Pushing the Limits

A search for the Lepton Flavour Violating decay $\tau \to 3\mu$ within Heavy Flavour Jets with the ATLAS Detector

Matteo Bedognetti

Pushing the Limits – A search for the Lepton Flavour Violating decay $\tau \to 3\mu$ within Heavy Flavour Jets with the ATLAS detector
Thesis, Radboud University Nijmegen

Cover design Caterina Scaramellini – *caterina.scaramellini@gmail.com*

Printing: Ipskamp Printing, Amsterdam

This work has been performed at the National Institute for Subatomic Physics (Nikhef) which is financed by the Dutch Research Council (NWO).

Pushing the Limits

A search for the Lepton Flavour Violating decay $\tau \to 3\mu$ within Heavy Flavour Jets with the ATLAS Detector

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op
donderdag 5 maart 2020
om 12:30 uur precies

door

Matteo Bedognetti

geboren op 27 juni 1987
te Chiavenna, Italië

Promotoren:	Prof. dr. N. de Groot
	Prof. dr. A.P. Colijn (Universiteit Utrecht)
Manuscriptcommissie:	Prof. dr. J.R. Hörandel
	Prof. dr. A.N.J.J. Schellekens
	Prof. dr. W. Verkerke (Universiteit van Amsterdam)

Contents

Introduction

> *"Bismillahi ar rahman ar rahim. That is the Arab phrase used before starting out on a journey. Eh bien, we too start on a journey. A journey into the past. A journey into the strange places of the human soul."*
>
> — Agatha Christie, *Murder in Mesopotamia*

High-energy physics is the study the tiniest knowable components of the physical world. It characterizes a number of fundamental particles and the forces that govern them. One requires "high-energies" to probe sub-nuclear particles because of the inverse proportionality between a particle's quantum wavelength and its energy (energies about 10 orders of magnitude higher than the ones carried by the photons of visible light are needed). The laws describing this field are condensed into the Standard Model of particle physics, which has been formulated in the second half of the twentieth century and has the form of a local gauge theory, i.e. it originates from a set of symmetries to which the fields representing particles and forces must obey.
Whilst successful in describing experimental observations to an astonishing degree, the Standard Model does neither account for the gravitational force, nor explain a handful of phenomena such as what constitutes dark matter or how anti-matter came to be virtually absent from the baryonic matter in the visible universe.
A fervent community of theorists has formulated a variety of extensions to the SM to describe the so far unexplained phenomena, but progress cannot be made without the active input from experiments. Now more than ever it is unclear which is the right way forward, making it crucial for experiments to provide concrete leads to what lies Beyond the Standard Model (BSM).
This thesis presents an analysis intended to probe physics outside the Standard Model: a search for the decay of a tau lepton into three muons ($\tau \to 3\mu$). This process is a so-called lepton-flavour violating decay that does not occur within the Standard Model, not because it is forbidden by any fundamental symmetries at its basis, but merely because none of the processes within the Standard Model can mediate it. Accordingly, if one were to see the decay, it would be a direct manifestation of BSM physics. And, on the other hand, failing to see the process would help constraining the parameter space of possible BSM models. This search uses the proton-on-proton collisions performed at the Large Hadron Collider at $\sqrt{s} = 13$ TeV as a source of tau leptons (taus can not be found in stable matter, being highly unstable particles) and is based on the data collected by

the ATLAS detector during the Run-2 years of operation.
The analysis targets the abundant leptons produced within hadronic decay chains (jets), in this it distinguishes itself from a previous ATLAS search for $\tau \to 3\mu$, which was focused on leptons produced through the decay of a W boson (and used Run-1 data collected in the year 2012).
This thesis is structured as follows:

- **Chapter 1** introduces the Standard Model of particle physics and the symmetries that accompany it. The concept of (charged) lepton flavour violation (LFV) and its connection to the origin of neutrino masses is described. A broad overview of the current experimental and theoretical status with respect to LFV processes is given, closing with a listing of the most viable theoretical models, specifically linked to the decay $\tau \to 3\mu$.
- **Chapter 2** describes the Large Hadron Collider, the ATLAS detector and its triggering system.
- **Chapter 3** illustrates the two dominant channels for τ production at ATLAS and the actions undertaken in preparation for the Run-2 $\tau \to 3\mu$ search, during the transition between Run-1 and Run-2 .
- **Chapter 4** describes the object reconstruction process at ATLAS, with particular attention for the reconstruction of quantities playing a leading role in the analysis.
- **Chapter 5** covers the $\tau \to 3\mu$ search at ATLAS on Run-2 data, focused on taus produced in hadronic jets. The proposed form of the analysis is preliminary and may undergo a a few changes prior to publication.

My personal contributions were the writing and maintenance of the *derivation* software (secondary vertex building) delineated in 5.3.1, the overlap removal improvement (4.3.2), the a_0^{xy} bias description (4.2.4), the signal normalization procedure and the HF-channel analysis presented in Chapter 5, except for the development of the trigger chains.

Chapter 1

About Neutrinos and Charged Leptons

> *"–I do mean it–, Peter said. –Whatever you think. I mean it. They only authorized you because I was so promising. But I didn't pan out. You did better. They think you're better. But I don't want a better little brother, Ender. I don't want a Third."*
>
> — Orson Scott Card, *Ender's Game*

Particle physics (or high-energy physics) describes the elementary particles constituting matter and radiation, together with the fundamental interactions that explain their behaviour. Our current understanding of the laws of particle physics has reached extraordinary precision, and is encoded in the Standard Model of Particle Physics. But what lies *beyond*? How can we connect our knowledge of the smallest to the understanding of the largest? Can we explain the universe around us by the laws governing the tiniest fragments of matter we can conceive? And how should we go about to learn more about it?

These questions are at the basis of a large variety of ongoing studies, examining nature under the "atto-meter scale microscope". Amongst all these, our group focusses on the behaviour of charged leptons and tries to answer (part of) these questions testing leptonic flavour symmetries. These symmetries arise from the Standard Model, i.e. processes or decays that do not respect them are vanishingly rare, according to the Standard Model. From this follows that the observation of any such mode would be definite evidence for what we call physics Beyond the Standard Model (BSM). In other words, we attempt to observe phenomena that could give us a direct handle onto physical processes that are not yet accounted for in our model.

This work centers on the search for a mixing amongst different flavours of charged leptons. Neutral leptons are known to be able to change from one flavour to another, thus the question arises whether charged leptons may, to some degree, do the same.

This chapter gives the reader a peek inside the rabbit hole, beginning from what is known

and continuing into what is possibly and impossibly awaiting us beyond the domain of current flavour physics.
This chapter begins with an introduction of the τ, the particle whose decay is the subject of this thesis (1.1). Then our starting point, the Standard Model, is introduced (1.2); the section closes with a listing of the shortcomings of this model and the flavour symmetries that form the basis for searches for flavour-violating decays. The next section (1.3) dives into the possible lead given by the masses of neutral leptons. And the last section tops it off with a view of the possible channels of flavour-violating decays, their current status and the implications of a search for the decay of τ into three muons (1.4).

1.1 Meet the τ

The tau lepton τ is a charged particle with a mass of 1776.8 MeV and spin 1/2. Its mass makes it such that it can not only decay to other, lighter leptons via the emission of two additional neutrinos, but as well hadronically, into various combinations of pions, kaons or lighter mesons and one neutrino. For this reason the τ has an average lifetime of about 2.9×10^{-13} seconds, much shorter than for the second heaviest charged lepton, the muon (about 2.2 μs). The τ's lifetime prevents it from flying through the whole thickness of earth's atmosphere, like muons do, however, if it is produced at relativistic speed, it can just cover distances of about 1 mm per each 10 GeV of boost it carries, due to time dilation. The existence of the τ was first discovered in the 1970s, at the Mark I detector operating at the SPEAR electron-positron collider at SLAC, California [1]. The evidence for the τ came from the observation of unexplained $e^+e^- \to e^\pm\mu^\mp + neutrals$ events. The identification of these events as the consequence of a third generation of charged leptons was made difficult by the closeness of its mass to that of the charm quark. The charm was being sought for in the same period to prove the Glashow-Weinberg-Salam Electroweak model, which could explain the suppression of strange meson decays [2]. Eventually, in 1978, the existence of the τ was confirmed and its mass and spin were accurately measured for the first time at DESY (DASP) [3] and SLAC (DELCO) [4]; the new particle was named τ after the Greek $\tau\rho\iota\tau\sigma\nu$ meaning third.
It is interesting to notice that at the time physicists were trying to determine whether there is any difference between the muons and electrons, besides their masses. There was no apparent reason why there should be multiple copies of lepton-neutrino pairs (ℓ, ν_ℓ), while for quarks the various flavours were hypothesized to participate in the GIM mechanism, explaining the suppression of flavour changing neutral currents [2]. In this spirit the expectation for a third flavour lepton-neutrino pair, was justified by the *sequential leptons* hypothesis, i.e. having found a second copy of (e, ν_e), namely (μ, ν_μ), it was deemed possible that there would be a whole sequence of them. The hope was that the discovery of a third family would help clarify an underlying logic explaining the mystery of lepton families.
Nowadays it is known from the extensive study of neutrino oscillations that lepton families are connected in a similar way as quarks, but the question whether there is some further distinction between three lepton flavours, or whether new processes have pref-

erential coupling to some families rather than others (Lepton Universality Violation) is regarded with renewed interest. There is as well a growing interest to know whether there are exotic interactions leading to so far unseen transitions within the lepton families (Lepton Flavour Violation). These are treated further in section 1.2.8.

1.2 The Standard Model of Particle Physics

The Standard Model (SM) is currently the most developed description of all known fundamental particles and their interactions via the electromagnetic, weak and strong force. Its current form has been crowned in 2012 with the discovery of the Brout-Englert-Higgs boson at the LHC (hence on called Higgs for short). This section gives a description of the SM by first introducing the fermions (matter particles), the bosons (force carriers) and, finally, the formalism and characterization of the SM's components.

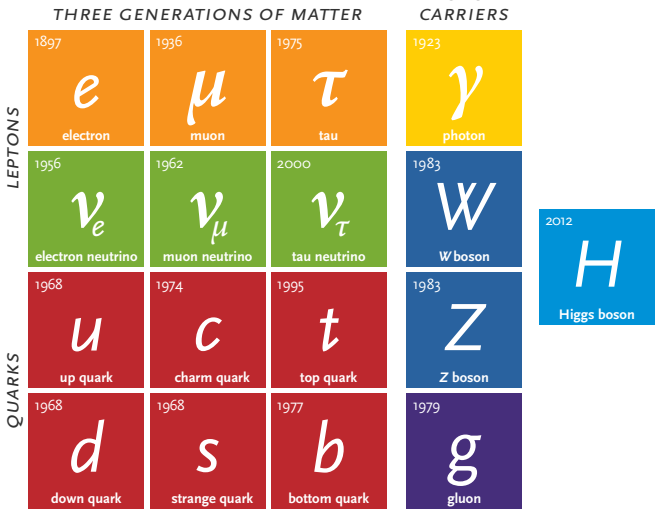


Figure 1.1: Illustration of the Standard Model: the three generations of fermions, the gauge bosons and the Higgs boson. The dates above each entry represent the relative year of discovery. Figure taken from [5].

1.2.1 Fermions

The SM spin $1/2$ fermions are conventionally grouped in leptons (color neutral) and quarks, carrying non-zero color charge. In both cases fermions are arranged in three generations or flavours, with essentially the same properties, but different masses.
The leptons count three charged particles: electron (e), muon (μ) and tau (τ) and three charge neutral neutrinos of the corresponding flavours: electron neutrino, muon neutrino and tau neutrino (ν_e, ν_μ and ν_τ respectively).
The lightest first-generation quarks are up (u) and down (d), followed by the second-generation charm (c) and strange (s) and the heaviest third-generation quarks top (t) and bottom (b). All quarks carry one of three color quantum numbers which makes

them susceptible to interaction via the strong force. The up-type quarks (u, c and t) carry an electric charge of $+2/3\ e$[1] and the remaining down-type quarks have $-1/3\ e$. For each fermion there is an anti-fermion, identical in all to its counterpart, but for its opposite electric and colour charge; they are indicated putting a bar over the corresponding particle's symbol.

1.2.2 Bosons

The forces described in the SM are mediated by spin-1 vector bosons. Of these the photon (γ) mediates the electromagnetic force, the charged $W^{\pm}$ and neutral Z bosons mediate the weak force and the neutral gluons (g) carry one unit of color charge and one of anti-color, thus mediating the strong force.
The Higgs boson (H) is the only elementary scalar in the SM (spin zero), it represents an excitation of the homonymous field, which interacts with the $W^{\pm}$ and Z bosons, as well as with most fermions and generates a mass for them [6].

1.2.3 The Underlying Symmetry

The SM is formulated as a quantum field theory with the imposition of local gauge symmetries

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

In the above group $SU(3)_C$ determines the strong interaction (subscript "C" for to *color*). $SU(2)_L$ operates in weak isospin space, the subscript "L" stands for left-handed, denoting that only the left-handed component[2] of spinors partake in this symmetry and in the derived weak interaction. $U(1)_Y$ acts on hypercharge ("Y"), a quantity which behaves mathematically identical to an electric charge. The physical fields associated with the weak and electromagnetic force arise only after the $SU(2)_L \otimes U(1)_Y$ group has been broken into $U(1)_{EM}$ via the Brout-Englert-Higgs mechanism [7].

1.2.4 How Local Gauge Theories Characterize Interactions

The way local gauge symmetries translate into interactions works as follows:
we begin from a $U(1)$ symmetry for simplicity. First one writes the Lagrangian for a fermion as the Lagrangian of a free Dirac-field:

$$\mathscr{L} = \overline{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi \tag{1.2}$$

[1] e is the elementary charge, it coincides with the electric charge of a proton and corresponds to about 1.602×10^{-19} C.

[2] Right/left handedness refers to the chirality of a particle, i.e. the corresponding spinors are eigenstates of the projection operator $\frac{1}{2}(1 \pm \gamma^5)$; for example $\psi_L = \frac{1}{2}(1-\gamma^5)\psi$. Here γ^5 is the matrix obtained multiplying the other four: $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$, often called the fifth gamma matrix.

where the index $^{\mu}$ denotes the four dimensions of Minkowski space, γ^{μ} are the four-by-four Dirac matrices (see [7]), ∂_{μ} is the partial derivative operator and m is the (real) mass of the field. The enclosing ψ and $\overline{\psi}$ are, respectively, the bi-spinors of the fermion field and its Dirac adjoint, constructed from the field's Hermitian adjoint and the time-like Dirac matrix: $\overline{\psi} = \psi^{\dagger}\gamma^{0}$.
Requiring local $U(1)$ symmetry means that the Lagrangian must be invariant under the transformation

$$\psi \rightarrow \psi' = e^{i\alpha(x)}\psi \tag{1.3}$$

with $\alpha(x)$ an arbitrary phase dependent on the local space-time coordinate x.
It is clear that changing ψ with ψ' in equation (1.2) one ends up with an extra term $-\gamma^{\mu}\partial_{\mu}\alpha(x)$. To circumvent this problem one introduces a new field A_{μ}, transforming like $A_{\mu} \rightarrow A_{\mu} + \frac{1}{g}\partial_{\mu}\alpha(x)$, and substitutes the derivative with the covariant derivative

$$D_{\mu} \equiv \partial_{\mu} - igA_{\mu} \tag{1.4}$$

which introduces a coupling constant g. For completeness one needs to add a kinetic term for the field A_{μ}, ending with a fully $U(1)$ invariant Lagrangian

$$\mathscr{L} = \overline{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \qquad F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \tag{1.5}$$

The found Lagrangian (1.5) is the Lagrangian of Quantum Electro Dynamics (QED) describing electromagnetic interactions and ultimately sub-part of the SM Lagrangian. In QED field A_{μ} is the photon field.
Note that (1.5) contains an interaction term proportional to $\overline{\psi}\gamma_{\mu}A^{\mu}\psi$, characterizing the interaction of fermions with photons. Hence demanding local $U(1)$ invariance leads to the description of a photon field which interacts with the fermions we started from.

1.2.5 The Electroweak Group

The weak force is associated with the group $SU(2)_L$; this group has three generators coinciding with the three Pauli matrices σ_a ($a = 1, 2, 3$) multiplied by $\frac{1}{2}$ (see e.g. [7] for the formulation of the Pauli matrices).
The chirality of fermions links to this group in the following way: left-handed fermions are isospin doublets, e.g. $(\nu_{\ell L}, \ell_L)^T$ or $(u_L, d_L)^T$, while their right-handed counterparts are singlets (ℓ_R), invariant under $SU(2)_L$ transformation.
Imposing $SU(2)_L$ symmetry on the Lagrangian, translates in invariance under the transformation

$$\psi \rightarrow \psi' = e^{i\tau^a(x)\frac{1}{2}\sigma_a}\psi \tag{1.6}$$

with $\tau^a(x)$ arbitrary values depending on x (summation over a is implied). Similarly to what is shown for the U(1) group, three gauge boson fields W^a_{μ} need to be introduced, one for each generator of SU(2). These gauge fields transform as

$$W^a_{\mu} \rightarrow W^a_{\mu} + \partial_{\mu}\tau^a(x) - g\ \epsilon_{abc}\tau^b(x)W^c_{\mu} \tag{1.7}$$

introducing the Levi-Civita symbol ϵ_{abc} ($\epsilon_{123} = 1$, any permutation of two indices flips its sign and it's identically zero whenever two indices coincide). The covariant derivative becomes

$$D_\mu = \partial_\mu + ig\frac{1}{2}\sigma_a W_\mu^a \tag{1.8}$$

Now, to obtain the physical fields of the SM, one needs to take the combined group $SU(2)_L \otimes U(1)_Y$. The effect of $U(1)_Y$ is identical to the one in 1.2.4, leading to

$$D_\mu = \partial_\mu + ig\frac{1}{2}\sigma_a W_\mu^a + ig'\frac{Y}{2}B_\mu \tag{1.9}$$

where in the additional term the field A_μ from equation (1.4) is renamed to B_μ, according to convention, and the coupling constant is re-written as $g'\frac{Y}{2}$, to distinguish from the weak coupling already present in (1.8), and to make explicit the dependence on the new quantum number hypercharge Y. The latter relates to electric charge Q and isospin T_3 via the relation $Q = T_3 + \frac{1}{2}Y$. At this point one can write down the Lagrangian term for each fermion field

$$\mathscr{L}_\psi = \overline{\psi}\gamma^\mu\Big(i\partial_\mu - g\frac{1}{2}\sigma_a W_\mu^a - ig'\frac{Y}{2}B_\mu\Big)\psi \tag{1.10}$$

and the remaining kinetic terms for the gauge bosons

$$\mathscr{L}_{Vkin} = -\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \tag{1.11}$$

The first thing to notice is that in this theory the mass term originally present in (1.2) must be zero, because

$$\frac{1}{2}m_\psi\overline{\psi}\psi = \frac{1}{2}m_\psi(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L) \tag{1.12}$$

is not invariant under (1.6), as the invariance of the ψ_R singlet under the transformation does not allow to cancel out the transformation affecting the ψ_L pairs. One cannot introduce a mass term for the W_μ bosons, either, as a bare term

$$\frac{1}{2}m_W W_\mu W^\mu \tag{1.13}$$

is not invariant under (1.7). In the SM the Electroweak theory is completed adding the Brout-Englert-Higgs mechanism and thus providing the missing masses to fermions and some of the gauge bosons.

1.2.6 The Brout-Englert-Higgs Mechanism

The Brout-Englert-Higgs mechanism consists in the introduction of a complex scalar field in the form of an $SU(2)_L$ doublet with unit hypercharge.

$$\phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} \tag{1.14}$$

For this field a specific potential is chosen, leading to a Lagrangian term of the form

$$\mathscr{L}_H = (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \tag{1.15}$$

Of the two real coefficients for the quadratic and quartic term of the potential, the value of μ^2 is chosen to be negative (and λ is positive), such that the potential has a "circular" minimum at $\phi^\dagger \phi = -\mu^2/2\lambda$. The key concept is that, while the potential, and thus the Lagrangian conserve their $SU(2)_L$ symmetry, the ground state of the field will lie somewhere on the minimum, thus breaking the symmetry. The universe is expected to have transitioned into the broken-symmetry ground state of the Higgs field just a fraction of a second after the birth of the universe. The field is usually parametrized as follows, around its ground state:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \tag{1.16}$$

where $v = \sqrt{\mu^2/\lambda}$ is the vacuum expectation value (VEV) of the field and h is the remaining real component of the Higgs field, representing excitations around the VEV. Using (1.16) in (1.15) and isolating all terms without the Higgs field h, one obtains the mass contribution for the vector boson fields

$$\mathscr{L}_{Vmass} = \left| \left(\frac{g}{2} \sigma_a W_\mu^a + \frac{g'}{2} B_\mu \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \right|^2 \tag{1.17}$$

In a few steps one obtains mass eigenstates from mixed states of the original fields

$$\mathscr{L}_{Vmass} = \frac{1}{8} v^2 \left[2 g^2 W_\mu^+ W^{-\mu} + (g^2 + g'^2) Z_\mu Z^\mu + 0 A_\mu A^\mu \right] \tag{1.18}$$

where the fields $W_\mu^\pm$ are positively and negatively charged W bosons (combinations of W^1 and W^2) and the remaining two are the massive Z^0 and massless photon (combinations of W^3 and B). Recombining the remaining terms in the Lagrangian to these new fields gives raise to interaction terms with each other and with the Higgs field.
The Higgs field is used to obtain lepton masses by adding to the Lagrangian the gauge invariant term

$$\mathscr{L}_{H\ell} = -y_\ell \left[(\overline{\nu}_\ell, \overline{\ell})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \ell_R + \overline{\ell}_R (\phi^-, \overline{\phi}^0) \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \right] \tag{1.19}$$

which, after symmetry breaking, turns into a mass term for the charged leptons and their coupling to the Higgs field:

$$\mathscr{L}_{H\ell} = -m_\ell \overline{\ell} \ell - \frac{m_\ell}{v} \overline{\ell} \ell h \tag{1.20}$$

where the mass of the lepton depends on the postulated Yukawa coupling y_ℓ: $m_\ell = v y_\ell / \sqrt{2}$. The theory does not predict Yukawa couplings, but experiments can test

whether the couplings to the Higgs field are proportional to the masses as expected.
The masses of down-type quarks are attained in precisely the same way for charged leptons, while for up-type quarks one requires a modified form of the Higgs doublet in (1.14)

$$\phi^{\mathcal{C}} = \begin{pmatrix} -\overline{\phi}^0 \\ \phi^- \end{pmatrix} = -i\sigma_2\phi^* \tag{1.21}$$

leading to the likewise gauge symmetric term

$$\mathscr{L}_{Hu} = -y_u\Big[(\overline{u},\overline{d})_L \begin{pmatrix} -\overline{\phi}^0 \\ \phi^- \end{pmatrix} u_R + \overline{u}_R(-\phi^0,\phi^+)\begin{pmatrix} u \\ d \end{pmatrix}_L\Big] \tag{1.22}$$

The same mechanism can be adopted to provide neutrinos with tiny Dirac masses, simply replacing the quark doublet with a lepton doublet. Experimentally it is still ambiguous whether neutrinos manifest Majorana or Dirac mass, though for as far as present time oscillation experiments are concerned, one may safely assume neutrinos to be Dirac particles (this topic is developed more in section 1.3.2).
The photon field is the sole particle to remain massless in the electroweak sector. Also massless are the gauge bosons associated to the strong interaction, because the Higgs field carries no color charge. The strong interaction is explained starting from the remaining $SU(3)_C$ symmetry in (1.1); the development is analogous to the one shown in 1.2.5, but far more lengthy, due to the complexity of the $SU(3)$ group, and goes beyond the purpose of this work.

