

Use this document as a templated

Example and documentation of the kaobook class

Customise this page according to your needs

Adrian Salvador Salas

June 29, 2022

An Awesome Publisher

Disclaimer

You can edit this page to suit your needs. For instance, here we have a no copyright statement, a colophon and some other information. This page is based on the corresponding page of Ken Arroyo Ohori's thesis, with minimal changes.

No copyright

© This book is released into the public domain using the CC0 code. To the extent possible under law, I waive all copyright and related or neighbouring rights to this work.

To view a copy of the CC0 code, visit:

<http://creativecommons.org/publicdomain/zero/1.0/>

Colophon

This document was typeset with the help of **KOMA-Script** and **L^AT_EX** using the **kaobook** class.

The source code of this book is available at:

<https://github.com/fmarotta/kaobook>

(You are welcome to contribute!)

Publisher

First printed in May 2019 by An Awesome Publisher

The harmony of the world is made manifest in Form and Number, and the heart and soul and all the poetry of Natural Philosophy are embodied in the concept of mathematical beauty.

– D'Arcy Wentworth Thompson

Preface

I am of the opinion that every \LaTeX geek, at least once during his life, feels the need to create his or her own class: this is what happened to me and here is the result, which, however, should be seen as a work still in progress. Actually, this class is not completely original, but it is a blend of all the best ideas that I have found in a number of guides, tutorials, blogs and tex.stackexchange.com posts. In particular, the main ideas come from two sources:

- ▶ [Ken Arroyo Ohori's Doctoral Thesis](#), which served, with the author's permission, as a backbone for the implementation of this class;
- ▶ The [Tufte-Latex Class](#), which was a model for the style.

The first chapter of this book is introductory and covers the most essential features of the class. Next, there is a bunch of chapters devoted to all the commands and environments that you may use in writing a book; in particular, it will be explained how to add notes, figures and tables, and references. The second part deals with the page layout and design, as well as additional features like coloured boxes and theorem environments.

I started writing this class as an experiment, and as such it should be regarded. Since it has always been intended for my personal use, it may not be perfect but I find it quite satisfactory for the use I want to make of it. I share this work in the hope that someone might find here the inspiration for writing his or her own class.

Federico Marotta

Contents

| | |
|---|------------|
| Preface | v |
| Contents | vii |
| 1 Introduction | 1 |
| THEORETICAL AND EXPERIMENTAL SETUP | 3 |
| 2 The Standard Model of Particle Physics and beyond | 5 |
| 2.1 The Standard Model of Particle Physics | 5 |
| 2.1.1 Particle content of the Standard Model | 5 |
| 2.1.2 Interactions of the Standard Model | 7 |
| 2.1.3 Quantum Chromodynamics | 9 |
| 2.1.4 ElectroWeak (EWs) theory | 11 |
| 2.1.5 Spontaneous symmetry breaking and the Higgs mechanism | 13 |
| 2.2 Successes and shortcomings of the Standard Model | 20 |
| 2.2.1 Experimental measurements | 20 |
| 2.2.2 Open questions | 21 |
| 2.3 Beyond the Standard Model | 23 |
| 3 The ATLAS experiment at the LHC | 25 |
| 3.1 The LHC | 25 |
| 3.1.1 Performance in Run 2 | 26 |
| 3.2 The ATLAS experiment | 27 |
| 3.2.1 Particle identification | 28 |
| 3.2.2 Coordinate System | 28 |
| 3.2.3 The Inner Detector | 29 |
| 3.2.4 The Calorimeter System | 30 |
| 3.2.5 Muon Spectrometer | 31 |
| 3.2.6 Magnet System | 32 |
| 3.2.7 Trigger System and Data Acquisition | 32 |
| 4 Physics simulation of proton collisions | 35 |
| 4.1 Event simulation | 35 |
| 4.1.1 Factorisation theorem | 35 |
| 4.1.2 Parton density function | 36 |
| 4.1.3 Matrix element | 36 |
| 4.1.4 Parton shower | 37 |
| 4.1.5 Hadronisation | 37 |
| 4.1.6 Pile-up and underlying event | 38 |
| 4.1.7 Monte Carlo simulation and generators | 38 |
| 4.2 Detector simulation | 39 |
| 5 Object reconstruction | 41 |
| 5.1 Base objects | 41 |
| 5.1.1 Tracks and vertices | 41 |
| 5.1.2 Topological clusters | 42 |

| | | |
|-----------|---------------------------------------|-----------|
| 5.2 | Jets | 42 |
| 5.2.1 | Reconstruction | 42 |
| 5.2.2 | Jet tagging | 44 |
| 5.3 | Leptons | 45 |
| 5.3.1 | Electrons | 45 |
| 5.3.2 | Muons | 46 |
| 5.4 | Taus | 47 |
| 5.5 | Missing transverse energy | 47 |
| 6 | Machine learning | 49 |
| 6.1 | Pipeline | 49 |
| 6.2 | Performance | 50 |
| 6.3 | Neural networks | 50 |
| 6.4 | BDTs if used | 50 |
| 7 | Margin Stuff | 51 |
| 7.1 | Sidenotes | 51 |
| 7.2 | Marginnotes | 51 |
| 7.3 | Footnotes | 52 |
| 7.4 | Margintoc | 52 |
| 7.5 | Marginlisting | 53 |
| 8 | Figures and Tables | 55 |
| 8.1 | Normal Figures and Tables | 55 |
| 8.2 | Margin Figures and Tables | 58 |
| 8.3 | Wide Figures and Tables | 58 |
| 9 | References | 61 |
| 9.1 | Citations | 61 |
| 9.2 | Glossaries and Indices | 62 |
| 9.3 | Hyperreferences | 63 |
| 9.4 | A Final Note on Compilation | 64 |
| | DESIGN AND ADDITIONAL FEATURES | 67 |
| 10 | Page Design | 69 |
| 10.1 | Headings | 69 |
| 10.2 | Headers & Footers | 70 |
| 10.3 | twoside mode | 70 |
| 10.4 | Table of Contents | 70 |
| 10.5 | Paper Size | 71 |
| 10.6 | Page Layout | 71 |
| 10.7 | Numbers & Counters | 72 |
| 10.8 | White Space | 73 |
| 11 | Mathematics and Boxes | 75 |
| 11.1 | Theorems | 75 |
| 11.2 | Boxes & Environments | 77 |
| 11.3 | Experiments | 77 |

| | |
|---|-----------|
| APPENDIX | 79 |
| A Heading on Level 0 (chapter) | 81 |
| A.1 Heading on Level 1 (section) | 81 |
| A.1.1 Heading on Level 2 (subsection) | 81 |
| A.2 Lists | 82 |
| A.2.1 Example for list (itemize) | 82 |
| A.2.2 Example for list (enumerate) | 82 |
| A.2.3 Example for list (description) | 83 |
| B Fonts Testing | 85 |
| B.1 Font Sizes | 85 |
| B.2 Font Families | 85 |
| Bibliography | 87 |
| Notation | 89 |
| Alphabetical Index | 91 |

List of Figures

| | | |
|-----|----------------------------|----|
| 8.1 | Mona Lisa, again | 57 |
| 8.2 | A wide seaside | 58 |

List of Tables

| | | |
|------|---|----|
| 8.1 | A useless table | 56 |
| 8.2 | Another useless table | 58 |
| 8.3 | A wide table with invented data about three people living in the UK. Note that wide figures and tables are centered and their caption also extends into the margin. | 59 |
| 8.4 | Caption of the longtable. | 60 |
| 10.1 | Commands to add a particular entry to the table of contents. | 71 |
| 10.2 | Some non-standard paper sizes supported by kaobook. | 71 |

The discovery of the Higgs boson in 2012 by ATLAS and CMS [1] is one of the most recent historic milestones in the field of particle physics. CERN hosts the LHC, which one of the main objectives in the physics program was to prove the existence of the Higgs boson. After this achievement, all particles predicted by the Standard Model are discovered, the theory which describes the fundamental particles and their interactions. Nevertheless, the ATLAS experiment continues to scrutinise the Standard Model by analysing the ever-increasing amount of particle collisions delivered by the LHC. There are many phenomena not covered by the current theory and any measurement that deviates from the predictions or the hint of a new particle will lay the foundation for a new path in particle physics. That is the reason for the precise measurement of the Higgs boson properties and searches related to found deviations in the leptonic sector.

Many theoretical expansions of the Standard Model contain additional scalar particles. For example, if the Higgs sector is built with one extra doublet (Two Higgs Doublet Models), a total of five scalars are predicted instead, including charged scalars. Another interesting possibility involves the presence of flavour-changing neutral current (FCNC) interactions, which are predicted to be heavily suppressed and any increased measurement would also point to new physics. Other theories predict higher contributions, in particular, FCNC interactions involving the top quark and the flavon, a new neutral scalar arising from a symmetry between quark families. In this thesis, a direct search for charged Higgs bosons heavier than the top quark and a direct search for neutral scalars lighter than the top quark are presented. The charged Higgs process is searched produced in association with top and bottom quarks and decaying into a top-bottom, while the neutral scalar is searched from the FCNC decay of a top quark involving a c - or a u -quark, and decaying to a pair of b -quarks. Both searches are performed using the full Run-2 proton-proton collisions collected by the ATLAS experiment from 2015 to 2018 at a center-of-mass energy of 13 TeV. Limits on the production of charged Higgs bosons have been previously obtained by ATLAS with only the data from 2015 and 2016 for $H^\pm \rightarrow tb$ in the 180 – 600 GeV mass range [2], and more recently by CMS in the 200 – 600 GeV mass range using the full Run-2, setting upper limits on the production cross section of X - X pb and X - X pb respectively. On the other hand, both ATLAS and CMS have searched for FCNC $t \rightarrow qH$ decays, with q either a c - or u -quark, setting upper limits on the branching fraction of XXX and XXX. However the generic signature presented involving a new scalar lighter than the top quark is uncovered in literature.

The $H^\pm \rightarrow tb$ search presented in this thesis is performed for a mass range from 200 to 2000 GeV produced in association with top and bottom quarks and in final states with a single lepton. The limits on the $H^\pm \rightarrow tb$ production and the $t \rightarrow qX(bb)$ are set by means of a binned profile likelihood fit of the different simulated signal and Standard Model backgrounds to the collected data. The fit is performed on a discriminant

obtained by combining several kinematic variables through different machine learning techniques, which allows to improve the sensitivity of the analysis by improving the separation of signal and background events. The largest part

THEORETICAL AND EXPERIMENTAL SETUP

The Standard Model of Particle Physics and beyond

2

Be great to say that the lagrangian we want has to describe reality, what we want is something Lorentz invariant, Pointcaré group?

The Standard Model (SM) of particle physics [1–3] is the theoretical framework that so far better describes subatomic particles and their interactions. It is a Quantum Field Theory (QFT) and since its initial development in the 1960's, the model has been overwhelmingly successful and guided many experimental achievements including the discovery of the top quark [4, 5] in 1995 and the Higgs boson at the LHC in 2012 [6, 7]. Regardless of the success of the model, there are known phenomena not covered in the model and other questions which clearly point to the need of a new theories.

This chapter presents an overview of the SM with a brief summary of the particle content as of yet without [going] into underlying theoretical details. *Next, the description of...*

Throughout this dissertation, natural units are used: the speed of light and the reduced Plank constant are set to unity ($c = \hbar = 1$), electric charges are expressed in units of the electron electric charge ($-e$) and masses are expressed in terms of energy (eV). Within the theoretical developments in this chapter, the Einstein's summation convention is used by default.

2.1 The Standard Model of Particle Physics

From the mathematical point of view, the SM is a renormalisable non-abelian gauge QFT based on the symmetry group,

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad (2.1)$$

where $SU(3)_C$ is the group described by Quantum Chromodynamics (QCD) [8] that represents the strong interactions of colored quarks and gluons (strong force), while $SU(2)_L \times U(1)_Y$ is the inclusive representation of both electromagnetic (EM) and weak interactions described by the ElectroWeak (EW) theory [1, 2, 9]. The SM describes all the interactions between elementary particles except gravity, for which no renormalisable QFT has been formulated so far. The following sections, introduce the particles of the SM and the theories that describe their interactions.

2.1.1 Particle content of the Standard Model

In the SM, elementary particles are described as excitations of quantum fields. There are two main classes of particles within the theory: *fermions* and *bosons*. The main difference between the two is the spin: fermions have

half-integer spin and therefore obey the Pauli exclusion principle [10], while bosons have integer spin.

Fermions

Fermions can be divided further into two categories: quarks and leptons, based on their interactions, or their charges. Both types manifest in EW interactions, having a weak isospin $T_3 = \pm 1/2$ although only the quarks experience the strong interaction in addition, so have strong charge referred to as *colour*. Quarks have a fractional electric charge $Q = 2/3$ or $1/3$, and the colour, which values are usually denoted as *red*, *green* and *blue*. Table [] presents a summary of the fundamental fermions and their characteristics.

There is a total of six quark types, named *flavours* and are split into three generations. The first generation consists in the *up* and the *down* quark, the former with $Q = +2/3$ and $T_3 = +1/2$, while the latter $Q = -1/3$, $T_3 = -1/2$ and a slightly lower mass. The next two generations are copies of the first one with increasing mass, with a pair of a *up*-type quark and a *down*-type quark. The second family consists in *charm* and *strange* quarks, and the third of *top* and *bottom* quarks. In addition, all of the six quark flavours have antimatter states with the same mass, but opposite quantum numbers, as an example, an anti-*up*-type quark has $Q = -2/3$, $T_3 = -1/2$ and can carry anti-*red* colour.

Leptons are also similarly divided in six different types and in three separate generations named *electron* (e), *muon* (μ) and *tau* (τ), also with increasing mass. Each generation contains a lepton with $Q = -1$ and $T_3 = +1/2$ named after its generation, and an associated electrically neutral lepton with $T_3 = -1/2$ named neutrino (ν). The neutrino is assumed to be massless in the formulation of the SM, however the phenomena of neutrino oscillations is experimental proof of them actually having very small, but non-zero, mass values. This apparent failure of the theory is discussed in Section[]. As before, the associated antimatter states have the same mass but opposite quantum numbers.

All the stable SM matter in the universe is constituted by the massive particles of the first generations of quarks and leptons, as the heavier versions eventually decay to lighter ones through their disclosed interactions. While it is possible to observe free leptons, quarks exist only in bound states, or hadrons, like the neutron or the proton. This is a feature of the strong interaction under the name of confinement, discussed in Section[]. Only colour-less bounded states are observable then, and can be built from three quarks with overall half-spin, named baryons, or by two quarks with integer spin, named mesons.

In the context of particle physics, the formulation of the classical Lagrangian, \mathcal{L} , is used to describe physics systems. A generic free fermion field ψ with mass m , can be described by the Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \quad (2.2)$$

where γ^μ are Dirac matrices and ∂_μ is the four-momentum derivative.

Bosons

Particles with integer spin are referred to as bosons. The bosonic sector with spin-1 gauge fields are force carriers that naturally follow from imposing the requirement of local gauge invariance on Eq. 2.2 under symmetry groups, in this case Eq. 2.1. In Section [] the nature and origin of the gauge bosons will be detailed. In summary, the photon (γ) is the carrier of the electromagnetic force, being a massless and electrically neutral particle. The weak force carriers are the W^+ , W^- and Z bosons, all massive with the Z boson being electrically neutral and the W^\pm with either $Q = \pm 1$. Gluons (g) are the strong force carriers which are massless and with no electric charge. Instead, there are eight different gluons representing each possible colour exchange. Table [] presents a summary of the gauge bosons that mediate the different interactions.

The SM also includes a neutral spin-0 particle, or *scalar*, the Higgs boson, with a mass of 125.25 ± 0.17 GeV [11]. The Higgs field is responsible for all SM particles acquiring mass through the Higgs mechanism, as described in Section []. The kinematics of a generic scalar, ϕ with mass m , is described by the Klein-Gordon Lagrangian,

$$\mathcal{L} = \frac{1}{2} \partial^\mu \phi \partial_\mu \phi - m^2 \phi^2 \quad (2.3)$$

Charged scalars can be described instead through a complex field and the expression of the Lagrangian is slightly modified,

$$\mathcal{L} = \partial^\mu \phi \partial_\mu \phi^* - m^2 \phi \phi^* \quad (2.4)$$

Vector fields A^μ , which represent spin-1 bosons, are described by the Proca Lagrangian,

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} m^2 A^\mu A_\mu \quad (2.5)$$

with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ the field strength tensor. In the case of massless particles, the previous expression with $m = 0$ is known as the Maxwell Lagrangian.

2.1.2 Interactions of the Standard Model

The Lagrangian of the SM is defined to be locally invariant to the Eq. 2.1 symmetry group, condition that generates and defines the interactions of the corresponding particles as representations of the symmetry transformations. For a generic Lagrangian, the physical system can have symmetries, so its Lagrangian is invariant under different kind of transformations. These transformations can be time-space independent, called global transformations, or dependent, called gauge or local transformations. Any invariant transformation of a Lagrangian describes a physical system which conserves a physical quantity, as described by the Noether theorem [12]. Then, the interactions are introduced in the Lagrangian as additional terms by promoting an already existing global symmetry, ϕ , of

the Lagrangian to a local gauge symmetry, $\phi(x)$. The physical motivation behind introducing gauge symmetries is to be able to describe vector bosons in QFT, as seen in [Section \[\]](#). The procedure expands the theory with additional fields that mediate the resulting interactions, which properties depend on the characteristics of the symmetry group.

An example of the process is shown in the following, to afterwards derive the SM interactions of the strong and electroweak sectors.

Gauging a symmetry to interaction

A general global transformation θ which acts on the field ψ is described as,

$$\psi \rightarrow e^{ig\theta^a T^a} \psi \quad (2.6)$$

with g the coupling constant and T^a the generators of the Lie group associated to the transformation (like $SU(n)$ or $U(n)$), with a ranging from 1 to $n^2 - 1$, for the corresponding number of the Lie algebra, $n > 1$. The generators can be characterised by their commutation relation,

$$[T^a, T^b] = if^{abc} T^c \quad (2.7)$$

where f^{abc} are the structure constants of the group. Following Noether's theorem, there are as many conserved quantities as generators of the Lagrangian's symmetries. As an example, it is straightforward to see that a Lagrangian like Equation 2.2 is invariant to a $U(1)$ transformation where θ is just a constant and hence, a constant phase change. One can obtain the current, j^μ , that is conserved, $\partial_\mu j^\mu = 0$,

$$j^\mu = \bar{\psi} \gamma^\mu \psi \quad (2.8)$$

and the conserved charge,

$$Q = \int d^3x j^0 = \int d^3x \psi^\dagger \psi \quad (2.9)$$

With some algebra and introducing solutions in momentum space, ψ can be interpreted as annihilating a fermion and creating an anti-fermion (ψ^\dagger the other way around) in the Fock space and then, this product becomes the difference of the number of fermion and anti-fermion leading to the conservation of the fermion number.

Promoting the global symmetry to a local symmetry is done by introducing locality in the θ transformation, $\theta \rightarrow \theta(x)$, which introduces new $\partial_\mu \theta$ terms in the Lagrangian. A way to counter the new terms and, hence, keep the Lagrangian invariant, is to introduce gauge vector fields A_μ^a , following Yang-Mills theory [13]. In the most generalised approach, there have to be as many A_μ^a as generators of the symmetry, that transform as,

$$A_\mu^a \rightarrow A_\mu^a + \partial_\mu \theta^a + g f^{abc} A_\mu^b \theta^c \quad (2.10)$$

Note that the last term proportional to the structure constant is relating the gauge field to the conserved symmetry charge. The next step is to replace the standard derivative in the Lagrangian by the covariant derivative,

$$D_\mu \equiv \partial_\mu - igT^a A_\mu^a \quad (2.11)$$

The final ingredient is to complete the Lagrangian with the kinematic Lagrangian for the massless vector fields, the Maxwell Lagrangian from Equation 2.5 with a slightly different field strength tensor,

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc} A_\mu^b A_\nu^c \quad (2.12)$$

The last term is present only for non-abelian symmetry groups, since it is proportional to the structure constants, and has huge consequences in the resulting interactions as discussed in [Section \[\]](#). Another remark is that the gauge fields have to be massless, as a mass term proportional to $A_\mu^c A^{\mu c}$ is not gauge invariant.

