s} = 8$ TeV”. In: *Journal of High Energy Physics* 2017.10 (). ISSN: 1029-8479. DOI: [10.1007/jhep10\(2017\)076](https://doi.org/10.1007/JHEP10(2017)076). URL: [http://dx.doi.org/10.1007/JHEP10\(2017\)076](http://dx.doi.org/10.1007/JHEP10(2017)076).
- [45] ATLAS Collaboration. *Search for the decay of the Higgs boson to a Z boson and a light pseudoscalar particle decaying to two photons*. 2023. arXiv: [2312.01942](https://arxiv.org/abs/2312.01942) [[hep-ex](#)].
- [46] Xu Feng et al. “Two-Photon Decay of the Neutral Pion in Lattice QCD”. In: *Physical Review Letters* 109.18 (Nov. 2012). ISSN: 1079-7114. DOI: [10.1103/physrevlett.109.182001](https://doi.org/10.1103/physrevlett.109.182001). URL: <http://dx.doi.org/10.1103/PhysRevLett.109.182001>.
- [47] Georges Aad et al. “Measurement of light-by-light scattering and search for axion-like particles with 2.2 nb⁻¹ of Pb+Pb data with the ATLAS detector”. In: *JHEP* 03 (2021). [Erratum: *JHEP* 11, 050 (2021)], p. 243. DOI: [10.1007/JHEP11\(2021\)050](https://doi.org/10.1007/JHEP11(2021)050). arXiv: [2008.05355](https://arxiv.org/abs/2008.05355) [[hep-ex](#)].
- [48] 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](https://doi.org/10.1007/JHEP08(2016)045). arXiv: [1606.02266](https://arxiv.org/abs/1606.02266) [[hep-ex](#)].
- [49] Charalampos Anastasiou et al. “High precision determination of the gluon fusion Higgs boson cross-section at the LHC”. In: *JHEP* 05 (2016), p. 058. DOI: [10.1007/JHEP05\(2016\)058](https://doi.org/10.1007/JHEP05(2016)058). arXiv: [1602.00695](https://arxiv.org/abs/1602.00695) [[hep-ph](#)].
- [50] Joerg Jaeckel and Michael Spannowsky. “Probing MeV to 90 GeV axion-like particles with LEP and LHC”. In: *Physics Letters B* 753 (Feb. 2016), pp. 482–487. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2015.12.037](https://doi.org/10.1016/j.physletb.2015.12.037). URL: <http://dx.doi.org/10.1016/j.physletb.2015.12.037>.
- [51] CERN. *CERN Annual report 2022*. Tech. rep. Geneva: CERN, 2023. DOI: [10.17181/AnnualReport2022](https://doi.org/10.17181/AnnualReport2022). URL: <https://cds.cern.ch/record/2857560>.
- [52] ATLAS Collaboration. “Luminosity determination in *pp* collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. In: *Eur. Phys. J. C* 83.10 (2023), p. 982. DOI: [10.1140/epjc/s10052-023-11747-w](https://doi.org/10.1140/epjc/s10052-023-11747-w). arXiv: [2212.09379](https://arxiv.org/abs/2212.09379) [[hep-ex](#)].
- [53] 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](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [54] Joao Pequeno and Paul Schaffner. “How ATLAS detects particles: diagram of particle paths in the detector”. 2013. URL: <https://cds.cern.ch/record/1505342>.