1.2.7 Limitations of the Standard Model and Physics Beyond

The SM has proven to be extremely accurate in its description of physics governed by the electroweak and strong force, but there is reason to suspect that its current form is not the final description of particle physics up to energies of the order of the Plank scale[3].
Some of the reasons why the SM is most likely incomplete are:

- **Matter abundance in the Universe:** it is known that the visible universe is matter dominated, i.e. to our knowledge there is virtually no antimatter in it. This state of affairs is impossible to explain on the basis of the SM and the current understanding of cosmology, as the SM lacks enough processes to cause such a pronounced discrepancy between the abundance of matter and antimatter. This points towards an extension of the SM in which ulterior interactions provide the needed violation of the matter-antimatter symmetry, to explain the observed imbalance.

[3]The Planck scale (or Planck energy) $E_P \simeq 10^{19}$GeV is the energy at which quantum gravity effects are believed to become relevant, and therefore a theory not including gravity, like the SM, ceases to be valid.

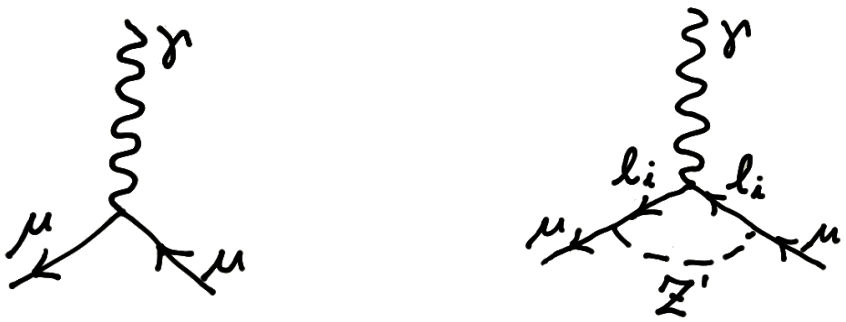
- **Dark Matter:** there is solid evidence that baryonic matter constitutes only a small portion of the mass in the visible universe; a widely embraced hypothesis is that the remainder is made up of so-called Dark Matter (DM). DM denotes stable matter with very weak or no interaction with ordinary matter or itself, besides gravitational attraction. Dominant evidence for the existence of Dark Matter comes from the explanation of density perturbations in the Cosmic Microwave Background [8], the large scale structure formation in the universe [9], and the rotational spectrum of interstellar gas in galaxies [10].

- **Neutrino masses:** the non-zero value of neutrino masses is a consolidated fact following from the observation of neutrino oscillations (described in section 1.3.1). Historically neutrinos were assumed to be massless, thus initial formulations of the SM neither include a mass term nor right-handed neutrinos ν_R. Attributing Dirac masses to neutrinos is perfectly legitimate, but because they are neutral, they are the only fermions that might be Majorana particles (i.e. be their own anti-particles), and for this reason it is important to establish what kind of mass term is the correct one for the SM Lagrangian (Dirac or Majorana). Further questions to be answered are whether we can observe interactions of ν_R, and, in the case that neutrinos should be Majorana, does this mean there is a new sector to be observed?

- **The anomalous gyromagnetic moment of the muon** (g_μ-2) features a discrepancy of about 3.5 σ between the calculated and the experimental value [11]. In QFT understanding the muon's gyromagnetic moment g_μ means calculating the higher-order corrections to the leading order coupling of muons to photons, shown in the Feynman diagram in Figure 1.2a. The quantity has been measured with highest precision at the E821 experiment at Brookhaven National Laboratory in 2006 and a fourfold increase in accuracy is expected imminently from the E989 experiment at Fermilab [12]. The magnitude of the experimentally observed deviation is quite large, this means that requires a one-loop diagram like in Figure 1.2b, to be explained. The case $\ell_i = \tau$ is preferred, due to a proportionality to m_{ℓ_i}. For this reason a Z' boson coupling to the μ-τ flavour changing current appears to be one of the most natural explanations [13]. Another remarkable possibility are leptoquarks, which can contribute to the effective coupling through a similar loop (see section 1.5.1).

- **Semi-leptonic meson decays** have shown the tendency to deviate from SM predictions. The most interesting results in this sense are the recent result at LHCb for the ratio $R(K^*) = \mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)/\mathcal{B}(B^0 \to K^{*0}e^+e^-)$ where a 2.5 σ deficit is observed [14], and the generally more trusted $R(D^*) = \mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$[4] 2.1 σ above the SM value [15]. The latter is in agreement with a 3.4 σ discrepancy observed at BaBar combining the measure-

[4] $\mathcal{B}(A \to B)$ denotes the Branching Fraction for a A to the specific final state B. It is nothing else but the fraction of times A decays to B with respect to all possible decays of A.



(a) (b)

Figure 1.2: Feynman diagrams of interest for the determination of g-2. (a) is the leading diagram affecting g. (b) shows the leading diagram introduced by the existence of a massive Z' boson with flavour-changing couplings.

ments of $R(D)$ and $R(D^*)$ [16], and the same measurement done at Belle, where a discrepancy smaller than 2 σ is seen [17].
The combined measurement of LHCb, BaBar and Belle, for $R(D^{(*)})$, has a 3.9 σ discrepancy with the SM prediction (HFLAV combination [18]). Several further discrepancies have been observed in other observables by LHCb as well, though they are more strongly affected by the lack of understanding of hadronic contributions, than the mentioned ratios. These results, although not enlightening by themselves, are suggestive as they seem to consistently point towards lepton universality violation (see 1.2.8) through a tree-level flavour violating contribution, mediated for example by leptoquarks or gauge bosons with flavour violating couplings [19].

For these and other reasons there may be the need for an extended model. There is an enormous literature of complete and partial models to explain particle physics Beyond the SM (BSM). Amongst these one counts are Supersymmetric models (SUSY) [20], left-right symmetric models (RLSM) [21], models with extended Higgs sectors [22] and many more.
Another possibility is "the bottom-up" approach of Effective Field Theory (EFT), consisting in adding individual extra terms to the Lagrangian, suppressing them through a denominator given by a cut-off scale Λ [23]. The larger the cut-off scale Λ of an extra term, the smaller its contribution to cross-sections of processes below this energy scale. The use or EFTs avoids the complication of having to deal with a high number of model parameters to be tuned as well as having to choose a specific global model on the basis of a few deviations from an otherwise very accurate SM. The down-side of EFTs is that the number of possible operators grows steeply with their order (related to the power of the cut-off scale), making their number too large to be constrained on the basis of the limited number of available experimental observables. This can result in inaccuracy when focussing only on a few "leading" operators, or make it difficult to judge from an

EFT study whether specific theories are factually compatible with data.
The analysis presented in this work has the focus of searching contributions to flavour physics beyond the SM, as a gateway into the nature of BSM physics. In the description of BSM processes of interest, reference is made to effective operators, as well as (partial) theoretical models, where possible.

1.2.8 Accidental Symmetries in the Standard Model

The SM Lagrangian features accidental global symmetries $U(1)_X$ conserving baryon number and lepton number[5] for each flavour individually ($X = B, L_e, L_\mu, L_\tau$). Focussing on the breaking of such symmetries is instructive because any indication that they occur is a proof of the existence of BSM physics, as no process can bring about related decays in the SM. The ulterior information of the relative rates of various symmetry breaking processes is then a lead on the nature of BSM processes [24].
The mentioned symmetries, combined, form the group

$$U(1)_B \otimes U(1)_{L_\tau} \otimes U(1)_{L_\mu} \otimes U(1)_{L_e} \tag{1.23}$$

this is equivalent, via basis change, to the following, more instructive form [24]

$$U(1)_{B+L} \otimes U(1)_{B-L} \otimes U(1)_{L_\mu - L_\tau} \otimes U(1)_{L_\mu + L_\tau - 2L_e} \tag{1.24}$$

where L is the aforementioned lepton number, i.e. $L = L_e + L_\mu + L_\tau$. Here below these symmetries are briefly treated, showing that most of them are not conserved in nature and introducing the processes they relate to.
The first term represents $B+L$ symmetry, and is broken at the quantum level by the non-perturbative process of Sphaleron transition, related to the Adler-Bell-Jackiw anomaly (which in turn conserves $B - L$) [25] [26].
This constitutes a transition of the vacuum state, which generates a variation of the number of leptons and baryons by $\Delta B = 3\Delta L$ affecting all flavours equally $\Delta L_e = \Delta L_\mu = \Delta L_\tau$, note that this implies that the overall lepton number is violated. To occur it requires the crossing of an energy barrier $E_S \sim v/g \approx 10$ TeV between two adjacent vacua [25].
It is hypothesized that Sphalerons could be produced at the LHC, where they would manifest as unstable energy bundles with a $\sim 10^{-2}$ fm extension, almost instantaneously decaying into a blast of ~ 100 particles, detectable as a number of highly energetic hadronic jets with a common origin, with a total energy of the order of 5-7 TeV. No experimental evidence of such events has so far been found from the recasting of the results of microscopic black hole searches [27].
The second term in (1.24) represents the $B - L$ symmetry. It is the only one of these symmetries that is not yet proven to be violated, but would be broken by the existence of Majorana neutrinos, which permit violation of lepton number by two units: $\Delta L = \pm 2$.

[5] These quantum numbers, of little importance in the SM, are defined as follows: quarks each carry Baryon number $+1/3$ and anti-quarks $-1/3$, while leptons each carry lepton number $+1$ and anti-leptons -1.

The remaining two symmetries embody Lepton Flavour Conservation. Their violation is denoted Lepton Flavour Violation (LFV) and strictly speaking is already observed, through the phenomenon of neutrino oscillations.
What still remains unproven is the existence of LFV in the charged sector (CLFV), i.e. an indication that LFV happens in tree diagrams involving charged leptons, or through any other process producing a rate of LFV interactions between charged leptons higher than the one attributed merely to neutrino mixing (more on this in sections 1.4 and 1.5). $\tau \to 3\mu$ represents such CLFV, causing $\Delta(L_\mu - L_\tau) = 2$.
Besides the ones already mentioned, one further symmetry for lepton flavours is found in the SM: Lepton Universality, i.e. lepton couplings are identical for all flavours leptons, with exception of their Yukawa couplings, as those are linked to their different masses. Lepton Universality Violation (LUV) is still not conclusively proven, but it is a very interesting field of research, becoming more in focus after LHCb's findings in the last years [28].

1.3 Neutrino Masses and More

1.3.1 Evidence for Neutrino Masses

The thorough observation of neutrino mixing proved beyond any doubt that neutrinos have non-zero masses, a discovery awarded with the Nobel prize for Physics in 2015 to researchers at Super-Kamiokande [29] and SNO [30].
Experiments have observed numerous oscillation phenomena: disappearance of electron neutrinos from the solar neutrino flux $\nu_e \not\to \nu_e$ (SNO [31]), anti-neutrino disappearance in reactor experiment $\overline{\nu}_e \not\to \overline{\nu}_e$ (KamLAND [32]), disappearance of muon neutrinos $\nu_\mu \not\to \nu_\mu$ (Super-Kamiokande [33]) oscillation of $\nu_\mu \to \nu_e$ in accelerator experiment (T2K [34]) and finally tau-neutrino appearance in atmospheric flux (Super-Kamiokande [35]) and in an accelerator experiment (OPERA [36]).
The phenomenon of neutrino oscillation is explained by neutrinos having non-vanishing distinct masses and non-coincidence of their weak flavour eigenstates with mass states, similar to what is known for the quark sector. This means that any weak interaction generates a superposition of mass eigenstates corresponding to a specific flavour state. As the neutrino then propagates through space the various mass eigenstates interfere periodically with each other. This leads to oscillating probabilities for the neutrino to belong to any specific flavour (for a complete description see e.g. [37]).
The base-change matrix between mass and flavour eigenstates, for three flavours of neutrino, is then a 3×3 unitary matrix U^{PMNS} and flavour eigenstates can be expressed as

$$\nu_\ell = \sum_{j=1}^{3} U_{\ell j} \nu_j \qquad (\ell = e, \mu, \tau) \tag{1.25}$$

where ν_1, ν_2, ν_3 are the postulated neutrino mass eigenstates.
The Pontecorvo-Maki-Nakagawa-Sakata matrix U^{PMNS} can be parametrized in terms of

3 mixing angles θ_{jk} and one complex phase δ:

$$U^{\text{PMNS}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \tag{1.26}$$

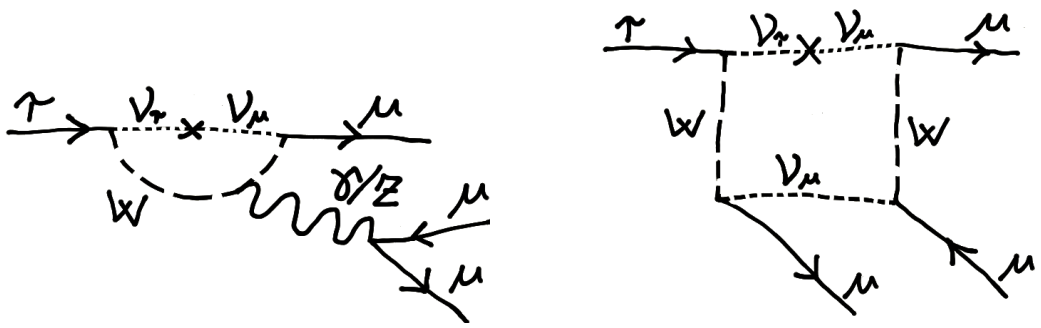
where the short forms mean $c_{jk} = cos(\theta_{jk})$ and $s_{jk} = sin(\theta_{jk})$. The periodicity of neutrino oscillations is, to first approximation, sensitive only to the difference of squared masses Δm^2_{jk}. This is how Δm^2_{21} and Δm^2_{31} are known; the two mass differences are historically associated to solar and atmospheric neutrinos, respectively. By convention $m_{\nu 1} > m_{\nu 2}$, and it still remains uncertain whether ν_3 is heavier (Normal mass Ordering) or lighter than the other two (Inverted mass Ordering); current global fits only marginally favour Normal Ordering [38], it is expected that it will take 10-20 more years before one of the orderings can be confidently rejected [39]. The current best estimate of all the mentioned parameters is reported in Table 1.1 for the two ordering scenarios. Constraints on the absolute masses of the neutrinos are obtained from cosmological analysis of the Cosmic Micro-wave Background (CMB) and other astrophysical observations: $\sum m_{\nu i} < 0.18eV$ at 95% confidence level [40].

Parameter	Best fit $\pm$ 1σ (NO)	Best fit $\pm$ 1σ (IO)
Δm^2_{21} [$\times 10^{-5} eV^2$]	$7.55^{+0.20}_{-0.16}$	$7.55^{+0.20}_{-0.16}$
$\lvert\Delta m^2_{31}\rvert$ [$\times 10^{-3} eV^2$]	$2.50^{\pm 0.03}_{\pm 0.03}$	$2.42^{+0.03}_{-0.04}$
$\sin^2\theta_{12}$ [$\times 10^{-1}$]	$3.20^{+0.20}_{-0.16}$	$3.20^{+0.20}_{-0.16}$
$\sin^2\theta_{23}$ [$\times 10^{-1}$]	$5.47^{+0.20}_{-0.30}$	$5.51^{+0.18}_{-0.30}$
$\sin^2\theta_{13}$ [$\times 10^{-2}$]	$2.160^{+0.083}_{-0.069}$	$2.220^{+0.074}_{-0.076}$
δ [°]	218^{+38}_{-27}	281^{+23}_{-27}

Table 1.1: A summary of neutrino mixing parameters and mass splittings, as obtained from a global fit. The left column reports the values obtained under the assumption of Normal Ordering (NO), the right one refers to Inverse Ordering (IO). The values are taken from [38].

As well noteworthy is the presence of the phase δ, which represents the only CP violating[6] term in the lepton sector, according to this model. Interestingly this phase appears to be maximally CP violating in the inverted scenario ($|sin(\delta)| \simeq 1$) and not suppressing CP violation in the normal scenario either. The fact that U^{PMNS} is not diagonal implies that L_ℓ are not conserved, though CLFV processes caused by this are strongly suppressed, through an amplitude suppression factor of the order m^2_ν/m^2_W, leading to expected branching rates of the order $10^{-55}-10^{-56}$ [41]. This means that any observation of LFV rare lepton decay in the near future would unmistakably be a discovery of BSM physics.

[6] CP violation refers to the non-invariance of wave functions under the combined application of charge



(a) (b)

Figure 1.3: Feynman diagrams dominating the decay $\tau \to 3\mu$ when allowing only for neutrino oscillation as a source of LFV in the leptonic sector.

1.3.2 Origin of Neutrino masses

It has to be pointed out that neutrinos can receive a Dirac mass through an interaction with the Higgs field, in quite the same way as has been shown for up-like quarks in equation (1.22). However, if neutrinos really have a Dirac mass then their Yukawa couplings need to be of the order of 10^{-12} or smaller, i.e. about seven orders of magnitude below the ones of charged leptons. It is therefore a question of "naturalness" to look for a model that motivates the tiny magnitude of neutrino masses in a more organic way.
As neutrinos are neutral it is common practice to assume their mass originating from a Majorana term proportional to $\overline{\nu}^C_{iL}\nu_{iL}$, where $\nu^C_{iL} \equiv \nu^T_{iL}\mathcal{C}^{-1}$ with $\mathcal{C}$ the charge conjugation matrix. This term is clearly not $SU(2)_L$ gauge invariant, thus it requires the introduction of a more complete theory, generally making use of the *seesaw* mechanism, as explained in the next section.

1.3.3 Seesaw Mechanism

The seesaw mechanism (Figure 1.4) can be introduced looking at the only type of dimension-5 EFT operator[7] (involving exclusively SM fields) that is Lorentz and gauge invariant [23] [42]:

$$\lambda_{\ell\ell'} L_\ell L_{\ell'} \phi\phi/\Lambda \tag{1.27}$$

where ϕ is the Higgs doublet, shown in (1.14), $L_\ell = (\nu_{\ell L}, \ell_L)^T$ is the leptonic isospin doublet, and $\lambda_{\ell\ell'}$ is a coupling constant depending on the combination of the flavours of the leptonic doublets. Equation (1.27) should be considered as a listing of relevant factors, and not an accurate Lagrangian term.

($\mathcal{C}$) and parity ($\mathcal{P}$) operators. Presence of CP violation in a model implies differentiation between particles and antiparticles.

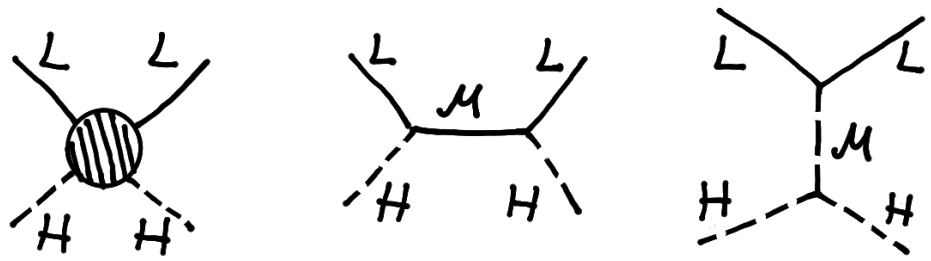
[7] Although the dimension 5 EFT operator is chosen here as a lead for the understanding of the three types of seesaw mechanism, the mechanism is not restricted to the EFT formulation. A type-I seesaw mechanism can for example be formed directly adding a (collection of) right-handed Majorana neutrinos

Depending on how currents are assembled, three types of tree-level interactions can be constructed with the form (1.27), which may derive from renormalizable SM extensions; these correspond to the three types of seesaw models [42]:

- **Type-I seesaw**: in these models the fields interact with a fermion which is identified as a right-handed Majorana neutrino, coupling to lepton-Higgs doublet pairs. The Lagrangian term takes the form
  $\mathscr{L}_{seesawI} \sim \lambda_{\ell\ell'}(L_\ell^T \phi)(L_{\ell'}^T \phi)/\Lambda$.
  Note that in this simplified Lagrangian form, no extra non-SM fields appear, as their effect is reduced to an effective coupling ($\lambda_{\ell\ell'}$) below the cut-off scale Λ. The attention here, and in the following formulae, is to the way the $SU(2)$ doublets are combined together into potential vertices, thus hinting at the different nature of the underlying propagator in a tree-level interpretation.

- **Type-II seesaw**: the mediator is given by three scalar fields, coupling to lepton and Higgs doublet currents. Note that a scalar singlet does not suffice here, as it would lead to a Mayorana mass term for charged leptons and neutrinos alike.
  $\mathscr{L}_{seesawII} \sim \lambda_{\ell\ell'}(L_\ell^T \sigma_a L_{\ell'})(\phi^T \sigma_a \phi)/\Lambda$.

- **Type-III seesaw**: the mediator is given by a multiplet of three or more fermions, each of them having the same couplings as in the case of the Type-I seesaw.
  $\mathscr{L}_{seesawIII} \sim \lambda_{\ell\ell'}(L_\ell^T \sigma_b \phi)(L_{\ell'}^T \sigma_b \phi)/\Lambda$.
  Identical to the Type-I structure, but for the matrices σ_b carrying the dimension of the fermion field multiplet.



(a) (b) (c)

Figure 1.4: A schematic view of the seesaw mechanism. Figure (a) corresponds to an effective interaction between $SU(2)_L$ lepton doublets and the Higgs field, as indicated in formula (1.27). Diagram (b) represents Type-I or Type-III seesaw, depending on the multiplet nature of the mediator(s) $\mathcal{M}$. The diagram (c) corresponds to Type-II seesaw mechanism and the mediator $\mathcal{M}$ is given by a triplet of scalar fields.