As an example, the promotion of the global $U(1)$ symmetry seen in Equation 2.2 results in the upgraded Lagrangian,

$$\begin{aligned} \mathcal{L} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ D_\mu &\equiv \partial_\mu - igA_\mu \\ F_{\mu\nu} &\equiv \partial_\mu A_\nu - \partial_\nu A_\mu \end{aligned} \quad (2.13)$$

introducing just one massless gauge field that interacts with the field ψ . The interaction term between the two fields is $g\bar{\psi}\gamma^\mu A_\mu\psi$, hidden in the covariant derivative definition and proportional to the coupling constant g .

The Lagrangian of the SM is built from imposing local invariance under $SU(3)_C$ transformations, which leads to strong interactions, and $SU(2)_L \times U(1)_Y$ transformations, which brings EW interactions,

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} \quad (2.14)$$

After this introduction on field theory, the theories of the two orthogonal sectors can now be described and then, the mechanism to introduce mass terms in the Lagrangian, the spontaneous symmetry breaking.

2.1.3 Quantum Chromodynamics

The quantum field theory that describes quarks and gluons interactions is named *quantum chromodynamics*, based on the $SU(3)_C$ symmetry group. Each quark has an internal degree of freedom, known as the colour charge, and it is defined by a triplet of fields,

$$q = \begin{pmatrix} q_{\text{red}} \\ q_{\text{blue}} \\ q_{\text{green}} \end{pmatrix} \quad (2.15)$$

where each of the components is a Dirac spinor associated to the corresponding colour state (red, blue and green). The colour In addition, there are a total of six quarks, so the fields are labelled as $q_{f\alpha}$ with f indicating the quark flavour ($f = u, d, c, s, t, b$) and α the colour. Note that there is an anti-quark of each flavour carrying an anti-colour charge.

The algebra of the $SU(3)$ group is characterised by the non-abelian commutation relation from Equation 2.7 with a total of eight generators, T^a . The generators can be written as $T^a = \lambda^a/2$ where λ^a denote the Gell-Mann matrices [14]. Because of the eight generators, the interaction is mediated by a total of eight gauge bosons, called gluons G_μ^a . There are different matrix representation for the colour states of the gluons, following with the Gell-Mann matrices, taking,

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.16)$$

and applying it to a general quark triplet like Equation 2.15, it can be seen that the transformation switches the red and blue charges. To do so, the gluon has to carry a colour/anti-colour pair, to be able to "remove" the red charge (r) and "add" the blue charge (b), and the other way around. There are nine possible combinations of colour/anti-colour pairs, which can be used to re-write the λ^1 transformation as,

$$\frac{r\bar{b} + b\bar{r}}{\sqrt{2}} \quad (2.17)$$

known as the first state of the gluon colour octet. The rest of the states are equivalent to the other Gell-Mann matrices and all conserve the three different colour flows.

The QCD Lagrangian can be obtained from modifying the the Dirac Lagrangian (Equation 2.2) to achieve gauge invariance under $SU(3)_C$ transformations, following the definitions from Section [1]. The resulting Lagrangian is,

$$\begin{aligned} \mathcal{L}_{QCD} &= i \sum_f \bar{q}_f \gamma^\mu D_\mu q_f - \frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu} \\ D_\mu &\equiv \partial_\mu - i g_s T^a G_\mu^a \\ G_{\mu\nu}^a &\equiv \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c \end{aligned} \quad (2.18)$$

with g_s being the strong force coupling constant and where the covariant derivative has been introduced with the G_μ^a gluons fields, together with the kinematic term for the gluons, introducing the gluon tensor, $G_{\mu\nu}^a$. As described in Section [1], gluons are massless because the term in the Lagrangian is not gauge invariant. Notice that the masses of the quarks are also not present, not because it would break the symmetry, but for convention. The masses in the SM come from the electro-weak sector. Another remark is that the addition of a charge conjugation and parity symmetry (CP) violating interaction term is allowed under local gauge

invariance, but such an interaction has been experimentally observed to be effectively zero^{referencethestrongCPexperiment}.

The resulting interactions in the Lagrangian are shown in [Figure\[\]](#), consisting of couplings between quarks and gluons¹, and three- and four-point gluon self-interactions. As foreshadowed in [Section\[\]](#), for non-abelian groups the gauge bosons have the self-interacting terms in the tensor.

There are two more important characteristics of this theory, that also arise from the non-abelian nature of the symmetry: asymptotic freedom and confinement [15, 16]. Asymptotic freedom refers to the fact that at very high energies (in momentum transfer), or short distances, quarks and gluons interact weakly with each other allowing predictions to be obtained using perturbation theory. Confinement is the name given to the impossibility of directly observing quarks, only confined in hadrons, which are colorless composite states². The idea is that for high distances, the strong coupling becomes larger, so when the distance between two quarks is increased, the energy of the gluon field is larger, up to the point to create from the vacuum a quark/anti-quark pair and thus forming a new hadron.

1: Equivalent to the interaction obtained from the gauge $U(1)$ symmetry.

2: Color singlets are quantum states that are invariant under all eight generators of $SU(3)$, and therefore carry vanishing values of all colour conserved charges.

Running coupling

Say somewhere that we abuse group theory ????????? introduce the beta function? just the coupling from griffits?

2.1.4 ElectroWeak (EWs) theory

The quantum field theory that describes both the electromagnetic and weak interactions is named *electroweak* theory, based on the $SU(2)_L \otimes U(1)_Y$ symmetry group³. The product is non-abelian, like the $SU(3)_C$ group, and chiral. It will spawn four mediators, as the number of generators. The symmetry spontaneously breaks down through *EW symmetry breaking*, giving rise to the electromagnetic interaction, mediated by the photon, and to the weak interaction, mediated by the Z and W^\pm bosons. This process occurs at ~ 100 GeV, defined as the EW scale, and after which only the $U(1)_Q$ symmetry is unbroken, described by *EW symmetry breaking* (QED). The process of the EWSB, and the resulting effects are described in more detail in [Section](#).

3: L refers to the left-handed chirality and Y to the weak hypercharge

The interactions for the EW sector can be obtained following the procedure described in general in [Section](#), already used in [Section](#) for QCD. First, only left-handed fermion fields interact via the weak interaction⁴, transforming as doublets under $SU(2)_L$, whereas right-handed fermion fields do not interact weakly and thus transform as singlets,

4: As a consequence, parity can be violated in weak interactions [17, 18].

$$\begin{aligned}\psi_L^i &= \begin{pmatrix} \ell_L^i \\ \nu_L^i \end{pmatrix}, \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix} \\ \psi_R^i &= \ell_R^i, u_R^i, d_R^i\end{aligned}\tag{2.19}$$

with i corresponding to the number of the generation. Fields with subscripts L/R are left- and right-handed fields that can be defined through the chirality operators P_L and P_R , projecting a generic field into only its left- and right-handed components, respectively,

$$\begin{aligned}\psi_L &= P_L \psi = \frac{1}{2}(1 - \gamma_5)\psi \\ \psi_R &= P_R \psi = \frac{1}{2}(1 + \gamma_5)\psi\end{aligned}\tag{2.20}$$

with γ_5 defined from the Dirac matrices $\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$.

The $SU(2)_L$ group consists of three generators, \hat{T}_i , which can be written as $\hat{T}_i = \sigma_i/2$ where σ_i denotes the Pauli matrices. Also, the quantum number associated is the weak isospin, T . On the other side, the $U(1)_Y$ group introduces the weak hypercharge quantum number, Y . After EWSB, the Gell-Mann-Nishijima equation relates Y to the third component of the weak isospin operator, T_3 and the electric charge Q ,

$$Q = Y + T_3\tag{2.21}$$

Regarding the EW Lagrangian, four gauge fields need to be introduced to achieve invariance under the $SU(2)_L \otimes U(1)_Y$, $W_{\mu\nu}^i$ ($i=1,2,3$) from $SU(2)_L$, and B_μ from $U(1)_Y$. The resulting Lagrangian is,

$$\begin{aligned}\mathcal{L}_{EW} &= i \sum_{f=l,q} \bar{f}(\gamma^\mu D_\mu)f - \frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ D_\mu &\equiv \partial_\mu - ig\frac{\sigma}{2}W_\mu^i - ig'YB_\mu \\ W_{\mu\nu}^i &\equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk}W_\mu^j W_\nu^k \\ B_{\mu\nu} &\equiv \partial_\mu B_\nu - \partial_\nu B_\mu\end{aligned}\tag{2.22}$$

with ϵ^{ijk} the Levi-Civita symbol, an antisymmetric tensor defined as $\epsilon^{ijk}\epsilon_{lmn} = \delta_m^j\delta_n^k - \delta_n^j\delta_m^k$ with $i, j, k, l, m, n \in [1, 2, 3]$. Also, the $W_{\mu\nu}^i$ and $B_{\mu\nu}$ field tensors are defined to introduce the additional kinetic terms to the Lagrangian. The former contains a quadratic piece, due to the non-abelian nature of $SU(2)_L$, hence the full Lagrangian contains cubic and quartic self-interactions, as seen for the gluons in QCD. In contrast, the coupling constant g increases rapidly with the energy scale.... As encountered before, mass terms for the gauge boson would break the gauge invariance. In this case, terms for the fermion masses would also break the symmetry as they would mix left- and right-handed fields, which transforms distinctively under $SU(2)_L$. Instead, the mass terms appear from the EWSB, described in Section[].

Summing all the interactions described, the SM Lagrangian for all the fermions before EWSB becomes,

$$\begin{aligned}
\mathcal{L}_{SM} = & \sum_f \sum_{\psi=L,e_R,Q_L,u_R,d_R} i\bar{\psi}^f \gamma^\mu D_\mu \psi^f \\
& - \frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\
D_\mu = & \partial_\mu - ig_s T^a G_\mu^a - ig \frac{\sigma^i}{2} W_\mu^i - ig' Y B_\mu
\end{aligned} \tag{2.23}$$

no right handed neutrino, masses mix in EW

2.1.5 Spontaneous symmetry breaking and the Higgs mechanism

The model described so far cannot reproduce measured results, first of all the different fermions and the weak force mediators have mass and second, the $SU(2)_L \times U(1)_Y$ symmetry is not preserved in nature. Even if somehow the EW gauge bosons are allowed to have mass, it leads to the lack of renormalisability and the violation of unitarity. Renormalisation is a collection of techniques that allows the computation of measurable observables in QFT, managing the different sources of infinities within the theory like those from self-interactions. Unitarity is needed more in general in quantum mechanics, to ensure proper time-evolution predictions of a quantum state. The longitudinal component of the massive boson is the cause of the problem, as in a boosted frame in which $p^\mu = (p^0, 0, 0, |\mathbf{p}|)$, the parallel polarisation component of a massive boson is $\epsilon_\mu = (|\mathbf{p}|/m, 0, 0, p^0)$, growing indefinitely with the energy of the system. When computing the cross-section of the corresponding boson scattering, the value will indefinitely grow breaking the mentioned unitarity. If computed explicitly for the W^\pm bosons, the energy scale where this happens is around the TeV scale, pointing to a fundamental problem in the theory to describe that scale.

The solution is provided by the EWSB and the Higgs-Englert-Brout mechanism, discussed next, after showing the spontaneous symmetry breaking process for a simple gauge theory.

How to break a symmetry

Spontaneous symmetry breaking is a phenomenon where a symmetry of the theory is unstable and the vacuum, or fundamental state, is degenerate. In the process, new interactions appear and a field obtains a non-zero vacuum expectation value.

The topic is broad as there are many symmetries and representations to potentially break, to illustrate the mechanism for the SM, let's consider a system with a scalar field ϕ , a gauge field A_μ , and the following Lagrangian with a gauge symmetry,

$$\begin{aligned}
\mathcal{L} &= (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\
D_\mu &\equiv \partial_\mu - ig A_\mu \\
F_{\mu\nu} &\equiv \partial_\mu A_\nu - \partial_\nu A_\mu
\end{aligned} \tag{2.24}$$

with a general potential $V(\phi)$ given by,

$$V(\phi) = \frac{1}{2} \mu^2 \phi^\dagger \phi + \frac{1}{4} \lambda (\phi^\dagger \phi)^2 \tag{2.25}$$

with the real parameters μ^2 and λ relating respectively to the mass term and the strength of the self-interaction. [INSERT FIGURE] There are two sensible ranges for these parameters, the first one is the case $\lambda, \mu^2 > 0$, similar to the previous seen theories and only one solution in the minimisation. The second one is for $\lambda > 0$ and $\mu^2 < 0$, where the $\mu^2 \phi^\dagger \phi$ term cannot be understood as a mass term and the solution $\phi = 0$ is a local maximum, physically unstable. The minimum of the potential is degenerate and identified by the complex plane circle, $\phi^\dagger \phi = v^2/2$ with $v^2 \equiv -\mu^2/\lambda$ and

$$\phi = v e^{-i\theta} \tag{2.26}$$

The symmetry is broken spontaneously when the system choses the fundamental state. Suppose $\phi = 0$, then the *Vacuum Expectation Value* (VEV) of ϕ is set to,

$$\langle 0 | \phi | 0 \rangle = \frac{v}{\sqrt{2}} \tag{2.27}$$

Next, lets suppose the following change of variables to center the new fundamental state,

$$\phi(x) = \left(\frac{v + \eta(x)}{\sqrt{2}} \right) e^{i\zeta(x)/v} \tag{2.28}$$

the Lagrangian can be expressed as,

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2} (\partial_\mu \eta)^2 + \frac{1}{2} (\partial_\mu \zeta)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\
&+ \mu^2 \eta^2 + \frac{1}{2} g^2 v^2 A_\mu A^\mu - g v A_\mu \partial^\mu \zeta + \text{interactions}
\end{aligned} \tag{2.29}$$

which now contains the η and ζ fields, additional to the gauge A_μ . Also, square terms appear for η and A_μ , which can be identified as mass terms, $\frac{m_\eta}{2} \eta^2$ and $\frac{m_A}{2} A_\mu A^\mu$, resulting in $m_\eta = \sqrt{-2\mu^2}$ and $m_A = gv$. $\zeta(x)$ is massless and a particular resulting type of field named *Goldstone boson*, which the *Goldstone theorem* predicts. The theorem states that a massless boson appears for every symmetry that the VEV spontaneously breaks. In this abelian case, the VEV is not invariant under the $U(1)$ transformation. $\zeta(x)$ does not appear explicitly in the potential, therefore

can take any value without affecting the energy of the system, which is not very physical. In addition, it appears in an estrange mixing term with A_μ , $-gvA_\mu\partial^\mu\zeta$. A way to remove this annoyance is to choose the gauge,

$$\begin{aligned}\phi &\rightarrow \phi' = e^{-i\zeta/v}\phi \\ A_\mu &\rightarrow A'_\mu = A_\mu - \frac{1}{gv}\partial_\mu\zeta\end{aligned}\quad (2.30)$$

together with the previous change of variable for ϕ . Essentially the gauge freedom of the Lagrangian is being used to remove ζ , which becomes the longitudinal component of the transformed gauge boson A_μ . The gauge chosen is the so-called *unitary gauge*, which makes the physical content of the Lagrangian explicit.

In summary, this process of acquiring mass by means of absorbing a Goldstone boson is known as the *Higgs mechanism*.

The Higgs-Englert-Brout Mechanism in the Electroweak Sector

The Higgs-Englert-Brout mechanism [19–21] solved the contradictions found between massive particles and the requirement of gauge invariance. The mechanism is based in a spontaneous symmetry breaking of the $SU(2)_L \otimes U(1)_Y$ to $U(1)_{EM}$, giving mass to the different particles involved in the EW interactions except the photon. A similar procedure can be applied to the EW Lagrangian derived in Equation 2.23, first introducing an isospin doublet ($Y=+1/2$) of complex scalar fields Φ , the Higgs field,

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (2.31)$$

where ϕ^+ corresponds to an electrically charged field ($T_3=+1/2$) and ϕ^0 to a neutral one ($T_3=-1/2$). This field transforms under $SU(2)_L$ and its Lagrangian, the Higgs Lagrangian,

$$\mathcal{L}_\Phi = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi) \quad (2.32)$$

with the same covariant derivative as in Equation 2.23 and the Higgs potential given by,

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \quad (2.33)$$

which shape depends on the parameters μ^2 and λ . As seen before, choosing the case where $\lambda > 0$ and $\mu^2 < 0$, the potential at $\Phi = 0$ is unstable and a continuous collection of possible minimum values appear, defined by the circle,

$$\Phi^\dagger\Phi = \frac{1-\mu^2}{\lambda} \equiv \frac{1}{2}v^2 \quad (2.34)$$

Following, the spontaneous symmetry breaking with the choice of the new vacuum state,

$$\langle 0|\Phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.35)$$

This vacuum is not invariant to any of the $SU(2)_L$ and the $U(1)$ transformations, however, the $Q = T_3 + Y$ transformation is not affected,

$$Q \langle 0|\Phi|0\rangle = \frac{1}{2\sqrt{2}} \sigma_3 \begin{pmatrix} 0 \\ v \end{pmatrix} + \frac{1}{2\sqrt{2}} Y \begin{pmatrix} 0 \\ v \end{pmatrix} = \frac{1}{2\sqrt{2}} \left[\begin{pmatrix} 0 \\ -v \end{pmatrix} + \begin{pmatrix} 0 \\ v \end{pmatrix} \right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (2.36)$$

The field is rewritten in the unitary gauge, which automatically removes the extra nonphysical Goldstone bosons,

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (2.37)$$

where $H(x)$ is centered around the vacuum state. With this change the Higgs potential becomes,

$$V(\Phi) = \frac{1}{4} \lambda v^2 H^2 + \frac{1}{4} \lambda v H^3 + \frac{1}{16} \lambda H^4 \quad (2.38)$$

spawning the Higgs boson mass $m_H^2 = \lambda v^2/2 = -\mu^2/2$, in the quadratic H term. The cubic and quartic terms constitute the three- and four-point Higgs boson self-interactions.

The EWSB generates new interactions and mass terms for the different particles involved in the EW interactions. Gluons are not affected as the scalar field is a doublet and does not transform under $SU(3)$. The effects on the boson and fermion sectors of the SM are discussed in the following, individually.

Boson sector

The gauge boson masses spawn from the covariant derivative, $(D_\mu \Phi)^\dagger (D^\mu \Phi)$, which includes the gauge fields. Expanding,

$$\mathcal{L}_{mass} = \frac{v^2}{8} V_\mu \begin{pmatrix} g^2 & 0 & & \\ 0 & g^2 & & \\ & & 0_{2 \times 2} & \\ 0_{2 \times 2} & & g^2 & -gg' \\ & & -gg' & g'^2 \end{pmatrix} V^\mu \quad (2.39)$$

with $V_\mu = \begin{pmatrix} W_\mu^1 & W_\mu^2 & W_\mu^3 & B_\mu \end{pmatrix}$. Diagonalising the matrix, the next eigenvectors are found,

$$\begin{aligned} A_\mu &\equiv \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \\ Z_\mu &\equiv \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \end{aligned} \quad (2.40)$$

where the Weinberg angle, or weak mixing angle, is defined by $\tan \theta_W \equiv g'/g$. The corresponding eigenvalues, the square masses, for the A_μ and Z_μ fields are zero and $v^2(g^2 + g'^2)/8$. On the other side, W_μ^1 and W_μ^2 are well defined mass states but not charge states. This is due T_1 and T_2 being not diagonal, connecting the different states of T_3 (hence of Q). The operator $T_\pm = T_1 \mp iT_2$ can be defined, which increases or decreases one unit of T_3 (hence of Q). In addition, the fields can be redefined,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (2.41)$$

In summary the Lagrangian in Equation 2.39 can now be written as

$$\mathcal{L}_{mass} = \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} - \frac{v^2}{8} (g^2 + g'^2) Z_\mu Z^\mu \quad (2.42)$$

where the mass terms of the different bosons can be identified,

$$\begin{aligned} m_A &= 0 \\ m_Z &= \frac{v}{2} \sqrt{g^2 + g'^2} \\ m_W &= \frac{vg}{2} = m_Z \cos \theta_W \end{aligned} \quad (2.43)$$

Note that the remaining symmetry after breaking $SU(2)_L \otimes U(1)_L$ is $U(1)_{EM}$. The associated A_μ field is massless, the photon, which is a combination of the W_μ^3 and B_μ fields. The associated the quantum number, the electric charge, has been defined previously in the chapter, $Q = T_3 - Y$.