- [55] B. Abbott et al. “Production and Integration of the ATLAS Insertable B-Layer”. In: *JINST* 13.05 (2018), T05008. DOI: [10.1088/1748-0221/13/05/T05008](https://doi.org/10.1088/1748-0221/13/05/T05008). arXiv: [1803.00844](https://arxiv.org/abs/1803.00844) [[physics.ins-det](#)].
- [56] Georges Aad et al. “Alignment of the ATLAS Inner Detector in Run-2”. In: *Eur. Phys. J. C* 80.12 (2020), p. 1194. DOI: [10.1140/epjc/s10052-020-08700-6](https://doi.org/10.1140/epjc/s10052-020-08700-6). arXiv: [2007.07624](https://arxiv.org/abs/2007.07624) [[hep-ex](#)].
- [57] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. DOI: [10.17181/CERN.FWRW.F00Q](https://doi.org/10.17181/CERN.FWRW.F00Q). URL: <https://cds.cern.ch/record/331061>.
- [58] ATLAS LAr Collaboration and Margherita Spalla. *ATLAS LAr Calorimeter Performance in LHC Run 2*. Tech. rep. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2707741>.
- [59] Morad Aaboud et al. “Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016”. In: *Eur. Phys. J. C* 79.3 (2019), p. 205. DOI: [10.1140/epjc/s10052-019-6650-6](https://doi.org/10.1140/epjc/s10052-019-6650-6). arXiv: [1810.05087](https://arxiv.org/abs/1810.05087) [[hep-ex](#)].
- [60] Peter Jenni et al. *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. Technical design report. ATLAS. Geneva: CERN, 2008. DOI: [10.17181/CERN-LHCC-2008-004](https://doi.org/10.17181/CERN-LHCC-2008-004). URL: <https://cds.cern.ch/record/1095847>.
- [61] M. Aaboud et al. “Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2”. In: *Eur. Phys. J. C* 77.10 (2017), p. 673. DOI: [10.1140/epjc/s10052-017-5225-7](https://doi.org/10.1140/epjc/s10052-017-5225-7). arXiv: [1704.07983](https://arxiv.org/abs/1704.07983) [[hep-ex](#)].
- [62] 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. C* 76.5 (2016), p. 292. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y). arXiv: [1603.05598](https://arxiv.org/abs/1603.05598) [[hep-ex](#)].
- [63] Georges Aad et al. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *JINST* 14.12 (2019), P12006. DOI: [10.1088/1748-0221/14/12/P12006](https://doi.org/10.1088/1748-0221/14/12/P12006). arXiv: [1908.00005](https://arxiv.org/abs/1908.00005) [[hep-ex](#)].
- [64] 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. C* 79.8 (2019), p. 639. DOI: [10.1140/epjc/s10052-019-7140-6](https://doi.org/10.1140/epjc/s10052-019-7140-6). arXiv: [1902.04655](https://arxiv.org/abs/1902.04655) [[physics.ins-det](#)].
- [65] Georges Aad et al. “Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data”. In: *JINST* 19.02 (2024), P02009. DOI: [10.1088/1748-0221/19/02/P02009](https://doi.org/10.1088/1748-0221/19/02/P02009). arXiv: [2309.05471](https://arxiv.org/abs/2309.05471) [[hep-ex](#)].

- [66] “Jet Reconstruction with the Seeded Cone algorithm in the CMS Phase-2 Level-1 Trigger”. In: (2023). URL: <https://cds.cern.ch/record/2859652>.
- [67] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-ktjet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). URL: <http://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [68] Georges Aad et al. “ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79.11 (2019), p. 970. DOI: [10.1140/epjc/s10052-019-7450-8](https://doi.org/10.1140/epjc/s10052-019-7450-8). arXiv: [1907.05120 \[hep-ex\]](https://arxiv.org/abs/1907.05120).
- [69] Morad Aaboud et al. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77.7 (2017), p. 466. DOI: [10.1140/epjc/s10052-017-5031-2](https://doi.org/10.1140/epjc/s10052-017-5031-2). arXiv: [1703.10485 \[hep-ex\]](https://arxiv.org/abs/1703.10485).
- [70] Georges Aad et al. “Operation of the ATLAS trigger system in Run 2”. In: *JINST* 15.10 (2020), P10004. DOI: [10.1088/1748-0221/15/10/P10004](https://doi.org/10.1088/1748-0221/15/10/P10004). arXiv: [2007.12539 \[physics.ins-det\]](https://arxiv.org/abs/2007.12539).
- [71] Georges Aad et al. “Performance of the ATLAS RPC detector and Level-1 muon barrel trigger at $\sqrt{s} = 13$ TeV”. In: *JINST* 16.07 (2021), P07029. DOI: [10.1088/1748-0221/16/07/P07029](https://doi.org/10.1088/1748-0221/16/07/P07029). arXiv: [2103.01029 \[physics.ins-det\]](https://arxiv.org/abs/2103.01029).
- [72] G. Aad et al. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70 (2010), pp. 823–874. DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9). arXiv: [1005.4568 \[physics.ins-det\]](https://arxiv.org/abs/1005.4568).
- [73] *Summary of ATLAS Pythia 8 tunes*. Tech. rep. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1474107>.
- [74] Gavin Bewick et al. “Herwig 7.3 Release Note”. In: (Dec. 2023). arXiv: [2312.05175 \[hep-ph\]](https://arxiv.org/abs/2312.05175).
- [75] Enrico Bothmann et al. “Event generation with Sherpa 2.2”. In: *SciPost Phys.* 7 (2019), p. 034. DOI: [10.21468/SciPostPhys.7.3.034](https://doi.org/10.21468/SciPostPhys.7.3.034). URL: <https://scipost.org/10.21468/SciPostPhys.7.3.034>.
- [76] GEANT4 Collaboration, 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](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [77] Aram Hayrapetyan et al. “Search for an exotic decay of the Higgs boson into a Z boson and a pseudoscalar particle in proton-proton collisions at $s=13$ TeV”. In: *Phys. Lett. B* 852 (2024), p. 138582. DOI: [10.1016/j.physletb.2024.138582](https://doi.org/10.1016/j.physletb.2024.138582). arXiv: [2311.00130 \[hep-ex\]](https://arxiv.org/abs/2311.00130).