The terms given are again rough simplifications, only meant to point out the $SU(2)_L$ structure of the terms and their multiplicity.
As an illustration of the mass generation in these models, a description of Type-I seesaw mechanism follows.

to the SM, without introducing higher order operators.

1.3.4 The working of the Type-I Seesaw

In this example $N_{\alpha R}$ is a Right-Handed (RH) Majorana neutrino, an $SU(2)_L$ singlet. Its mass is linked to the cut-off scale $\Lambda \simeq M^2$ in the low-momentum transfer approximation ($q^2 \ll M^2$). The RH Majorana neutrino is coupling to the Higgs field and the $\nu_{\ell L}$ by a cross-chirality "Dirac mass term" arising from the formula shown earlier, by explicitly formulating the right-handed neutrino states $N_{\alpha R}$

$$\begin{aligned} \mathscr{L}_{L\phi N} &= \lambda_{\ell\alpha}\overline{L}_\ell \phi^c N_{\alpha R} + h.c. \\ &= \lambda_{\ell\alpha}(\overline{\nu}_\ell, \overline{\ell})_L \begin{pmatrix} \overline{\phi}^0 \\ \phi^- \end{pmatrix} N_{\alpha R} + h.c. \end{aligned} \tag{1.28}$$

$\lambda_{\ell\alpha}$ is the coupling constant for the vertex; the indices represent the left-handed lepton doublet's flavour (ℓ) and α is the index of a possible flavour structure in the RH neutrino sector. This term is gauge invariant, in compete equivalence with (1.22) and it is easy to see how it corresponds to the vertices in Figure 1.4b. After symmetry breaking (1.28) gives rise to

$$\mathscr{L}_{L\phi N} \Rightarrow \mathscr{L}_{m_D} = -\frac{1}{2} M_{\ell\alpha D} \overline{\nu}_{\ell L} N_{\alpha R} + h.c. \tag{1.29}$$

$M_{\ell\alpha D}$ being the Dirac mass term due to the non-zero Higgs VEV (hence the subscript "D"). where a second term describing the coupling to the Higgs field has been dropped because it has no relevance here. The right-handed neutrino itself is provided with a Majorana mass

$$\mathscr{L}_{m_N} = -\frac{1}{2} M_{\alpha\beta R} \overline{N}^C_{\alpha R} N_{\beta R} + h.c. \tag{1.30}$$

The processes behind the origin of the Majorana mass are explained within ultra-violet complete theories (LRSM, $SU(5)$, $SO(10)$, E_6 and other GUT theories) and are in no way required to be linked to the Higgs field. All the information given can then be condensed in a single neutrino mass matrix, referring to the flavour basis $\{\nu^C_{\ell L}, N_{\alpha R}\}$ (this basis has dimension 3+n_R, where n_R is the number of RH flavour states)

$$\mathcal{M}_\nu = \begin{pmatrix} 0 & M_D \\ M_D^T & M_R \end{pmatrix} \tag{1.31}$$

note that the 3×3 entry on top left has to be null, because it refers to the trivial Majorana terms for left-handed neutrinos, which are not $SU(2)_L$ invariant, as mentioned in section 1.3.2. Diagonalizing this matrix one obtains, in the approximation[8] $||M_D M_R^{-1}|| \ll 1$, an effective mass matrix for left-handed neutrinos.

$$M_\nu \simeq -M_D M_R^{-1} M_D^T \tag{1.32}$$

This means that left-handed neutrinos have masses inversely proportional to the Majorana neutrino's masses due to M_R^{-1} in (1.32), while RH neutrinos prevalently maintain

[8] Here $||M||$ refers to the Frobenius norm $\sqrt{Tr(M^\dagger M)}$.

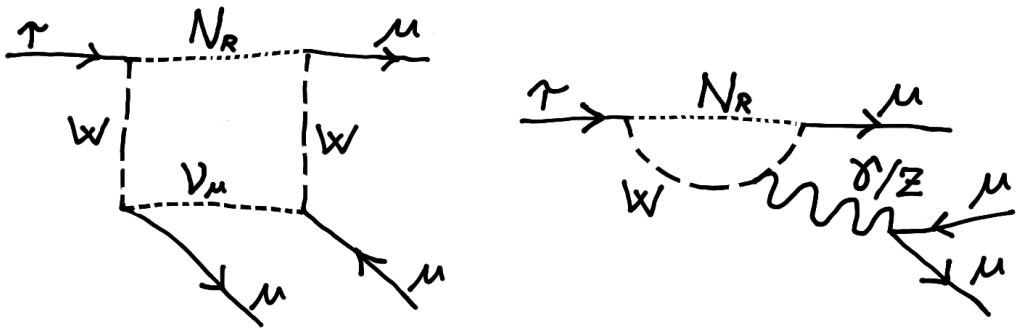
their Majorana nature and have masses very close to the large (eigen)values of M_R.
A simple example would be taking three flavours of RH neutrinos, each coupling only to the corresponding flavour of light neutrinos, and having M_R diagonal. In this case the various families decouple and one may take the matrices M_D and M_R to be numbers (different for each flavour).

1.3.5 Implications of Massive RH Neutrinos

It is often mistakenly believed that RH Majorana neutrinos are too heavy to ever be observed, for the following reason: assuming a single RH neutrino, and the coupling $\lambda_{\ell 1} \sim 1$ in equation (1.28), light neutrinos with $m_\nu \sim 0.1$ eV entail $M_N \simeq M_R \sim 10^{14}$-$10^{15}$ GeV [43]. However this does not have to be the case for (some of) the RH neutrinos, if there are multiple flavours of them, and of course $\lambda_{\ell\alpha}$ can vary as well.
Indeed there have been studies showing that models with masses as low as a few GeV or even keV could not only explain the lightness of the "active" neutrino flavours, but as well be a candidate for Dark Matter [44] [45] [46]. Even stronger than this, the *naturalness* requirement of limiting the loop corrections to the Higgs potential quadratic term translates in an *upper bound* on the mass of the lightest RH neutrino mass of $\sim 10^7$ GeV [47].
Various scales for RH neutrino masses can be justified based on several valid arguments [48]. Because the mass of possible RH neutrinos is not strictly constrained by theory, it may be possible to produce them at collider experiments. Searches have been accordingly performed at ATLAS [49] [50] and CMS [51] in dilepton channels, but no direct evidence has been found. The limits, set by these searches, reach masses of ~ 2.5 TeV depend on modelling parameters, most notably the effective coupling strength between RH neutrinos and charged leptons. For this reason these searches do not entirely exclude the existence of RH neutrinos even with such masses. Seesaw models, mostly with 3 RH neutrinos are very attractive ($SO(10)$ and E_6, for instance, lead automatically to 3 RH neutrinos), not only because of their "natural" modelling of neutrino masses, but as well because they fit with leptogenesis scenarios.
Leptogenesis denotes a series of models in which the matter-antimatter asymmetry of the universe is explained through an imbalance in the leptonic sector, through substantial CP violation in the decay of heavy leptonic states (Majorana neutrinos or oscillating sneutrinos, in the case of SUSY models), then translated in baryon asymmetry through sphalerons and frozen in place by the ceasing of the underlying process due to a sudden transition during the universe's evolution [52] [53].
Focussing on more complete seesaw theories it becomes evident that their parameter space is considerably constrained by current limits on LFV and LNV processes, such as neutrino-less double-beta decay[9] $0\nu\beta\beta$ and $\mu \to e\gamma$ (reported in section 1.4).

[9]Neutrinoless double beta decay is a, so far unconfirmed, nuclear decay that could happen if there are processes violating lepton number. The decay itself should happen in unstable even-even nuclei (nuclei with an even number of neutrons and protons), when a single beta decay is energetically forbidden. In these cases a double beta decay is generally observed, entailing the simultaneous emission of two positrons and two electronic neutrinos. If possible, in these cases a neutrinoless double beta decay should have

In particular minimal extensions of the SM with ∼GeV-TeV RH neutrinos appear to be beyond the reach (in the near future) of $\tau \to 3\mu$ searches, as diagrams such as the ones depicted in Figure 1.5 appear to be sub-dominant. For these scenarios the study of processes involving μ and e or semi-hadronic channels like $K^+ \to \mu^+\mu^+\pi^-$ seem more promising [55] [56]. This said, there is no reason why nature should not have chosen



(a) (b)

Figure 1.5: Two Feynman diagrams responsible for the possible contribution of right-handed neutrinos to the decay $\tau \to 3\mu$. From diagram (b) it is evident that an analogous diagram for $\tau \to \mu\gamma$, without the photon converting to two muons, has a larger amplitude, as it does not require the photon to be virtual.

a more exotic twist: [57] investigates such an approach, looking at a Type-III seesaw scenario with scalar and fermion $SU(2)_L$ multiplets and an additional $U(1)_{L_\mu - L_\tau}$ gauge symmetry. In this model the smallness of the neutrino masses is achieved through the double suppression effect of Majorana masses of the fermion multiplet and a small VEV for the additional scalar fields. The added gauge symmetry allows to explain the g-2 anomaly through the required low-order contribution of Z' and it is shown that it may even fit the observed deviations in semileptonic meson decays mentioned in section 1.2.7 (more on this in the next section). This extremely encompassing scenario can produce $\mathcal{B}(\tau \to 3\mu)$ just below the current best upper limit of $\mathcal{B}(\tau \to 3\mu) < 2.1 \times 10^{-8}$ [58], mainly relying on the diagram in Figure 1.6 [59] [60].

a wider phase space, partially compensating for its smaller couplings. The process could happen if positrons couple to Majorana neutrinos. As these would be their own antiparticle there would be no need to emit neutrinos. Such a decay violates lepton number $\Delta L_e = -2$ and is commonly considered the exemplary test for Majorana nature of neutrinos [54].
Although very significant, the non-observation of this decay only sets limits on the effective coupling between electrons and Majorana neutrinos, while couplings of other lepton flavours and Majoranas undergo indirect model-dependent constraints.

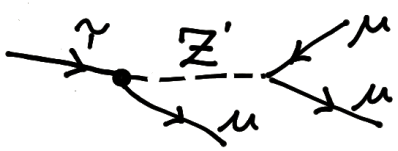


Figure 1.6: Diagram of the LFV decay $\tau \to 3\mu$ mediated by a massive Z' boson. The neutral boson is supposed to have a LFV coupling in the marked vertex.

1.4 Experimental Status of CLFV Searches

The study of CLFV and LNV processes and their interrelations is an extended field of research. Their high sensitivity to BSM processes motivates these searches, but even if new physics should be first discovered on the basis of heavy resonances at accelerator experiments, CLFV searches represent an invaluable starting point to pin down the precise nature of BSM physics and it remains imperative to study all possible CLFV channels (for an overview see [61] or [62]). Table 1.2 reports the current best upper limits set on

Process	Limit $\mathcal{B}()$	C.L.	Experiment
$\tau \to 3\mu$	2.1×10^{-8} [63]	90%	Belle
	3.8×10^{-7} [64]	90%	ATLAS
$\tau \to \mu\gamma$	4.4×10^{-8} [65]	90%	BaBar
$\mu \to 3e$	1.0×10^{-12} [66]	90%	SINDRUM
$\mu \to e\gamma$	4.2×10^{-13} [67]	90%	MEG
$\mu N \to eN$	7.0×10^{-13} [68]	90%	SINDRUM II
$Z \to e\mu$	7.3×10^{-7} [69]	95%	CMS
	7.5×10^{-7} [70]	95%	ATLAS
$Z \to e\tau$	9.8×10^{-6} [71]	95%	OPAL
	4.7×10^{-5} [72]	95%	ATLAS
$Z \to \mu\tau$	1.2×10^{-5} [73]	95%	DELPHI
	1.7×10^{-5} [74]	95%	ATLAS
$H \to e\mu$	0.036% [75]	95%	CMS
$H \to e\tau$	0.61% [76]	95%	CMS
	1.04% [74]	95%	ATLAS
$H \to \mu\tau$	0.25% [76]	95%	CMS (13 TeV)
	1.43% [74]	95%	ATLAS (8 TeV)

Table 1.2: A selection of CLFV processes with the best limits set on them by experiments and ATLAS Run-1 limits.

the branching fractions of CLFV processes, it is clearly visible that the highest sensitivity is attained for the mentioned processes. The leading experiments in this field are the ones dealing with muon decays and muon conversion in nuclear environment. The reason for this is the huge number of muons that are produced at proton accelerators, combined with the simplicity of their final states, permitting an ease of construction of effective and efficient experiments. The signature of $\mu^+ \to e^+\gamma$ is, for instance, the emission of a positron and a photon, back-to-back in the rest frame of the initial anti-muon; similarly $\mu \to 3e$ is recognized by three electrons/positrons, this time linked by the common origin and an invariant mass corresponding to the muon mass. Muon conversion: $\mu^- N \to e^- N$ represents the conversion of a muon bound to a heavy nucleus, into an electron, via the exchange of a virtual photon with the nucleus. The resulting electron carries the bulk of the muon's energy, thus being freed from the nucleus. The process is distinguished from normal muon decay $\mu \to e\nu_\mu\bar{\nu_e}$ by the sharp value of the energy of the emitted electron, caused by the presence of a two-body decay.
CLFV decays of Z and Higgs bosons are of great interest, as well, and are gaining importance as an increasing number of them are produced at the LHC. These decays count $Z/H \to e\mu$, $Z/H \to e\tau$ and $Z/H \to \mu\tau$ (see Table 1.2).
ATLAS further performed a generic search for a heavy state (possibly a Z') decaying into a different-flavour lepton pair and the strongest limit has been set on the product of production cross section and branching fraction into an e-μ pair, of the order $\sigma(X)\mathcal{B}(X \to e\mu) \sim 3 \times 10^{-4} pb$ for masses in the range $1 - 5$ TeV [77]. Tau decays are of course of great interest for the study of CLFV as well; due to its mass the possible final states to study are numerous. Figure 1.7 shows all the sought final states together with the best limits set on the relative branching fraction. The majority of these limits were set by "B factories" BaBar [78] and Belle [79], however LHC experiments LHCb and ATLAS have researched some of the channels as well.
Of particular interest for this analysis are the limits set on $\tau \to 3\mu$ and $\tau \to \mu\gamma$, as well reported in Table 1.2. Currently the lowest limit on $\mathcal{B}(\tau \to 3\mu)$ is set by Belle at $2.1 \pm \times 10^{-8}$ [63] and the lowest limit on $\mathcal{B}(\tau \to \mu\gamma)$ is set by BaBar at 4.4×10^{-8} [65].

Generally speaking experimental sensitivities to rare μ decays are higher than for τ ones, in connection with cleaner final states and ease of muon production and handling, but several new physics scenarios predict an enhancement of rare τ decays, compared to μ decays. Such enhancements are, of course, model dependent, but may be as large as $(m_\tau/m_\mu)^2 \sim 280$ if linked to chirality effects, $|U_{\tau 3}/U_{\mu 3}|^2 \sim 20$ if dependent on lepton mixing factors or $\Delta m^2_{13}/\Delta m^2_{12} \sim 30$ if coming from neutrino masses [61]. A more trivial enhancement of LFV in tauonic channels can of course be achieved in the case of BSM scenarios which are predominantly coupled to the third generations [81].
For this reason studying rare $\tau \to 3\mu$ decays is an important ingredient for the understanding of the nature of BSM physics. Knowing the (excluded) rate of $\tau \to 3\mu$, in comparison with processes such as $\tau \to \mu\gamma$ will give a handle on the higher-order operators involved. These two decays need to be compared because of the different operators they may depend upon. In principle $\tau \to 3\mu$ may just occur through the same process as $\tau \to \mu\gamma$, with roughly two orders of magnitude separating the rates of the two; however

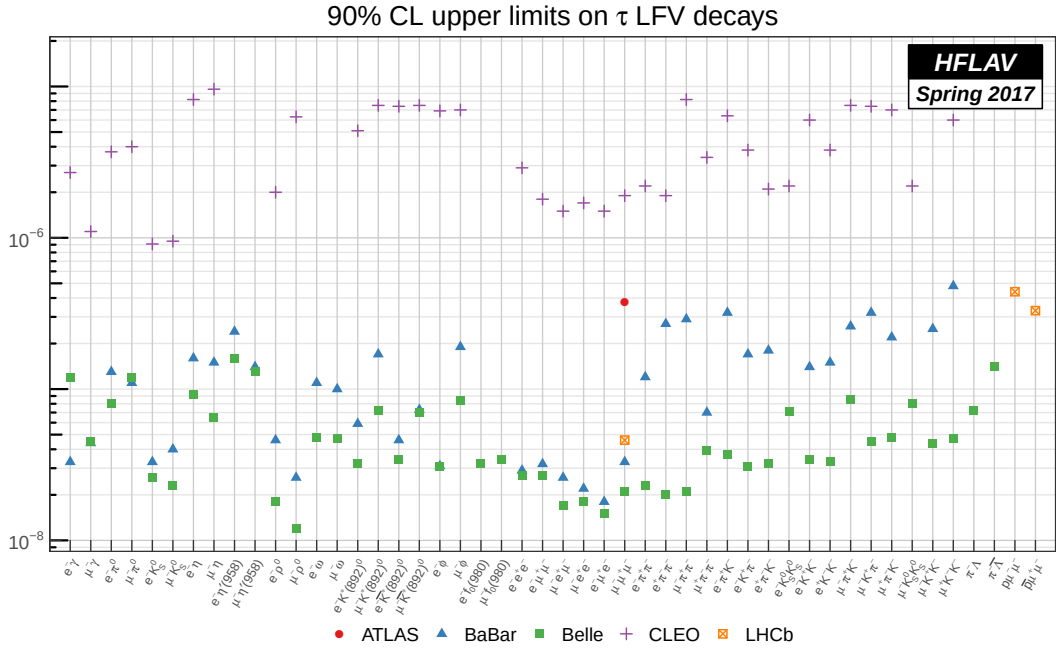


Figure 1.7: Summary plot of the upper limits set on branching fractions for rare τ decays by various experiments. Image taken from [80].

a far more interesting situation would be if the two τ decays were governed by different diagrams, in which case we would learn much from them, and it might be possible to discover $\tau \to 3\mu$ at rates just below the current limits, as suggested by the models below. And once $\tau \to 3\mu$ will be observed, further valuable information will come from the polarizations of the decay products [82].

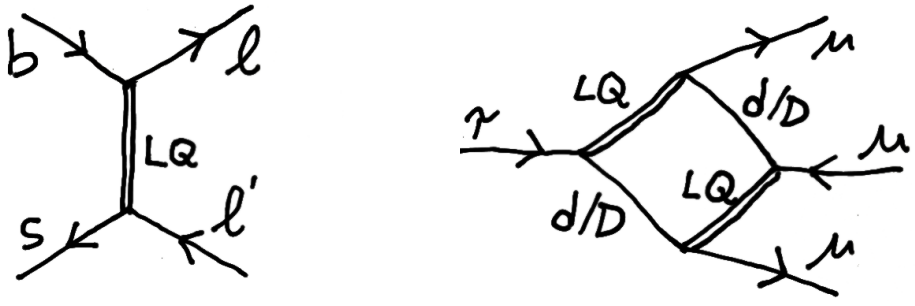
1.5 Theoretical Models for $\tau \to 3\mu$

This section summarises four theoretical BSM models mechanisms that can play a role in an enhancement of $\tau \to 3\mu$. These particular models are chosen because of their historical relevance or because of recent favour amongst the theoretical community, as well as for their relevance for $\tau \to 3\mu$ and LFV in general.

1.5.1 Leptoquarks

Leptoquarks (LQ) are bosons that carry both lepton and baryon number, they are color triplets and obey flavour structuring. They naturally arise in Pati-Salam models, where the SM $SU(3)_C$ originates from a larger group in which leptons play the role of a fourth color. Leptoquarks seem to be indicated as a natural explanation for the deviations seen in $R(K^*)$ and $R(D^{(*)})$, which they affect through loop-level or even tree-level diagrams

like the one in Figure 1.8. The same figure shows the loop diagram through which LQ can cause a signal in $\tau \to 3\mu$. Recent minimal, but complete, models for SM extensions with LQ can be found described in [83] and [84]. The article [83] focusses on the group $SU(4)_C \times SU(2)_L \times U(1)_R$; within this model any impact on rare τ decays is far from the current limits, while the strongest constraint from rare lepton decays comes from the limit on $\mu \to 3e$. It is though noteworthy that this model is unable to describe the observation in $R(D^*)$, due to constraints from LFV observables; thus it is reasonable to assume that a comprehensive model should be slightly more complex. The model described in [84] is similar, but centred on accounting for both the deviations observed in semi-leptonic decays. Strong constraints from LFV channels (for example $\mu \to e\gamma$) are integrated assuming the absence of mixing with the first generation, while rare decays involving third generation leptons are significantly enhanced with respect to the SM, in particular $B_S \to \tau\tau$ and $b \to s\mu\tau$; from this follows that $\tau \to 3\mu$ should receive a non-negligible contribution as well. In such models the anomalous magnetic moment of the muon can be easily accounted for, through the contribution of a loop involving LQ and heavy vector-like fermions (as well representation of $SU(4)_C$).
Another very interesting model is the mentioned PS^3 model [81]. Its peculiarity is to triplicate the high-scale group, explicitly splitting the three generations. At the low-energy scale, the three generations are united, while leptons and strong colors become independent. This separation of the three flavours permits for the BSM physics to couple predominantly to third generation particles, at accessible energies. The model also manifests a distinctive enhancement of LFV processes like $\tau \to 3\mu$ or $\tau \to \mu\gamma$, which can fall broadly within the range of rates testable in the near future; the magnitude of these rates is not predicted by the model, but there seems to be an anti-correlation between these modes and LFV b decays such as $B_s \to K\tau\mu$ and $B_s \to \tau\mu$, thus future measurements of both classes of rare decays are of great importance to test the model.



(a) (b)

Figure 1.8: Coupling of leptoquarks to SM quarks and leptons. Figure (a) shows the minimal tree diagram responsible for decays like $B_S \to \tau\tau$ or $B_S \to \tau\mu$; figure (b) shows the loop through which LQ can cause $\tau \to 3\mu$. In the propagators labelling D denotes vector-like down-type fermions, while d denotes a SM down-type quark (in both cases the main contribution should come from a second generation fermion).