Regarding interactions, the covariant derivative can be expressed in terms of the new bosons,

$$\partial_\mu - igW_\mu^3 = \partial_\mu - ig \sin \theta_W A_\mu - ig \cos \theta_W Z_\mu \quad (2.44)$$

where the electromagnetic coupling constant e can be defined as $e = g \sin \theta_W$. In addition, the field tensors can be rewritten as,

$$\begin{aligned} W_{\mu\nu}^3 &= \partial_\mu W_\nu^3 - \partial_\nu W_\mu^3 - ig(W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) \\ &= \sin \theta_W F_{\mu\nu} + \cos \theta_W Z_{\mu\nu} - ig(W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) \\ B_{\mu\nu} &= \cos \theta_W F_{\mu\nu} - \sin \theta_W Z_{\mu\nu} \end{aligned} \quad (2.45)$$

where the field strength tensors for the photons and the Z boson, $F_{\mu\nu}$ and $Z_{\mu\nu}$ are defined.

Fermion sector

The procedure required to acquire the fermion masses is more complicated than for the gauge bosons. Instead of just expanding the kinematic term with the new Higgs field, Yukawa [22] interactions that couple left- and right-handed fermions with the Higgs need to be introduced.

As seen in this chapter, only $q_{\alpha L}^i$ and l_L^i fields are $SU(2)_L$ doublets,

$$q_{\alpha L}^i = \begin{pmatrix} u_{\alpha L}^i \\ d_{\alpha L}^i \end{pmatrix}, \quad l_L^i = \begin{pmatrix} \nu_L^i \\ \ell_L^i \end{pmatrix} \quad (2.46)$$

where the i refers to the generation and α to the colour. It has been already pointed out that is not possible to construct a well defined $m f^\dagger f$ term that transforms under the SM group, necessary for gauge invariance.

The solution is provided by introducing Yukawa interactions between the fermion fields and the Higgs field Φ , also a doublet under $SU(2)$,

$$\mathcal{L}_{Yukawa} = -y^{ab} \bar{q}_{\alpha L}^a \Phi d_{\alpha R}^b - y'^{ab} \bar{q}_{\alpha L}^a \tilde{\Phi} u_{\alpha R}^b - y''^{ab} \bar{l}_L^a \Phi \ell_R^b + \text{h.c} \quad (2.47)$$

where y , y' and y'' are the Yukawa matrices, 3×3 matrices with one dimension for each generation. Also, $\tilde{\Phi} \equiv i\sigma_2 \Phi^*$. Note that there is no second term for the leptons, as the SM does not contemplate the right handed neutrino, ν_R . Also, this Lagrangian breaks explicitly the chiral symmetry but yields a singlet representation, safe for gauge invariance. Next, writing the field Φ in terms of the unitary gauge as in the EWSB, $\phi^0(x) = v + H(x)$,

$$\begin{aligned} \mathcal{L}_{Yukawa} &= -\frac{1}{\sqrt{2}}(v + H)y^{ab} \bar{q}_{\alpha L}^a d_{\alpha R}^b - \frac{1}{\sqrt{2}}(v + H)y'^{ab} \bar{q}_{\alpha L}^a u_{\alpha R}^b \\ &\quad - \frac{1}{\sqrt{2}}(v + H)y''^{ab} \bar{l}_L^a \ell_R^b + \text{h.c} \\ &= -\frac{1}{\sqrt{2}}(v + H)y^{ab} \bar{D}_\alpha^a D_\alpha^b - \frac{1}{\sqrt{2}}(v + H)y'^{ab} \bar{U}_\alpha^a U_\alpha^b \\ &\quad - \frac{1}{\sqrt{2}}(v + H)y''^{ab} \bar{L}^a L^b + \text{h.c} \end{aligned} \quad (2.48)$$

where the expression has been rearranged to define Dirac fields in spinor notation,

$$D_\alpha^a = \begin{pmatrix} d_\alpha^a \\ \bar{d}_\alpha^{+a} \end{pmatrix}, \quad U_\alpha^a = \begin{pmatrix} u_\alpha^a \\ \bar{u}_\alpha^{+a} \end{pmatrix}, \quad L_\alpha^a = \begin{pmatrix} \ell_\alpha^a \\ \bar{\ell}_\alpha^{+a} \end{pmatrix} \quad (2.49)$$

After diagonalising the three Yukawa matrices, the eigenvalues terms are related to the masses, which can be identified for each generation as,

$$\begin{aligned}
m_{d^i} &= y^{ii} v / \sqrt{2} \\
m_{u^i} &= y'^{ii} v / \sqrt{2} \\
m_{\ell^i} &= y''^{ii} v / \sqrt{2} \\
m_{\nu^i} &= 0
\end{aligned} \tag{2.50}$$

There is a major consequence from the differences between the representation in generation space (Equation 2.46, $SU(2)_L$ doublets), and in mass space, after diagonalising the Yukawa matrices. D_α^a and U_α^a rotated to diagonalise their corresponding Yukawa matrix are affected by different transformations, however the individual $d_{\alpha L}^a$ and $u_{\alpha L}^a$ fields are part of the same $SU(2)_L$ doublet. The effect can be seen writing the W^\pm interactions in the mass state representation of the fields which become off-diagonal,

$$\frac{-g}{\sqrt{2}} (\bar{u}_L \quad \bar{c}_L \quad \bar{t}_L) \gamma^\mu W_\mu^+ V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c} \tag{2.51}$$

$$\begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} = V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} \tag{2.52}$$

where the superscript ' denotes the mass representation and V_{CKM} is the Cabibbo-Kobayashi-Maskawa matrix [Cabibbo, 23]. This unitary matrix is the product of the transformations that diagonalise the y and y' Yukawa matrices, which encodes the mixing of the different generations of fields in charged-mediated weak interactions. This is known as flavour violation, where a weak interaction of a quark can result on changing its flavour⁵. The neutral current interactions, mediated by the Z boson, relate the fields with the same charge, affected by the same transformation, hence not spawning a mixing matrix. The reason for not having *Flavour Changing Neutral Currents* (FCNC) explicitly in the SM Lagrangian. On the other side, the leptons are represented with the same $SU(2)_L$ doublet, so any mixing of lepton generations is not present in the theory.

5: Maybe add a feynman diagram Rafael notes

There is still another interesting feature that arises from the CKM matrix. The standard representation [24] of the matrix takes into account invariant phase rotations of the fields, leaving as free parameters three angles θ_{12} , θ_{23} and θ_{13} (chosen to lie in the first quadrant so $\sin \theta, \cos \theta \geq 0$), and a single complex phase δ that cannot be rotated to zero. The matrix reads,

$$\begin{aligned}
V_{CKM} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}
\end{aligned} \tag{2.53}$$

6: The three symmetries are related as the combination, CPT symmetry, must always be respected in theory.

where $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$. The presence of the complex phase leads to different couplings for anti-matter, as the complex phase will switch sign, leading to matter/anti-matter asymmetry. This asymmetry in flavour-changing processes is the only source in the SM of CP violation, or T violation (from the time-reversal symmetry⁶) however, as discussed in Section[], fails to describe the current matter/anti-matter content of the universe. The CKM matrix is predicted and measured to be almost diagonal, with very small sources of CP violation, or V_{ub} and V_{td} . The current matrix as in 2022 [11] reads,

$$V_{CKM} = \begin{pmatrix} 0.97401 \pm 0.00011 & 0.22650 \pm 0.00048 & 0.00361^{+0.00011}_{-0.00009} \\ 0.22636 \pm 0.00048 & 0.97320 \pm 0.00011 & 0.04053^{+0.00083}_{-0.00061} \\ 0.00854^{+0.00023}_{-0.00016} & 0.03978^{+0.00082}_{-0.00060} & 0.999172^{+0.000024}_{-0.000035} \end{pmatrix} \quad (2.54)$$

2.2 Successes and shortcomings of the Standard Model

Since the formulation of the SM, most experimental observations and measurements have been described successfully by the model. Throughout the years, predicted particles have been found and multiple precision measurements have tested its validity. However, there are theoretical and experimental issues not solved by the theory, leading to the conclusion that the SM is an effective theory and there is a more complete theory that can explain the whole range of observations. In this section, a brief summary of the measurements of the SM parameters is presented, followed by an overview of the main open questions.

2.2.1 Experimental measurements

Decades of experiments have performed measurements into parameters that define the SM. The SM can be summarised with nineteen parameters, which have been described in this chapter: nine fermion masses (six for quarks and three for leptons), the three gauge couplings (g_s , g and g'), the Higgs vacuum expectation value (v), the Higgs mass, four parameters of the CKM matrix (three angles and the complex phase), the QCD CP violating phase⁷. There is no underlying relation between these parameters, only being set from experimental observations. With these parameters measured, theoretical predictions of observables can be tested with experimental data in order to explore new physics.

7: This has not been described

One typical observable in particle physics is the cross-section σ , the expected interaction rate between two interacting particles in terms of the effective surface area measured in pb (picobarn, $1pb = 10^{-40} m^2$). The cross-section of a process depends on the interacting forces involved, as well as the energy and momentum of the interacting particles, which can be calculated from the S-matrix (scattering matrix) using relativistic mechanics. Feynman diagrams are a tool to translate a visual description of a process to a mathematical expression, the matrix amplitude, which is proportional to the probability of the specific process happening and

needed for the computation. The decay width, Γ , can be computed in similar fashion to obtain another common observable, the Branching Ratio (BR). The BR of an unstable particle is the probability for it to decay into specific particles among all possible states. It is computed dividing the Γ of the specific process with respect to the sum of all the possible process. Both σ and Γ are calculated from perturbation approximations, as the actual process is not the product of just one Feynman diagram, but all the possible interactions that lead to the same final state including loops, interferences and radiative corrections, refereed to as high order corrections. However, each particle interaction is proportional to the probability making higher order corrections become less important. Typically, *leading-order* (LO) calculations use only the leading order terms from the perturbation expansion, while if complemented by higher order corrections are referred to next-to-leading-order (NLO) or next-to-NLO (NNLO) calculations.

Figure[] shows a summary of a wide range of cross-section measurements by the ATLAS Collaboration, compared to the theoretical predictions, showing an excellent agreement between data and theory. In addition, the Higgs boson has been scrutinised since its discovery in 2012 [refs], both to characterise all its properties and because the uniqueness of its interaction with massive particles. Figure[] shows a summary of Higgs boson production cross-sections and measurements, including the coupling strengths to other SM particles, showing that the coupling is proportional with the mass of the resulting particle as expected from the Higgs mechanism. As the Higgs couples with any particle that acquires mass through its field, it is an excellent candidate to study any other particle still to be discovered.

On another note, the top quark is the most massive known particle known to date since its discovery in 1995. Such characteristic makes the top quark the only one that decays before hadronisation and an excellent candidate to study new particles, from much more massive ones that might decay to the top quark to lighter exotic particles that might decay into. Figure[] cross-section?. The top quark is also used to measure possible FCNC processes, for which there is no direct coupling in the SM as explained in Section[]. Nevertheless, higher order processes involving W^\pm are possible making FCNC processes very unlikely. As the top quark can decay to any other SM particle, observation of FCNC processes would imply the need to expand the SM. Table with measurements

Interest in top physics

2.2.2 Open questions

Observed neutrino oscillations [25] are only possible if there are mass differences between the three neutrino generations, which implies non-zero masses for some neutrinos although not measured directly. A mass term for neutrinos could be added in the SM in different ways, through adding right-handed neutrinos or as describing neutrinos as Majorana particles [ref]. Nevertheless, the description of the SM particles needs at least seven additional parameters: three for the neutrino masses, three

for their mixing angles and one CP violating phase for the neutrino mixing Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [], similar to the CKM quark flavor matrix.

The SM also fails to describe the other known fundamental force in nature, gravity. There is no renormalisable quantum field theory for gravity successfully described as general relativity only describes macroscopic systems, with the observation of gravitational waves as the latest achievement []. There are theories like string theory that provide alternatives although difficult to test experimentally. The SM is understood as an effective theory of a more complete unified theory, hence only valid at low energies as it breaks in the most extreme scenario around the Plank scale ($M_P = \sqrt{\hbar/(8\pi G_N)} \sim 2.4 \cdot 10^{18}$ GeV), where gravitational effects are supposed to become as important as the other forces in the SM.

The SM describes what is known as baryonic matter, which accounts for about 5% of the energy density of the universe. Cosmology, which studies the composition of the universe, estimates huge amounts of *dark matter* (DM) and *dark energy*, phenomena not contemplated by the SM. The existence of DM was postulated as extra non-luminous matter needed to explain observed rotation curves of galaxies not matching the gravitational pull of observed stars []. In addition, gravitational lensing effects observed in some galaxy collisions [] also provide the need of huge invisible mass concretations. More recently, the WMAP and Planck collaborations have studied anisotropies in the cosmic microwave background (CMB) [], postulating cold DM. On the other side, the universe is observed to be expanding at an accelerated rate compatible with dark energy, understood to be the product of an intrinsic space-time energy density, or cosmological constant, that causes the expansion. Observation of red-shift of light from supernovae, used as standard candles, points that cosmological objects are moving away at faster rate with the distance []. Further, studies of the CMB provide another measurements of the accelerated expansion []. The most updated results [] points that baryonic matter accounts for a mere % of the total energy density of the universe, dark matter for %, while dark energy for %.

The universe seems to be completely made up by matter. To explain the imbalance in abundance of matter and anti-matter, often referred as matter/anti-matter assymetry, the SM only provides one not nearly sufficient source of CP violation in the quark weak interactions, as mentioned in Section[]. Additional sources are added such as the complex phase in the PMNS matrix, however more phenomena is needed to have generated the current net balance of matter, like possible baryon number violating effects at high energy scales.

Besides the natural phenomena uncovered by the SM, there are also what are known as naturalness problems. Those are aesthetic concerns about the precise different values of some of the SM parameters, which seem "unnatural" if there is no hidden mechanism behind. The general consensus is that the fewer fine tuning is needed in a theory, the more natural it is. Although these matters are completely subjective, these unexplained features could be a hint for the existence of a new underlying mechanism that could complement the SM. The first problem is commonly

named as the hierarchy problem, where the cutoff energy of the SM (Λ_{SM}) is commonly set to the Planck scale, $\sim 10^{18}$ GeV, but in contrast the EW scale ($v \sim 246$ GeV) is very small. The problem can be read as there is no apparent reason for the EWSB to happen at its scale, orders of magnitude smaller than the Plank scale. When calculating high-order corrections from the SM like the computation of loops, Λ_{SM} has to be introduced to manage ill-defined divergent integrals in a process called regularisation []. As particles are endlessly interacting with virtual particles, represented by self-energy diagrams like Figure[], their mass are effectively diverging. This is solved by interpreting the physical mass measured in experiments as the resulting after dealing with the divergent terms. Leading radiation corrections for the fermion masses are of the order of $\log \Lambda_{\text{SM}}$, sensitive to the scale but the fine-tuning is considered small. On the other side, the physical Higgs mass including radiation corrections reads,

$$m_H^2 = m_0^2 + \frac{3}{8\pi^2 v^2} \Lambda_{\text{SM}}^2 [m_0^2 + 2m_W^2 + M_Z^2 - 4m_t^2] + \mathcal{O}(\ln \frac{\Lambda_{\text{SM}}}{m_0}) \quad (2.55)$$

with m_0 the bare Higgs mass. The nature of the hierarchy problem is evident in the correction as the Higgs mass is more sensitive to the cutoff scale and requires huge tuning to counter the Λ_{SM} term and achieve such a low measured physical mass. It can also be observed that the most important correction is given by the top quark, and it is often questioned whether the reason for the huge mass of this quark could hide a solution. Although the Higgs mass and the EW scale are difficult to justify, it can be argued that the appearance of the Λ_{SM} is related to the chosen regularisation scheme and cut-offs play no physical role. Some Feynman diagrams?

Another related problem is the fermion mass hierarchy, as the fact that the masses of the SM particles range from ~ 1 MeV to ~ 173 GeV (of the top quark [11]), it is not understood. Additionally, there is also not an apparent reason for the three mass families of quarks and leptons. It might be related again to renormalisation, since fermion masses also have correction terms with the logarithm of the cut-off scale as stated before.

There is also the problem known as the strong CP problem. The most general QCD Lagrangian could include a CP-violating angle without breaking any symmetry or the renormalisability of the theory. This would lead to the prediction of axion particles and the neutron having non-zero electric dipole moment. Measures of the former in ultracold neutrons and mercury, constrain the CP-violating angle to $|\theta| \lesssim 10^{-10}$ [], and supposes an incredibly low value for a parameter that can have any value in the theory.

2.3 Beyond the Standard Model

Where to look? SM right?

Statistical tools?

The study of particle physics is performed at the TeV scale, in the order of 10^{-15} m, which requires large and complex machines only possible within international collaborations. CERN is one of the biggest and most respected scientific collaborations and, since its origins in the 1950's, has hosted many groundbreaking experiments. The Large Hadron Collider (LHC) is the world's largest particle accelerator, situated underground in the France-Swiss border and in operation since September 2008. The A Toroidal LHC ApparatuS (ATLAS) detector is one of the experiments hosted within the LHC and records the collisions for further data analysis. The work in this thesis is based on the recorded proton-proton collision data at a center-of-mass energy, \sqrt{s} , of 13 TeV between 2015 and 2018.

3.1 The LHC

The LHC is a circular particle accelerator with a circumference of 27 km, situated on an average of 100 m underground. The primary activity is the collisions of protons, however proton-Pb and Pb-Pb collisions are also typically performed for one month a year. Particles are steered, collimated and boosted by different types of superconducting magnets and structures along the accelerator ring.

Proton beams circulate over different accelerators before reaching the LHC and the optimal energy. [Figure](#) shows a schematic view of the CERN accelerator complex. First, protons are extracted from ionised hydrogen and accelerated up to 50 MeV with LINAC2, a linear accelerator. Then, protons are injected into the Proton Synchrotron Booster (PSB), an accelerator made of four synchrotron rings of 157 m in circumference, increasing the energy up to 1.4 GeV. Similarly, the protons are accelerated in sequence to 26 GeV and 450 GeV by the Proton Synchrotron (PS), a circular accelerator of 628 m in circumference, and the Super Proton Synchrotron (SPS), of 6.9 km in circumference. Finally, the protons are injected to the two pipes of the LHC and boosted up to 6.5 TeV before collision. For the Pb operations, the extraction and accelerators prior to the SPS are performed instead using LINAC3 and the Low Energy Ion Ring (LEIR).

Inside the LHC, two particle beams travel close to the speed of light before they are made to collide. The two separated particle beam pipes are designed to operate at 6.5 TeV in opposite directions and kept at ultrahigh vacuum, down below 10^{-13} atmospheres. Surrounding the pipes, superconducting magnets built from niobium-titanium alloy coils generate strong magnetic fields of the order of 8 T through an electric current of 11.8 kA. The magnet coils are surrounded by the magnet yoke, tonnes of solid steel sheets designed to keep the wiring firmly in place and stabilise the temperature of the magnets. The magnets are cooled down to 1.9 K with superfluid helium provided by a cryogenic system requiring 120 tonnes of helium. The rest of external layers are dedicated to shield

the particle radiation, insulate the magnet or maintain the vacuum and the whole structure of up to 28 tonnes. Different varieties and sizes of magnets constitute the accelerator, mainly 1232 dipole magnets 15 m in length which bend the particle beams to follow the circular trajectory, and 392 quadrupole magnets, each 5-7 m long, which focus the beam. Other types of magnets are used to correct the beam shape or to align them for collision. [Figure](#) shows the cross-section of a dipole magnet of the LHC, with the different components.