- [78] Tanju Gleisberg and Stefan Hache. “Comix, a new matrix element generator”. In: *Journal of High Energy Physics* 2008.12 (Dec. 2008), p. 039. DOI: [10.1088/1126-6708/2008/12/039](https://doi.org/10.1088/1126-6708/2008/12/039). URL: <https://dx.doi.org/10.1088/1126-6708/2008/12/039>.
- [79] Federico Buccioni et al. “OpenLoops 2”. In: *The European Physical Journal C* 79.10 (Oct. 2019). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-019-7306-2](https://doi.org/10.1140/epjc/s10052-019-7306-2). URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7306-2>.
- [80] F. Cascioli, P. Maierhöfer, and S. Pozzorini. “Scattering Amplitudes with Open Loops”. In: *Phys. Rev. Lett.* 108 (11 Mar. 2012), p. 111601. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.108.111601>.
- [81] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. “Collier: A fortran-based complex one-loop library in extended regularizations”. In: *Computer Physics Communications* 212 (2017), pp. 220–238. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2016.10.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0010465516303320>.
- [82] S. Schumann and F. Krauss. “A parton shower algorithm based on Catani-Seymour dipole factorisation”. In: *Journal of High Energy Physics* 2008.03 (Mar. 2008), p. 038. DOI: [10.1088/1126-6708/2008/03/038](https://doi.org/10.1088/1126-6708/2008/03/038). URL: <https://dx.doi.org/10.1088/1126-6708/2008/03/038>.
- [83] Stefan Höche et al. “A critical appraisal of NLO+PS matching methods”. In: *Journal of High Energy Physics* 2012.9 (Sept. 2012). ISSN: 1029-8479. DOI: [10.1007/jhep09\(2012\)049](https://doi.org/10.1007/jhep09(2012)049). URL: [http://dx.doi.org/10.1007/JHEP09\(2012\)049](http://dx.doi.org/10.1007/JHEP09(2012)049).
- [84] Stefan Höche et al. “QCD matrix elements + parton showers. The NLO case”. In: *Journal of High Energy Physics* 2013.4 (Apr. 2013). ISSN: 1029-8479. DOI: [10.1007/jhep04\(2013\)027](https://doi.org/10.1007/jhep04(2013)027). URL: [http://dx.doi.org/10.1007/JHEP04\(2013\)027](http://dx.doi.org/10.1007/JHEP04(2013)027).
- [85] Stefano Catani et al. “QCD Matrix Elements + Parton Showers”. In: *Journal of High Energy Physics* 2001.11 (Jan. 2002), p. 063. DOI: [10.1088/1126-6708/2001/11/063](https://doi.org/10.1088/1126-6708/2001/11/063). URL: <https://dx.doi.org/10.1088/1126-6708/2001/11/063>.
- [86] Stefan Höche et al. “QCD matrix elements and truncated showers”. In: *Journal of High Energy Physics* 2009.05 (May 2009), p. 053. DOI: [10.1088/1126-6708/2009/05/053](https://doi.org/10.1088/1126-6708/2009/05/053). URL: <https://dx.doi.org/10.1088/1126-6708/2009/05/053>.
- [87] Richard D. Ball et al. “Parton distributions for the LHC run II”. In: *Journal of High Energy Physics* 2015.4 (Apr. 2015). ISSN: 1029-8479. DOI: [10.1007/jhep04\(2015\)040](https://doi.org/10.1007/jhep04(2015)040). URL: [http://dx.doi.org/10.1007/JHEP04\(2015\)040](http://dx.doi.org/10.1007/JHEP04(2015)040).