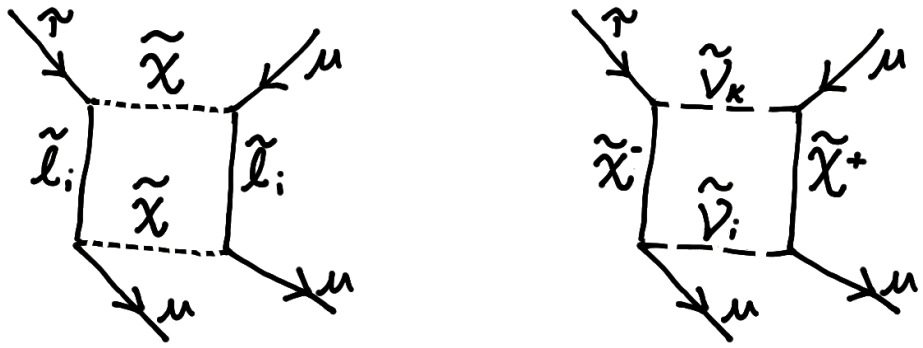
1.5.2 Extended Electroweak Models

The mentioned anomalies observed by LHCb in $R(K^{(*)})$ and $R(D^{(*)})$ have led to considerable interest in theoretical research, the understanding is that the observed effects could be caused by LUV contributions modelled as operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ [85], similar to the effect of a massive Z' boson [86].
Interesting about this is the fact that it has been shown that BSM LUV processes unavoidably induce enhancement in LFV as well (fine-tuning would be required to escape this) [87]. For instance it is calculated that the interplay of the various operators explaining the anomalies may give $\mathcal{B}(\tau \to 3\mu) \sim 4 \times 10^{-9}$ [88], not far from the current limit.
For a more complete encompassing model, one may see the example of a Pati-Salam model, where the arising Z' has an analogous LFV impact [81]. In this model, high-scale breaking of the group produces TeV-scale weak vector bosons Z' and W'.

1.5.3 Supersymmetry

Considerable contributions to (C)LFV can be obtained in Supersymmetric models, as well. SUSY is historically one of the major players when it comes to BSM modelling.
The fundament of Supersymmetric models is the introduction of a super-partner for each of the SM fermions and gauge bosons. Super-partners of fermions are bosons and super-partners of bosons are fermions; super-partners share all quantum numbers with their SM counterparts, except their spin. The dominant argument for SUSY is to provide a symmetry to protect the Higgs potential against enormous loop corrections due to the close-to-unity coupling with the top quark. The non-observation of super-partners with the same masses as SM particles, though implies that the symmetry must be broken, leading to the assumption that their masses should lie in the ∼TeV range. SUSY models are characterized by a high number of free parameters and are therefore not very predictive and difficult to constrain.
In these models LFV is originated through the non-alignment of mass states in the two sectors (super-partners and SM) in diagrams such as in Figure 1.9. Again minimal models seem connected to highest sensitivity in $N\mu \to Ne$ and $\mu \to e\gamma$, while the vast parameter space in non-minimal models may lead to very diversified signatures, in particular when considering R-parity[10] violating models. For instance [89] shows that CLFV in τ decays can be enhanced sufficiently to justify $\mathcal{B}(\tau \to 3\mu) \sim 10^{-8}$. R-parity violation, opens the way for further CLFV diagrams, for instance Figure 1.10 shows a possible tree-level source for this decay involving a sneutrino [90].

[10] R-parity refers to the conservation of the quantity $P_R = (-1)^{3B+L+2s}$ with B the boson number, L the lepton number and s the spin. This imposes conservation of Baryon and Lepton number. SM particles have $P_R = +1$ while super-partners have $P_R = -1$. This requires super-partners to be pair-produced, and the lightest of them to be stable.
R-parity violation allows for more complex models, though particular care is needed in these models to prevent proton decay.



(a) (b)

Figure 1.9: Two R-parity conserving supersymmetric contributions to the $\tau \to 3\mu$ decay. (a) involves the exchange of sleptons $\tilde{\ell}_i$ and neutralinos $\tilde{\chi}^0_k$, while (b) involves a loop of sneutrinos $\tilde{\nu}_i$ and charginos $\tilde{\chi}^\pm_k$.

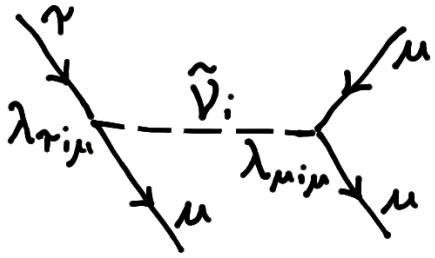


Figure 1.10: The Feynman diagram of a tree-level contribution to $\tau \to 3\mu$ obtained by the R-parity violating exchange of sneutrinos.

1.5.4 Extended Higgs Sector

The SM includes a single Higgs doublet, but it is possible that there may exist more, not unlike left-handed fermion doublets. Although there is no theoretical constraint on the number of such scalar doublets, the simplest, and therefore most studied case concerns two Higgs doublets (2HDM). The best known example of 2HDM is minimal Supersymmetry, where they produce the well known five physical bosons (h, H, A, $H^\pm$). In SUSY models many channels are possible for the introduction of LFV contributions [91], including the one coming from the Higgs doublets shown in Figure 1.11 [92]. It has been shown that a necessary and sufficient condition for having no FCNC arising in models with more than one Higgs doublet, is that all fermions with the same charge need to couple exclusively to one scalar field [93]. Whenever this is not the case, the Yukawa matrices relative to each scalar cannot be diagonalised simultaneously, thus inducing FCNCs.
2HDM are characterised by a large number of degrees of freedom, and thus are constrained on the basis of many processes amongst which $B_S \leftrightarrow \overline{B}_S$, $h \to \ell\ell'$ and LFV lepton decays. A recent review [94] identified $\mu \to e\gamma$ as the source of strongest con-

straint from LFV decays on flavour violating Higgs couplings, but recognises that the interplay of other bounds are more model dependent, and thus might, given the right model, be the first ones to manifest a contribution from extra scalars. So far constraints from $h \to \mu\tau$ decays at the LHC are weaker than the expected branching fractions derived from rare leptonic decays, therefore the input from leptonic decays like $\tau \to 3\mu$ remains in any case fundamental to study models with multiple Higgs doublets.

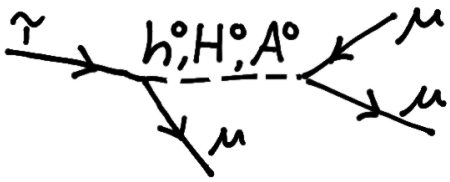


Figure 1.11: Feynman diagram showing the possible tree-level contribution of 2HDM to $\tau \to 3\mu$.

Chapter 2

The LHC and ATLAS

> *"They are very near the Big Combination now. Ty has seen it, a great criss-crossing confusion of metal rising into the clouds from a smoking pit about half a mile east. It looks like a madman's conception of a skyscraper, a Rube Goldberg collection of chutes and cables and belts and platforms, everything run by the marching, staggering children who roll the belts and pull the great levers."*
>
> — Stephen King, Peter Straub, *Black House*

This chapter outlines the experimental setup used in the presented $\tau \to 3\mu$ analysis. First a description of the Large Hadron Collider (LHC) is given and its performance is introduced through the concept of delivered *Luminosity* (2.1). Then the ATLAS experiment is described in general (2.3), followed by a more detailed view of the various subdetectors: the Inner Detector (2.4), the Calorimeters (2.5) and the Muon Spectrometer (2.6). The chapter closes on the topic of the ATLAS Trigger System (2.7).

2.1 The LHC

To date the most powerful man-made accelerator, the LHC is placed in the 27 km long circular tunnel originally built for LEP [95], at about 100 m underground at CERN, near Geneva (Switzerland). The accelerator stores and accelerates two beams of protons or heavier nuclei revolving in opposite direction, colliding them at the locations of the four main experiments: ATLAS, CMS, Alice and LHCb.
A system of 1232 superconducting dipole magnets, generating an 8 T field, is employed to steer the beams in their circular path. The dipole magnets are kept at 1.9 K and are designed such that the magnetic dipole field switches direction over a short distance, providing the field for both positively charged beams at the same time, a clever way for the magnets to be compact enough to fit in the tunnel. For the acceleration, the LHC uses 16 radio-frequency (RF) cavities where a 40 MHz alternated electric field of 5 MV/m boosts protons in each bunch to 6.5 TeV and keep that energy against a nominal per-beam energy loss of 3.7 kW in the form of cyclotron radiation.

As can be seen in the schematic view of the various accelerating facilities at CERN (Figure 2.1), protons undergo a series of progressive boosts before being injected into the LHC. The procedure, which accelerates protons to more than 99.99% of the speed of light, is as follows: protons reach a momentum of 50 MeV in Linac 2, then 1.4 GeV in the PS Booster, at this point they are divided into 25 ns separated bunches and reach 26 GeV in the Proton Synchrotron (PS) before being injected into the Super Proton Synchrotron (SPS), where they attain the energy of 450 GeV; only then they are ready to be injected in opposite directions into the LHC [96].
Maintaining the beams collimated inside the LHC for several hours is an extremely complex task, and a whole system of magnetic multipoles is used to preserve the beams against coulomb repulsion and the non-linear effects of beam-beam interaction, as well as the (much weaker) gravitational pull. Most notably triplets of quadrupole magnets are set at both ends of the straight sections in which the experiments are placed, to achieve an extra "squeeze" of the beams as they collide with each other [97]. This is important for the delivery of *Luminosity* to the experiments, which is the topic of the next section.

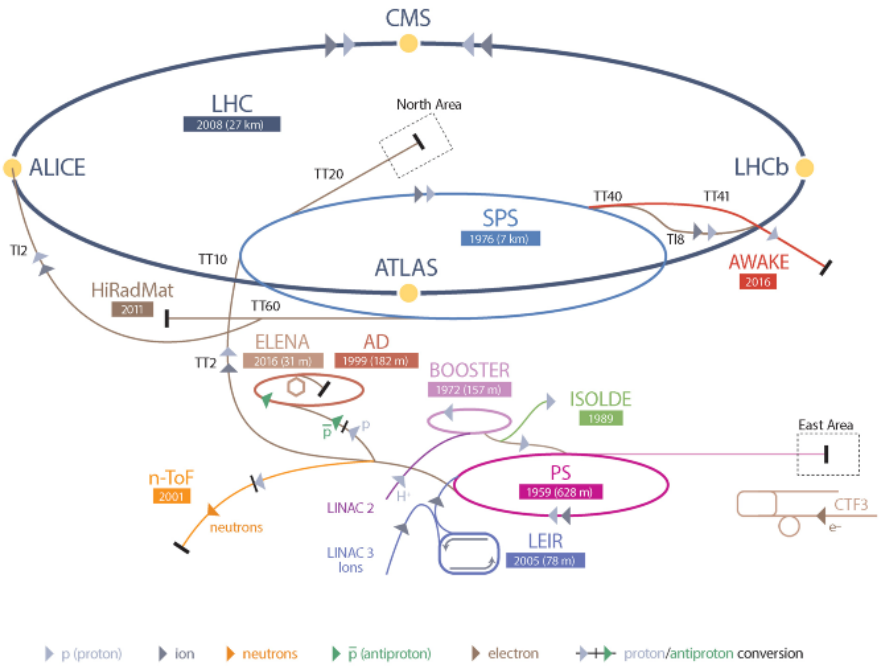


Figure 2.1: Schematic representation of the accelerator facilities at CERN. Image taken from [96].

2.2 Luminosity and Pileup

The quantities related to collider physics that are of importance to the experiments, and specifically ATLAS, are Centre-of-Mass (CM) energy, luminosity and pile-up.

2.2.1 Centre-of-Mass energy

During data taking, relevant for this work, proton bunches were colliding at the LHC at 13 TeV CM energy (the LHC's design energy is 14 TeV). The CM energy directly relates to the cross section of interactions and makes it energetically possible to generate high-energy states. For this reason the LHC can boast the largest production of top quarks, and, since it crosses the energy threshold for the production of TeV states, it allows the search and study of processes well into the TeV scale.
A peculiarity of high-energy proton collisions is that the colliding particles are not the protons themselves, but partons (gluons and quarks) inside them, which are asymptotically free at such energies. Thus the colliding particles carry only a fraction of the proton's energy, making the CM energy of colliding partons unknown a priori. Nevertheless, so called *hard scatter* events do still reach multiple TeV of CM energy.
This analysis is based exclusively on $\sqrt{s} = 13$ TeV p-p collisions, produced in the years 2015-2018.

2.2.2 Luminosity

The instantaneous Luminosity (L) is a measure of how many particles are brought into sufficient proximity that they may collide, per unit of time. It is related to the rate of specific events under study through the relation:

$$N_{event} = L\sigma_{event} \tag{2.1}$$

where N_{event} is the occurrence rate for the type of events under study and σ_{event} the relative cross section, commonly expressed in barn ($1b = 10^{-24}cm^2$). L can be calculated for gaussian shaped bunches to be:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\sigma^{*2}} F \qquad\qquad F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)^{-1/2} \tag{2.2}$$

Here N_b is the number of protons in a bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency of the beams (11.2455 kHz), γ_r is the relativistic gamma factor (6930 at 6.5 TeV), and $\sigma^* = \sqrt{\varepsilon_n \beta^*}$ the RMS of the transverse beam size at the interaction point, depending on the normalized transverse beam emittance ε_n and the beta function at the collision point β^*. The factor F represents the reduction in luminosity due to the crossing angle between the beams at the collision point, the expression for F shown in (2.2) is valid for identical gaussian bunches: θ_c is the crossing angle between the beams, σ_z is the RMS bunch length. Table 2.1 reports some instructive typical LHC parameters for the years 2015, 2016, 2017 and 2018, as well as their design values.
Figure 2.2a shows the time-integrated luminosity $\mathcal{L} = \int L dt$ recorded by ATLAS in the period 2015-2018; $\mathcal{L}$, measured in units of inverse femtobarn ($1fb^{-1} = 10^{15}b^{-1}$), ultimately determines the number of produced events. In the case of the $\tau \to 3\mu$ search, the branching fraction is tiny (see Chapter 3), thus a very high luminosity is fundamental to attain sensitivity to the process.

Parameter	2015	2016	2017	2018	Design
Maximum number of filled bunches	1380	2232	2208	2556	2808
Typical number of protons per bunch $[\times 10^{11}]$	1.2	1.18	1.3	1.1	1.15
Peak instantaneous luminosity $[\times 10^{33} cm^{-2} s^{-1}]$	5.0	13.8	20.9	21	10
Delivered luminosity, integrated $[fb^{-1}]$	4.2	38.5	50.2	63.3	-
Average pileup for the year	13.4	25.1	37.8	36.1	25.0

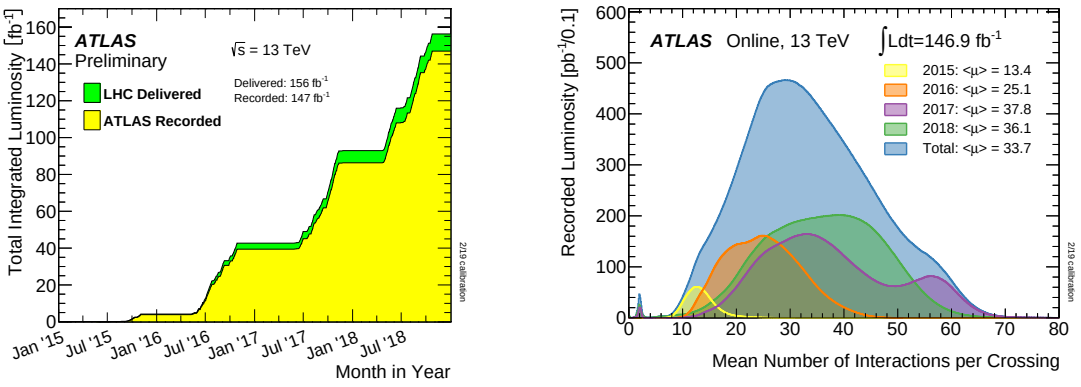
Table 2.1: Performance of the LHC for the $p-p$ runs at $\sqrt{s} = 13$ TeV in the years 2015-2018 taken from [98] and [99] compared with the design parameters extracted from [97] and [96].

2.2.3 Pileup

Pileup denotes the overlap of the products of multiple interactions in the ATLAS detector. It can be of two types: *out-of-time* pileup, given by the products of consecutive bunches or asynchronous events, or *in-time* pile up, namely the concurrence of multiple inelastic p-p interactions during a single bunch crossing. *In-time* pileup is generally referred to as just *pileup* and is shown in Figure 2.2b. *In-time* pileup follows a Poisson distribution around a mean given by:

$$\mu = \frac{\sum L_{bunch}}{n_b f_{rev}} \sigma_{inel} \tag{2.3}$$

Where the fraction represents the instantaneous average per-bunch Luminosity (the sum-



(a) (b)

Figure 2.2: LHC performance in terms of luminosity and pileup. Figure (a) shows the integrated luminosity delivered and recorded by ATLAS between 2015 and 2018. Figure (b) shows the mean pileup distribution weighed by luminosity, for the years 2015, 2016, 2017 and 2018 and their combination. The average pileup is determined from equation (2.3), with $\sigma_{inel} = 80$ mb and the bunch-by-bunch luminosity measured by ATLAS' luminometers BCM [100] and LUCID [101], calibrated with February 2018 calibrations determined with the van der Meer scan method [102].

mation symbol indicates a sum over the bunches) and σ_{inel} is the total inelastic p-p cross section; σ_{inel} is determined for $\sqrt{s} = 13$ TeV using MC generator PYTHIA8 [103] [104], as well as being measured by ATLAS, TOTEM [105] [106] and LHCb. The ATLAS measurement is obtained through an extrapolation of a fiducial measurement (i.e. extended over a limited solid angle) performed on the basis of counts of the forward minimum-bias trigger scintillators and led to the estimate 78.1 ± 2.9 mb [107], while the MC derived value is 78.4 mb; the measurements by TOTEM and LHCb produce, respectively, the values 75.4 ± 5.4 mb [108] and 79.5 ± 1.8 mb [109]. Keeping in mind (2.3) and (2.2) one can understand that the mean pileup varies during a run, as the protons inside each bunch N_b are progressively depleted, thereby reducing the per-bunch Luminosity.

2.3 ATLAS

2.3.1 The Detector

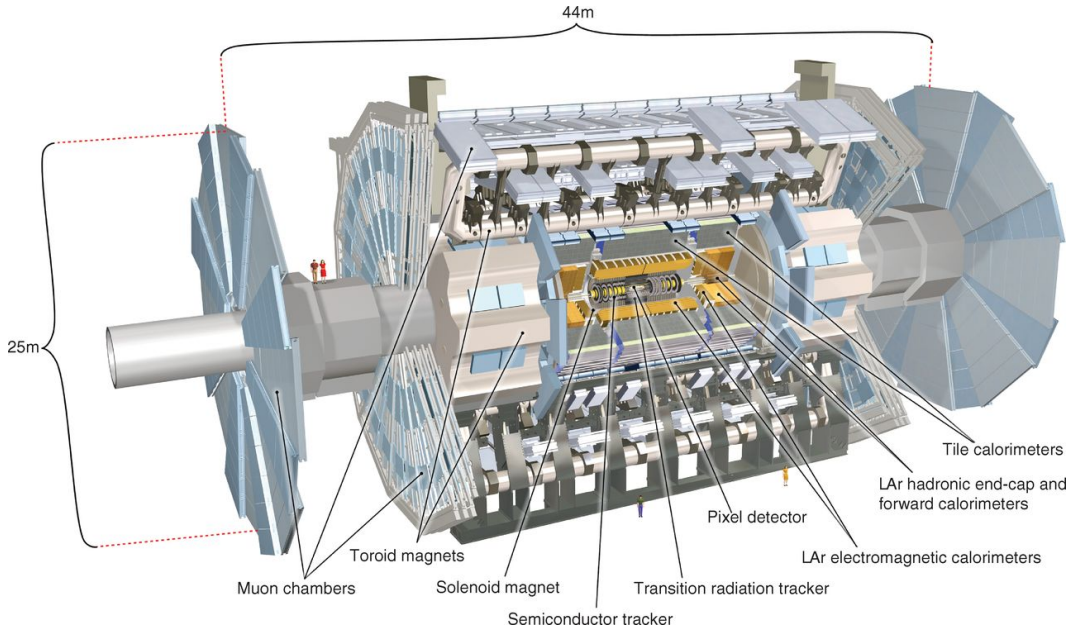


Figure 2.3: A cut-away view of the ATLAS detector.

The ATLAS detector [110] is a so-called "general purpose" detector built with the search for the Higgs and SUSY final states in mind; it has been optimised for the detection of hadrons, photons and charged leptons in the GeV to TeV energy range. The detector has a cylindrical symmetry and is built around one of the four beam crossing points of the LHC. It has an almost hermetic coverage: the Inner Detector covering an angle with the beam direction down to 9.4° and the Hadronic Calorimeter even down to 1°. An overall view of the detector is shown in Figure 2.3: one can distinguish, from the inside out, the Inner Detector tracker (ID), the 2 T solenoid magnet, the liquid Argon (LAr)

Electromagnetic Calorimeter (EMCal), the Hadronic Calorimeter (HCal) and finally the Muon Spectrometer (MS), throughout which the barrel and endcap toriod magnets extend.