Particles in each beam pipe are accelerated by 8 superconducting radiofrequency (RF) cavities, metallic chambers with alternating electric fields housed in cryogenic chambers, which also space the particles into compact groups named bunches. When protons are accelerated, the bunches contain more than 10^{11} protons spaced every 25 ns (around 7 meters).

The particles are brought to collision at interaction points (IPs) by superconducting multiple magnets focusing the beams. Four detectors are situated at the different IPs: ATLAS, CMS, LHCb and ALICE. The first two are multi-purpose experiments that study a wide range of physics, comprising precision measurements of the SM as well as searches for beyond the SM such as Supersymmetry, exotic particles or dark matter. Both collaborations are formed by around 3000 scientist each, the two largest at CERN. The LHCb experiment is dedicated to explore hadrons containing b- or c-quarks, especially investigating CP-violating processes. The ALICE experiment is the only experiment fully focused on heavy-ion collisions and therefore specialised on QCD physics.

The Large Electron-Positron Collider (LEP) was the previous main experiment at CERN and its operations finished in 2000 to start the LHC installation in the same tunnel, replacing the predecessor. LEP was designed to collide e^+e^- beams and operated at a maximum of $\sqrt{s} = 209$ GeV. LHC was designed to accelerate protons or lead ions, which in comparison, are easier to accelerate to higher energies and provide more collision data, although harder to be detected and studied. LEP explored the EW scale and provided precision measurements of the SM and set a lower bound for the mass of the Higgs boson, later discovered during LHC operation in 2012. In September 2008, the first LHC operations started and in November 2009 the first collision was produced.

3.1.1 Performance in Run 2

The number of events of a certain process is key for its study and can be written as

$$N = \sigma \mathcal{L} = \sigma \int \mathcal{L} dt \quad (3.1)$$

where σ is the event cross-section for the given process, \mathcal{L} the integrated luminosity and \mathcal{L} the instantaneous luminosity. The cross-section highly depends on the center-of-mass energy, \sqrt{s} , one of the main characteristics of a particle collider. As a general rule, the higher the \sqrt{s} , the higher is the σ for rare SM processes, interesting for precision measurements, or searches for new massive particles.

The instantaneous luminosity is another of the main characteristics of a particle collider, which for the LHC can be approximated to

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} F \quad (3.2)$$

with f the revolution frequency, $n_{1,2}$ the total number of protons in each beam and F a reducing factor accounting for the beams are not colliding exactly head-on and other geometric and beam effects. The first parameter can be approximated to $f = c/27 \text{ km} = 11 \text{ kHz}$ and the total number of protons from the nominal number of bunches, 2808, which can contain up to 10^{11} protons. Finally, the denominator is the approximated transverse beam area with transverse beam size $\sigma_{x,y} = 16.6 \text{ }\mu\text{m}$. With these assumptions, the instantaneous luminosity is of $\mathcal{O}(10^{34} \text{ cm}^{-2}\text{s}^{-1})$.

Talk more in detail about the effects?

During 2010 and 2011, the LHC delivered proton collisions at $\sqrt{s} = 7 \text{ TeV}$, while in 2012 were at $\sqrt{s} = 8 \text{ TeV}$. The first proton physics run, namely Run 1, ended in February 2013, which enabled the discovery of the Higgs boson with $\mathcal{L} \sim 30 \text{ fb}^{-1}$. The evolution of the integrated luminosity delivered of the Run 2 to the ATLAS experiment is shown in Figure for a total of $\mathcal{L} = 139 \text{ fb}^{-1}$, used in this thesis.

Another parameter of interest is the *pile-up*, which are the additional expected inelastic collisions that occur when crossing bunches of protons. The main source are the collisions that appear within a single bunch crossing, called in-time pile-up. In addition, out-of-time pile-up is referred to interactions from neighboring bunch crossings not resolved fast enough by the detectors. Pile-up effects are a challenge for physics analysis and is inherent to the increase of instantaneous luminosity. The mean interactions per crossing, $\langle\mu\rangle$, is a measure to quantify the pile-up and changes throughout the data taking periods, as shown in Figure.

should I include the formula?

The LHC operation parameters have changed every year and some are even above the original designed. Table summarises some of the parameters of the LHC original design and through different years.

3.2 The ATLAS experiment

The ATLAS detector is a multi-purpose particle detector, used to study a wide range of physics topics. It is installed 100 m underground at IP-1 of the LHC. Being 25 m in diameter, 44 m in length and 7000 tonnes in weight, ATLAS is the largest particle detector ever built at a collider. As illustrated in Figure, the detector has a cylindrical shape and is composed of several detector layers built around the collision point of the particles, with an almost full solid angle coverage.

The data recorded by the detector is used by the collaboration in an extensive physics program. One of the main branches is the measurement of SM processes, as the large luminosity provided by the LHC allows both precision measurements or to target uncovered phenomena. The data collected during Run 2 allows the study of the Higgs boson and its properties, heavily scrutinised since the first observation of the particle. In addition, measurements of the interactions and processes involving the top-quark are particularly interesting to probe the SM and BSM

theories. A large portion is dedicated to a wide range of BSM theories that includes supersymmetry, dark matter or additional resonances among others. Finally, the physics program also includes the study of the physics involving b -/ c -quarks and heavy ions.

All detector systems are designed such that they provide optimal performance to fulfill the different physics targets. For that, it is important that the detector can identify particles ranging from few GeV to several TeV with outstanding efficiency and resolution, providing accurate measurements of position and momentum. On the other hand, the conditions of the LHC are extreme and the electronics installed have to be highly resistant to radiation and fast for readout (25 ns between interactions).

3.2.1 Particle identification

3.2.2 Coordinate System

The convention to describe the particles recorded with the ATLAS detector is a right-handed coordinate system as illustrated in [Figure](#), with the origin at the center of the detector which is also the interaction point. The z -axis is defined in the counterclockwise direction along the LHC beamline, the y -axis points towards the surface and the x -axis towards the center of the ring defined by the accelerator. To describe the physics objects within the detector, spherical coordinates are used instead, with the polar angle, θ , measured from the z -axis while the azimuthal angle, ϕ , measured from the x -axis in the $x - y$ plane. The pseudorapidity, η , is usually used instead of the polar angle, as it transforms easily under relativistic boosts along the z -axis. For particles with energy E and forward momentum p_z , the expression of η can be found from the rapidity, y , in the high-energy approximation,

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \xrightarrow{\frac{m}{E} \rightarrow 0} -\ln \left(\tan \frac{\theta}{2} \right) \equiv \eta \quad (3.3)$$

ATLAS covers the pseudorapidity region up to $|\eta| < 4.9$, although physics analyses typically consider objects restricted to $|\eta| < 2.5$. In addition, the difference in η between two points, $\Delta\eta$, is invariant under Lorentz transformation, thus angular distances can be described in the $\eta - \phi$ plane as,

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (3.4)$$

Another useful expression is the momentum in the $x - y$ plane,

$$\vec{p}_T = \begin{pmatrix} p_x \\ p_y \end{pmatrix}, p_T = \sqrt{(p_x)^2 + (p_y)^2} \quad (3.5)$$

as at the time of the collision the particles are made to collide along the z -axis, the initial momentum of the transverse plane is known to be zero due to conservation of energy.

3.2.3 The Inner Detector

The Inner Detector (ID) [54, 55] is the innermost detector system, which encloses the beam pipe. This detector system provides precise tracking information of charged particles with momentum as low as 100 MeV with a $|\eta| < 2.5$ coverage. [Figure](#) shows an overview of the system, which is structured into three sub-detectors: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT).

Pixel Detector

The inner part of the ID is the silicon pixel detector comprising 4 cylindrical layers and 2 end-caps with 3 disc layers each. The layers are located between 33.25 mm to 122.5 mm around the beam pipe with a coverage of $|\eta| < 2.5$. A single 3D pixel is a radiation-hard silicon detector that produces a small measurable current when a charged particles passes through. The detector is especially important for the reconstruction of tracks, path of charged particles; the primary vertex reconstruction, the position of the main energetic collision; as well as for secondary vertex finding. The insertable b-layer (IBL) is the innermost layer, installed in-between Run 1 and Run 2, having the highest granularity with a pixel size of $50 \times 250 \mu\text{m}$ ($50 \mu\text{m}$ in the ϕ -direction and $250 \mu\text{m}$ in z-direction) for a total of 12 M pixels. In particular, the IBL is very efficient to reconstruct secondary vertices, which are key signatures of long-lived particles decays and crucial for the identification of b -hadrons. Furthermore, the three remaining layers have a pixel size of $50 \times 400 \mu\text{m}$. Overall, the pixel detector contains 86 M pixels with an expected hit resolution of $8 \times 40 \mu\text{m}$ for the IBL and $10 \times 115 \mu\text{m}$ for the rest of pixel layers. In addition, the system makes up around 50% of all ATLAS readout channels. For the next upgrade, the High Luminosity LHC (HL-LHC), a new fully silicon-based Inner Tracker (ITk) will replace the full ID.

Semiconductor tracker

The semiconductor tracker (SCT) is a silicon strip detector comprising 4 double layers in the barrel region and nine planar end-cap discs on each side, installed around the pixel detector. The planar strips technology is simpler compared to the silicon pixels, for lower resolution covering a larger area. The strips have a size of $80 \mu\text{m} \times 12 \text{ cm}$ and cover a region up to $|\eta| < 2.5$. The two layers within one layer-module are tilted by 40 mrad. Overall, the SCT has a resolution of $17 \times 580 \mu\text{m}$ with a total of 6.3 M readout channels. In general, the semiconductor-based detectors in ATLAS operate at a temperature between -10°C and -5°C to suppress different types of electronic noise.

Transition Radiation Tracker

The outermost part of the ID is the transition radiation tracker (TRT). In contrast to the others, the TRT is not based on silicon but is a gaseous detector system. It consists of around 300 k straw tubes with a diameter

of 4 mm filled with a gas mixture * of Xe (70%), CO₂ (27 %) and O₂ (3 %) and with a gold-plated tungsten wire in the tube centre with a potential different to the tube surface of 1.5 kV. When a charged particle hits the tube, the ionisation of the gas is detected as the signal. The straws have a length of 144 cm in the barrel region and 37 cm in the end cap, while the single hit resolution is 120 μm and 130 μm , respectively. The TRT only provides tracking information in the ϕ direction, as the tubes are parallel to the beamline. Besides, the TRT provides particle identification from emitted transition radiation at the material boundaries, since the straws are interleaved with polypropylene. Especially, electrons can be distinguished from charged pions due to larger transition radiation.

3.2.4 The Calorimeter System

The calorimeter system is responsible for the precise measurement of the energy carried by both charged and neutral particles as well as measuring shower properties to allow for particle identification. Showers are cascades of secondary particles which are formed when a highly energetic particle interacts with dense material. ATLAS uses sampling calorimeters which consist of alternating layers of active material (liquid argon and plastic scintillators) and passive detector material (copper, iron, tungsten and lead). While the active material measures the energy deposit of the particles going through, the passive material is designed to interact and absorb particles, thus induces the showering. The calorimeter system covers the region $|\eta| < 4.9$ and is placed between the central solenoid and the muon spectrometers. [Figure](#) shows an overview of the system, which is composed of two sub-detectors: the electromagnetic [61, 62] and the hadronic calorimeter [63, 64].

Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter encloses the ID and is a high granularity calorimeter based on liquid argon (LAr) technology with absorber plates made out of lead. To provide full coverage in ϕ , the EM calorimeter has an accordion-shaped structure where the active material is placed in the gaps between the lead absorber plates and the Kapton electrodes. The detector operates at -183 °C with a total of 170 k readout channels. The barrel region of the EM calorimeter covers the region $|\eta| < 1.475$ and consists of three layers with a 4 mm gap between them and a length of 3.2 m each, with decreasing granularity. The layer closest to the ID has a granularity of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$, while the second layer $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ and the outermost $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$. In addition, the two end-caps cover the region $|\eta| < 3.2$ with a slightly coarser granularity. In general, the absorption power at high energies of a calorimeter is quantified by means of the radiation length X_0 of its medium. It is defined as the distance over which the particle energy is reduced via radiation losses by a factor $1/e$. The thickness of the barrel region, given in terms of the radiation length, is $22 X_0$ and $24 X_0$ for the end-caps. Moreover, the intrinsic energy resolution of the EM calorimeter is,

* In Run 2, a mixture of Ar (80%) and CO₂ (20%) was used instead in modules with gas leaks

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{17\%}{E} \oplus 0.7\% \quad (3.6)$$

Hadronic Calorimeter

The second calorimeter system is the hadronic calorimeter located around the EM calorimeter and consists of three components providing around 19000 readout channels. First, the tile calorimeter is made out of alternating layers of steel as absorber material and scintillator plastic tiles as active material, being read out via photomultiplier tubes. The first two layers have the highest granularity with $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The barrel part of the tile calorimeter covers a region with $|\eta| < 1.0$ and the two extended barrels a range of $0.8 < |\eta| < 1.7$. The resolution of the tile calorimeter is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ [66, p. 3]. Next, the end-cap calorimeters are directly outside the EM calorimeter and are based on the LAr technology. The end-caps use copper as passive material and cover a region of $1.5 < |\eta| < 3.2$ with their highest granularity of also 0.1×0.1 within $|\eta| < 2.5$. Finally, the forward calorimeter is also LAr based and its first layer uses copper as absorber, which provides information for both electromagnetic and hadronic particles. The other two layers make use of tungsten as absorber which is better suitable for hadronic measurements. In total the forward calorimeter covers a region of $3.2 < |\eta| < 4.9$. The overall resolution of the LAr based hadronic calorimeters is,

check this numbers

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus \frac{1\%}{E} \oplus 3\% \quad (3.7)$$

$$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\% \text{ [67, p. 2].}$$

3.2.5 Muon Spectrometer

The muon spectrometer (MS) is the outermost detector system of ATLAS, designed to identify and measure the energy of muons. **Figure** shows an overview of the system, which is composed of four detector systems grouped into trigger and precision muon tracking chambers. In total, the MS has more than one million readout channels and is embedded in three superconducting toroidal magnets, that provide a magnetic field in the ϕ -direction. Muons mostly reach the MS without losing energy, and the strong magnetic fields is design for their precise measurement. The p_T resolution is around 3% for 10-200 GeV and 10% for 1 TeV muons.

Muon Trigger Chambers

The muon trigger chambers are designed for a fast readout to provide energetic muon identification in a timescale compatible with every bunch crossing. In the barrel region, $|\eta| < 1.05$, three layers of resistive plate chambers (RPCs) which consists of two parallel plates with high resistivity and filled with a gas mixture (94.7% $C_2H_2F_4$, 5% Iso- C_4H_{10} , 0.3% SF_6). The RPCs provides an $\eta - \phi$ measurement with a spatial resolution of 10 mm and time resolution of 1.5 ns. thin gap chambers (TGCs) are installed in the end-caps, $1.05 < |\eta| < 2.4$, which are multi-wire chambers

filled with a gas mixture (55% CO₂ and 45% n-C₅H₁₂) with the wires separated by 1.8 mm. Besides trigger information, the TGCs provide ϕ information with a resolution of 5 mm.

Precision Muon Tracking Chambers

The precision muon tracking chambers are designed to provide high resolution and precision tracking information, usually after the trigger decision of the trigger muon chambers. The system is mainly composed of monitored drift tubes (MDTs), installed in the barrel and end-cap region covering $|\eta| < 2.7$. MDTs are aluminum drift tubes with 3 cm of diameter filled with a gas mixture (95% Ar and 5%). Each chamber contains 3-8 layers of drift tubes with a spatial resolution of 35 μm . The forward region of the system, $2.0 < |\eta| < 2.7$, is covered by cathod strip chambers (CSCs) and provide a resolution of 40 μm in the radial direction and 5 mm in the ϕ direction. These chambers are proportional multi-wire chambers, similar to the TGCs, with lower response time.

3.2.6 Magnet System

The magnet system is of major importance to allow momenta and charge measurements, bending the trajectory of charged particles depending on these properties. **Figure** shows an overview of the system, which consists of two sub-systems: the central solenoid magnet, which is located between the ID and the calorimeters, and the toroidal magnet system, embedded within the MS. The solenoid generates a constant magnetic field of 2 T, with a superconducting magnet made out of NbTi cooled via liquid helium to a temperature of 1.8 K. There is one barrel toroid magnet and two end-cap toroid magnets with eight coils each, where each deliver an inhomogeneous magnetic field of roughly 0.5 T and 1 T, respectively.

3.2.7 Trigger System and Data Acquisition

With a bunch crossing every 25 ns, the LHC produces collisions at a frequency of 40 MHz at nominal operation conditions. For ATLAS is translated to an unmanageable rate of more than 60 TB/s of data. The TDAQ system is designed to select and record the events considered interesting for analysis for an average storage of 1 kHz.

The trigger system, summarised in **Figure**, is structured into two parts since Run II: the Level-1 (L1) hardware-based trigger system and the software-based high level trigger (HLT). The L1 trigger uses reduced-granularity information from the calorimeters and from the muon RPCs and TGCs to select events with interesting objects (normally high p_T electrons, muons, photons, jets or high missing transverse momentum). The latency is of 2.5 μs and reduces the rate from 40 MHz to about 100 kHz, which corresponds to 100 collisions. The information of the collisions is stored in buffers and the Central Trigger Processor (CTP) performs the decision based on the inputs of the other L1 sub-systems. The L1 trigger output are regions of interest (RoIs) in η and ϕ , which are passed to the HLT. The HLT uses the full detector information within

the RoIs to reduce the event rate down to approximately 1 kHz with a latency of 200 ms.

After, the data is transferred to a computing centre for further processing and storage. An offline data quality monitoring system performs checks on fully reconstructed events, to ensure their quality for use in physics analyses. Some of the criteria to validate are requirements on the condition and performance of the beams and different ATLAS sub-detectors at the time of operation. As summarised in [Figure](#), from the 156 fb^{-1} of integrated luminosity delivered by the LHC during Run 2, 147 fb^{-1} were recorded by the detector and 139 fb^{-1} cataloged as good-quality data.

Physics simulation of proton collisions

4

Proton collisions are complex processes and their understanding is essential to interpret the experimental data from the LHC. Normally, physics analyses rely on the ability to accurately simulate the various processes of proton-proton collisions and the interactions with the detector in order to perform comparisons with the recorded data and quantify its the level of agreement of the SM. The simulation is usually performed with Monte Carlo Monte Carlo (MC) generators, which are stochastic tools that incorporate both theoretical predictions and empirical results to describe the statistical processes. Sections bla and bla

4.1 Event simulation

The typical proton-proton collision at the LHC is depicted in [Figure](#). The inelastic scattering is the main interesting process, where the energy of the system is large enough so a constituent of each proton (partons) interact and allow the production of additional particles. The interaction that involves any of the other partons, normally at lower energies, is referred to as underlying event. A key phenomenon is the parton shower, a processes where due to the strong interaction, particles loose energy due the radiation of gluons which further generate quark-antiquark pairs, which in turn radiate gluons again in a chain reaction. These generated particles loose energy progressively down to the point where QCD leaves the perturbative regime (~ 1 GeV) and the hadronisation occurs, when quarks and gluons form hadrons, colorless bound states. To complete the simulation of the collision, the pile-up is included which adds the effects from the other proton collisions that originate from the same or previous bunch-crossing.