- [88] Charalampos Anastasiou et al. “High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order”. In: *Phys. Rev. D* 69 (9 May 2004), p. 094008. DOI: [10.1103/PhysRevD.69.094008](https://doi.org/10.1103/PhysRevD.69.094008). URL: <https://link.aps.org/doi/10.1103/PhysRevD.69.094008>.
- [89] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nuclear Physics B* 867.2 (2013), pp. 244–289. ISSN: 0550-3213. DOI: <https://doi.org/10.1016/j.nuclphysb.2012.10.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0550321312005500>.
- [90] ATLAS Collaboration. “ATLAS data quality operations and performance for 2015–2018 data-taking”. In: *JINST* 15 (2020), P04003. DOI: [10.1088/1748-0221/15/04/P04003](https://doi.org/10.1088/1748-0221/15/04/P04003). arXiv: [1911.04632](https://arxiv.org/abs/1911.04632) [[physics.ins-det](https://arxiv.org/abs/1911.04632)].
- [91] Georges Aad et al. “ATLAS data quality operations and performance for 2015–2018 data-taking”. In: *JINST* 15.04 (2020), P04003. DOI: [10.1088/1748-0221/15/04/P04003](https://doi.org/10.1088/1748-0221/15/04/P04003). arXiv: [1911.04632](https://arxiv.org/abs/1911.04632) [[physics.ins-det](https://arxiv.org/abs/1911.04632)].
- [92] Kristof Schmieden et al. *Search for axions in $H \rightarrow aa \rightarrow 4\text{photon}$ decays with the ATLAS experiment at the LHC*. Tech. rep. Geneva: CERN, 2022. URL: <https://cds.cern.ch/record/2824797>.
- [93] ATLAS Collaboration. *Update on the jet energy scale systematic uncertainty for jets produced in proton–proton collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector*. ATLAS-CONF-2011-007. 2011. URL: <https://cds.cern.ch/record/1330713>.
- [94] Johannes Krause and Frank Siegert. “NLO QCD predictions for $Z + \gamma +$ jets production with Sherpa”. In: *The European Physical Journal C* 78.2 (Feb. 2018). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-018-5627-1](https://doi.org/10.1140/epjc/s10052-018-5627-1). URL: <http://dx.doi.org/10.1140/epjc/s10052-018-5627-1>.
- [95] *Recommendations for the Modeling of Smooth Backgrounds*. Tech. rep. All figures available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2020-028>. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2743717>.
- [96] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-021>. Geneva: CERN, 2019. URL: <https://cds.cern.ch/record/2677054>.
- [97] Lukas Heinrich et al. “pyhf: pure-Python implementation of HistFactory statistical models”. In: *Journal of Open Source Software* 6.58 (2021), p. 2823. DOI: [10.21105/joss.02823](https://doi.org/10.21105/joss.02823). URL: <https://doi.org/10.21105/joss.02823>.