2.3.2 The ATLAS Coordinate System

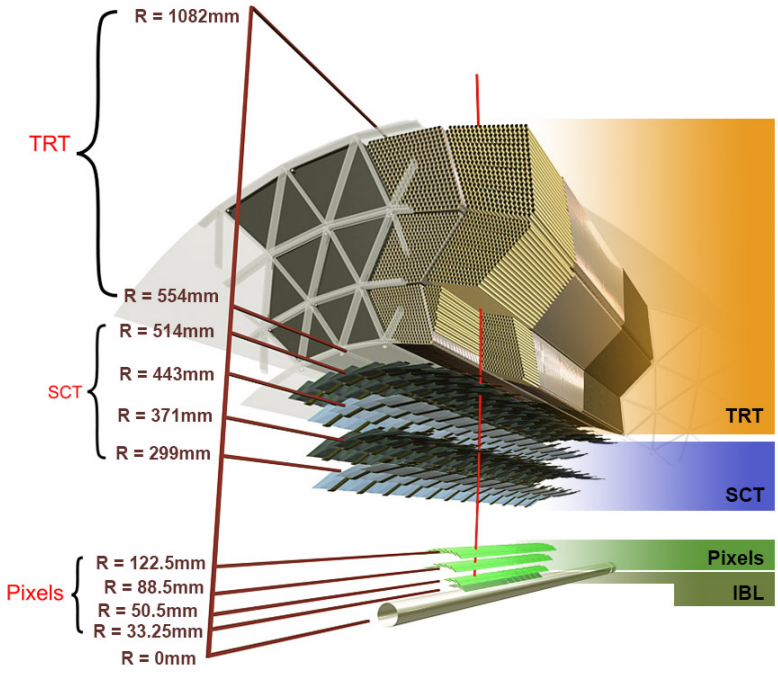
The coordinate system used for ATLAS is a set of right-handed Cartesian coordinates, with the origin at the centre of the detector; the x-axis points towards the centre of the LHC, y points upwards, and z is along the direction of the proton beams.
Often used are polar coordinates: θ, ϕ and r. ϕ is the azimuthal angle, i.e. the angle from the x axis measured around z, the polar angle θ is the angle from z and r is the radial distance from the z-axis.
Commonly used derived quantities are the pseudorapidity coordinate $\eta = -\ln\tan(\theta/2)$ [1] and the relation ΔR: the distance in pseudorapidity-azimuthal angle space, defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.
The ATLAS detector is not static at the μm scale, because of deformation due to gravity, strong magnetic fields and even displacement of the cavern floor. Hence its global reference system is chosen to be essentially co-moving with the Inner Detector. This reference frame has been determined in the pre-data taking stage and adapted to assure the z direction to be parallel to the average beam direction.
In the reference frame the position of the beam's axis varies, and is therefore recalculated throughout each run, from sets of a few thousand collisions. This information is encoded as a Beam Spot (BS) position through a maximum likelihood fit to the ellipsoidal distribution of the highest Σp_T primary vertex (see 4.2.2) of each event [111]. The design size of the BS is of 11.8 μm in the transverse direction ($x - y$) and 4.5 cm in the longitudinal direction (z). Due to the BS's elongation, calculations often use just the *beam line* , i.e. the parallel to z passing through the center of the BS.

2.4 The Inner Detector

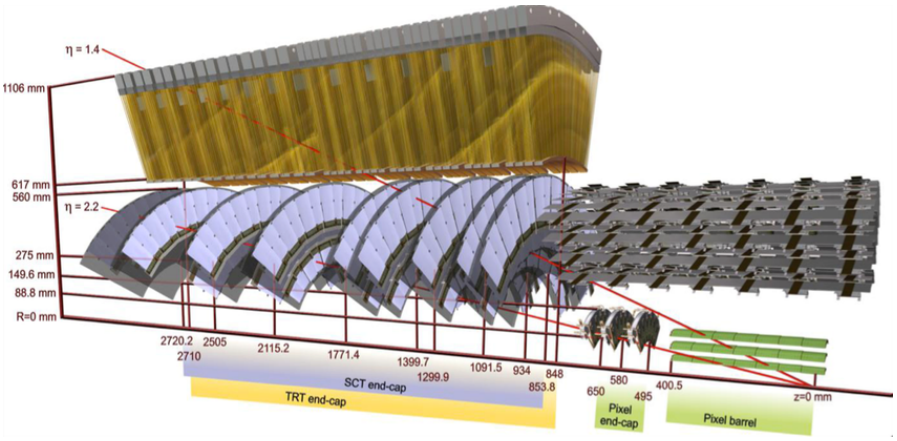
The Inner Detector (ID) reconstructs the path of charged particles, which allows to both determine the position from which they originate and their momentum. Several layers of active material ionize (or free electrons) at the passage of charged particles, while these propagate away from the interaction point. The ID covers a volume of about 25 m^3 and is immersed in a 2 T solenoidal magnetic field generated by a superconducting magnet. As the path of the particles is curved by the Lorentz force in the magnetic field, their momentum can be estimated from the curvature of their tracks (4.1.2).
The ID has a high granularity, and extends as close as possible to the interaction point, to allow for accurate identification of the vertices from which the particles originate. A low material density is essential to maintain the tracks unaltered, preventing sudden

[1] The use of pseudorapidity is justified by the number of particles produced by the bunch crossings being roughly constant as a function of η, as well as its closeness to the relativistic rapidity, which means that $\Delta\eta$ is conserved under Lorentz boosts along the z direction.

energy loss or emission of energetic secondary particles. The detector is composed by four separate parts, shown in Figure 2.4, which are (ordered from closest to the beam pipe to furthest away): Insertable B-Layer (IBL), Pixel Detector, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT) [112] [113].



(a)



(b)

Figure 2.4: A 3D view of he ATLAS Inner Detector. All the subdetectors are shown for the barrel (a) and end-cap (b) region.

2.4.1 Insertable B-Layer

The IBL was added to the ID in May 2014, during the Long Shutdown following the completion of LHC's Run-1. It constitutes an extra silicon Pixel B-Layer placed in the $\sim$ 1 cm wide free space between the pre-existing Pixel B-Layer and the Beam Pipe. To make space for it the section of beam pipe crossing the detector was replaced with a 4 mm thinner one. The IBL's modules are mounted on 14 *staves*, that are radially inclined such as to overlap by a few degrees in the ϕ direction, though due to space constraints its modules do not have overlap in η. The pixels are located at an average radius of 33.25 mm and are 50x250 μm in surface (ϕ-z plane), for an intrinsic resolution of 8x40 μm [114] [115]. In the period 2015-2018 approximately 99.3% of the IBL channels were operational during physics data taking [116].

2.4.2 Pixel Detector

The high granularity Pixel Detector consists of modules with 50x400 μm pixels, distributed in 3 layers with cylindrical symmetry in the barrel region and 3 layers of disks in the end-cap region. Modules are disposed such that there are no gaps, thus leading to 3 readings for virtually any charged particle travelling through the detector. The Pixel's intrinsic resolution is 10 μm in the transverse coordinates (ϕ-R in barrel, ϕ-z in end-caps) and 115 μm for the longitudinal direction (z and R respectively). In 2015-2018 $\sim$94.8% of the detector's channels were operational.

2.4.3 Semiconductor Tracker

The SCT comprises of 4 layers of silicon sensors in the barrel region and 9 more disks in the end-cap region. Its modules are composed of narrow micro-strips, with a width of 80 μm in the barrel and an average width of 80 μm or 60 μm (depending on the radius), in the end-cap. The length of the individual strips varies slightly depending on the position in the detector, but the uninterrupted active length of each module is about 50 mm. On each module the sensors are glued back to back at an angle of 40 mrad, to enable an intrinsic resolution of 17 μm in the transverse coordinates, 580 μm in the longitudinal one. Its design makes it almost hermetic, such that charged tracks will cross at least four modules in the detector [117]. In 2015-2018 $\sim$98.6% of SCT channels were operational.

2.4.4 Transition Radiation Tracker

The TRT constitutes the outermost structure of the ID. It combines straw tubes with Transition Radiation (TR) inducing medium to provide between 22 and 36 readings per track and give separation power between pions and electrons through the higher intensity of TR given by electrons in the polypropylene radiator. The straws are oriented axially in the barrel region 56 cm $< R <$ 107 cm, and radially in the end-caps in the region 64 cm $< R <$ 103 cm (48 cm $< R <$ 103 cm outside z of 280 cm). The detector's design meets the need for a large volume coverage with contained cost and limited radiation

length (10% X_0)[2], though it is prone to have high occupancy as the experiment is run beyond the initial design luminosity of 10^{34} cm^{-2} s^{-1}. In the period 2015-2018 ∼97.2% of TRT channels were operational.

2.5 The Calorimeters

ATLAS is equipped with an Electromagnetic Calorimeter (EMCal) and a Hadronic Calorimeter (HCal), to absorb and reconstruct the energy of electrons, photons and neutral pions, or hadronic jets (section 4.4) respectively. The calorimeters are shown in Figure 2.5 and are described separately in 2.5.1 (EMCal) and 2.5.2 (HCal).
For the $\tau \to 3\mu$ analysis the calorimeters are relevant for three reasons: the reconstruction of jets (recoiling from the muon triplet or denoting hadronic activity, typical of background events), the calculation of E_T^{miss} (section 4.5) and the stopping of the majority of non-muonic collision products (together with the calorimeter's external support structure), which allows the identification of muons as one of the few particles that reach the MS.

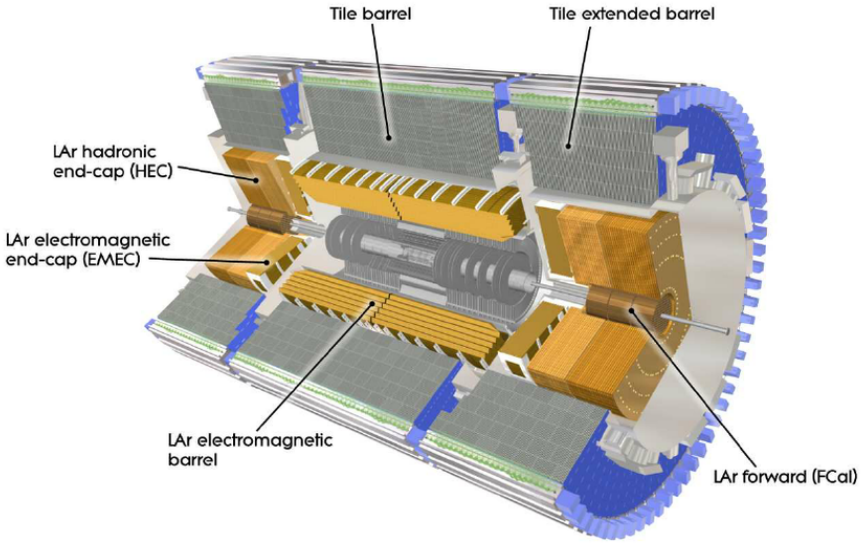


Figure 2.5: A view of the Electromagnetic and Hadronic Calorimeters of ATLAS.

[2] X_0 corresponds to one radiation length: the length scale for high-energy electromagnetic cascades. It corresponds to the thickness of material in which a high-energy electron loses 1- 1/e of its energy through bremsstrahlung, as well as 7/9 of the mean free path of a high-energy photon before it pair-produces an electron-positron pair [118].

2.5.1 Electromagnetic Calorimeter

The EMCal [119] is composed of a barrel section covering $|\eta| < 1.475$ and two end-caps extending in $1.375 < |\eta| < 3.2$. It uses liquid argon (LAr) as active material, Kapton electrodes and lead absorber plates. The lead plates have a distinctive accordion shape, which allows a crack-less coverage in the azimuthal direction. During physics data taking in the years 2015-2018 the EMCal was 100% operational.
Figure 2.6a shows the structure of the EMCal up to $|\eta| < 2.5$:

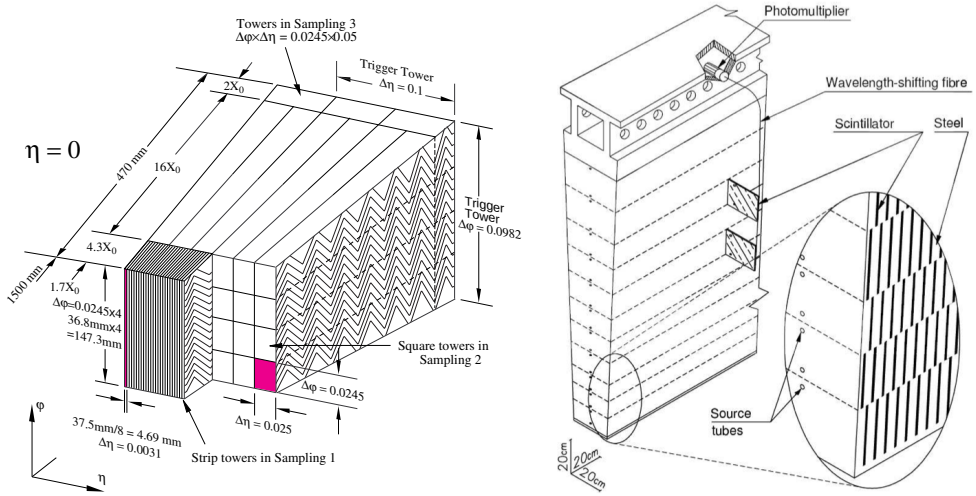
- The first (innermost) sampling layer is composed by 4 mm narrow strip cells which act as a *preshower* for precise η measurement and separation of $\pi^0(2\gamma)$ from prompt γs and charged pions from electrons.
- The second sampling layer contains the bulk of the calorimeter's material and is subdivided in cells with a $\Delta\eta \times \Delta\phi$ section of 0.025×0.025.
- The third layer has a coarser granularity (0.05×0.05) and its thickness varies strongly with η, as it fills the remaining space in the regular shapes of barrel and end-cap. Above $\eta = 2.5$ the EMCal is not used for more than reconstructing jets or E_T^{miss}, hence it is composed by only 2 layers, with coarser granularity (0.1×0.1).

To account for the energy loss of particles before they reach the calorimeter, a *presampler*, formed by a thin layer of LAr and a slab of scintillator, is added before the bulk of the EMCal's cryostats up to $|\eta| < 1.8$. In total the radiation thickness of the EMCal is > 24 X_0 (> 26 X_0) in barrel (end-cap) while the material preceding it has a thickness of 2-4 X_0 (~ 2 X_0) in the barrel (end-cap), with a local maximum of ~ 7 X_0 at the transition from barrel to end-cap.

2.5.2 Hadronic Calorimeter

The HCal consists of three parts: a Tile Calorimeter divided a barrel section ($|\eta| < 1.0$) and two extended-barrel sections ($0.8 < |\eta| < 1.7$), two hadronic LAr end-cap calorimeters ($1.5 < |\eta| < 3.2$) and two forward calorimeters ($3.1 < |\eta| < 4.9$), also using LAr as active material. The HCal had overall $\sim$99.5% of operational channels in the Run-2 periods used in the present analysis.

The Tile Calorimeter [120] uses 14 mm iron plates as absorber and scintillator tiles as active material as shown in Figure 2.6b. The tiles are oriented on the $x - y$ plane and stacked radially in a staggered pattern. The scintillators are read out by photomultipliers via fibers that are grouped to obtain a "pseudo-projective" cell structure, i.e. towers of tiles pointing to the centre of the detector (not unlike the cells of the EMCalo). The resulting granularity is of 0.1×0.1 ($\Delta\eta \times \Delta\phi$) for the whole calorimeter except for the outermost $\sim$45 cm thick layer, where the "cell size" is 0.2×0.1. It is interesting to note that the Tile Calorimeter is operated at 15 °C ("hot" if compared to the LAr



(a) (b)

Figure 2.6: View of a section of the EMCal barrel (a) and of the hadronic Tile Calorimeter (b).

calorimeters) and it acts as a return yoke for the ID's solenoidal magnetic field.

The hadronic LAr end-cap calorimeters are composed of two wheels with, respectively, 25 mm and 50 mm thick copper plates as absorber. The 8.5 mm gaps contain the sensitive liquid argon and three electrodes that are operated as electrostatic transformers [121], a technique that permits the use of a contained high-voltage and to limit ion build-up in the gaps. The read-out combines chains of modules in a similar way as for the Tile Calorimeter, defining "pseudo-projective" cells of size 0.1×0.1 up to $|\eta| = 2.5$ and 0.2×0.2 beyond.

The Forward Calorimeters are made up of three segments, each of which has a structural matrix holding in place axially oriented grounded tubes and inside these positively charged rods which are used for read-out. In the innermost segment uses copper structures as the absorption material, while the other two use tungsten. The sensitive material is LAr, permeating the 250 μm (375 μm) narrow gap between tube and rod in the innermost (two outer) segment.

2.6 The Muon Spectrometer

The Muon Spectrometer (MS) is the outermost and by far the most voluminous subdetector of ATLAS; it ultimately determines the detector's height of 25 m and length of 44 m. Its goal is to measure direction, momentum and charge of muons by bending their trajectories in an extensive magnetic field.
At relativistic energies muons are minimally ionizing, this means that they can traverse

far more material than other particles, without being stopped. Hence muons can be identified by their tracks in the MS, provided that to a good approximation all other detectable particles are absorbed before reaching the MS. Figure 2.7 shows the material crossed by particle before reaching the MS as a function of pseudorapidity; over the whole spectrum the material amounts to more than 10 hadronic interaction lengths [3]. This leads to a misidentification rate in the order of one percent of hadronic tracks as muons in the MS; this number is further reduced to about 10^{-5} building combined muons (see section 4.3.1) and applying isolation criteria to reject muons originating within energetic jets, for which the change of leading to punch-through is the highest [122] [123].

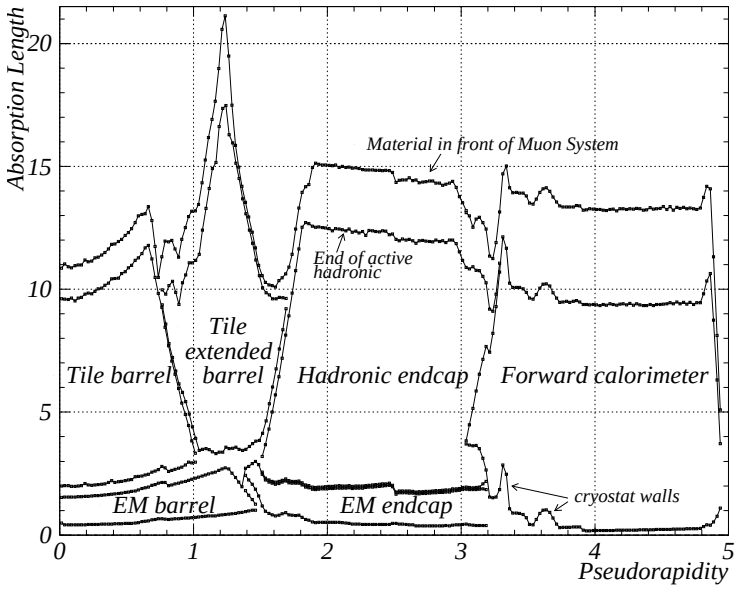


Figure 2.7: The amount of material (expressed in hadronic interaction lengths) traversed by particles before reaching the MS as a function of pseudorapidity. Image taken from [119].

As can be seen on Figure 2.8, the MS consists of various *stations* (layers) of drift chambers placed throughout three systems of 8 superconducting coils which generate a toroidal magnetic field.
The largest air-core magnet surrounds the barrel region and has a length of 25.3 m, corresponding to $|\eta| < 1.4$. Its central field is of approximately 0.5 T. The remaining two end-cap toroid magnets are inserted into the barrel toroid at each side, cover the range $1.0 < |\eta| < 2.7$ and have a central field of about 1 T. While the barrel magnet has each coil cooled separately, the end-cap toroids are wholly encased in a cryostat, hence the end-cap stations are located before and after the bending has occurred. The

[3] The hadronic interaction length is the length scale for high-energy hadronic cascades, governed by the strong force. It is based on the mean free path of relativistic protons in a material, which can be approximated as $\lambda \simeq \frac{A^{1/3}}{\rho} 35 g/cm^2$, where A is the material's mass number and ρ its density [118].

resulting magnetic field is approximately cylindrical, following the symmetry of ATLAS itself, this makes it mostly orthogonal to the trajectory of the outgoing muons, giving it maximum bending power.
In the remaining part of this section the MS's detector chambers are described, subdivided in Precision Chambers and Trigger Chambers.

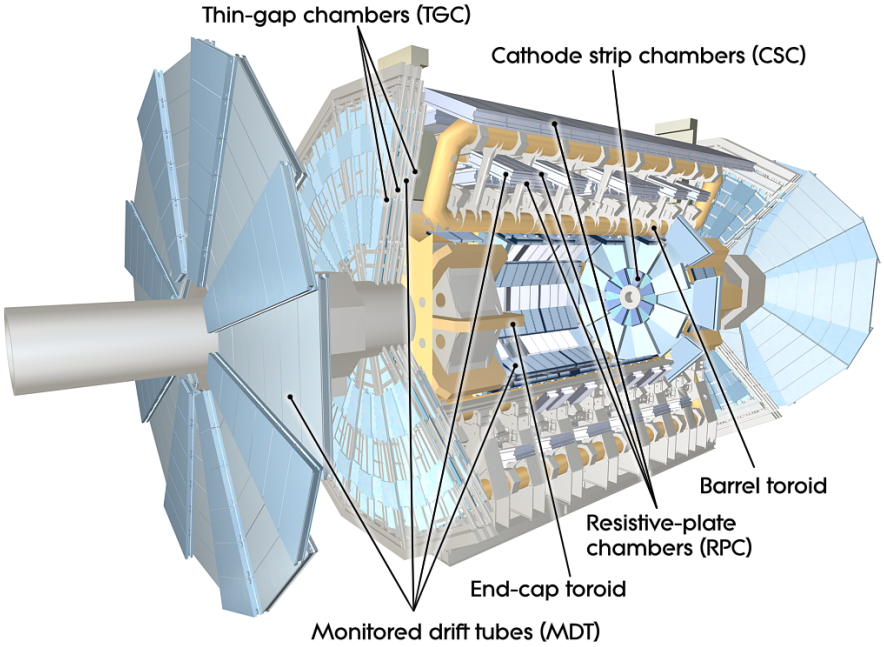


Figure 2.8: The ATLAS Muon Spectrometer, figure from [124].

2.6.1 Precision Chambers

The bulk of the MS consists of Monitored Drift Tube (MDT) stations. These are mounted in three layers (Inner, Middle and Outer) both for the barrel region ($|\eta| < 1.0$) as for the end-caps ($1.0 < |\eta| < 2.7$). In the barrel stations are organized in 16 $\phi - z$ planes with radius dependent on the azimuthal position, to make space for the magnet's coils. In the end-cap the stations are mounted on three wheels of which the inner one (small wheel) is placed before the end-cap magnet. Cathode Strip Chambers (CSC) are used instead of MDTs on the inner radius of the small wheel, because of their ability to operate at a very high hit rate.

MDTs are aluminium tubes with a 30 mm diameter put at a high voltage and filled with a mixture of argon and carbon-dioxide. Each MDT contains a tungsten anode wire which reads out the charge of electrons released at the ionization of the gas (93% Ar, 7% CO_2) by traversing muons. The drift time of the electrons is used to determine the position at which the muons crossed, with a resolution of about 80 μm (radial po-

sition with respect to each MDT's axis). In each station MDTs are stacked in three to four layers per side, such that the position is known with a 35 μm resolution in the z (r) direction in the barrel (end-cap). The MDTs are oriented in the direction of ϕ (approximately parallel to the magnetic field), to give a precise handle on the curvature of the muons. The muon position along the length of the tubes (ϕ direction) cannot be determined from the MDTs. The MDTs had ∼99.7% of operational channels in the period 2015-2018.