4.1.1 Factorisation theorem

The cross-section to produce a final stat X from the hard scattering of two protons, $\sigma_{pp \rightarrow X}$ can be factorised into two components in perturbation theory, as the strong coupling constant, α_s , is small at high energy kinematic regimes. Using the factorisation theorem [cite](#),

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \cdot \hat{\sigma}_{ab \rightarrow X}(x_a p_a, x_b p_b, \mu_F^2, \mu_R^2), \quad (4.1)$$

where $f_i(x_i, \mu_F^2)$ are the parton distribution function (PDF) for partons $i = a, b \in \{g, u, \bar{u}, d, \dots\}$ and encodes the probability of finding a parton of type i within the proton carrying a fraction of the proton's momentum x_i , at the factorisation scale μ_F . The dependence of the scale appears from performing only fixed-order calculations and the value is typically set comparable to the energy of the process, for example, to the total transverse mass of the final-state particles. The partonic cross-section,

$\hat{\sigma}_{ab \rightarrow X}(x_a p_a, x_b p_b, \mu_F^2, \mu_R^2)$, is calculated at finite perturbative order, hence the additional dependence on the renormalisation scale, μ_R , at which to evaluate α_s .

4.1.2 Parton density function

The PDFs are crucial for the accurate description of the partons that form the protons. The first type of partons are the valence quarks which determine the quantum numbers of the proton. In addition, gluons and virtual quark-antiquark pairs (sea-quarks) are also part of the proton and come from the vacuum fluctuations. A PDF, $f_i^A(x_i, Q^2)$ describes the probability density of a parton of a certain type, i , inside a given hadron, A to carry a certain momentum fraction, $x = p_i/p_A$ evaluated at a specific momentum transfer Q^2 . In general, PDF are extracted from empirical measurements performed at a specific scale. Then, the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations are used to extrapolate the PDF to different scales. Other alternatives to extract the functions like lattice QCD are possible, but very computationally challenging. There are dedicated collaborations such as the NNPDF, CTEQ and MSTW that provide PDFs for physics analyses. Figure shows the NNPDF3.0NLO PDF set for the different proton partons and two different factorisation scales.

reference equation of the running

There are two main factorisation schemes to describe processes involving b -quarks: the four-flavour scheme (4FS) and the five-flavour scheme (5FS). The 4FS treats the b -quarks massive ($m_b > \mu_R$) and since $m_b > m_p$, they are not included in the sea of quarks and do not have an associated PDF. In the context of QCD perturbative evolution, one of the consequences is that calculations at lower scales $\mu_R < m_b$ are especially impacted as the α_s running depends on the number of quark flavours in the initial state, $n_f = 4$. On the other hand, at high scales the mass effects are negligible and usually described by the 5FS, in which the b -quark is considered massless, included in the initial state and treated as the other quarks, $n_f = 5$.

4.1.3 Matrix element

The computation of the partonic cross-section of partons i, j into an arbitrary final state X , is related to the ME amplitude as,

$$\hat{\sigma}_{ij \rightarrow X} \sim \sum_{k=0}^{\infty} \int d\Phi_{X+k} \left| \sum_{l=0}^{\infty} M_{X+k}^l \right|^2 (\Phi_F, \mu_F, \mu_R) \quad (4.2)$$

where PDFs and other normalisation factors are removed for compactness. M_{X+k}^l is the ME amplitude for the production of X in association with k additional final-state partons, or legs, and with l additional loop corrections. In a perturbative regime, the ME amplitudes for increasingly complex processes (diagrams with additional legs and loops) tend to decrease. As a result, the cross-section is generally computed at a perturbative order, without the sum computed to infinity and for a choice of μ_F and μ_R . The Leading Order (LO) is the lowest possible order for

the calculation, with $k = l = 0$. Next, $l = 0, k = n$ provides the LO computation for the production of $X + n$ jets. Finally, $k + l \leq n$ corresponds to a N^n LO prediction for the production of X , while also to a N^{n-k} LO prediction for the production of X in association of k jets.

4.1.4 Parton shower

One problem that arises in the fixed order computations of the differential cross-section is the appearance of logarithmic divergences from collinear splitting that originate from the integration of the phase space, Φ , of the additional k partons. For an inclusive cross-section computation, these divergencies cancel out with virtual corrections order by order, following the KLN theorem. In this case, the basic event is simulated at fixed order while the QCD emission process (splitting) is computed with the PS algorithm, which generates a sequence of emissions with decreasing angle or energy. The algorithm recursively produces the typical splitting processes ($g \rightarrow q\bar{q}$, $g \rightarrow gg$ and $q \rightarrow qg$) for each parton until the energy of the shower reaches ~ 1 GeV, the hadronisation scale. This showering process that is applied to the final products after the hard-scattering is referred to as final state radiation (FSR), while the simulation of the initial state radiation (ISR) is performed to the incoming partons. In the case of ISR, the subsequent emissions grow on energy and are modelled instead with a backwards-evolution algorithm.

There is an incompatibility with ME and PS for a full cross-section computation at order $n > 1$, as there is a possible overlap in the phase space of the extra partons that are considered for the ME at order n with the ones considered in the splitting at order $n - 1$. There are different approaches to solve the double counting, known as ME-PS matching. The most common strategy is known as slicing, which defines a matching scale where the higher energy region is covered only by the ME while any additional parton with energy below the scale is vetoed and only covered with the PS algorithm. With this strategy, both energy regions are described with the corresponding optimal algorithm.

4.1.5 Hadronisation

The hadronisation process starts when the energy of the PS emissions is low enough to reach the hadronisation scale (~ 1 GeV), where the perturbative regime of QCD is not valid. At that point, the partons from the PS have defined momentum, flavour and color and further description of the emissions has to rely on phenomenological models. The process consists on a reconstruction algorithm that groups together the partons into different hadrons, that can further split, until all partons are confined into stable hadrons.

The two most widely used models are: the Lund string model and the cluster model. In the first, the quark-antiquark pair colour interaction is described as a string with a potential assumed to be linearly increasing with the distance, emulating the QCD potential. The string then splits forming new quark-antiquark pairs when the energy stored passes the quark-antiquark total mass, forming hadrons which momentum is determined from the initial momentum by a fragmentation function.

The momentum of the hadron as a fraction x . On the otherhand, the second model is based on forcing the final state gluons to split into quark-antiquark pairs and then grouping all quarks in colour-singlet clusters, allowed to decay and split into smaller clusters or hadrons. For both models, the process is repeated iteratively until only stable hadrons remain.

4.1.6 Pile-up and underlying event

Other interactions aside the hard-scattering event have to be included in the MC simulations to properly model the LHC collisions, the pile-up and the underlying event. Both sources mainly consist of soft QCD interactions, the first arising from other protons colliding in the same or previous bunch-crossing while the second being the interaction of the other partons that do not originate the hard-scattering process. Both mainly consist of soft QCD interactions in the forward regime, close to the beam axis, and the description is based on the combination of phenomenological models and the configuration of the LHC beam. In the especial case of out-of-time pileup (interactions from previous bunch-crossings), the simulation has to take into account the time response of the detector.

4.1.7 Monte Carlo simulation and generators

MC generators are dedicated software tools to perform the MC simulations, based on pseudo-random numbers to generate the events from the predicted distributions. They are generally classified according to which of the steps of the simulation can perform, with general purpose generators being capable of simulating the whole event process, while dedicated generators target specific parts of the chain, such as the ME or the PS computation.

The full process involving ME generation, PS, underlying event, hadronisation and fragmentation can be simulated by MC generators like PYTHIA 8, HERWIG 7 or SHERPA. However, PYTHIA 8 provides leading order calculations which are often not sufficient and hence, the generator is typically used to compute the PS process, which is based on the Lund string model. On the other hand, HERWIG 7 provides many ME at NLO, however since the fraction of negative event weights can be quite large (up to $\sim 40\%$ for certain generator setups), the generator is also typically used for PS computation, based on the cluster model. POWHEG-Box and MADGRAPH5_aMC@NLO are examples of other generators that are especially designed to provide accurate high-order ME calculations which are typically interfaced with PYTHIA 8 or HERWIG 7 for the simulation of PS and hadronisation.

More in detail, these tools have parameters to describe the non-perturbative processes that can be tuned using collision data. The most common tunes used by the ATLAS collaboration are the A14 parameters combined with NNPDF3.0LO PDFs set for PYTHIA 8 and the H7UE set with the MMHT2014LO PDFs sets for HERWIG 7. Throughout this thesis different combinations of MC generators and settings, which are detailed in the corresponding chapters and, if not stated otherwise, share the same

parameters. The mass of the top quark is set to $m_t = 172.5$ GeV, the mass of the Higgs boson to $m_H = 125$ GeV and the mass of the b -quark to $m_b = 4.8$ GeV for PYTHIA 8, to $m_b = 4.5$ GeV for HERWIG 7 and to $m_b = 4.75$ GeV for SHERPA. The simulation involving b - and c -hadron decays is performed with EVTGEN (except for SHERPA).

4.2 Detector simulation

With the proton-proton collisions simulated and the final-state stable particles defined, the remaining step is to simulate the interactions with the detector. The full ATLAS detector simulation is performed in two steps: first, the ATLAS detector response of the MC output is simulated and then the signals are reconstructed using the same algorithms used in real data. [Figure](#) depicts the different steps both for data and simulated MC events.

The GEANT 4 package is a widely used in physics to simulate the propagation and interaction of particles with matter. The simulation that includes all the geometry of the ATLAS sub-detector systems with GEANT 4 is referred to *Full Simulation* (FullSim), which is computationally expensive (several minutes per event) but gives the most accurate result. As more than 90% of the dedicated CPU time is spent on the calorimeter simulations, alternatives are used in practice. The *AtlFast-II* (AF-II) simulation is performed with faster simulation algorithms for the calorimeter simulation, ATLAS Fast Calorimeter Simulation (FASTCALOSIM), and for the ID simulation, Fast ATLAS Tracking Simulation (FATRAS). The rest of systems are simulated with GEANT 4 adding to significantly less CPU time while maintaining an adequate level of accuracy. Finally, the normalisation of a SM process is normally chosen according to the cross-section at the highest-order available and other corrections are applied in the form of scale factors (SFs), derived from the ratio between data and MC in specific calibration regions.

The concept of *reconstruction* refers to the use of algorithms for the identification of physics objects from the signals recorded in the different sub-systems of the detector. The physics processes described in this thesis produce electrons, muons, taus, photons, neutrinos and quarks in the final state. However, not all of these listed particles can be directly observed, as quarks form cascades of hadronic particles, neutrinos leave without interacting with ATLAS and tau leptons decay before reaching the detector. **Figure** illustrates the interaction of different particles with the ATLAS detector. Charged particles produce a track in the ID, electrons and photons shower in the EM calorimeter, hadrons shower in the hadronic calorimeter and muons leave signals in the muon spectrometer.

The reconstruction of the different physics objects used in this thesis analyses is described in the following chapter.

5.1 Base objects

The fundamental blocks used in the reconstruction algorithms are tracks, vertices and topo-clusters (or calorimeter energy clusters). All physics objects are composed by these blocks and introduced in the following section.

5.1.1 Tracks and vertices

Tracks are objects produced by charged particles interacting in the ID and used to identify their trajectory. The reconstruction consists in grouping hits in the different tracking sub-systems and requiring different criteria to ensure the quality of the tracks. The tracks that originate from the hard-scattering are referred to as primary tracks, and the origin of the track (vertex) is referred to as the primary vertex (PV) (PV)

As a first step, hits are built from groups of pixels and strips that reach a threshold energy deposit starting from the inner layers ID. The seed to reconstruct a track consists in three hits in the silicon detector, and then hits from the outer layers of the tracker compatible with the trajectory are added iteratively. When adding points, a score is assigned to the track to quantify the correctness of the track trajectory and suppresses the contribution of random collections of hits (or fake tracks). Then, a dedicated algorithm evaluates the different seeds to limit shared hits, which typically indicate wrong assignments. In addition, quality criteria are applied where tracks are required to have $p_T > 500$ MeV, $|\eta| < 2.5$, minimum of seven pixel and SCT clusters, a maximum of either one shared pixel cluster or two SCT on the same layer, no more than one missing expected hit (or hole) in the pixel detector and a maximum of two holes in both pixel and SCT. Also, requirements in the transverse impact parameter calculated with respect to the beam position, $|d_0| < 2$ mm, and related to z_0 , the longitudinal difference between the PV and d_0 along

the beam, $|z_0 \sin \theta| < 3$ mm. As a last step, TRT hits are added to the tracks after extrapolation.

Vertices are of particular interest as they are the origin of the charged particles or interactions. The PV is the most important, as denotes the origin of the hard-scattering interaction, but secondary vertices are also characteristic of long-lived particles or for heavy-flavour tagging.

not talked yet

For a given event, the PVs are reconstructed iteratively from tracks using a dedicated vertex finding algorithm. From a set of quality tracks, a candidate position is defined and the compatibility with the set of tracks in terms of weights is evaluated in order to recompute the vertex position. In each step then, the tracks that are less compatible are given smaller weights and, after the convergence of the optimal vertex position, are left unassigned and remain as input for the following vertex. The PV is defined as the vertex with the largest p_T^2 sum.

5.1.2 Topological clusters

Topological cell clusters, or topo-clusters, are objects reconstructed iteratively from calorimeter information and are the first step in the reconstruction of electrons, photons and hadrons. The seed consists of calorimeter cells which readout signal is four times higher than the background noise, and neighbour cells are added if the ratio is higher than two. As a last step, an extra layer is added regardless of the signal-to-background ratio.

a little bit weak?

5.2 Jets

Jets are the cone-shaped collimated showers formed by the hadronic cascades that originate from the complex interactions of quarks and gluons when travelling through the detector. These objects are essential for physics analyses with partons in the final state, especially b -quarks, which jets have particular properties that can be used to characterise them with great efficiency. Nevertheless, the kinematic properties of the cascades are challenging to define, as they can contain information from one or multiple final state partons and from the hard-scattering or other radiation processes. There are different possible definitions that depend of dedicated algorithms which group calorimeter information and do not depend on common QCD effects. Jet algorithms are collinear safe, referred to the jet not changing if two constituents are merged forming one with double the momentum (or vice-versa), and infrared safe, meaning that the reconstruction is not affected by adding low p_T particles.

5.2.1 Reconstruction

The jet reconstruction is typically performed using the anti- k_t algorithm. This family of algorithms merges clusters based on a relative distance defined as,

$$d_{i,j} = \min(p_{T,i}^{2n}, p_{T,j}^{2n}) \frac{\Delta R_{i,j}}{R^2} \quad (5.1)$$

with $p_{T,i/j}$ the p_T of the cluster i and j , $\Delta R_{i,j}$ the angle separation between them, R the chosen radius parameter that sets the size of the jet and n chosen integer that defines the p_T dependance of $d_{i,j}$. The decision to combine clusters or to define a cluster as a jet comes from comparing the $d_{i,j}$ value with the beamspot distance, $d_{i,B} = p_{T,i}^{2n}$. Clusters are grouped if $d_{i,j} < d_{i,B}$, otherwise the cluster i is defined as a jet, in an iterative process until all input clusters are used. The anti- k_t algorithm is defined by setting $n = -1$, which groups with higher priority the high energy clusters, and leads to a cone-shape around the highest object. This feature can be observed in [Figure](#).

Various jet collections based on the anti- k_t algorithm are used in ATLAS, two of them are used in this thesis: EMTopo jets and Pflow jets.

EMTopo jets

The so-called EMTopo jets were the primary jet collection used in physics analyses in ATLAS before the end of Run 2. The reconstruction is performed at the EM energy scale only using topo clusters with the anti- k_t algorithm implemented in the *FASTJET* software package. The jets used in this thesis are reconstructed with the radius parameter $R = 0.4$ with requirements in $p_T > 25$ GeV and $|\eta| < 2.5$. The EMtopo jets are calibrated in several steps summarised in [Figure](#). After the jet reconstruction, the jet direction is modified such that the jet originates from the primary vertex. Then, energy corrections based on pile-up are applied subtracting the average energy due to in-time pile-up and other residual corrections that depend on the number of PV and bunch crossings. After, absolute calibrations are applied to the jet energy scale JES and η derived from dedicated dijet MC events. Then, a global sequential calibration is set to improve the p_T resolution and the associated uncertainties from the jet fluctuations that can arise from various initial factors like the flavour or energy of the original parton. The final step is the in-situ calibration which is only applied to data, extracted from p_T and η comparisons to known well-modelled MC that include central jets in dijet events, $\gamma/Z + jets$ or multijet events.

PFlow jets

Particle Flow jets, known as PFlow jets, were introduced during Run 2 and combine tracking and calorimeter information. This collection of jets has improved energy and angular resolution compared to EMTopo jets and enhanced reconstruction and stability against pile-up. The reconstruction is also based on the anti- k_t algorithm with $R = 0.4$, and the first step consists in matching the tracks (from the ID) from charged particles to the topo-clusters. The energy deposits of the matched topo-clusters are replaced by the corresponding track momentum. Then, the resulting topo-clusters and the tracks matched to the PV are used as input of the anti- k_t algorithm. The jets are calibrated like the EMTopo jets in the range $20 \text{ GeV} < p_T < 1500 \text{ GeV}$.

5.2.2 Jet tagging

Jet or flavour tagging consists in identifying the parton flavour that generated the signal reconstructed as the jet. Efficient tagging is essential for analyses studying processes with b - or c -quarks in their final state (known as heavy flavour quarks), as it is additional information which can be used to select events based on the flavour of their jets and improve the selection of the signal.

Jets originating from the hadronisation of b -quarks, or b -tagged jets, leave a distinct signal due to the properties of b -hadrons: lifetime of $\sim 10^{-12}$ s (decay after 2.5 mm with a momentum of 30 GeV), mass of ~ 5 GeV and high decay multiplicity (including semi-leptonic decays). Figure shows a scheme of a typical signal, that includes displaced tracks from the PV with large d_0 . The signal of the c -hadrons is similar but not identical as the lifetime, mass and decay multiplicity are lower, which makes the distinction between these two kinds of jets difficult. The last type of jet is referred to light-flavour jets, which signal originates directly from quark fragmentation and can be easily separated from b -jets. However, other phenomena like long-lived particles, photon conversion or low quality tracks can also prompt displaced vertices and tracks.

Algorithms

Flavour tagging algorithms use the properties of a given jet to return a score, referred to as output discriminant, which indicates how likely the input jet is considered to be a b -, c - or *light*-jet. Two main taggers are used in ATLAS: the MV2c10 tagger which was the default option for EMTopo jets, and the DL1r tagger that is the recommendation for PFlow jets.

The MV2c10 tagger is based on the MV2 algorithm, which relies on boosted decision trees (BDTs) trained with several kinematic and other trigger taggers as inputs. This particular tagger was trained with $t\bar{t}$ and Z' events, to cover a large p_T spectrum, and b -jets defined as signal while the background was defined to consist of 7% c -jets and 93% light-jets.

The DL1r tagger is a multi-class Deep Neural Network (DNN) model, with three output nodes corresponding to the classification of the input jet to be a b -, c - or light jet. The final discriminant is given as a function of the three probabilities. The input consists in the MV2c10 tagger input, additional variables for c -jet identification used in a jet vertex finder algorithm and flavour probabilities provided by a recursive NN. The training set consists of the same $t\bar{t} + Z'$ events, weighted to have an equal mix of quark flavour jets.

NN defined at this point?

Comparing the two algorithms used in this thesis, the DL1r tagger is the most recent recommendation and relies on more advanced machine learning techniques than the MV2c10 tagger. The multi-class output together with the possibility of tuning the final discriminant computation, makes the DL1r tagger more flexible than the binary classification of the MV2c10. Regarding performance, the efficiency of both algorithms to tag true b -jets is comparable, while the rejection rates of c - and light jets is greater for the DL1r tagger. The improvement in rejection for the 60% WP, detailed below, is by up to 70% for c -jets and 120% for light jets.

Working points

The full spectrum of the final b -tagging discriminant is not directly used in physics analyses due to the complexity of the calibration. Instead, four different b -tagging working points (WP) are defined based on the b -jet acceptance efficiency evaluated on a $t\bar{t}$ sample: 60%, 70%, 77% and 85%, which are often referred to as *Very Tight*, *Tight*, *Medium* and *Loose* operating points, respectively. Most of the c - and light-jets do not pass the 85% WP, ending up in the b -tagging efficiency between 85% and 100%. Meanwhile, the jets that pass the 60% WP mainly consists in b -jets. This criteria is important when defining the b -jets for event selection, as the b -jets misidentification, so c - and *light*-jet acceptance inefficiency, improves for lower b -jet efficiency working points, therefore rejecting more background but with lower signal statistics. On the other hand, the pseudo-continuous b -tagging WP, so the WP that a jet passes, is additional information that can be used to further refine the selection or in multivariate methods.