- [98] W. Heisenberg and H. Euler. *Consequences of Dirac Theory of the Positron*. 2006. arXiv: [physics/0605038](https://arxiv.org/abs/physics/0605038) [[physics.hist-ph](#)]. URL: <https://arxiv.org/abs/physics/0605038>.
- [99] Barry M. Dillon et al. “KK graviton resonance and cascade decays in warped gravity”. In: *Int. J. Mod. Phys. A* 32.33 (2017), p. 1745006. DOI: [10.1142/S0217751X17450063](https://doi.org/10.1142/S0217751X17450063). arXiv: [1606.07171](https://arxiv.org/abs/1606.07171) [[hep-ph](#)].
- [100] John Ellis, Nick E. Mavromatos, and Tevong You. “Light-by-Light Scattering Constraint on Born-Infeld Theory”. In: *Physical Review Letters* 118.26 (June 2017). ISSN: 1079-7114. DOI: [10.1103/physrevlett.118.261802](https://doi.org/10.1103/physrevlett.118.261802). URL: <http://dx.doi.org/10.1103/PhysRevLett.118.261802>.
- [101] E. J. Williams. “Nature of the High Energy Particles of Penetrating Radiation and Status of Ionization and Radiation Formulae”. In: *Phys. Rev.* 45 (10 May 1934), pp. 729–730. DOI: [10.1103/PhysRev.45.729](https://doi.org/10.1103/PhysRev.45.729). URL: <https://link.aps.org/doi/10.1103/PhysRev.45.729>.
- [102] L. A. Harland-Lang. “Exciting ions: A systematic treatment of ultraperipheral heavy ion collisions with nuclear breakup”. In: *Phys. Rev. D* 107.9 (2023), p. 093004. DOI: [10.1103/PhysRevD.107.093004](https://doi.org/10.1103/PhysRevD.107.093004). arXiv: [2303.04826](https://arxiv.org/abs/2303.04826) [[hep-ph](#)].
- [103] L. Frankfurt et al. “Generalized parton distributions and rapidity gap survival in exclusive diffractive pp scattering”. In: *Phys. Rev. D* 75 (2007), p. 054009. DOI: [10.1103/PhysRevD.75.054009](https://doi.org/10.1103/PhysRevD.75.054009). arXiv: [hep-ph/0608271](https://arxiv.org/abs/hep-ph/0608271) [[hep-ph](#)].
- [104] Spencer R. Klein and Joakim Nystrand. “Photoproduction of quarkonium in proton proton and nucleus nucleus collisions”. In: *Phys. Rev. Lett.* 92 (2004), p. 142003. DOI: [10.1103/PhysRevLett.92.142003](https://doi.org/10.1103/PhysRevLett.92.142003). arXiv: [hep-ph/0311164](https://arxiv.org/abs/hep-ph/0311164) [[hep-ph](#)].
- [105] Spencer R. Klein et al. “STARlight: A Monte Carlo simulation program for ultraperipheral collisions of relativistic ions”. In: *Comput. Phys. Commun.* 212 (2017), pp. 258–268. DOI: [10.1016/j.cpc.2016.10.016](https://doi.org/10.1016/j.cpc.2016.10.016). arXiv: [1607.03838](https://arxiv.org/abs/1607.03838) [[hep-ph](#)].
- [106] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70 (2010), pp. 823–874. DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9). arXiv: [1005.4568](https://arxiv.org/abs/1005.4568) [[physics.ins-det](#)].
- [107] Iwona Grabowska-Bold et al. *Measurement of light-by-light scattering with 2.2 nb¹ of Pb+Pb data in the ATLAS detector – Supporting Note*. Tech. rep. Geneva: CERN, 2019. URL: <https://cds.cern.ch/record/2703499>.
- [108] ATLAS Collaboration. *Reconstruction of collinear final-state-radiation photons in Z decays to muons in $\sqrt{s} = 7$ TeV proton–proton collisions*. ATLAS-CONF-2012-143. 2012. URL: <https://cds.cern.ch/record/1491697>.