CSCs are multi-wire proportional chambers, they feature central anode wires with a pitch of 2.54 mm surrounded by two orthogonal layers of cathode strips. They are compatible with a hit rate of 1 kHz/cm^2 and have a spatial resolution of 40 μm × 5 mm. The higher spatial resolution corresponds to the bending direction, while in the orthogonal direction the strip segmentation is coarser. The narrow drift space between the planes allows a 7 ns time resolution, and, together with the specially selected choice of gas mixture (80% argon, 20% carbon-dioxide), ensures a low sensitivity to neutrons (through the absence of hydrogen). The CSCs had ∼96.5% of operational channels in the period 2015-2018.

2.6.2 Trigger Chambers

Trigger chambers are obviously important for the fast identification of muons in the triggering step (3.4), but they also integrate the information provided by the MDTs with accurate timing information and resolution to the muon's ϕ coordinate. The trigger chambers are of two types: Resistive Plate Chambers (RPCs) mounted on the barrel MDTs in $|\eta|<$ 1.05 and Thin Gap Chambers (TGCs) surrounding the middle end-cap wheel over $1.05 < |\eta| < 2.4$.

RPCs are mounted around the middle layer and outside the outer layer of barrel MDTs, for a total of three layers. The core structure of RPCs are pairs of resistive bakelite plates, separated by a 2 mm layer of an admixture of gases with a 4.5 kV/mm electric field (the gas admixture consists of 94.7% $C_2H_2F_4$, 5% Iso-C_4H_{10} and 0.3% SF_6). At the crossing of a muon an avalanche is triggered in the gas; this leads to short-lived localized charges on the plates, which are read out through capacitive coupling of copper strips on their outer face. The copper strips on the outside of a resistive plate pair are oriented perpendicularly to each other to attain a 1 cm resolution in both directions and a time resolution in the order of 1 ns. Each station of RPC contains two pairs of plates and two layers of read-out strips in each orientation. The RPCs had ∼97% of operational channels in the period 2015-2018.

TGCs are multi-wire proportional chambers similar to the CSCs. The narrow gap between its cathode planes permits a time resolution of 4 ns and spatial resolution of 2-7 mm. Thin gap chambers have the peculiarity of operating in saturated mode: their response has little dependence on the amount of charge deposited by particles and consequently their angle of impact. The gas used in the TGCs is 55% CO_2 and 45%

n-pentane (straight chains of C_5H_{12}). The TGCs had ~99.5% of operational channels in the period 2015-2018.

2.7 The Trigger System

While ATLAS sees a bunch crossing rate of 40 MHz, data recording and processing is not possible for more than about 1000 events per second in Run-2. Fortunately only a small fraction of the collisions exhibit particle physics processes that are studied at the LHC. The ATLAS Trigger System (TS) is responsible of quickly identifying and selecting the few events of interest to be written to disk, thus meeting the current limitation in data transmission and recording speeds. The TS has undergone a substantial evolution from its original design applied in Run-1 [125] [126], to account for the increase in interaction rates and in the allowed throughput (from 100 Hz to 1kHz) [127]. As shown in Figure 2.9, the TS in Run-2 consists of two main blocks: a Level-1 (L1) trigger and a High Level Trigger (HLT) [128].

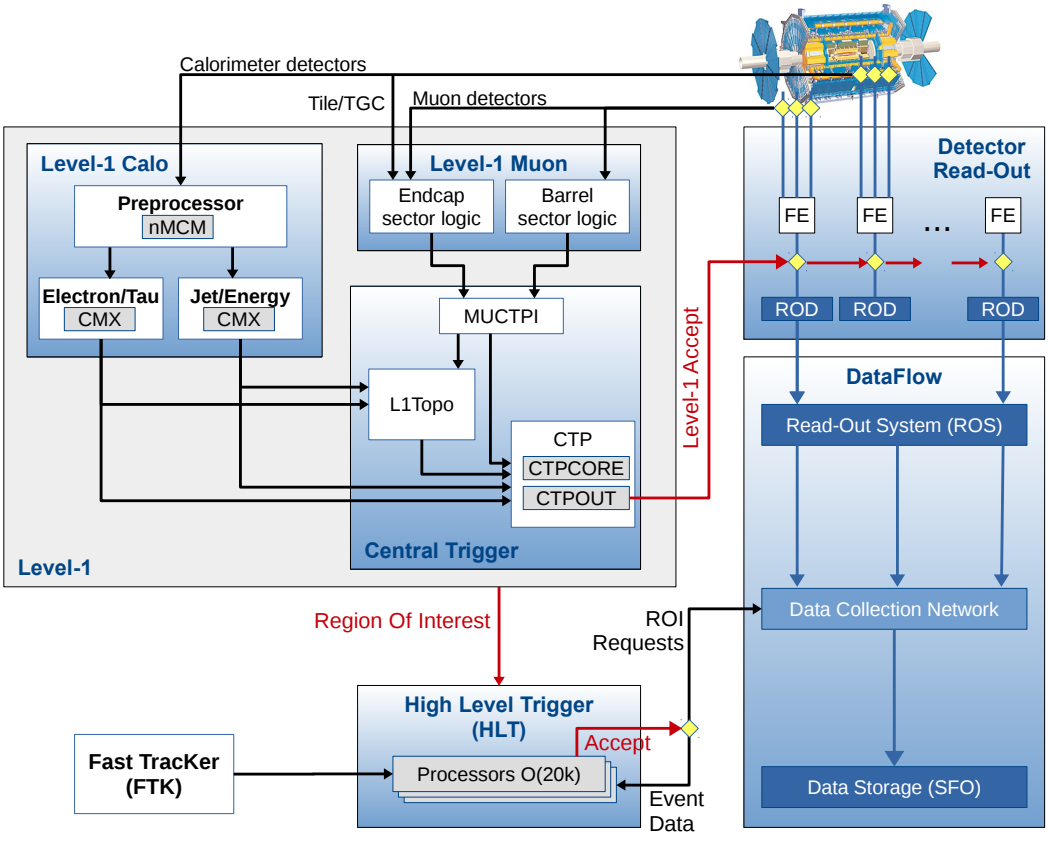


Figure 2.9: An illustration of the block-logic of the ATLAS trigger system in Run-2.

2.7.1 L1 Trigger

The L1 trigger is hardware-based and its aim is to reduce the event rate to about 100 kHz in a decision time of about 2.5 μs. It relies on reduced granularity information from the calorimeters to select on energetic photons, electrons, jets and E_T^{miss}. It also uses MS Trigger Chambers signals to find muons.
A fundamental change with respect to Run-1 is the introduction of the L1 Topological trigger processor (L1Topo), which produces kinematic information about the calorimeter and muon trigger objects already at L1. This is used, amongst other purposes, to attain angular information and invariant mass estimates which permit to keep sensitivity to states with low energy such as $\tau \to 3\mu$, in spite of the increasing luminosity. If an event is accepted by any of the L1 triggers, its information is read out and temporarily stored in the computers of the Read-Out System (ROS), while information on the triggered event is sent to the HLT in the form of Regions of Interest (RoI).

2.7.2 HLT

The HLT runs on a computer farm where multiple events are processed in parallel. Its purpose is to reduce the rate of passing events to an average 1 kHz with an average decision time of about 200 ms (individual events can take more than 5s). The HLT is based on full-granularity information either of the full detector or part of it, as determined by the RoIs received from L1. The trigger objects are reconstructed with software as close as possible to the offline reconstruction. The HLT unifies in a single system what in Run-1 was split in Level-2 trigger and Event Filter; a change that reduced the trigger's complexity as well as the required per-event bandwidth as it avoids the duplication of data transferring. Figure 2.9 shows as well the Fast Tracker, an extremely promising system to generate full tracking information in a L1 time scale without being affected by increases in pileup. This will give extra information and processing time to the HLT; unfortunately it is only being introduced starting in 2018 [129].

Chapter 3

Concept of the $\tau \to 3\mu$ Analysis

> *"The size of a proton is about 10^{-15} m. So the ratio of the height of Mount Everest to the size of a proton is only a thousand times smaller than the ratio of the observable part of the Universe to Everest. Notice that the whole Universe can have a finite size. What occupies the space outside the boundaries of the observable part of the Universe, if anything, we do not know."*
>
> — Prof. Olya Igonkina, *Inaugural Speech*

This thesis presents a search for $\tau \to 3\mu$ at the ATLAS experiment; its goal is to either discover the so far unseen decay or set a limit on the branching fraction $\mathcal{B}(\tau \to 3\mu)$. While the motivation and current best limit are presented in Chapter 1 (specifically in sections 1.4 and 1.5), this chapter gives an overview of the basics and peculiarities of the $\tau \to 3\mu$ search, focussing upon the τ production in p-p collisions and the distinctive characteristics of ATLAS that make it viable to target this specific final state.
The chapter begins with a brief outline of the analysis itself (3.1), followed by a description of the available τ production channels at ATLAS and a justification of the choice for the two channels used in the analysis (3.2). The third section focuses on the particular problems that were recognized during the Run-1 analysis and the solutions adopted in Run-2 (3.3). The last section handles the devised triggering strategy (3.4).

3.1 Analysis Structure

Most searches for LFV τ decays have been performed at "B factories", colliding electrons and positrons at a Centre of Mass (CM) energy corresponding to the Υ resonance (~ 10.6 GeV). These take advantage of the high number of $\tau^+\tau^-$ produced (just under 1M per fb^{-1} of luminosity [130]) and the clarity with which the final decay products can be observed.
At the LHC there is a number of ways in which τ leptons are produced, both as product of QCD interactions and EW processes. Two channels are most suited for ATLAS: the EW $W \to \tau\nu_\tau$ channel (referred to as "W-channel") and the QCD production of D_s

mesons and their successive decay ($D_s \to \tau\nu_\tau$), as well as decay of a b quark to τX (X denoting further decay products), further referred to as "HF-channel" (HF stands for Heavy Flavour: a reference to the charm content of D_s, as well as the decay of b quarks). Why precisely these channels are the best is the argument of the next section. Of these two channels, about 1.2×10^9 τ were produced via the W-channel during Run-2 within ATLAS acceptance (about 40% of produced W are within detector acceptance), while the HF-channel produced about a factor thousand more before applying any p_T restriction on the τ[1].

Amongst the many possibilities to search for CLFV in tau decays, the analysis presented here investigates decays of tau into three muons, capitalizing on the strong muon reconstruction and identification capabilities of the ATLAS detector.

ATLAS can reconstruct muons with transverse energies as low as 2.5 GeV and, for tracks with $|\eta| < 2.5$, the ID permits to determine whether such muons originate from a common vertex.

During Run-1 an analysis was performed focussing exclusively on the W-channel, which is less plagued by backgrounds, as a proof of concept [64]. The Run-2 $\tau \to 3\mu$ analysis aims at combining the sensitivity of the two channels in an effort of reaching sensitivity to $\mathcal{B}(\tau \to 3\mu)$ of the order of few times 10^{-8}, comparable with the current best limit [63].

The analysis is constructed around the selection of events which contain a system of three muons consistent with a $\tau \to 3\mu$ decay (the event characteristics are extracted from MC). The selections follow distinct but similar strategies for W-channel and HF-channel. In both cases the reconstruction of the common vertex of origin of the three muons is paramount. For the W-channel, selections focus on the momentum boost of the τ, its relative orientation with respect to the neutrino (reconstructed as E_T^{miss}) and the compatibility with a W decay of the transverse mass of τ and E_T^{miss}. The absence of intense hadronic activity, typical for W production is used to reject the bulk of combinatoric muon triplets coming from hadronic jets. In the case of the HF-channel a tighter selection is applied to the displacement of the three-muon vertex from the beam line, taking advantage of the fact that both D_s and τ have significant decay length.

It must be said at this point that muon reconstruction efficiency only reaches a plateau of about 90% around $p_T \simeq 5$ GeV (see Figure 4.14a), so it is by no means possible to reconstruct every $\tau \to 3\mu$ candidate, and in the case of the HF-channel, where no significant boost is associated with the initial decay, only a tail of the total produced p_T spectrum of τ can be used for the analysis (compare Figures 5.1a and 3.1a). Numerically this means that only about three times more τ decaying to three muons are expected in the HF-channel compared to the W one [132]. For both channels, the most sensitive variables are fed to a multivariate classifier, BDT, after the cut-based selections, to attain the maximal possible separation of signal-like events from backgrounds. For the training, data in Side-Band regions, described here below, is used to characterize the backgrounds.

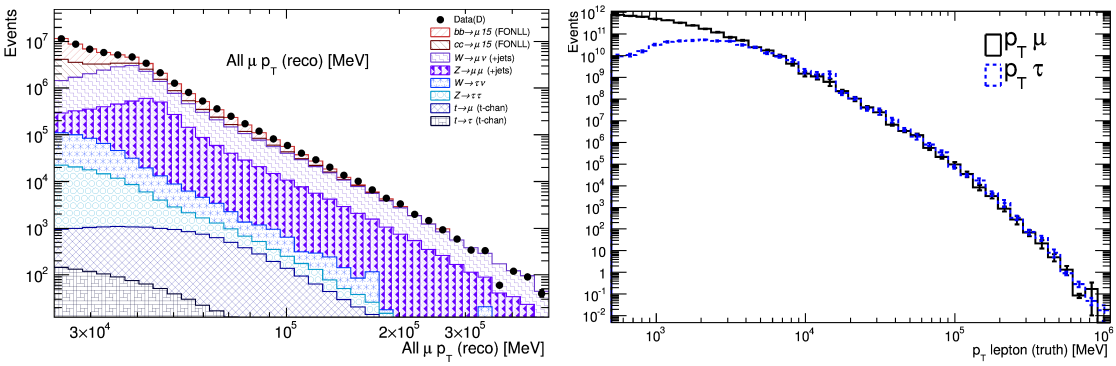
For both channels the final discriminant is M^{triplet}: a Signal Region (SR) is defined as a narrow window around M_τ and events in it are blinded. Around the SR window, two

[1] Estimate based on the $\sqrt{s} = 8$ TeV cross sections reported in [131]

Side-Bands (SB) are chosen and from these the expected number of background events in the SR is extrapolated. At this point the the cut value to be used on the multivariate classifier response is picked to maximize the sensitivity to signal events in the SR, assuming only background events are found (minimal correlation between the classifier and μ is verified). The backgrounds in question are mainly given by hadronic jets producing two or more muons in the final state combined with π/K mesons decaying in flight. It is though very rare for such decays to pass the selections, especially because of an explicit rejection of events with di-muon mass close to one of the known resonances. This makes it computationally impossible to fully characterize the observed backgrounds through simulation. Consequently background simulation is employed only to show that the selections efficiently reject the few known backgrounds.

3.2 τ Production Channels

Proton-proton collisions at the LHC produce a great variety of states, and a number of these contain τ leptons. To be sensitive to rare decays at the level of branching fractions of the order of 10^{-8}, a τ sample of order 10^8 and higher is needed (10^{-8} being the order of the current best limit set on $\tau \to 3\mu$ [63]).
At $\sqrt{s} = 8$ TeV p-p collisions, $W \to \tau\nu_\tau$ has a cross section of about 4.75 nb (after acceptance), which means that in the 20.3 fb^{-1} of Run-1 data collected in 2012 there were $\sim 10^8$ $W \to \tau\nu_\tau$ originated τ leptons, justifying a first search.
For Run-2, an overview of the channels generating τ leptons can be seen in Figure 3.1a, as a function of generated τ p_T. The figure actually refers to μ production, but the two are virtually identical for $p_T > 5$ GeV as can be seen from Figure 3.1b (note that correspondence in the hadronic channels is no trivial consequence of lepton universality, but rather a fortuitous balancing out between kinematic suppression of τ production and helicity suppression of μ production). While the W channel is definitely relevant as a source, for p_T of a few tens of GeV, almost the same amount of high momentum leptons are found in cc and bb originated di-jet events. The QED process $Z \to \tau\tau$ is also a non-negligible τ source, but it is clear from the figure that this produces significantly fewer τ leptons, compared to W, and would thus not lead to comparable sensitivity. At lower energies than the ones displayed on Figure 3.1a, the di-jet (QCD) sources become by far dominating over EW ones. During Run-2 the production of τ through W boson decay has increased by almost a factor two with respect to Run-1, as a consequence of the increase of the $W^\pm$ production cross-section with the collision CM energy (this can be seen from the theoretical prediction in Figure 3.2). This variation in the W cross-section has as well been measured by ATLAS: [134] [135]. Hence with 140 fb^{-1} of 13 TeV data (full Run-2), the sample of $W \to \tau\nu_\tau$ events is approximately of 1.2×10^9 (140 fb$^{-1}\times$ 8 nb, after acceptance).
For the HF-channel, τ are predominantly produced from D_s decays (in spite of $\mathcal{B}(D_s \to \tau\nu_\tau)$ being only about 5.6% [136]); the fiducial production cross section of τ within ATLAS is of $\mathcal{O}(10^{10})$ fb (see details in Appendix A). A comprehensive list of major QCD τ production channels for $\sqrt{s} = 7$ TeV are listed in Table 3.1, and their



(a) (b)

Figure 3.1: Muon production spectrum at $\sqrt{s} = 8$ TeV as a function of the lepton's p_T (a) and comparison of τ and μ production in simulated inclusive di-jet events, as a function of p_T (b). (a) shows the distribution of muons produced at ATLAS in a fraction of Run-1 (3.5 fb^{-1}), overlayed with MC of the various sources. The simulation for inclusive di-jet events in Figure (a) has a generator filter at 15 GeV for the muon p_T. Data events are selected with a single muon 24 GeV trigger (range beginning at 25 GeV). The distribution is indicative for τ production, as a consequence of lepton universality (the image can be taken as a measure for the relative contributions in Run-2 as well). The contribution of of $W \to \tau\nu$ or $Z \to \tau\tau$ to μ production (visible in (a)) constitutes a difference between the two generations, but does not have a big impact. The differences in the QCD channels between τ and μ are evaluated with the use of Pythia8 MC (b), clearly differences in hadronic production cross section are only relevant for $p_T < 5$ GeV.

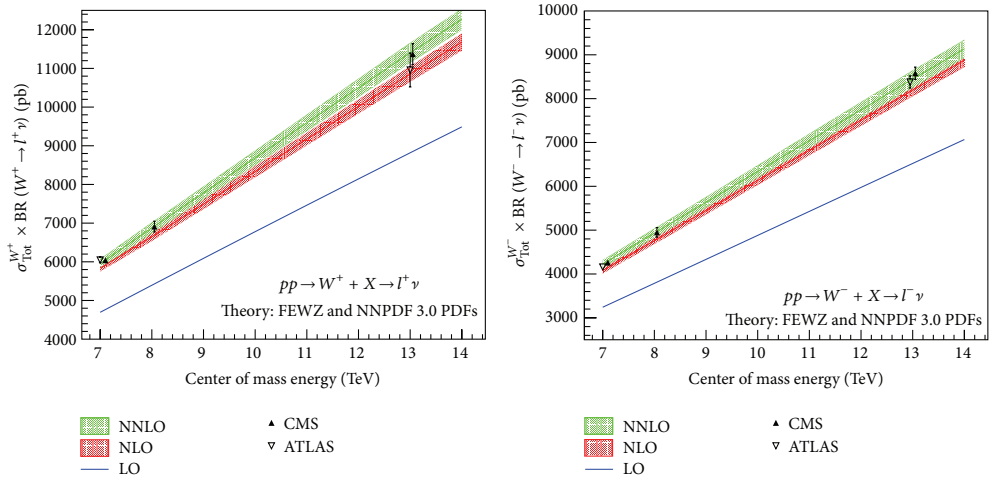


Figure 3.2: Theoretical prediction of the total production cross-section of W^+ and W^- times $\mathcal{B}(W \to \ell\nu_\ell)$ in p-p collisions, as a function of CM energy. Overlaid are the corresponding (preliminary) measurements at ATLAS and CMS. Image taken from [133]

relative cross sections are substantially unaltered at $\sqrt{s} = 13$ TeV. It is clear from it that the main contribution to τ production is prompt and non-prompt D_s production (first and third entry), immediately followed by direct production in b decay (second entry). These are the channels considered in this analysis. Considering $\sim$140 fb^{-1} of luminosity collected in Run-2 the HF channel counts for a few times 10^{12} τ produced within ATLAS acceptance (considering only the three major QCD channels and the cross section estimate of Appendix A). This magnitude makes it definitely interesting to pursue an analysis targeting this channel.

Channel	Cross-Section [μb]
$pp \to D_s \to \tau\nu$	10.7 $\pm$ 1.8
$pp \to bb \to \tau X$	2.4 $\pm$ 0.4
$pp \to bb \to D_s X \to \tau\nu X$	1.3 $\pm$ 0.4
$pp \to D^{\pm} \to \tau\nu$	0.7 $\pm$ 0.1
$pp \to bb \to D^{\pm} X \to \tau\nu X$	0.023 $\pm$ 0.004
$pp \to$ charmonium/bottomonium $\to \tau X$	<0.005

Table 3.1: QCD production cross sections for τ within LHCb acceptance ($2 < |\eta| < 4.5$) for $\sqrt{s} = 7$ TeV p-p collisions. The relative magnitude of these cross sections is indicative of the relative τ production channels at ATLAS at $\sqrt{s} = 13$ TeV, because the various channels have comparable η dependence. The values are collected from [131].