5.3 Leptons

5.3.1 Electrons

Electrons interact with the ID and the EM calorimeter system. The typical signature is a track in the ID and electromagnetic shower in the EM calorimeter. Overall, the performance in terms of identification and reconstruction of electrons is high.

First, topo-clusters are selected and matched to ID tracks in the region $|\eta| < 2.47$ excluding the transition region of the barrel and end-cap ($1.37 < |\eta| < 1.52$). Next, the matched clusters are grouped to form superclusters, which are variable-size clusters, using a dynamic clustering algorithm. After a first energy and position calibration, tracks are matched to the electron superclusters. The calibration of the energy scale and resolution of electrons is computed from $Z \rightarrow ee$ decays and validated in $Z \rightarrow \ell\ell\gamma$. In addition, the energy resolution of the reconstructed electron is optimised using a multivariate algorithm based on the properties of showers in the EM calorimeter.

Further identification criteria are required for an electron candidate, passing a selection to increase the purity. The prompt electrons are evaluated with a likelihood discriminant to define three operating points with different purities: *Tight*, *Medium* and *Loose*. The discriminant is computed using variables measured in the ID and the EM calorimeter, chosen such that they discriminate prompt isolated electrons from other signals deposits (jets, converted photons or other electrons from heavy-flavoured hadron decays). The most important quantities are based on the track quality, the lateral and longitudinal development of the electromagnetic shower as well as the particle identification in the TRT. The probability density function to build the likelihood are derived from $Z \rightarrow ee$ ($E^T > 15$ GeV) and $J/\psi \rightarrow ee$ ($E^T < 15$ GeV) events.

Another requirement is the isolation criteria, to require the electron signal to be separated from other particles. Electrons are typically required to

be spatially separated from other particles, based on two quantities: a maximum value for the sum of transverse energy of topo-clusters in a $\Delta R = 0.2$ cone surrounding the electron and of the sum of transverse momentum of tracks around the electron, with a ΔR cone that decreases with p_T . Effects of leakage and pile-up are taken into account and also tracks are required to satisfy $p_T > 1$ GeV, $|\eta| < 2.5$ and quality criteria. In this thesis, the criteria used is the Gradient isolation which has an efficiency of 90% at $p_T = 25$ GeV and 99% at $p_T = 60$ GeV.

5.3.2 Muons

Muons leave the ATLAS detector without significant energy loss. The typical signal consists on a track in the ID and MS sub-detectors. There are different types of muons depending on which ID, MS or calorimeter information is available .

As a summary, the muon reconstruction has two stages: tracks are reconstructed independently in the ID and MS, and then are combined to form the muon tracks. The muon track candidates are built from track segments found in the different MS sub-systems. In the muon trigger chambers and MDTss, segments are reconstructed with a straight line to fit the hits of each detector layer after an alignment to the trajectory in the bending plane of the detector. The RPCs, TGCs and CSCs hits provide measurements in the orthogonal direction and the forward region of the detector to build additional track segments. The muon track candidates are then built from the track segments fit together using a global χ^2 fit. With that information, different types of muons can be defined.

The combined (CB) muons are the muon candidates obtained from using combined information from MS tracks that are extrapolated to the tracks of ID (an inside-out approach is also used). The segment-tagged (ST) muons are reconstructed from tracks in the ID extrapolated to typically one track segment in the MDTss and CSCs. ST muons are normally low in p_T and in regions with low acceptance. Calorimeter-tagged (CT) muons are built from an ID track that is instead matched to an energy deposit in the calorimeter compatible with a minimal ionising particle. The CT muon strategy outputs the lowest purity, although proves useful for detector regions not fully covered by MS, optimised for $15 \text{ GeV} < p_T < 100 \text{ GeV}$ and $|\eta| < 0.1$ is optimised. The fourth type, extrapolated (ME) muons, are only reconstructed using the MS with an acceptance of $2.5 < |\eta| < 2.7$.

The muon identification criteria (similar to the electron identification) is performed applying quality criteria to increase the purity of the selection. In order to identify prompt muons with high efficiency and a good momentum resolution, a requirement is done for amount of hits in the ID and the MS systems. Four different muon operating points are defined: *Tight*, *Medium*, *Loose*, *high p_T* and *low p_T* . The Medium and Loose working points are used in this thesis. The first one is widely used in physics analyses and is designed to minimise muon reconstruction and calibration systematic uncertainties. It consists of combined and extrapolated muons with three or more hits in at least two MDTs layers, or just one hit for $|\eta| < 0.1$ with no more than one hole in the MS. On the other hand, the Loose working point maximises the reconstruction

efficiency and accounts all types of muons, adding the segmented- and calorimeter-tagged muons for $|\eta| < 0.1$. The reconstruction efficiency for muons with $p_T > 20$ GeV at the Medium and Loose working points is 96.1% and 98.1%, respectively. beginning

The isolation criteria is based on track and calorimeter variables, similar to the electron case. The criteria improves the efficiency removing non-prompt muons, the ones not generated in the hard-scattering but in other parton shower processes for example, which are usually close to jets and other objects. The track related variable, $p_T^{\text{varcone30}}$ is the scalar p_T sum of the additional tracks in a cone $\Delta R = 10 \text{ GeV}/p_T^\mu$ (maximum of 0.3), that depends on the muon transverse momentum p_T^μ . The calorimeter related variable is the same as for electrons, build from the sum of energies around the muon track. In this thesis, the *FixedCutTightTrackOnly* working point is used, which is defined only with track isolation: $p_T^{\text{varcone30}}/p_T^\mu < 0.06$.

5.4 Taus

The τ leptons typically decay before reaching active electronics of the ATLAS detector and have to be identified via their decay products. The decay can be either leptonically (into electrons or muons) or hadronically. The leptonic decay represents the 35% of the cases and is covered by the reconstruction of the produced electron or muon. The hadronic decays represent 65%, which contain one or three charged pions in 72% and 22% of the cases, respectively. In addition, at least one associated neutral pion is also produced in 68% of the hadronic decays. The dedicated τ reconstruction and identification algorithms in ATLAS target the hadronic decay, with the main background being jets from energetic hadrons produced in the fragmentation of quarks and gluons, known as the QCD background. Therefore, the τ objects in ATLAS mentioned in this thesis refer to hadronically decaying τ leptons.

The candidates are seeded by jets which are required to have $p_T > 10$ GeV and $|\eta| < 2.5$ excluding the barrel-end-cap transition region. The tau identification is based on a machine learning classifier which is trained using the calorimeter information and the tracks associated to the jet candidate. A trained BDT is used for EMTopo jets while a recurrent neural network (RNN) is used for PFlow jets. Three different efficiency working points are defined: *Loose*, *Medium* and *Tight*. The τ leptons used in this thesis are defined with the medium working point, required to have $p_T > 25$ GeV and an isolation criteria of $\Delta R < 0.2$ between the τ and any selected electron or muon.

5.5 Missing transverse energy

The missing transverse momentum, also denoted as E_T^{miss} is the transverse component of the negative vector sum of the fully calibrated objects (electrons, muons, photons, τ leptons and jets) as well as soft objects associated to the PV. In an ideal detector, the the sum of four-momenta of all particles produced is equal to the net momentum of the initial

collision, implying that the net momentum in the transverse plane of the collision has to be zero, $E_T^{\text{miss}} = 0$. Nevertheless, the net momentum is not null as particles like neutrinos leave the detector without depositing energy or others can interact with the detector in regions not covered by electronics. For analyses with neutrinos in the final state, it is typical to consider that the transverse energy carried by the neutrinos is the E_T^{miss} , which allows their reconstruction.

Machine Learning (ML) is one of the core developing fields in computer science allowing the analysis of large and complex datasets, offering sophisticated techniques with a broad range of possible applications. Regarding high energy physics, the large amount of ML simulations or data that is being recorded is well suited for the application of ML techniques.

The deployment of these methods is already reaching crucial tasks as online data recording in ATLAS, from neural networks in calorimetry FPGAs to particle reconstruction in trigger algorithms, which benefit from faster and more efficient response than previous filters. For those cases, a neural network is trained to reduce background signal, to offer a high-level discriminating variable for a classification problem or to provide a prediction of a certain quantity. These methods can outperform conventional algorithms as the inference is usually performed from multi-dimensional inputs, providing large amounts of information to the learning algorithm. Regarding simulation, the detector simulation is one of the most computational intensive tasks within ATLAS and solutions involving adversarial networks and auto-encoders are being studied to output faster output, specially from the calorimeter simulation. Regarding particle reconstruction and identification, examples of implementations can be found within the τ identification or b -tagging algorithms. In physics analyses, the use of ML is already standardised to typically reconstruct or discriminate the signal process.

6.1 Pipeline

Machine Learning is a very broad umbrella term covering all kinds of algorithms which are not per se optimised for a specific task but are flexible enough to adapt to different problem sets by tuning (training) their parameter set. ML requires besides the model itself also preparation and follow-up processing steps. In which extent they are necessary always depends on the available data, the model and its later application. Figure 7.1 shows such an example workflow (the single steps are explained in more detail in the dedicated sections e.g. sec. 9.2).

Generally, two types of machine learning are distinguished: Supervised learning requiring fully labelled training data and Unsupervised learning not requiring any labelled data. There are also

intermediate approaches called Semi-Supervised learning. In the context of this thesis supervised approaches are used based on Neural Networks (NNs) and Boosted Decision Trees (BDTs). In the following, a statistical parametric (ML) model is denoted as $P_{\text{model}}(x_i; \theta)$ parametrised with the parameters θ while P_{data} is the true but unknown distribution. A data set of length N is given as $X = (x_1, x_2, \dots, x_N)$, in which each data point i has a feature set $x_i = (x_{1i}, x_{2i}, \dots, x_{Mi})$ with M features and true labels y_i in case of supervised learning.

It is important to ensure an unbiased training process. For this purpose, at least three orthogonal datasets are needed as indicated in Figure 7.2. The training sample is utilised for the actual algorithm training. The validation set is typically used to choose between different models and to optimise the model further such as hyperparameter optimisation. While for the training itself mostly a loss function is used (see sec. 7.2.3) to find the best parameter set, on the validation set, the performance measures dedicated for the problem set are evaluated (e.g. signal over background ratio) to fine-tune the model choice. The testing sample is only used to evaluate the final performance and is not involved in the training process. In the case of samples with low statistics, one can use cross-validation or also called k-folding [139] where pairs of training and test-/validation sets are partitioned into k subsets.

Typically, in particle physics the event number variable is used to split the dataset into the training and testing set. The advantage is that at every point it is clear which events were used for the training. The event number is a unique integer number associated to each event not correlated with any physical observable.

6.2 Performance

Even though every ML application is different, the model performance is the decisive measure in the end. Depending on the task, different metrics are used to judge the performance. In the following, the most common approaches are discussed.

Likelihood Discriminant

Loss Function

6.3 Neural networks

Optimiser

Backpropagation

Activation Functions

Regularisation

6.4 BDTs if used

Sidenotes are a distinctive feature of all 1.5-column-layout books. Indeed, having wide margins means that some material can be displayed there. We use margins for all kind of stuff: sidenotes, marginnotes, small tables of contents, citations, and, why not?, special boxes and environments.

| | |
|-----------------------------|----|
| 7.1 Sidenotes | 51 |
| 7.2 Marginnotes | 51 |
| 7.3 Footnotes | 52 |
| 7.4 Margintoc | 52 |
| 7.5 Marginlisting | 53 |

7.1 Sidenotes

Sidenotes are like footnotes, except that they go in the margin, where they are more readable. To insert a sidenote, just use the command `\sidenote{Text of the note}`. You can specify a mark^O with `\sidenote[mark]{Text}`, but you can also specify an offset, which moves the sidenote upwards or downwards, so that the full syntax is:

```
\sidenote[mark][offset]{Text}
```

If you use an offset, you always have to add the brackets for the mark, but they can be empty.¹

In kaobook we copied a feature from the `snotez` package: the possibility to specify a multiple of `\baselineskip` as an offset. For example, if you want to enter a sidenote with the normal mark and move it upwards one line, type:

```
\sidenote[][*-1]{Text of the sidenote.}
```

As we said, sidenotes are handled through the `sidenotes` package, which in turn relies on the `marginnote` package.

O: This sidenote has a special mark, a big O!

1: If you want to know more about the usage of the `\sidenote` command, read the documentation of the `sidenotes` package.

7.2 Marginnotes

This command is very similar to the previous one. You can create a marginnote with `\marginnote[offset]{Text}`, where the offset argument can be left out, or it can be a multiple of `\baselineskip`, *e.g.*

```
\marginnote[-12pt]{Text} or \marginnote[*-3]{Text}
```

To Do

A small thing that needs to be done is to renew the `\sidenote` command so that it takes only one optional argument, the offset. The special mark argument can go somewhere else. In other words, we want the syntax of `\sidenote` to resemble that of `\marginnote`.

While the command for margin notes comes from the `marginnote` package, it has been redefined in order to change the position of the optional offset argument, which now precedes the text of the note, whereas in the original version it was at the end. We have also added the possibility to use a multiple of `\baselineskip` as offset. These things were made only to make everything more consistent, so that you have to remember less things!

We load the packages `marginnote`, `marginfix` and `placeins`. Since `sidenotes` uses `marginnote`, what we said for marginnotes is also valid for sidenotes. Side- and margin- notes are shifted slightly upwards (`\renewcommand{\marginnotevadjust}{3pt}`) in order to align them to the bottom of the line of text where the note is issued. Importantly,

both sidenotes and marginnotes are defined as floating if the optional argument (*i.e.* the vertical offset) is left blank, but if the offset is specified they are not floating. Recall that floats cannot be nested, so in some rare cases you may encounter errors about lost floats; in those cases, remember that sidenotes and marginnotes are floats. To solve the problem, it may be possible to transform them into non-floating elements by specifying an offset of 0pt.

7.3 Footnotes

Even though they are not displayed in the margin, we will discuss about footnotes here, since sidenotes are mainly intended to be a replacement of them. Footnotes force the reader to constantly move from one area of the page to the other. Arguably, marginnotes solve this issue, so you should not use footnotes. Nevertheless, for completeness, we have left the standard command `\footnote`, just in case you want to put a footnote once in a while.*

7.4 Margintoc

Since we are talking about margins, we introduce here the `\margintoc` command, which allows one to put small table of contents in the margin. Like other commands we have discussed, `\margintoc` accepts a parameter for the vertical offset, like so: `\margintoc[offset]`.

The command can be used in any point of the document, but we think it makes sense to use it just at the beginning of chapters or parts. In this document I make use of a KOMA-Script feature and put it in the chapter preamble, with the following code:

```
\setchapterpreamble[u]{\margintoc}
\chapter{Chapter title}
```

As the space in the margin is a valuable resource, there is the possibility to print a shorter version of the title in the margin toc. Thus, there are in total three possible versions for the title of a section (or subsection): the one for the main text, the one for the main table of contents, and the one for the margintoc. These versions can be specified at the same time when the section is created in the source \TeX file:

```
\section[alternative-title-for-toc]{title-as-written-in-text}[
  alternative-title-for-margintoc]
```

By default, the margintoc includes sections and subsections. If you only want to show sections, add

```
\setcounter{margintocdepth}{\sectiontocdepth}
```

somewhere in your preamble.

The font used in the margintoc is the same as the one for the chapter entries in the main table of contents at the beginning of the document.

* And this is how they look like. Notice that in the PDF file there is a back reference to the text; pretty cool, uh?

7.5 Marginlisting

On some occasions it may happen that you have a very short piece of code that doesn't look good in the body of the text because it breaks the flow of narration: for that occasions, you can use a `marginlisting`. The support for this feature is still limited, especially for the captions, but you can try the following code:

```
\begin{marginlisting}[-0.5cm]
\caption{My caption}
\vspace{0.2cm}
\begin{lstlisting}[language=Python,style=kaolstplain]
... code ...
\end{lstlisting}
\end{marginlisting}
```

Unfortunately, the space between the caption and the listing must be adjusted manually; if you find a better way, please let me know.

Not only textual stuff can be displayed in the margin, but also figures. Those will be the focus of the next chapter.

Listing 7.1: An example of a margin listing.

```
print("Hello World!")
```




8 Figures and Tables

8.1 Normal Figures and Tables

Figures and tables can be inserted just like in any standard \LaTeX document. The `graphicx` package is already loaded and configured in such a way that the figure width is equal to the `textwidth` and the height is adjusted in order to maintain the original aspect ratio. As you may have imagined, the captions will be positioned. . . well, in the margins. This is achieved with the help of the `floatrow` package.

Here is a picture of Mona Lisa (Figure 8.1), as an example. The captions are formatted as the margin- and the side-notes; If you want to change something about captions you can use the command `\captsetup` from the `caption` package. Remember that if you want to reference a figure, the label must come *after* the caption!

While the format of the caption is managed by `caption`, its position is handled by the `floatrow` package. Achieving this result has been quite hard, but now I am pretty satisfied. In two-side mode, the captions are printed in the correct margin.

Tables can be inserted just as easily as figures, as exemplified by the following code:

```
1 \begin{table}
2 \begin{tabular}{c c c c }
3   \toprule
4   col1 & col2 & col3 & col 4 \\\
5   \midrule
6   \multirow{3}{4em}{Multiple row} & cell2 & cell3 & cell4\\ &
7   cell5 & cell6 & cell7 \\\
8   cell8 & cell9 & cell10 \\\
9   \multirow{3}{4em}{Multiple row} & cell2 & cell3 & cell4 \\\
10  cell5 & cell6 & cell7 \\\ &
```

8.1 Normal Figures and Tables 55

8.2 Margin Figures and Tables 58

8.3 Wide Figures and Tables . 58

Listing 8.1: Caption of a listing.

The credits for the image above the chapter title go to: Bushra Feroz, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=68724647>

```
11 |      cell8 & cell9 & cell10 \\
12 |      \bottomrule
13 | \end{tabular}
14 | \end{table}
```

which results in the useless Table Table 8.1.

Table 8.1: A useless table.

| col1 | col2 | col3 | col 4 |
|-----------------|-------|-------|--------|
| Multiple row | cell2 | cell3 | cell4 |
| | cell5 | cell6 | cell7 |
| | cell8 | cell9 | cell10 |
| Multiple row | cell2 | cell3 | cell4 |
| | cell5 | cell6 | cell7 |
| | cell8 | cell9 | cell10 |

I don't have much else to say, so I will just insert some blind text. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



Figure 8.1: It's Mona Lisa again. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

8.2 Margin Figures and Tables

Marginfigures can be inserted with the environment `marginfigure`. In this case, the whole picture is confined to the margin and the caption is below it. Figure ?? is obtained with something like this:

Listing 8.2: Another caption.

```
1 \begin{marginfigure}
2   \includegraphics{monalisa}
3   \caption[The Mona Lisa]{The Mona Lisa.}
4   \labfig{marginmonalisa}
5 \end{marginfigure}
```

There is also the `marginfigure` environment, of which Table 8.2 is an example. Notice how you can place the caption above the table by just placing the `\caption` command before beginning the `tabular` environment. Usually, figure captions are below, while table captions are above. This rule is also respected for normal figures and tables: the captions are always on the side, but for figure they are aligned to the bottom, while for tables to the top.

Table 8.2: Another useless table.

| col1 | col2 | col3 |
|----------|-------|-------|
| Multiple | cell2 | cell3 |
| row | cell5 | cell6 |
| | cell8 | cell9 |

Marginfigures and tables can be positioned with an optional offset command, like so:

```
1 \begin{marginfigure}[offset]
2   \includegraphics{seaside}
3 \end{marginfigure}
```

Improve this part.

Offset can be either a measure or a multiple of `\baselineskip`, much like with `\sidenote`, `\marginnote` and `\margintoc`. If you are wondering how I inserted this orange bubble, have a look at the `todo` package.