- [109] Andreas Hocker et al. “TMVA - Toolkit for Multivariate Data Analysis”. In: (Mar. 2007). arXiv: [physics/0703039](https://arxiv.org/abs/physics/0703039).
- [110] Mariola Klusek-Gawenda, Piotr Lebiedowicz, and Antoni Szczurek. “Light-by-light scattering in ultraperipheral Pb-Pb collisions at energies available at the CERN Large Hadron Collider”. In: *Phys. Rev. C* 93 (4 Apr. 2016), p. 044907. DOI: [10.1103/PhysRevC.93.044907](https://doi.org/10.1103/PhysRevC.93.044907). URL: <https://link.aps.org/doi/10.1103/PhysRevC.93.044907>.
- [111] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Mar. 2014. ISBN: 978-1-107-03473-0, 978-1-107-03473-0.
- [112] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. “Broken Symmetries”. In: *Phys. Rev.* 127 (3 Aug. 1962), pp. 965–970. DOI: [10.1103/PhysRev.127.965](https://doi.org/10.1103/PhysRev.127.965). URL: <https://link.aps.org/doi/10.1103/PhysRev.127.965>.
- [113] John M. Cornwall, David N. Levin, and George Tiktopoulos. “Erratum: Derivation of gauge invariance from high-energy unitarity bounds on the S matrix”. In: *Phys. Rev. D* 11 (4 Feb. 1975), pp. 972–972. DOI: [10.1103/PhysRevD.11.972](https://doi.org/10.1103/PhysRevD.11.972). URL: <https://link.aps.org/doi/10.1103/PhysRevD.11.972>.
- [114] Rafael Delgado et al. “Production of vector resonances at the LHC via WZ-scattering: a unitarized EChL analysis”. In: *Journal of High Energy Physics* 2017 (July 2017). DOI: [10.1007/JHEP11\(2017\)098](https://doi.org/10.1007/JHEP11(2017)098).
- [115] Benjamin W. Lee, C. Quigg, and H. B. Thacker. “Weak interactions at very high energies: The role of the Higgs-boson mass”. In: *Phys. Rev. D* 16 (5 Sept. 1977), pp. 1519–1531. DOI: [10.1103/PhysRevD.16.1519](https://doi.org/10.1103/PhysRevD.16.1519). URL: <https://link.aps.org/doi/10.1103/PhysRevD.16.1519>.
- [116] Matthias Neubert. *Les Houches Lectures on Renormalization Theory and Effective Field Theories*. 2020. arXiv: [1901.06573](https://arxiv.org/abs/1901.06573) [[hep-ph](#)].
- [117] Aneesh V. Manohar. *Introduction to Effective Field Theories*. 2018. arXiv: [1804.05863](https://arxiv.org/abs/1804.05863) [[hep-ph](#)].
- [118] José Ignacio Illana and Alejandro Jiménez Cano. *Quantum field theory and the structure of the Standard Model*. 2022. arXiv: [2211.14636](https://arxiv.org/abs/2211.14636) [[hep-ph](#)].
- [119] R. D. Peccei and Helen R. Quinn. “Constraints imposed by CP conservation in the presence of pseudoparticles”. In: *Phys. Rev. D* 16 (6 Sept. 1977), pp. 1791–1797. DOI: [10.1103/PhysRevD.16.1791](https://doi.org/10.1103/PhysRevD.16.1791). URL: <https://link.aps.org/doi/10.1103/PhysRevD.16.1791>.

- [120] M.A. Shifman, A.I. Vainshtein, and V.I. Zakharov. “Can confinement ensure natural CP invariance of strong interactions?” In: *Nuclear Physics B* 166.3 (1980), pp. 493–506. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(80\)90209-6](https://doi.org/10.1016/0550-3213(80)90209-6). URL: <https://www.sciencedirect.com/science/article/pii/0550321380902096>.
- [121] Michael Dine, Willy Fischler, and Mark Srednicki. “A simple solution to the strong CP problem with a harmless axion”. In: *Physics Letters B* 104.3 (1981), pp. 199–202. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(81\)90590-6](https://doi.org/10.1016/0370-2693(81)90590-6). URL: <https://www.sciencedirect.com/science/article/pii/0370269381905906>.
- [122] Luca Di Luzio et al. “The landscape of QCD axion models”. In: *Physics Reports* 870 (July 2020), pp. 1–117. ISSN: 0370-1573. DOI: [10.1016/j.physrep.2020.06.002](https://doi.org/10.1016/j.physrep.2020.06.002). URL: <http://dx.doi.org/10.1016/j.physrep.2020.06.002>.
- [123] Giovanni Grilli di Cortona et al. “The QCD axion, precisely”. In: *Journal of High Energy Physics* 2016.1 (Jan. 2016). ISSN: 1029-8479. DOI: [10.1007/jhep01\(2016\)034](https://doi.org/10.1007/jhep01(2016)034). URL: [http://dx.doi.org/10.1007/JHEP01\(2016\)034](http://dx.doi.org/10.1007/JHEP01(2016)034).
- [124] Ciaran O’HARE. URL: <https://doi.org/10.5281/zenodo.3932430>.
- [125] Howard Georgi, David B. Kaplan, and Lisa Randall. “Manifesting the Invisible Axion at Low-energies”. In: *Phys. Lett. B* 169 (1986), pp. 73–78. DOI: [10.1016/0370-2693\(86\)90688-X](https://doi.org/10.1016/0370-2693(86)90688-X).
- [126] Martin Bauer, Matthias Neubert, and Andrea Thamm. “Analyzing the CP Nature of a New Scalar Particle via $S \rightarrow Z$ Decays”. In: *Physical Review Letters* 117.18 (Oct. 2016). ISSN: 1079-7114. DOI: [10.1103/physrevlett.117.181801](https://doi.org/10.1103/physrevlett.117.181801). URL: <http://dx.doi.org/10.1103/PhysRevLett.117.181801>.
- [127] A. Boccaletti et al. *High precision calculation of the hadronic vacuum polarisation contribution to the muon anomaly*. 2024. arXiv: [2407.10913 \[hep-lat\]](https://arxiv.org/abs/2407.10913). URL: <https://arxiv.org/abs/2407.10913>.
- [128] W. J. Marciano et al. “Contributions of axionlike particles to lepton dipole moments”. In: *Physical Review D* 94.11 (Dec. 2016). ISSN: 2470-0029. DOI: [10.1103/physrevd.94.115033](https://doi.org/10.1103/physrevd.94.115033). URL: <http://dx.doi.org/10.1103/PhysRevD.94.115033>.
- [129] Pierre Baldi, Peter Sadowski, and Daniel Whiteson. “Searching for Exotic Particles in High-Energy Physics with Deep Learning”. In: *Nature Commun.* 5 (2014), p. 4308. DOI: [10.1038/ncomms5308](https://doi.org/10.1038/ncomms5308). arXiv: [1402.4735 \[hep-ph\]](https://arxiv.org/abs/1402.4735).
- [130] Bossy Mostafa et al. “Machine and Deep Learning Approaches in Genome: Review Article”. In: *Alfarama Journal of Basic Applied Sciences* (Aug. 2020). DOI: [10.21608/ajbas.2020.34160.1023](https://doi.org/10.21608/ajbas.2020.34160.1023).