3.3 Lessons Learned from the Run-1 Analysis

The Run-1 $\tau \to 3\mu$ analysis at ATLAS was performed on the W-channel [64]. It was based on 20.3 fb^{-1} of p-p data collected in 2012, at centre-of-mass energy of 8 TeV. This pilot analysis showed that such an analysis is possible with ATLAS. Furthermore it gave valuable insights into problematic aspects of ATLAS trigger and reconstruction software. On the reconstruction side it became evident that a considerable fraction of the signal events had only two of the generated muons reconstructed as such. For this reason, in the Run-1 analysis, events were used which had an incomplete muon candidate (a *muon trackparticle*). Figure 3.3 shows various evaluated possibilities for the choice of muon candidates: the impact of accepting muon trackparticles is seen by the increase of events with three or more muons reconstructed between the full dot distribution (CB muons only) and the hollow dots (muon trackparticles added). The use of the incomplete muons added $\sim$13% of signal muon triplet reconstruction efficiency, as shown in Figure 3.3 (20% after successive quality selections).
In preparation for Run-2, improvements in the muon reconstruction software achieved an initial increase in the muon triplet reconstruction analogous to the use of incomplete muons in Run-1, mainly due to a better reconstruction of low energetic muons (2 GeV $< p_T <$ 6 GeV), visible in Figure 3.4. On top of this, a targeted effort dealt with the penalization of candidate triplets by the muon reconstruction *overlap removal* step, already observed in Run-1, attaining an ulterior 26% signal reconstruction efficiency (20% after quality selections). This lead to an overall $\sim$40% increase in signal reconstruction

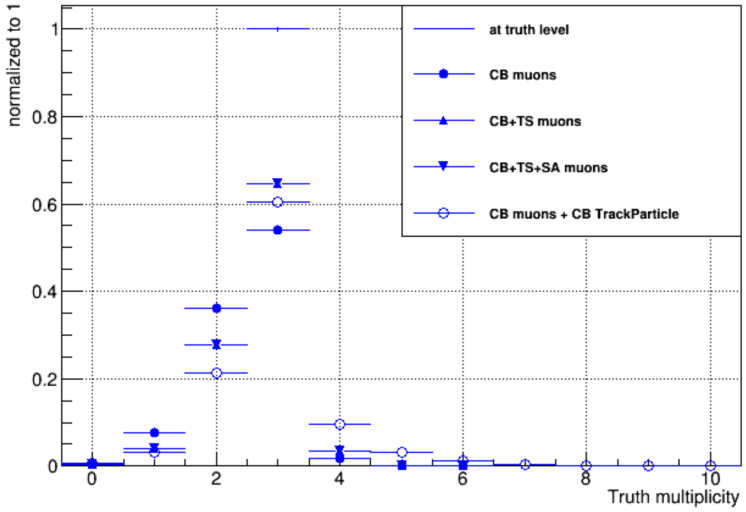


Figure 3.3: [132] Normalized distribution of $W \to \tau(3\mu)\nu$ events as a function of reconstructed muon multiplicity for Run-1. Events are subject to truth level selection requiring exactly 3 muons with $|\eta| < 2.5$ and $p_T > 2.5$ GeV. Full circles show the Run-1 performance of CB muon reconstruction, while the hollow circles show the effect of combining muon trackparticles and CB muons.

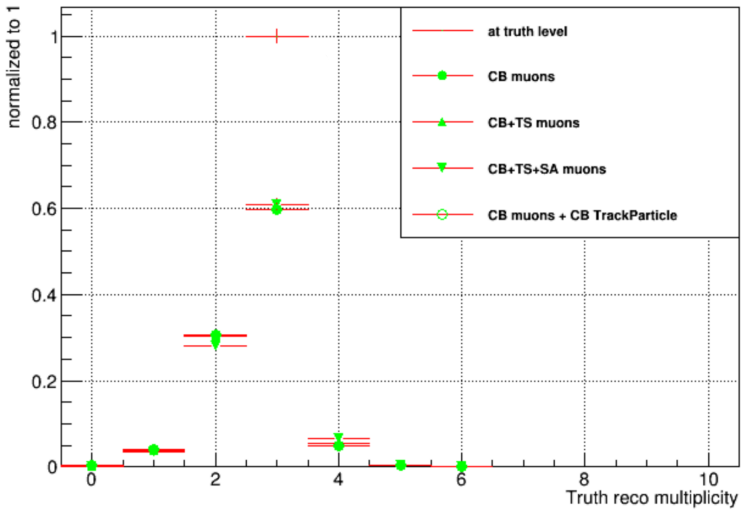
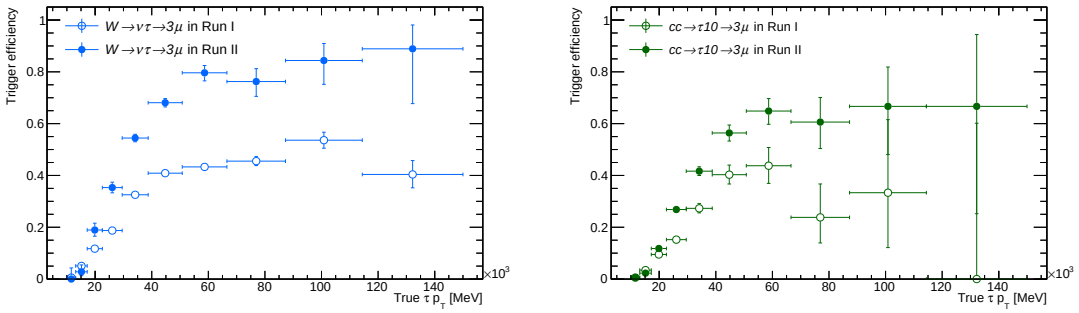


Figure 3.4: [132] Muon reconstruction efficiency in $W \to \tau(3\mu)\nu$ events as a function of reconstructed muon multiplicity for Run-2. Events are subject to truth level selection requiring exactly 3 muons with $|\eta| < 2.5$ and $p_T > 2.5$ GeV. Comparing to Figure 3.3 it shows the increased performance of CB muons. The efficiency does not yet include the improvement in overlap removal mentioned in the text.

efficiency between Run-1 and Run-2, using only CB muons (a full 76% of simulated $W \to \tau(3\mu)\nu$ events, with muons obeying $|\eta|<2.5$ and $p_T>2.5$ GeV, are reconstructed for Run-2). The mentioned problem with muon overlap-removal was due to a sub-optimal rejection criterion, so far overlooked because affecting only closely packed muons systems, such as in boosted $\tau \to 3\mu$ candidates, where a considerable fraction of the hits in the MS of the individual muons overlaps. The explanation of this correction is given in section 4.3.2, after introducing the ATLAS muon reconstruction (due to its slightly technical nature).

3.4 Event Triggering for Run-2

Inefficiencies, similar to the ones found in muon reconstruction, were identified within triggering software as well. Particularly difficult to control was the interplay of the three trigger levels (L1, L2 and HLT). Every stage of trigger would not only apply different track rejection, but had as well trouble matching the reconstructed muons in one level with the muon candidates in the preceding level, due to their narrow separation. Major effort was put in the development of triggers viable for Run-2 data taking, for the selection of $\tau \to 3\mu$ events. Figure 3.5 renders the impact of the new triggers on the acceptance of signal events: triggering efficiency increased by $\sim$65% for events in the W-channel and by $\sim$50% for HF-events.
The triggers used for the Run-1 analysis were seven multi-lepton and missing transverse energy triggers, which had not been optimized for the analysis. Of these one was a two-muon trigger with a 13 GeV threshold, one was an associated trigger requiring two muons above 8 GeV and estimated E_T^{miss} above 30 GeV, three required three muons, with 4 and 6 GeV thresholds and the last two were so-called full-scan triggers. The latter required a single $p_T>18$ GeV muon at L1, but then employed a higher-granularity reconstruction for the MS at high level to identify the presence of further muons in the event, and thus to effectively trigger on events with 2 or 3 muons.
For Run-2 dedicated triggers have been introduced. Inefficiencies due to overlap of the muons inside the MS have been addressed: some of the new triggers require fewer muons at L1, and more at HLT, similar to what happened for full-scan triggers; this implies the finding of more MS tracks only when using full MS information. The difference, with respect to full-scan, lies in the fact that these new triggers demand high granularity reconstruction only in a small η-ϕ cone surrounding the seed determined at L1. Other triggers are more straight forward and merely introduce specific kinematic selections at HLT, to reduce unwanted throughput. The new triggers, at L1, use a single-muon threshold of 11 GeV, a two-muon threshold of 6 GeV or a three-muon threshold of 4 GeV. Besides these "tailored" triggers a few basic two-muon and three-muon triggers are used as well. Further inefficiencies linked to the low momentum of the muons of interest and the matching of L1 ROIs (see 2.7) with HLT have been studied, and progressively addressed by the end of 2015. This allows for a better performance of any triggers based on two or three muons at L1.
A comprehensive list of the employed triggers is reported in section 5.2.



(a) (b)

Figure 3.5: Triggering efficiency of $\tau \to 3\mu$ events as a function of the true τ transverse momentum relative to the triggers employed in Run-1 (empty dots) and the ones designed for Run-2 (full dots). The figure on the left shows the trigger efficiency for W-channel $\tau \to 3\mu$ events (a) and to the right is the efficiency for prompt $D_s^{\pm} \to \tau(3\mu)\nu$ events (b). The latter is representative of the trigger performances on HF-channel events, the trigger efficiency spectrum of non-prompt candidates being substantially the same (and constituting only 25% of the HF-channel signal). The signal generated for figure (b) only contains candidates with a τ $p_T > 10$ GeV. The efficiencies are built from all triplet candidates that "fire" any of the Run-1 or Run-2 triggers respectively with the requirement that all signal true muons be within acceptance ($|\eta| < 2.5$ and $p_T > 2.5$ GeV). The increase in efficiency seen includes the small effect ($\sim$10%) coming from the improvement of the muon reconstruction efficiency between Run-1 and Run-2, referred to in section 3.3 (the change in overlap removal is not included). The source of the images is [132].

Chapter 4

Event Reconstruction

> *"–What is that, besides beautiful? –As I understand it–, Elphie said, –a sort of encyclopedia of things numinous. Magic; and of the spirit world; and of things seen and unseen; and of things once and future. I can only make out a line here or there."*
>
> — Gregory Maguire, *Wicked*

In this chapter the reconstruction of objects important for this analysis is described, along with the definition of several quantities that are fundamental for the distinction of collision events corresponding to a $\tau \to 3\mu$ decay from SM backgrounds.
Key quantities in determining the performance of the reconstruction are reconstruction efficiency and the fake rate. The first is the ratio of particles that are correctly reconstructed by the algorithms employed, with respect to the particles effectively present, while the latter is the relative fraction of reconstructions that do not represent a real particle. Because reconstruction is never perfect, the product of reconstruction is best referred to as *particle candidates*, and it is the role of analysis selections to handle the trade-off between keeping all events with genuine particles and reducing the contamination by mis-reconstructions.

4.1 Tracking in the Inner Detector

4.1.1 Track Reconstruction

For muons with a p_T in the range 5-15 GeV, like in the case of $\tau \to 3\mu$, the most accurate information about their momentum is provided by the InnerDetector (ID, see 2.4). The tracking procedure uses the information gathered by the ID to reconstruct the trajectory of charged particles, and determine their charge and momentum.
Tracking begins with the positions of clusters in the silicon detectors, referred to as *hits*, and organized on the base of the layers in which they occurred. The tracking process happens in three steps: the creation of seeds, their extension into full tracks through a combinatorial Kalman filter [137] and the resolution of ambiguities [138].

Seeds are formed from groups of three hits in distinct silicon layers (exclusively for this step, for SCT hits information from modules on both sides is required), to have the parameters for an initial helix (see 4.1.2). The production of tracks from the seeds happens sequentially adding hits from successive layers and accordingly adjusting the track parameters (Kalman filter), as illustrated in Fig. 4.1. This step efficiently excludes

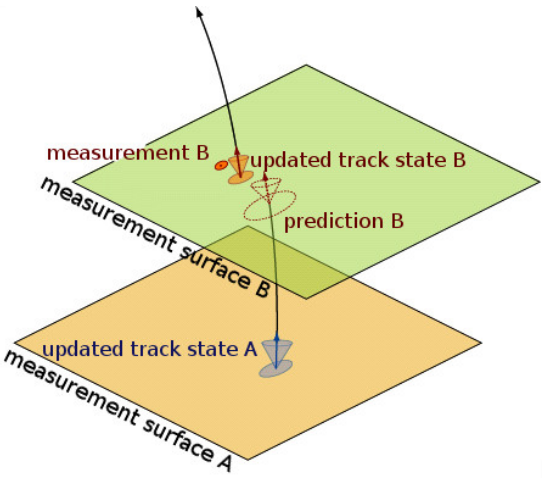


Figure 4.1: Visual representation of an extrapolation and update step of the Kalman Filter technique. The ellipses indicate the uncertainty of the track on a detector plane.

the majority of the incorrect seeds, while guaranteeing a high reconstruction efficiency for genuine tracks. For example the reconstruction of muon tracks above 5 GeV has been measured to be about 99% efficient within coverage of the MS ($0.1 < |\eta| < 2.5$) [122] . The last step in the pattern recognition consists of uniquely assigning hits that are *shared* amongst multiple tracks in favour of high quality and high momentum tracks. This relies on a *track score* system which is based on the χ^2 of the track fit, the resolution of assigned clusters, the logarithm of the momentum and the occurrence of *holes* (the crossing of sections of layers with defective modules or no active material). Particular attention has been put into the identification of *shared clusters* as *merged clusters*, in view of the increase of occupancy between Run-1 and Run-2. The term *merged cluster* denotes a single cluster which is likely to have originated from the crossing of more than one track, and as such should not penalize any track crossing it. Recognition of "merged clusters" is achieved in Run 2 through a Neural Network [139].

The performance of track reconstruction is dependent on the presence of close-by tracks (tracks within a pseudorapidity-azimuthal opening or less than $\Delta R < 0.02$), as this increases the chance of hit sharing. While the selection criteria have been chosen to attain a >99% reconstruction efficiency for tracks with $p_T > 1$ GeV and eight or more (uniquely) associated silicon hits, it can be seen in Figure 4.2 that single-track reconstruction efficiencies vary with the number of charged particles produced with a common origin and the boost of a particle system. The case of $\tau \to 3\mu$ candidates is comparable to $\tau \to \nu_\tau 3\pi^\pm$ in Figure 4.2, momenta of interest being in the range 5-60 GeV, where the

reconstruction efficiency is approximately stable at ∼95%.

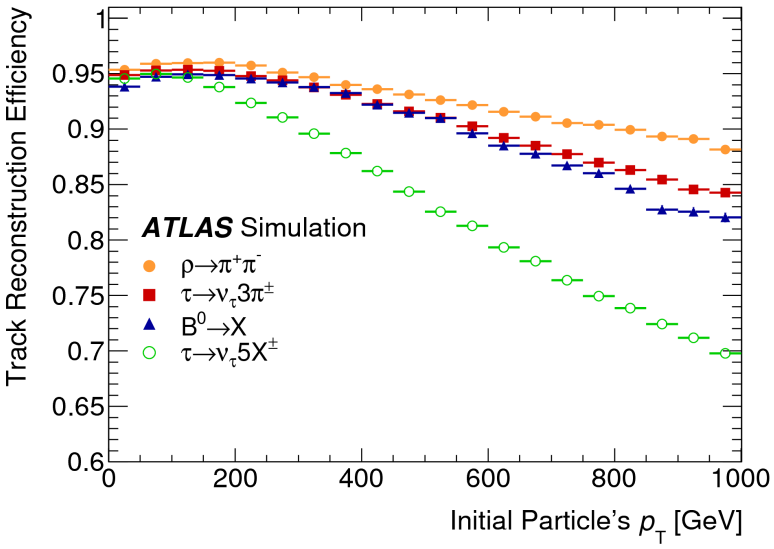


Figure 4.2: Reconstruction efficiency of single tracks as a function of the initial particle's p_T [138]. It is required that the parent particle decays before the IBL. The different colors denote simulated events of $\rho \to \pi^+\pi^-$, τ decaying into three or five charged particles and any decay of B^0 mesons. The efficiencies decrease for higher values of the initial particle's p_T, due to the consequent smaller angular separation ΔR between the products and thus the higher chance of hits-sharing amongst their tracks.

4.1.2 Track Parametrization

Tracks are parametrized locally as helices with the use of perigee parameters as shown in Figure 4.3 (d_0, z_0, ϕ_0, θ, q/p). A description of the track at its point of closest approach (perigee) to the beam-line (see 2.3.2) is parametrized as follows:

- $\boldsymbol{d_0}$ is the signed distance from the beam-line of the Position of Closest Approach (PCA). For particles originating from a primary p-p interaction it only deviates marginally from zero. It is not exactly zero for two reasons: resolution effects in tracking and the fact that the beam line is a geometric reference, not going exactly through every primary vertex. While for a particle coming from the displaced vertex of a secondary decay, d_0 may significantly deviate from zero. The sign is positive if the particle's angular momentum, with respect to the beam line, points along the z-axis and negative if it points in the opposite direction.

- $\boldsymbol{z_0}$ represents the axial coordinate of the PCA, with respect to the detector's coordinate system (see 2.3). It is essential to determine to which Primary Vertex (PV) a particle belongs.

- $\boldsymbol{\phi_0}$ and $\boldsymbol{\theta}$ are the azimuthal and polar angle at the PCA

- $\boldsymbol{q/p}$ represents the charge of the particle divided by its momentum. It is used as it is directly related to the curvature of the track due to the Lorentz force, as well as embodying the fundamental information of the particle's momentum.[1]

It should be clear that for a particle which didn't originate at its PCA, the perigee parameters do not contain the accurate direction of the particle's momentum at its formation. For such particles one would like the track's four-momentum at the point where the particle really came from. This so-called *refitted momentum* is determined through vertexing.

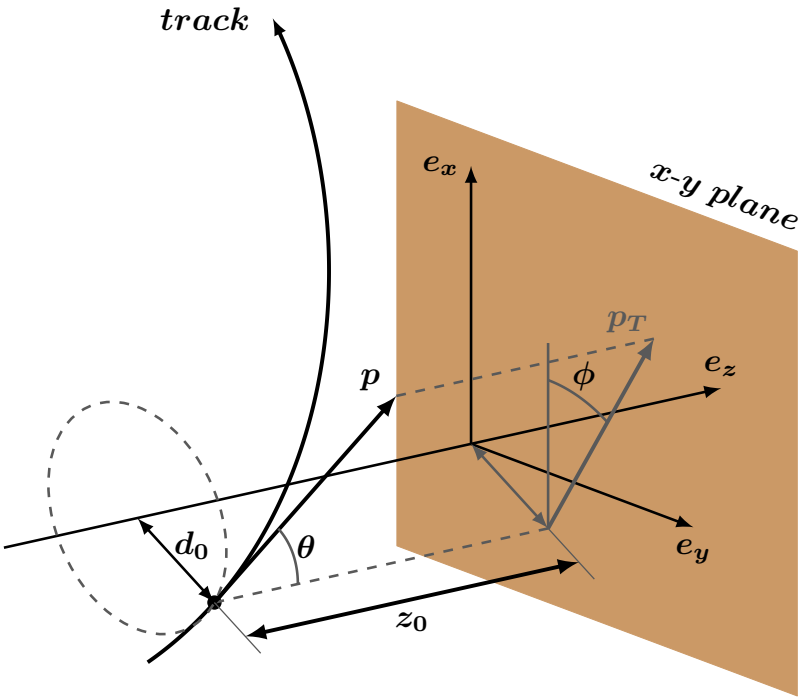


Figure 4.3: Visualization of the Perigee parameters for a track.

[1]Focusing on the x-y components of track's hits, for any set of three hits in the ID, one can determine the sagitta s (deviation of the central point from the straight line L connecting the other two). The sagitta is safely approximated as $s \simeq \frac{L^2 B}{8} \frac{q}{p_T}$ in ATLAS. This not shows why q/p is an intuitive quantity for the track's momentum. The uncertainty on s (Δs) is roughly constant because proportional to the ID's granularity. Deriving the expression for s with respect to p_T one obtains that the fractional uncertainty on particle's momentum is proportional to the momentum itself. From here one can easily return to q/p dividing by $cos\theta$, which is determined from the hits' radii and z coordinates.

4.2 Vertexing

Vertexing refers to the identification of the position from which multiple tracks have originated. There is a distinction between how this is achieved for Primary Vertices (PV) and Secondary Vertices (SV).

4.2.1 Secondary Vertices

SV are the vertices where long-lived particles decay, so they are displaced from the primary vertex at which the long-lived particle originated. In the case of $D_s^{\pm} \to \tau(3\mu)\nu$ the total time of flight is very short (covering only about 2.7 mm for a typical τ momentum of 30 GeV), much shorter than the radial distance of the IBL (30 mm). Vertexing is used to find the SV at which the muon's tracks meet.
This procedure is straightforward: in most events of interest only a few *combined* muons are found (see 4.3.1). So only a handful of three-muon combinations are possible, and for each of these a common vertex is fitted.
The calculation itself follows a minimisation fit for which the tracks are parametrised in proximity of the vertex (Billoir method [140]). In just a few steps the position of a vertex is determined, minimizing χ^2: the sum of the squared distances of all tracks, opportunely weighted by the uncertainty on each track fit.

$$\chi^2(\boldsymbol{v}) = \sum_i \chi_i^2(\boldsymbol{v}) = \sum_i \frac{d_i^2(\boldsymbol{v})}{\sigma_i^2} \tag{4.1}$$

In this equation $\boldsymbol{v}$ is the vertex position, χ_i^2 is the contribution of the i-th track, $d_i(\boldsymbol{v})$ is the distance of the i-th track from $\boldsymbol{v}$ and σ_i is the uncertainty on $d_i(\boldsymbol{v})$. The minimum is determined as the point with vanishing first derivative:

$$\frac{1}{2}\frac{\partial\chi^2(\boldsymbol{v})}{\partial\boldsymbol{v}} = \sum_i \chi_i(\boldsymbol{v})\frac{\partial\chi_i(\boldsymbol{v})}{\partial\boldsymbol{v}} = 0 \tag{4.2}$$

4.2.2 Primary Vertices

PV are the vertices where the partons collided. It is important to identify the PV of a specific interaction to separate it from pile-up interactions occurring in the same event. Figure 4.4 illustrates this showing the 26 reconstructed PVs of an event with a Higgs decay candidate, collected in 2016. The goal is to accurately and efficiently estimate the position of vertices of interest, which is non-trivial due to the high density of tracks. The challenge is to avoid splitting of vertices and the erroneous construction of numerous fake vertices from random associations of tracks. In Run-2 as in Run-1, the algorithm reconstructing PVs uses as seeds a series of z positions extracted iteratively from the reconstructed tracks. This approach is possible because the overlap of two colliding proton bunches is distributed as an extremely elongated ellipsoid. To first approximation PVs can be considered distributed along the beam line.

The reconstruction of PVs is done through adaptive fitting [141], a procedure that combines annealing with the Kalman Filter technique for vertexing [142] [143]. The main variation from the fit described in 4.2.1 is that weights are introduced to the contribution of each track to the χ^2. This means that equation (4.2) is changed into

$$\sum_i w_i(\chi_i^2)\chi_i \frac{\partial \chi_i}{\partial \boldsymbol{v}} = 0 \tag{4.3}$$

The weights are a direct function of each track's χ_i^2 but as well of a "temperature" parameter T:

$$w_i(\chi_i^2) = \frac{\exp(-\chi_i^2/2T)}{\exp(-\chi_i^2/2T) + \exp(-\chi_c^2/2T)} \tag{4.4}$$

where χ_c^2 is a constant, representing the "critical" value for χ_i^2 above which (4.4) is equivalent to removing the i-th track ($w_i \equiv 0$) in the limit $T \to 0$. This makes it such that tracks further from the vertex influence less the outcome of the fit. At each step T is lowered, such that the impact of the weights is gradual, preventing the fit from ending in a local minimum, biased by the initial z position. After each fit, tracks which are more than seven standard deviations away from the produced vertex are removed and may be used in a different vertex.