8.3 Wide Figures and Tables

With the environments `figure*` and `table*` you can insert figures which span the whole page width. For example, here are a wide figure and a wide table.



Figure 8.2: A wide seaside, and a wide caption. Credits: By Bushra Feroz, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=68724647>

Table 8.3: A wide table with invented data about three people living in the UK. Note that wide figures and tables are centered and their caption also extends into the margin.

| Name | Surname | Job | Salary | Age | Height | Country |
|-------|---------|-----------|---------|-----|--------|----------|
| Alice | Red | Writer | 4.000 £ | 34 | 167 cm | England |
| Bob | White | Bartender | 2.000 £ | 24 | 180 cm | Scotland |
| Drake | Green | Scientist | 4.000 £ | 26 | 175 cm | Wales |

It is the user’s responsibility to adjust the width of the table, if necessary, until it is aesthetically pleasing. The previous table was obtained with the following code:

```
1 \begin{table*}[h!]
2   \caption{A wide table with invented data about three people
3     living in the UK. Note that wide figures and tables are
4     centered and their caption also extends into the margin.}
5   \begin{tabular}{p{2.0cm} p{2.0cm} p{2.0cm} p{2.0cm} p{2.0cm} p
6     {2.0cm} p{1.5cm}}
7     \toprule
8     Name    & Surname  & Job      & Salary    & Age
9     & Height   & Country  & \\\
10    \midrule
11    Alice   & Red      & Writer   & 4.000 \pounds & 34
12    & 167 cm  & England  & \\\
13    Bob     & White    & Bartender & 2.000 \pounds & 24
14    & 180 cm  & Scotland & \\\
15    Drake   & Green    & Scientist & 4.000 \pounds & 26
16    & 175 cm  & Wales    & \\\
17    \bottomrule
18  \end{tabular}
19 \end{table*}
```

Listing 8.3: How to typeset a wide table

The floatrow package provides the ‘H’ specifier to instruct L^AT_EX to position the figure (or table) in precisely the same position it occupies in the source code. However, this specifier does not work with wide figures or tables: you should use ‘h!’ instead, like so: `\begin{figure*}[h!]`.

You may have noticed the full width image at the very beginning of this chapter: that, however, is set up in an entirely different way, which you’ll read about in Chapter 10 on page 69.

kaobook also supports paginated tables (have a look at the longtable package). The longtable¹ environment behaves a bit differently from table, in that longtable encompasses both table and tabular, so that you can write, *e.g.*,

1: Interestingly, longtables may require up to four rounds of compilation before they are typeset correctly.

```
1 \begin{longtable}{|l c c|}
2   \hline
3   One & Two & Three \\\
4   Left & Center & Center \\\
5   \hline
6   \caption{Caption of the longtable.}
7 \end{longtable}
```

Listing 8.4: Example of a longtable

to obtain the following table:

| | | |
|------|--------|--------|
| One | Two | Three |
| Left | Center | Center |

Table 8.4: Caption of the longtable.

The caption of a `longtable` is always positioned below the table, and it has the same width as the text (it doesn't extend into the margin). However, sometimes you may need a `longtable` that is so wide that it trespass into the margins; in those cases, you may want to also increase the width of the caption. To do so, you'll have to write two additional commands, one before and one after the `longtable`:

Listing 8.5: Increasing the width of the caption of a `longtable`.

```

1 \floatsetup[longtable]{margins=centering,LTcapwidth=table} % Add
   this line before the longtable to increase the caption width
2 \begin{longtable}{lp{8cm}p{5cm}p{2cm}}
3 ...
4 \end{longtable}
5 \floatsetup[longtable]{margins=raggedright,LTcapwidth=\textwidth}
   % Add this line after the longtable to revert the previous
   change

```

Having seen figures and tables, it is now time to tackle hyperrefferences.

9.1 Citations

To cite someone [Visscher2008, James2013] is very simple: just use the `\sidecite` command. It does not have an offset argument yet, but it probably will in the future. This command supports multiple entries, as you can see, and by default it prints the reference on the margin as well as adding it to the bibliography at the end of the document. Note that the citations have nothing to do with the text, [James2013] but they are completely random as they only serve the purpose to illustrate the feature.

Visscher2008, James2013

James2013

For this setup I wrote a separate package, `kaobiblio`, which you can find in the `styles` directory and include in your main `tex` file. This package accepts all the options that you can pass to `biblatex`, and actually it passes them to `biblatex` under the hood. Moreover, it also defines some commands, like `\sidecite`, and environments that can be used within a `kao` book.¹

1: For this reason you should always use `kaobiblio` instead of `biblatex`, but the syntax and the options are exactly the same.

If you want to use `bibtex` instead of `biblatex`, pass the option `backend=bibtex` to `kaobiblio`. `kaobiblio` also supports two options that are not shared with `biblatex`: `addspace` and `linkeverything`, both of which are boolean options, meaning that they can take either ‘true’ or ‘false’ as a value. If you pass `addspace=true` when loading `kaobiblio`, a space will be automatically added before the citation marks. If you pass `linkeverything=true`, the author’s name in the `authoryear-*` and `authortitle-*` styles will be a hyperlink like the year.²

2: The fact that the author name is not a hyperlink bothers more than one `biblatex` user. There are *strong arguments against* hyperlinking the author name, but in my personal opinion, linking the author’s name does not result in any problems in most practical cases.

As you have seen, the `\sidecite` command will print a citation in the margin. However, this command would be useless without a way to customise the format of the citation, so the `kaobook` provides also the `\formatmargincitation` command. By ‘renewing’ that command, you can choose which items will be printed in the margins. The best way to understand how it works is to see the actual definition of this command.

```
\newcommand{\formatmargincitation}[1]{%
  \parencite{#1}: \citeauthor*{#1} (\citeyear{#1}), \citetitle{#1}%
}
```

Thus, the `\formatmargincitation` accepts one parameter, which is the citation key, and prints the `parencite` followed by a colon, then the author, then the year (in brackets), and finally the title. [Battle2014] Now, suppose that you wish the margin citation to display the year and the author, followed by the title, and finally a fixed arbitrary string; you would add to your document:

Battle2014

```
\renewcommand{\formatmargincitation}[1]{%
  \citeyear{#1}, \citeauthor*{#1}: \citetitle{#1}; very interesting!%
}
```

The above code results in citations that look like the following. [Zou2005] Of course, changing the format is most useful when you also change the

Zou2005

default bibliography style. For instance, if you want to use the ‘philosophy-modern’ style for your bibliography, you might have something like this in the preamble:

```
\usepackage[style=philosophy-modern]{styles/kaobiblio}
\renewcommand{\formatmargincitation}[1]{%
  \sdcite{#1}%
}
\addbibresource{main.bib}
```

The commands like `\citeyear`, `\parencite` and `\sdcite` are just examples. A full reference of the available commands can be found in this [cheatsheet](#), under the ‘Citations’ section.

Finally, to compile a document containing citations, you need to use an external tool, which for this class is `biber`. You need to run the following (assuming that your tex file is called `main.tex`):

```
$ pdflatex main
$ biber main
$ pdflatex main
```

9.2 Glossaries and Indices

The `kaobook` class loads the packages `glossaries` and `imakeidx`, with which you can add glossaries and indices to your book. For instance, I previously defined some glossary entries and now I am going to use them, like this: `computer`. `glossaries` also allows you to use acronyms, like the following: this is the full version, Frame per Second (FPS), and this is the short one FPS. These entries will appear in the glossary in the backmatter.

Unless you use [Overleaf](#) or some other fancy IDE for \LaTeX , you need to run an external command from your terminal in order to compile a document with a glossary. In particular, the commands required are:³

```
$ pdflatex main
$ makeglossaries main
$ pdflatex main
```

Note that you need not run `makeglossaries` every time you compile your document, but only when you change the glossary entries.

To create an index, you need to insert the command `\index{subject}` whenever you are talking about ‘subject’ in the text. For instance, at the start of this paragraph I would write `index{index}`, and an entry would be added to the Index in the backmatter. Check it out!

A nomenclature is just a special kind of index; you can find one at the end of this book. To insert a nomenclature, we use the package `nomencl` and add the terms with the command `\nomenclature`. We put then a `\printnomenclature` where we want it to appear.

Also with this package we need to run an external command to compile the document, otherwise the nomenclature will not appear:

```
$ pdflatex main
$ makeindex main.nlo -s nomencl.ist -o main.nls
$ pdflatex main
```

3: These are the commands you would run in a UNIX system, but see also Section 9.4 (A Final Note on Compilation); I have no idea about how it works in Windows.

In theory, you would need to run an external command for the index as well, but luckily the package we suggested, `imakeidx`, can compile the index automatically.

These packages are all loaded in `packages.sty`, one of the files that come with this class. However, the configuration of the elements is best done in the `main.tex` file, since each book will have different entries and styles.

Note that the `nomencl` package caused problems when the document was compiled, so, to make a long story short, I had to prevent `scrhack` to load the hack-file for `nomencl`. When compiling the document on Overleaf, however, this problem seem to vanish.

This brief section was by no means a complete reference on the subject, therefore you should consult the documentation of the above package to gain a full understanding of how they work.

9.3 Hyperreferences

Together with this class we provide a handy package to help you referencing the same elements always in the same way, for consistency across the book. First, you can label each element with a specific command. For instance, should you want to label a chapter, you would put `\labch{chapter-title}` right after the `\chapter` directive. This is just a convenience, because `\labch` is actually just an alias to `\label{ch:chapter-title}`, so it spares you the writing of ‘ch:’. We defined similar commands for many typically labeled elements, including:

- | | |
|------------------------------------|---------------------------------------|
| ▶ Page: <code>\labpage</code> | ▶ Assumption: <code>\labassum</code> |
| ▶ Part: <code>\labpart</code> | ▶ Theorem: <code>\labthm</code> |
| ▶ Chapter: <code>\labch</code> | ▶ Proposition: <code>\labprop</code> |
| ▶ Section: <code>\labsec</code> | ▶ Lemma: <code>\lablemma</code> |
| ▶ Figure: <code>\labfig</code> | ▶ Remark: <code>\labremark</code> |
| ▶ Table: <code>\labtab</code> | ▶ Example: <code>\labexample</code> |
| ▶ Definition: <code>\labdef</code> | ▶ Exercise: <code>\labexercise</code> |

Of course, we have similar commands for referencing those elements. However, since the style of the reference should depend on the context, we provide different commands to reference the same thing. For instance, in some occasions you may want to reference the chapter by name, but other times you want to reference it only by number. In general, there are four reference style, which we call plain, vario, name, and full.

The plain style references only by number. It is accessed, for chapters, with `\refch{chapter-title}` (for other elements, the syntax is analogous). Such a reference results in: Chapter 9.

The vario and name styles rest upon the `varioref` package. Their syntax is `\vrefch{chapter-title}` and `\nrefch{chapter-title}`, and they result in: Chapter 9 on page 61, for the vario style, and: Chapter 9 (References), for the name style. As you can see, the page is referenced in `varioref` style.

The full style references everything. You can use it with `\frefch{chapter-title}` and it looks like this: Chapter 9 (References) on page 61.

Of course, all the other elements have similar commands (*e.g.* for parts you would use `\vrefpart{part-title}` or something like that). However, not all elements implement all the four styles. The commands provided should be enough, but if you want to see what is available or to add the missing ones, have a look at the [attached package](#).

In order to have access to all these features, the `kaorefs` should be loaded in the preamble of your document. It should be loaded last, or at least after `babel` (or `polyglossia`) and `plaintheorems` (or `mdftheorems`). Options can be passed to it like to any other package; in particular, it is possible to specify the language of the captions. For instance, if you specify `'italian'` as an option, instead of `'Chapter'` it will be printed `'Capitolo'`, the Italian analog. If you know other languages, you are welcome to contribute the translations of these captions! Feel free to contact the author of the class for further details.

The `kaorefs` package also include `cleveref`, so it is possible to use `\cref` in addition to all the previously described referencing commands.

9.4 A Final Note on Compilation

Probably the easiest way to compile a latex document is with the `latexmk` script, as it can take care of everything, if properly configured, from the bibliography to the glossary. The command to issue, in general, is:

```
1 | latexmk [latexmk-options] [filename ...]
```

`latexmk` can be extensively configured (see <https://mg.readthedocs.io/latexmk.html>). For convenience, I print here an example configuration that would cover all the steps described above.

```
1 | # By default compile only the file called 'main.tex'
2 | @default_files = ('main.tex');
3 |
4 | # Compile the glossary and acronyms list (package 'glossaries')
5 | add_cus_dep( 'acn', 'acr', 0, 'makeglossaries' );
6 | add_cus_dep( 'glo', 'gls', 0, 'makeglossaries' );
7 | $clean_ext .= " acr acn alg glo gls glg";
8 | sub makeglossaries {
9 |     my ($base_name, $path) = fileparse( $_[0] );
10 |    pushd $path;
11 |    my $return = system "makeglossaries", $base_name;
12 |    popd;
13 |    return $return;
14 | }
15 |
16 | # Compile the nomenclature (package 'nomencl')
17 | add_cus_dep( 'nlo', 'nls', 0, 'makenlo2nls' );
18 | sub makenlo2nls {
19 |     system( "makeindex -s nomencl.ist -o \"$_[0].nls\" \"$_[0].nlo\" \"\" );
20 | }
```

However, if you'd rather not use an external package and want to do everything manually, here are some tips.⁴

Compiling the examples in the kaobook repository

To compile the examples, and in particular the documentation, that are in the `examples` directory of the [kaobook repository](#) on GitHub, do as follows. `cd` into the root directory of the repository, and run `pdflatex -output-directory examples/documentation main.tex`. With this trick, you

4: As the author only uses Linux and compiles everything from the command line, he doesn't know how the compilation works in Windows or Mac. The tips, therefore, refer to the usage with Linux from the command line.

can compile the documentation using the class files pertaining to the repository (and not, say, those in your texmf tree). The ‘-output-directory’ option works with the other L^AT_EX-related commands such as biber and makeglossaries.

A note of warning: sometimes L^AT_EX needs more than one run to get the correct position of each element; this is true in particular for the positioning of floating elements like figures, tables, and margin notes. Occasionally, L^AT_EX can need up to four re-runs, so If the alignment of margin elements looks odd, or if they bleed into ther main text, try running pdflatex one more time.

DESIGN AND ADDITIONAL FEATURES



10 Page Design

10.1 Headings

So far, in this document I used two different styles for the chapter headings: one has the chapter name, a rule and, in the margin, the chapter number; the other has an image at the top of the page, and the chapter title is printed in a box (like for this chapter). There is one additional style, which I used only in the Chapter 11.3 (Appendix); there, the chapter title is enclosed in two horizontal rules, and the chapter number (or letter, in the case of the appendix) is above it.¹

Every book is unique, so it makes sense to have different styles from which to choose. Actually, it would be awesome if whenever a kao-user designs a new heading style, he or she added it to the three styles already present, so that it will be available for new users and new books.

The choice of the style is made simple by the `\setchapterstyle` command. It accepts one option, the name of the style, which can be: ‘plain’, ‘kao’, ‘bar’, or ‘lines’.² If instead you want the image style, you have to use the command `\setchapterimage`, which accepts the path to the image as argument; you can also provide an optional parameter in square brackets to specify the height of the image. `\setchapterimage` automatically sets the chapter style to ‘bar’ for that chapter (and also for subsequent chapters).

Let us make some examples. In this book, I begin a normal chapter with the lines:

```
1 \setchapterstyle{kao}
2 \setchapterpreamble[u]{\margintoc}
3 \chapter{Title of the Chapter}
4 \labch{title}
```

In Line 1 I choose the style for the title to be ‘kao’. Then, I specify that I want the margin toc. The rest is ordinary administration in \LaTeX , except that I use my own `\labch` to label the chapter. Actually, the `\setchapterpreamble` is a standard KOMA-Script one, so I invite you to read about it in the KOMA documentation. Once the chapter style is set, it holds until you change it.³ Whenever I want to start a chapter with an

| | |
|-----------------------------------|----|
| 10.1 Headings | 69 |
| 10.2 Headers & Footers | 70 |
| 10.3 twoside mode | 70 |
| 10.4 Table of Contents | 70 |
| 10.5 Paper Size | 71 |
| 10.6 Page Layout | 71 |
| 10.7 Numbers & Counters | 72 |
| 10.8 White Space | 73 |

1: To be honest, I do not think that mixing heading styles like this is a wise choice, but in this document I did it only to show you how they look.

2: Plain is the default \LaTeX title style; the other ones are self explanatory.

3: The `\margintoc` has to be specified at every chapter. Perhaps in the future this may change; it all depends on how this feature will be welcomed by the users, so keep in touch with me if you have preferences!

image, I simply write:

```
1 \setchapterimage[7cm]{path/to/image.png} % Optionally specify the
   height
2 \setchapterpreamble[u]{\margintoc}
3 \chapter{Catchy Title} % No need to set a chapter style
4 \labch{catchy}
```

If you prefer, you can also specify the style at the beginning of the main document, and that style will hold until you change it again.

10.2 Headers & Footers

Headers and footers in KOMA-Script are handled by the `scrlayer-scrpage` package. There are two basic style: ‘`scrheadings`’ and ‘`plain.scrheadings`’. The former is used for normal pages, whereas the latter is used in title pages (those where a new chapter starts, for instance) and, at least in this book, in the front matter. At any rate, the style can be changed with the `\pagestyle` command, *e.g.* `\pagestyle{plain.scrheadings}`.

In both styles, the footer is completely empty. In `plain.scrheadings`, also the header is absent (otherwise it wouldn’t be so plain. . .), but in the normal style the design is reminiscent of the ‘`kao`’ style for chapter titles.

10.3 twoside mode

To Do

The `twoside` class option is still unstable and may lead to unexpected behaviours. Great strides have been done since the first version of `kaobook`, but some work still needs to be done. As always, any help will be greatly appreciated.

By passing the `twoside` option to the `kaobook`, the style of left and right pages will be different, similarly to a printed book. In digital books, having a symmetrical layout for left and right pages is less important, and you may be tempted to use the `twoside=false` option. However, keep in mind that in ‘`oneside`’ mode the `\uppertitleback` and `\lowertitleback` commands are not available.⁴ If you want to have the upper/lower titleback in a one-side document, just add manually the contents that you’d put using the upper/lower titleback commands.

4: Another useful thing to keep in mind is that, when `twoside=true`, an extra white page will be added to the front-matter.

10.4 Table of Contents

Another important part of a book is the table of contents. By default, in `kaobook` there is an entry for everything: list of figures, list of tables, bibliographies, and even the table of contents itself. Not everybody might like this, so we will provide a description of the changes you need to do in order to enable or disable each of these entries. In the following Table 10.1, each item corresponds to a possible entry in the TOC, and

| Entry | Command to Activate |
|-------------------|---|
| Table of Contents | <code>\setuptoc{toc}{totoc}</code> |
| List of Figs/Tabs | <code>\PassOptionsToClass{toc=listof}{\@baseclass}</code> |
| Bibliography | <code>\PassOptionsToClass{toc=bibliography}{\@baseclass}</code> |

its description is the command you need to provide to have such entry. These commands are specified in the attached [style package](#),⁵ so if you don't want the entries, just comment the corresponding lines.

Of course, some packages, like those for glossaries and indices, will try to add their own entries. In such cases, you have to follow the instructions specific to that package. Here, since we have talked about glossaries and notations in Chapter 9, we will briefly see how to configure them.

For the glossaries package, use the 'toc' option when you load it: `\usepackage[toc]{glossaries}`. For nomencl, pass the 'intoc' option at the moment of loading the package. Both glossaries and nomencl are loaded in the attached ['packages' package](#).

Additional configuration of the table of contents can be performed through the packages etoc, which is loaded because it is needed for the margintocs, or the more traditional tocbase. Read the respective documentations if you want to be able to change the default TOC style.⁶

Table 10.1: Commands to add a particular entry to the table of contents.

5: In the same file, you can also choose the titles of these entries.

In a later section, we will see how you can define your own floating environment, and endow it with an entry in the TOC.