- [131] Y. Bengio. “Learning Deep Architectures for AI”. In: *Foundations 2* (Jan. 2009), pp. 1–55. DOI: [10.1561/22000000006](https://doi.org/10.1561/22000000006).
- [132] Irene Rivas-Blanco et al. “A Review on Deep Learning in Minimally Invasive Surgery”. In: *IEEE Access* PP (Mar. 2021), pp. 1–1. DOI: [10.1109/ACCESS.2021.3068852](https://doi.org/10.1109/ACCESS.2021.3068852).
- [133] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: [1412.6980](https://arxiv.org/abs/1412.6980) [cs.LG]. URL: <https://arxiv.org/abs/1412.6980>.
- [134] Leo Breiman. *Classification and regression trees*. Routledge, 2017.
- [135] Yann Coadou. “Boosted Decision Trees”. In: *Artificial Intelligence for High Energy Physics*. WORLD SCIENTIFIC, Feb. 2022, pp. 9–58. ISBN: 9789811234033. DOI: [10.1142/9789811234033_0002](https://doi.org/10.1142/9789811234033_0002). URL: http://dx.doi.org/10.1142/9789811234033_0002.

GEORGIOS LAMPRINOUDIS

giorgos.lampr812@gmail.com / glamprin@uni-mainz.de

ACADEMIC CAREER

Johannes Gutenberg University of Mainz *February 2022 - February 2025*

PhD in experimental particle physics
Department of Physics, Mathematics and computing
Grade: Magna cum laude

University of Glasgow *September 2019 - September 2020*

Master in Theoretical physics
Department of Physics and Astronomy
Grade: Distinction

National Technical University of Athens *September 2013 - October 2018*

Master of Electrical and Computer Engineering.
Grade: 7.44/10

CURRENT CAREER OBJECTIVE

Particle physics is a field where CERN research is at the forefront. The Run 3 era of ATLAS analysis should serve as an impetus for the establishment and search for signatures that challenge the traditional understanding of particle physics, combining phenomenology with new ideas. I would be extremely honored to contribute to this effort.

WORK EXPERIENCE

PRISMA+ internship on HGTD - University of Mainz *September 2021 - December 2021*

This internship was on the HL ATLAS High Granularity Timing Detector. It was supervised by Prof. Lucia Masetti (lucia.masetti@cern.ch) and Dr. Maria Manzano. My responsibilities included data analysis, measurements on the flex tails of the modules and presentation of results in ATLAS meetings.