6: (And please, send me a copy of what you have done, I'm so curious!)

10.5 Paper Size

Recent versions of Kaobook support paper sizes different from the default A4. It is possible to pass the name of the paper as an option to the class, as we are accustomed for any other L^AT_EX class. For example, the class option b5paper would set the paper size to the B5 format.

We also support the paper sizes specified in [this web page](#) and some additional sizes requested by the users, with the option names specified in Table 10.2.

For instance, to use the 'smallpocketpaper' add the correct description at the beginning of the documentclass instruction:

```
1 \documentclass[
2     smallpocketpaper,
3     fontsize=10pt,
4     twoside=false,
5     %open=any,
6     secnumdepth=1,
7 ]{kaobook}
```

Table 10.2: Some non-standard paper sizes supported by kaobook.

| Dimension | Option name |
|-----------------|------------------|
| 12.0cm x 19.0cm | smallpocketpaper |
| 13.5cm x 21.5cm | pocketpaper |
| 14.8cm x 21.0cm | a5paper |
| 15.5cm x 22.0cm | juvenilepaper |
| 17.0cm x 17.0cm | smallphotopaper |
| 21.0cm x 15.0cm | appendixpaper |
| 17.0cm x 22.0cm | cookpaper |
| 19.0cm x 27.0cm | illustratedpaper |
| 17.0cm x 17.0cm | photopaper |
| 16.0cm x 24.0cm | f24paper |

10.6 Page Layout

Besides the page style, you can also change the width of the content of a page. This is particularly useful for pages dedicated to part titles, where having the 1.5-column layout might be a little awkward, or for pages where you only put figures, where it is important to exploit all the available space.

Sometimes it is desirable to increase the width for just one or a few paragraphs; the `widepar` environment does that: wrap your paragraphs in this environment, and they will occupy the full width of the page.

In practice, there are two layouts: ‘wide’ and ‘margin’. The former suppresses the margins and allocates the full page for contents, while the latter is the layout used in most of the pages of this book, including this one. The wide layout is also used automatically in the front and back matters.

To change page layout, use the `\pagelayout` command. For example, when I start a new part, I write:

```
1 \pagelayout{wide}
2 \addpart{Title of the New Part}
3 \pagelayout{margin}
```

Beyond these two basic layouts, it is also possible to finely tune the page layout by redefining the `\marginlayout` command. This command is called internally by the higher-level `\pagelayout`, and it is responsible for setting the width of the margins and of the text. The default definition is:

```
1 \newcommand{\marginlayout}{%
2   \newgeometry{
3     top=27.4mm,           % height of the top margin
4     bottom=27.4mm,       % height of the bottom margin
5     inner=24.8mm,        % width of the inner margin
6     textwidth=107mm,     % width of the text
7     marginparsep=8.2mm,  % width between text and margin
8     marginparwidth=49.4mm, % width of the margin
9   }%
10 }
```

so if you want to, say, decrease the width of the margin while increasing the width of the text, you could write in the preamble of your document something like:

```
1 \renewcommand{\marginlayout}{%
2   \newgeometry{
3     top=27.4mm,           % height of the top margin
4     bottom=27.4mm,       % height of the bottom margin
5     inner=24.8mm,        % width of the inner margin
6     textwidth=117mm,     % width of the text
7     marginparsep=8.2mm,  % width between text and margin
8     marginparwidth=39.4mm, % width of the margin
9   }%
10 }
```

where the text width has been increased by 10mm and the margin width has been decreased by 10mm.

10.7 Numbers & Counters

In this short section we shall see how dispositions, sidenotes and figures are numbered in the `kaobook` class.

By default, dispositions are numbered up to the section in `kaobook` and up to the subsection in `kaohandt`. This can be changed by passing the option `secnumdepth tokaobook` or `kaohandt` (e.g. 1 corresponds to section and 2 corresponds to subsections).

The sidenotes counter is the same across all the document, but if you want it to reset at each chapter, just uncomment the line

```
\counterwithin*{sidenote}{chapter}
```

in the `styles/style.sty` package provided by this class.

Figure and Table numbering is also per-chapter; to change that, use something like:

```
\renewcommand{\thefigure}{\arabic{section}.\arabic{figure}}
```

10.8 White Space

One of the things that I find most hard in \LaTeX is to finely tune the white space around objects. There are not fixed rules, each object needs its own adjustment. Here we shall see how some spaces are defined at the moment in this class.

Attention! This section may be incomplete.

Space around sidenotes and citations marks

There should be no space before or after sidenotes and citation marks, like so:

```
sidenote7sidenote  
citation[James2013]citation
```

7: This paragraph can be used to diagnose any problems: if you see whitespace around sidenotes or citation marks, probably a % sign is missing somewhere in the definitions of the class macros.

Space around figures and tables

```
\renewcommand\FBskip{.4\topskip}  
\renewcommand\FBbskip{\FBskip}
```

Space around captions

```
\captionsetup{  
  aboveskip=6pt,  
  belowskip=6pt  
}
```

Space around displays (e.g. equations)

```
\setlength\abovedisplayskip{6pt plus 2pt minus 4pt}  
\setlength\belowdisplayskip{6pt plus 2pt minus 4pt}  
\abovedisplayskip 10\p@ \@plus2\p@ \@minus5\p@  
\abovedisplayshortskip \z@ \@plus3\p@  
\belowdisplayskip \abovedisplayskip  
\belowdisplayshortskip 6\p@ \@plus3\p@ \@minus3\p@
```


11.1 Theorems

11.1 Theorems 75
11.2 Boxes & Environments . . 77
11.3 Experiments 77

Despite most people complain at the sight of a book full of equations, mathematics is an important part of many books. Here, we shall illustrate some of the possibilities. We believe that theorems, definitions, remarks and examples should be emphasised with a shaded background; however, the colour should not be to heavy on the eyes, so we have chosen a sort of light yellow.¹

Definition 11.1.1 *Let (X, d) be a metric space. A subset $U \subset X$ is an open set if, for any $x \in U$ there exists $r > 0$ such that $B(x, r) \subset U$. We call the topology associated to d the set τ_d of all the open subsets of (X, d) .*

Definition 11.1.1 is very important. I am not joking, but I have inserted this phrase only to show how to reference definitions. The following statement is repeated over and over in different environments.

Theorem 11.1.1 *A finite intersection of open sets of (X, d) is an open set of (X, d) , i.e τ_d is closed under finite intersections. Any union of open sets of (X, d) is an open set of (X, d) .*

Proposition 11.1.2 *A finite intersection of open sets of (X, d) is an open set of (X, d) , i.e τ_d is closed under finite intersections. Any union of open sets of (X, d) is an open set of (X, d) .*

Lemma 11.1.3 *A finite intersection^a of open sets of (X, d) is an open set of (X, d) , i.e τ_d is closed under finite intersections. Any union of open sets of (X, d) is an open set of (X, d) .*

^a I'm a footnote

You can safely ignore the content of the theorems. . . I assume that if you are interested in having theorems in your book, you already know something about the classical way to add them. These example should just showcase all the things you can do within this class.

Corollary 11.1.4 (Finite Intersection, Countable Union) *A finite intersection of open sets of (X, d) is an open set of (X, d) , i.e τ_d is closed under finite intersections. Any union of open sets of (X, d) is an open set of (X, d) .*

Proof. The proof is left to the reader as a trivial exercise. Hint: Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift –

1: The boxes are all of the same colour here, because we did not want our document to look like Harlequin.

You can even insert footnotes inside the theorem environments; they will be displayed at the bottom of the box.

not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. \square

Here is a random equation, just because we can:

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4}}}}$$

Definition 11.1.2 Let (X, d) be a metric space. A subset $U \subset X$ is an open set if, for any $x \in U$ there exists $r > 0$ such that $B(x, r) \subset U$. We call the topology associated to d the set τ_d of all the open subsets of (X, d) .

Example 11.1.1 Let (X, d) be a metric space. A subset $U \subset X$ is an open set if, for any $x \in U$ there exists $r > 0$ such that $B(x, r) \subset U$. We call the topology associated to d the set τ_d of all the open subsets of (X, d) .

Remark 11.1.1 Let (X, d) be a metric space. A subset $U \subset X$ is an open set if, for any $x \in U$ there exists $r > 0$ such that $B(x, r) \subset U$. We call the topology associated to d the set τ_d of all the open subsets of (X, d) .

As you may have noticed, definitions, example and remarks have independent counters; theorems, propositions, lemmas and corollaries share the same counter.

Remark 11.1.2 Here is how an integral looks like inline: $\int_a^b x^2 dx$, and here is the same integral displayed in its own paragraph:

$$\int_a^b x^2 dx$$

There is also an environment for exercises.

Exercise 11.1.1 Prove (or disprove) the Riemann hypothesis.

2: The styles without framed are not showed, but actually the only difference is that they don't have the yellow boxes.

We provide one package for the theorem styles: `kaotheorems.sty`, to which you can pass the `framed` option you do want coloured boxes around theorems, like in this document.² You may want to edit this files according to your taste and the general style of the book. However, there is an option to customise the background colour of the boxes if you use the `framed` option: when you load this package, you can pass it the `background=mycolour` option (replace 'mycolour' with the actual colour, for instance, 'red!35!white'). This will change the colour of all the boxes, but it is also possible to override the default colour only for some elements. For instance, the `propositionbackground=mycolour` option will change the colour for propositions only. There are similar options for theorem, definition, lemma, corollary, remark, and example.

11.2 Boxes & Custom Environments ³

Say you want to insert a special section, an optional content or just something you want to emphasise. We think that nothing works better than a box in these cases. We used `mdf framed` to construct the ones shown below. You can create and modify such environments by editing the provided file `kao.sty`.

3: Notice that in the table of contents and in the header, the name of this section is ‘Boxes & Environments’; we achieved this with the optional argument of the section command.

Title of the box

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

If you set up a counter, you can even create your own numbered environment.

Comment 11.2.1

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

11.3 Experiments

It is possible to wrap marginnotes inside boxes, too. Audacious readers are encouraged to try their own experiments and let me know the outcomes.

I believe that many other special things are possible with the `kaobook` class. During its development, I struggled to keep it as flexible as possible, so that new features could be added without too great an effort. Therefore, I hope that you can find the optimal way to express yourselves in writing a book, report or thesis with this class, and I am eager to see the outcomes of any experiment that you may try.

title of margin note

Margin note inside a kaobox.
(Actually, kaobox inside a margin-note!)

APPENDIX



Heading on Level 0 (chapter)

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

A.1 Heading on Level 1 (section)

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

A.1.1 Heading on Level 2 (subsection)

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Heading on Level 3 (subsubsection)

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected

font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Heading on Level 4 (paragraph) Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

A.2 Lists

A.2.1 Example for list (itemize)

- ▶ First item in a list
- ▶ Second item in a list
- ▶ Third item in a list
- ▶ Fourth item in a list
- ▶ Fifth item in a list

Example for list (4*itemize)

- ▶ First item in a list
 - First item in a list
 - * First item in a list
 - First item in a list
 - Second item in a list
 - * Second item in a list
 - Second item in a list
- ▶ Second item in a list

A.2.2 Example for list (enumerate)

1. First item in a list
2. Second item in a list
3. Third item in a list
4. Fourth item in a list
5. Fifth item in a list

Example for list (4*enumerate)

1. First item in a list
 - a) First item in a list
 - i. First item in a list
 - A. First item in a list
 - B. Second item in a list
 - ii. Second item in a list
 - b) Second item in a list
2. Second item in a list

A.2.3 Example for list (description)

First item in a list
Second item in a list
Third item in a list
Fourth item in a list
Fifth item in a list

Example for list (4*description)

First item in a list
 First item in a list
 First item in a list
 Second item in a list
 Second item in a list
 Second item in a list
Second item in a list

B

Fonts Testing

B.1 Font Sizes

B.1 Font Sizes 85

B.2 Font Families 85

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

The quick brown fox jumps over the lazy dog.

B.2 Font Families

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

The quick brown fox jumps over the lazy dog. Medium.

The quick brown fox jumps over the lazy dog. Bold.

The quick brown fox jumps over the lazy dog. Upright.

The quick brown fox jumps over the lazy dog. Italics.

The quick brown fox jumps over the lazy dog. Slanted.

THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG. SMALL CAPS.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift - not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

The quick brown fox jumps over the lazy dog. Medium.

The quick brown fox jumps over the lazy dog. Bold.

The quick brown fox jumps over the lazy dog. Upright.

The quick brown fox jumps over the lazy dog. Italics.

The quick brown fox jumps over the lazy dog. Slanted.

The quick brown fox jumps over the lazy dog. Small Caps.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

The quick brown fox jumps over the lazy dog. Medium.

The quick brown fox jumps over the lazy dog. Bold.

The quick brown fox jumps over the lazy dog. Upright.

The quick brown fox jumps over the lazy dog. Italics.

The quick brown fox jumps over the lazy dog. Slanted.

THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG. SMALL CAPS.

Bibliography

Here are the references in citation order.

- [1] Steven Weinberg. 'A Model of Leptons'. In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. doi: [10.1103/PhysRevLett.19.1264](#) (cited on page 5).
- [2] Sheldon L. Glashow. 'Partial-symmetries of weak interactions'. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. doi: [https://doi.org/10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2) (cited on page 5).
- [3] Abdus Salam. 'Gauge unification of fundamental forces'. In: *Rev. Mod. Phys.* 52 (3 1980), pp. 525–538. doi: [10.1103/RevModPhys.52.525](#) (cited on page 5).
- [4] S. Abachi et al. 'Search for High Mass Top Quark Production in pp Collisions at $\sqrt{s}=1.8$ TeV'. In: *Physical Review Letters* 74.13 (1995), 2422–2426. doi: [10.1103/physrevlett.74.2422](#) (cited on page 5).
- [5] 'Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab'. In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2626–2631. doi: [10.1103/PhysRevLett.74.2626](#) (cited on page 5).
- [6] G. Aad et al. 'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'. In: *Physics Letters B* 716.1 (2012), 1–29. doi: [10.1016/j.physletb.2012.08.020](#) (cited on page 5).
- [7] S. Chatrchyan et al. 'Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC'. In: *Physics Letters B* 716.1 (2012), 30–61. doi: [10.1016/j.physletb.2012.08.021](#) (cited on page 5).
- [8] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. 'Advantages of the color octet gluon picture'. In: *Physics Letters B* 47.4 (1973), pp. 365–368. doi: [https://doi.org/10.1016/0370-2693\(73\)90625-4](https://doi.org/10.1016/0370-2693(73)90625-4) (cited on page 5).
- [9] Abdus Salam. 'Weak and Electromagnetic Interactions'. In: *Conf. Proc. C* 680519 (1968), pp. 367–377. doi: [10.1142/9789812795915_0034](#) (cited on page 5).
- [10] W. Pauli. 'Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren'. In: *Zeitschrift für Physik* 31.1 (1925), pp. 765–783. doi: [10.1007/BF02980631](#) (cited on page 6).
- [11] P.A. Zyla et al. 'Review of Particle Physics'. In: *PTEP* 2020.8 (2020), p. 083C01. doi: [10.1093/ptep/ptaa104](#) (cited on pages 7, 20, 23).
- [12] Emmy Noether. 'Invariant variation problems'. In: *Transport Theory and Statistical Physics* 1.3 (1971), pp. 186–207. doi: [10.1080/00411457108231446](#) (cited on page 7).
- [13] C. N. Yang and R. L. Mills. 'Conservation of Isotopic Spin and Isotopic Gauge Invariance'. In: *Phys. Rev.* 96 (1 1954), pp. 191–195. doi: [10.1103/PhysRev.96.191](#) (cited on page 8).
- [14] Murray Gell-Mann. 'Symmetries of Baryons and Mesons'. In: *Phys. Rev.* 125 (3 1962), pp. 1067–1084. doi: [10.1103/PhysRev.125.1067](#) (cited on page 10).
- [15] H. David Politzer. 'Reliable Perturbative Results for Strong Interactions?' In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1346–1349. doi: [10.1103/PhysRevLett.30.1346](#) (cited on page 11).
- [16] David J. Gross and Frank Wilczek. 'Ultraviolet Behavior of Non-Abelian Gauge Theories'. In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346. doi: [10.1103/PhysRevLett.30.1343](#) (cited on page 11).
- [17] T. D. Lee and C. N. Yang. 'Question of Parity Conservation in Weak Interactions'. In: *Phys. Rev.* 104 (1 1956), pp. 254–258. doi: [10.1103/PhysRev.104.254](#) (cited on page 11).
- [18] C. S. Wu et al. 'Experimental Test of Parity Conservation in Beta Decay'. In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415. doi: [10.1103/PhysRev.105.1413](#) (cited on page 11).

- [19] Peter W. Higgs. 'Broken Symmetries and the Masses of Gauge Bosons'. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. doi: [10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508) (cited on page 15).
- [20] P.W. Higgs. 'Broken symmetries, massless particles and gauge fields'. In: *Physics Letters* 12.2 (1964), pp. 132–133. doi: [https://doi.org/10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9) (cited on page 15).
- [21] F. Englert and R. Brout. 'Broken Symmetry and the Mass of Gauge Vector Mesons'. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. doi: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321) (cited on page 15).
- [22] Hideki Yukawa. 'On the Interaction of Elementary Particles. I'. In: *Progress of Theoretical Physics Supplement* 1 (Jan. 1955), pp. 1–10. doi: [10.1143/PTPS.1.1](https://doi.org/10.1143/PTPS.1.1) (cited on page 18).
- [23] Makoto Kobayashi and Toshihide Maskawa. 'CP Violation in the Renormalizable Theory of Weak Interaction'. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. doi: [10.1143/PTP.49.652](https://doi.org/10.1143/PTP.49.652) (cited on page 19).
- [24] Ling-Lie Chau and Wai-Yee Keung. 'Comments on the Parametrization of the Kobayashi-Maskawa Matrix'. In: *Phys. Rev. Lett.* 53 (19 1984), pp. 1802–1805. doi: [10.1103/PhysRevLett.53.1802](https://doi.org/10.1103/PhysRevLett.53.1802) (cited on page 19).
- [25] K. Abe et al. 'Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV'. In: *Physical Review D* 97.7 (2018). doi: [10.1103/physrevd.97.072001](https://doi.org/10.1103/physrevd.97.072001) (cited on page 21).

Notation

The next list describes several symbols that will be later used within the body of the document.

c Speed of light in a vacuum inertial frame

h Planck constant

Greek Letters with Pronunciations

| Character | Name | Character | Name |
|--------------------|---------------------------|----------------------|----------------------------------|
| α | alpha <i>AL-fuh</i> | ν | nu <i>NEW</i> |
| β | beta <i>BAY-tuh</i> | ξ, Ξ | xi <i>KSIGH</i> |
| γ, Γ | gamma <i>GAM-muh</i> | \omicron | omicron <i>OM-uh-CRON</i> |
| δ, Δ | delta <i>DEL-tuh</i> | π, Π | pi <i>PIE</i> |
| ϵ | epsilon <i>EP-suh-lon</i> | ρ | rho <i>ROW</i> |
| ζ | zeta <i>ZAY-tuh</i> | σ, Σ | sigma <i>SIG-muh</i> |
| η | eta <i>AY-tuh</i> | τ | tau <i>TOW (as in cow)</i> |
| θ, Θ | theta <i>THAY-tuh</i> | υ, Υ | upsilon <i>OOP-suh-LON</i> |
| ι | iota <i>eye-OH-tuh</i> | ϕ, Φ | phi <i>FEE, or FI (as in hi)</i> |
| κ | kappa <i>KAP-uh</i> | χ | chi <i>KI (as in hi)</i> |
| λ, Λ | lambda <i>LAM-duh</i> | ψ, Ψ | psi <i>SIGH, or PSIGH</i> |
| μ | mu <i>MEW</i> | ω, Ω | omega <i>oh-MAY-guh</i> |

Capitals shown are the ones that differ from Roman capitals.

Alphabetical Index

\sidecite, 61

citations, 61

glossary, 62

hyperreferences, 63

index, 62

nomenclature, 62

preface, v