PHD in Particle physics - University of Mainz *February 2022 - February 2025*

The PhD topic is the search for axion-like particles using the ATLAS experiment. The supervisors are Prof. Matthias Schott and Dr. Kristof Schmieden. This opportunity provided me with insight into the daily life of a particle physicist, and I gained valuable knowledge by attending and teaching lectures during the PhD. The main projects pursued during the PhD are outlined below.

RESEARCH PROJECTS

Grand Unification - MSc final project

This MSc project was a brief description of the Standard Model and the SU(5), Pati-Salam and SO(10) unification schemes. The project was supervised by Dr. David Miller (david.j.miller@glasgow.ac.uk).

ATLAS authorship qualification task - PhD

This project was supervised by Dr. Gerardo Vasquez. The main task was to include the NSW STG detector histograms into the Tier0 web display. Various flag algorithms were designed and implemented to improve the automation of the Muon data quality Tier0 shifting. The configuration is still used for Run 3 data. After the project, I continued to shift for MuonDQ at regular intervals.

hZa displaced vertices analysis - PhD

The main project of the PhD thesis. A reinterpretation of the ATLAS prompt hZa ($Z \rightarrow l^+l^-$, $a \rightarrow \gamma\gamma$) analysis to extend the sensitivity to long-lived ALPs. The analysis required experience with the ntuple framework which operated with C++ and the histfitter network in Python. A PUB note is nearing publication, including also results on the haa channel.

LbL HI analysis - PhD

The secondary project of the PhD thesis focuses on ultra-peripheral collisions of lead atoms recorded during Run 3 of ATLAS, which are sensitive to light-by-light scattering. Since the candidate light-by-light events produce photons with lower energy than the photons recommended by ATLAS, a classification is vital for selecting photons of good quality. Classifiers were trained using the TMVA framework and successfully tested on background-rich data regions. The ATLAS paper is nearing publication.

PUBLICATIONS

ATL-COM-PHYS-2024-982 - Reinterpretation of the ATLAS search for the decay of the Higgs boson to a Z boson and a light pseudoscalar particle decaying to two photons in terms of long-lived axion-like particles

ATL-COM-PHYS-2024-1025 - Reinterpretation of the ATLAS search for the decay of the Higgs boson to two light pseudoscalar particles decaying to two photons in terms of long-lived axion-like particles with masses in the range 10-100 MeV

PUB-HMBS-2024-69 - Reinterpretations of ATLAS searches for exotic decays of the Higgs boson to pseudo-scalar particles in terms of long-lived axion-like particles

ATL-COM-PHYS-2024-1027 - Measurement of light-by-light scattering with 1.67 nb^{-1} of Pb+Pb data in the ATLAS detector

CONFERENCES

DPG 2023 - Dresden, 20.03. - 24.03.2023

DPG 2024 - Karlsruhe, 04.03. - 08.03.2024

ICHEP 2024 - Prague, 18.07. - 24.07.2024

TECHNICAL STRENGTHS

Modeling and Analysis	Matlab, Root
Software & Tools	MS Office, Latex, Linux
Programming Languages	C, C++, Python

ACADEMIC ACHIEVEMENTS

Participated in various non compulsory lectures like EFT, Advanced QFT and machine learning at the university of Mainz and was successfully examined.

Was graded with distinction in the modules of Relativistic quantum mechanics and statistical mechanics in the MSc.

Was in the top 50 of graduates of year 2018 for the Electrical and computer engineering school, NTUA, Athens

Was examined nationally, at the end high school, in physics and mathematics and scored 20/20 and 18.5/20 respectively, the top grades in the district.

EXTRA-CURRICULAR

Member of the university basketball team.

Member of the Omicron Firefighter Volunteers in Chios, Greece.

B1 level in German language.

PERSONAL TRAITS

Highly motivated and eager to learn new things.

Strong motivational and leadership skills.

Ability to work as an individual and in a group.

Eager to travel globally to cultivate my knowledge and skills.

Able to adapt to new environments and conditions.

REFERENCES

Dr. Matthias Schott - University of Bonn - Matthias.Schott@cern.ch

Dr. Kristof Schmieden - University of Mainz - Kristof.Schmieden@cern.ch

Dr. Iwona Grabowska-Bold - University of Krakow - Iwona.grabowska@cern.ch