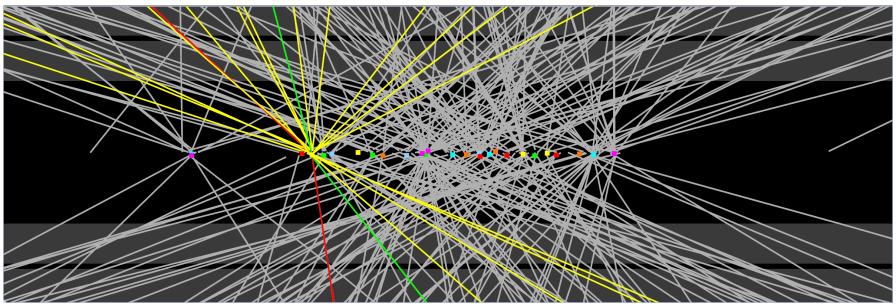


Figure 4.4: Event display of event number 2206548301 collected by ATLAS on July 25$^{\text{th}}$, 2016. In this event 26 primary vertices are reconstructed, as indicated by coloured dots. Image taken from [144].

4.2.3 Muon-Triplet Vertex Related Quantities

The largest source of information about $\tau \to 3\mu$ candidates comes from the correct reconstruction of the τ decay vertex. This section focuses on the description of the related quantities that are used for the $\tau \to 3\mu$ analysis.

χ^2 of fit The χ^2 (equation (4.1)) describes the "goodness of fit" for the vertex. It is primarily used to reject wrong vertices constructed from random combinations of tracks

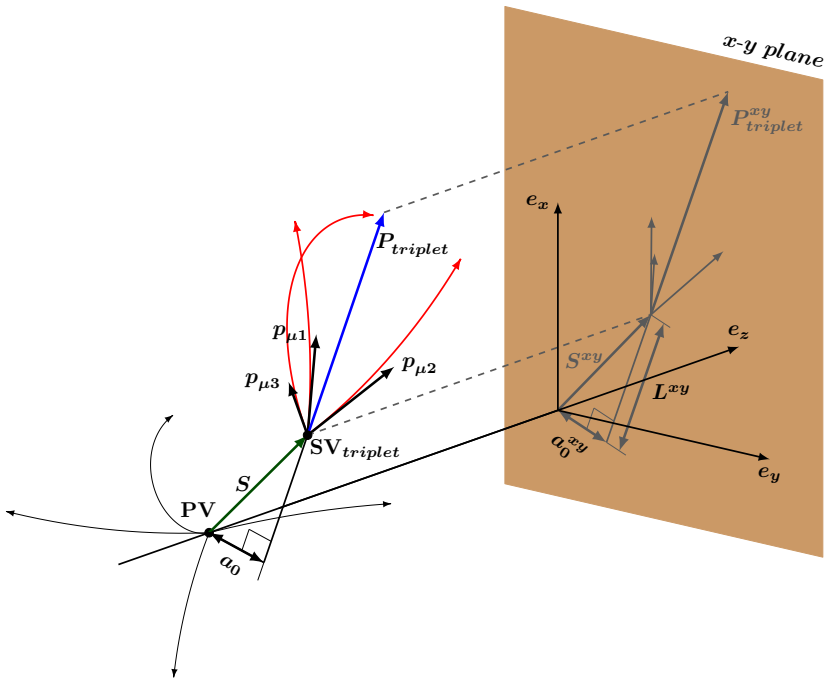


Figure 4.5: Graphical representation of the triplet-related vertex quantities, described in the text. The PV is taken to lie on the z-axis. For extra clarity the curvature of tracks is exaggerated and the projections on the x-y plane are coloured gray. The tracks of the three muons are indicated in red, their refitted momenta are shown in black, and the candidate muon triplet momentum is shown in blue. A few tracks are shown to indicate the (refitted) PV, and the displacement vector $\boldsymbol{S}$ is highlighted in green. 90° angles are indicated, to show that, while $\boldsymbol{a}_0$ forms a 90° angle with the direction of $\boldsymbol{P}_{triplet}$ (three dimensional), $\boldsymbol{a}_0^{xy}$ forms a 90° angle with the transverse component $\boldsymbol{P}_T^{triplet}$ (in two dimensions). Hence a_0^{xy} is not coinciding with the transverse component of a_0 (see also Figure 4.6).

Refitted Track Momenta Once a SV is reconstructed for a triplet candidate, it is possible to recalculate the track parameters with respect to this vertex. The parameters thus undergo a slight angular variation due to extrapolation to the new position on the tracks (bending in the magnetic filed); the magnitude of the momentum can vary as well as a track deforms to pass through the SV [145]. The resulting momentum of the "refitted" tracks is denoted p^{Ref} and its transverse component ${p_T}^{Ref}$. From these refitted momenta the candidate muon triplet momentum is calculated: $\boldsymbol{P}_{triplet} = \sum_{i=1}^{3} \boldsymbol{p}_i^{Ref}$, as well as its transverse component $\boldsymbol{P}_T^{triplet}$.

Refitted Primary Vertex After the SV is chosen, PVs are recalculated omitting the tracks coming from this secondary vertex. As there can be multiple choices of tracks for the construction of a SV, refitted PVs are defined for each SV candidate. Of all

these, only one refitted PV is associated to each SV, choosing the one with the smallest impact parameter a_0 (see the next entry). Unless differently stated, whenever a triplet candidate's PV is mentioned, this has been recalculated without the SV tracks.

a_0 The variable a_0 quantifies how closely the collective momentum of the tracks associated to a SV point back at the PV. a_0 is the component of the displacement between SV and PV orthogonal to the combined momentum of all tracks in the SV. As shown in Figure 4.5, denoting $\boldsymbol{S}$ the vector connecting PV and SV and $\boldsymbol{P}_{triplet}$ one can verify that $a_0 = \|\boldsymbol{S} \times \hat{\boldsymbol{P}}_{triplet}\|$ [2].

a_0^{xy} Because the ATLAS detector has higher azimuthal granularity, it is often better to focus on the transverse projection of momenta and displacements ($x-y$ plane). Taking on the transverse components of the displacement vector $\boldsymbol{S}^{xy}$ and the combined momentum of tracks at the SV $\boldsymbol{P}^{xy}$ one builds the transverse impact parameter of the muon-triplet with respect to the PV called a_0^{xy}. Denoting $\Delta\phi_{S-P}$ the angle between $\boldsymbol{S}^{xy}$ and $\boldsymbol{P}^{xy}$, one can write $a_0^{xy} = \|\boldsymbol{S}^{xy}\| \sin\Delta\phi_{S-P}$. Note that a_0^{xy} does not coincide with the transverse projection of a_0 (Figure 4.6).

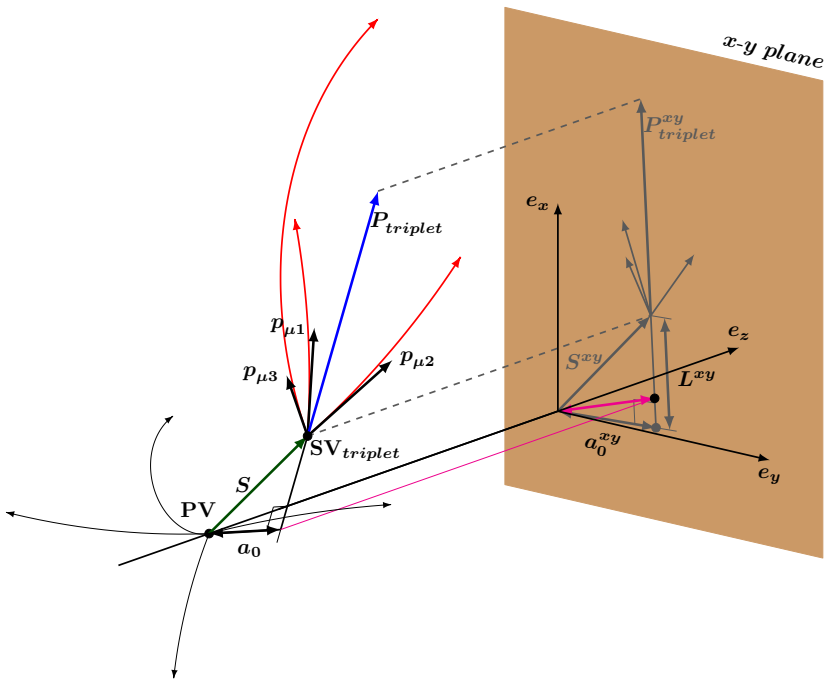


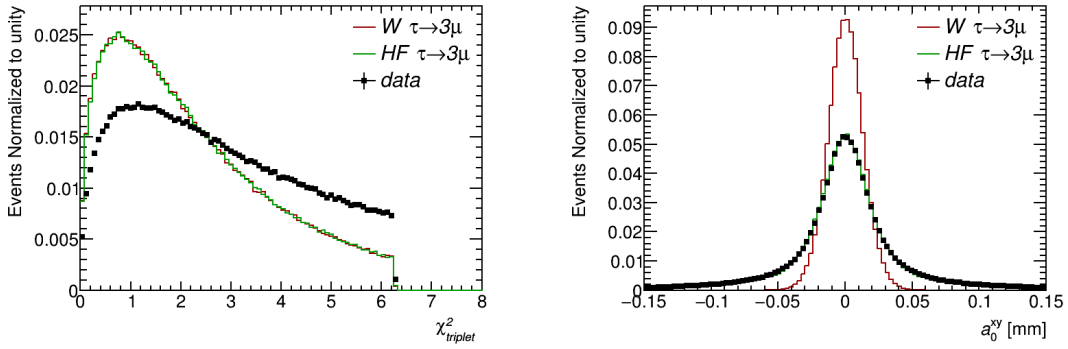
Figure 4.6: Graphical representation of triplet-related vertex quantities in a topology where the angle between $\boldsymbol{S}$ and $\boldsymbol{P}_{triplet}$ is particularly emphasized. This makes the difference between a_0^{xy} and the transverse projection of a_0 (shown in magenta) tangible. The colour scheme is analogous to Figure 4.5.

[2] The hat ($\hat{}$) is used to denote the unit vector with the direction of the given vector ($\hat{\boldsymbol{v}} = \boldsymbol{v}/\|\boldsymbol{v}\|$) and $\|\boldsymbol{v}\| \equiv v$ indicates a vector's magnitude ($\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}$)

$\boldsymbol{L^{xy}}$ Analogous to a_0^{xy}, L^{xy} is built from transverse components. It represents the longitudinal impact parameter of the triplet with respect to the PV: $L^{xy} = \|\boldsymbol{S}^{xy}\| \cos\Delta\phi_{S-P}$. Since in the case of long-lived particles their displacement must happen in the direction of their momentum, in the case of a $\tau \to 3\mu$ signal the distribution of L^{xy} shows an asymmetry, with mostly positive values. In so far L^{xy} is strongly correlated with the τ lifetime. Conversely, in the case of accidental combination of tracks into SV candidates, the distribution of L^{xy} remains symmetric around 0; negative values of L^{xy} are then cases in which the combined momentum at the position of the best SV fit points back in the general direction of the PV.

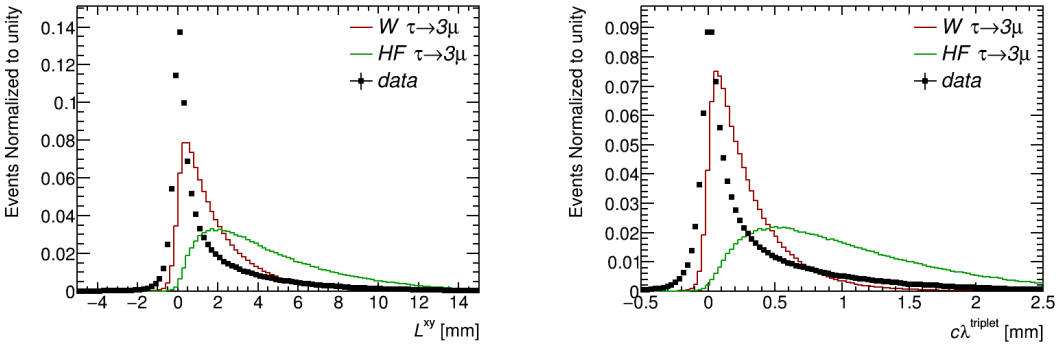
$\boldsymbol{\lambda^{\text{triplet}}}$ Due to time dilation L^{xy} remains correlated with the momentum carried by the decaying particle. This is addressed by calculating the proper lifetime of the decaying particle. This is calculated as:
$\lambda^{\text{triplet}} c = M_\tau \frac{L^{xy}}{P_T^{triplet}}$.
In the case of $D_s^\pm \to \tau(3\mu)\nu$ the value departs from the actual lifetime of the decaying τ, as it is affected by the displacement of the initial charmed meson, which cannot be disentangled from the τ's path.

For illustration purposes, distributions of four of the most characteristic vertex-related variables are displayed in Figure 4.7 and 4.8. The figure shows the distributions of full Run-2 data, $W \to \tau(3\mu)\nu$ and $D_s^\pm \to \tau(3\mu)\nu$ simulated events.



(a) (b)

Figure 4.7: Distribution of $\tau \to 3\mu$ events and data events representative of backgrounds, as a function of χ^2 (a) and a_0^{xy} (b). The distributions are normalized to unity integral and show events passing the loose selection adopted for the HF analysis (Table 5.6). In image (a) the effect of the $P\text{-}value^{triplet}$ cut is visible.

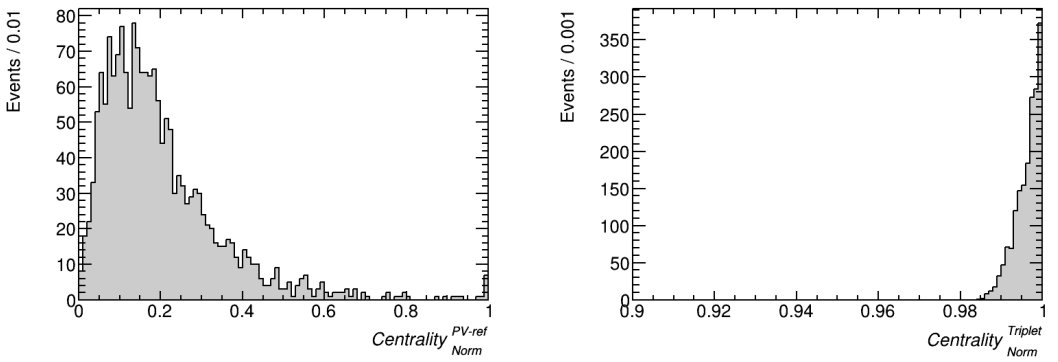


(a) (b)

Figure 4.8: Distribution of $\tau \to 3\mu$ events and data events representative of backgrounds, as a function of L^{xy} (a) and λ^{triplet} (b). Normalisation and selections are the same as in Figure 4.7.

4.2.4 a_0^{xy} bias and correction

Of the introduced muon-triplet variables, a_0^{xy} in particular has been found to display a time-dependent bias (the order of the time scale here being a few weeks). This section illustrates the cause of said bias and the action undertaken to counter it.
Two main alignment efforts have taken place in 2015, employing cosmic muons [146] and the products of p-p collisions [147]. These campaigns successfully accounted for major



(a) (b)

Figure 4.9: Centrality distributions for the refitted primary vertex (a) and the muon triplet vertices (b) in simulated $\tau \to 3\mu$ events. The abscissa in figure (b) begins at 0.9 for illustration purposes, as events form a narrow peak just below 1. Centrality is defined as $1/N_{\text{Tracks}} \sqrt{(\sum_{\text{Tracks}} \sin\phi)^2 + (\sum_{\text{Tracks}} \cos\phi)^2}$ where ϕ denotes the azimuthal angle of the tracks assigned to the vertex. Low centrality scores are indicative of vertices with tracks uniformly distributed in ϕ.

global and local displacements of ID modules, even solving the notorious problem of a significant temperature-dependent deformation of the newly installed IBL, due to a mismatch between the thermal expansion coefficients of its components [148]. Unfortunately several factors concurred in the delay of a fully accurate alignment being applied to central event reconstruction, leading to a systematic bias of specific track parameters, most notably in 2015 and 2016 data.
The residual misalignment are so called "weak modes": deformations (with respect to the design geometry) that do not affect the χ^2 of tracks, and can thus not be addressed

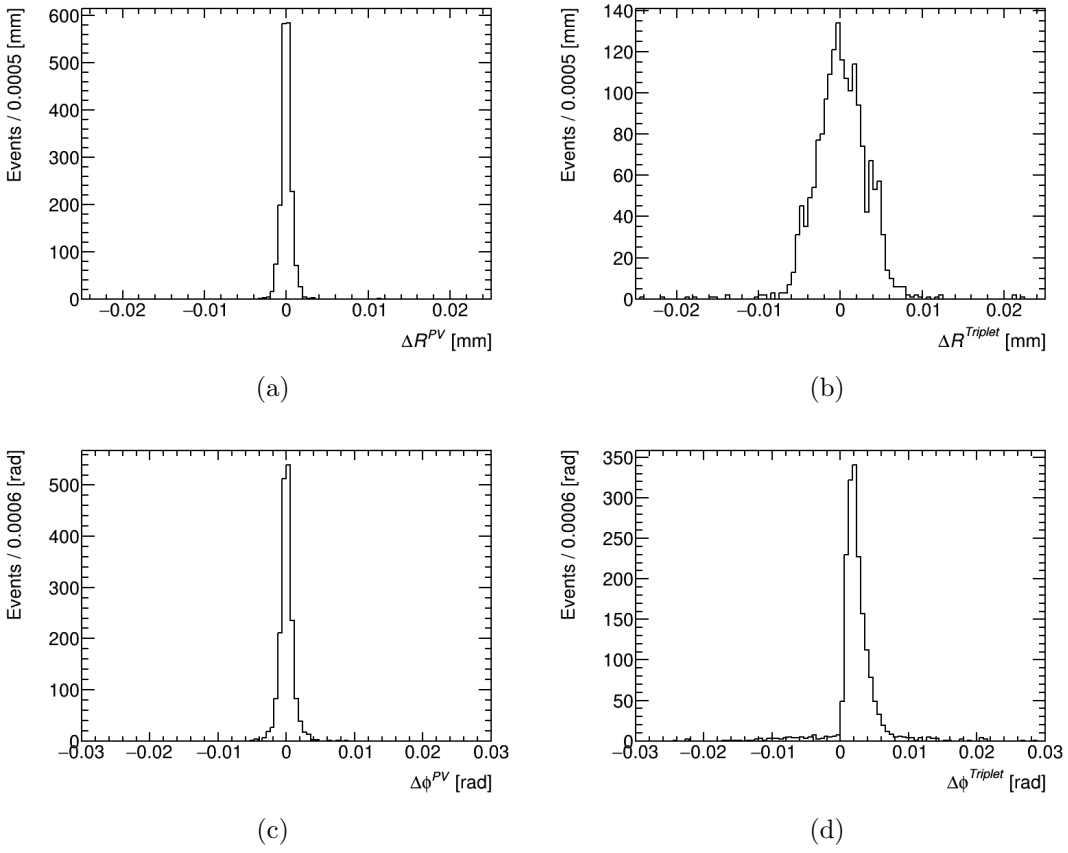
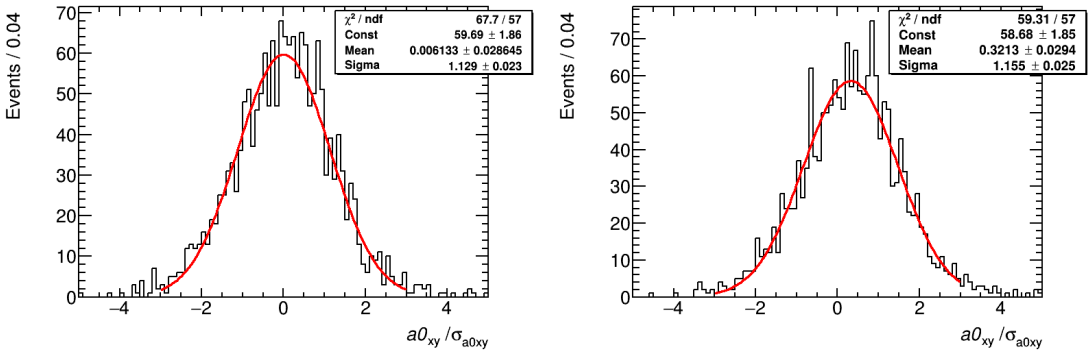


Figure 4.10: Display of the impact in terms of variation of position of refitted PV and muon-triplet vertex of the weak mode bias. The images show the variation of the vertex positions determined from simulation, by applying the deformation measured for the post-TS1 data taking in 2016. The deformation is introduced at reconstruction level and the parameters are the differences calculated on an event-by-event basis. Figures (a) and (b) represent the variation in radial distance from the z axis, of refitted primary vertex and muon-triplet vertex, respectively. Figures (c) and (d) show their variation in azimuthal angle, instead. It is noteworthy that only the azimuthal position of the secondary vertex is significantly and coherently shifted in the positive direction by the bias, while all other coordinates remain substantially unaltered.

through a sequential procedure of minimization like the one described in [146]. The nature of the relevant deformation is of a curl in the ϕ direction (i.e. a radius-dependent rotation around the z-axis) and a slight radial deviation. The misalignment affects mainly tracks' impact parameter (d_0), at the few μm level, while the typical resolution of d_0 is one order of magnitude larger.
Amongst the vertexing quantities used for the $\tau \to 3\mu$ analysis, a_0^{xy} showed particular correlation with tracks' d_0, thus displaying a bias itself. All other variables of interest are found to be substantially unaltered by the "weak modes".
The misalignment follows two separate regimes: first, in 2015 and 2016, before the technical stop of early June (TS1) when a negative shift was affecting tracks' reconstructed d_0 coordinate; for the remainder of 2016 a shift towards positive values affected the reconstruction, instead. The reason why the deformation affects specifically a_0^{xy} can be understood observing the nature of the deformation and its impact on the quantities involved in the a_0^{xy} computation, shown in Figure 4.10: the component of the deformation that generates the bias is the effective curl of the IBL modules in the ϕ direction. This translates in an effective ϕ-shift of the IBL hits, which at track reconstruction generates the d_0 bias. This bias has a different impact on the PV and on the muon triplet vertex. Due to the distribution of tracks determining the PV, over all azimuthal angles, no substantial net effect is seen on the position of the PV. While the three tracks of the muon triplet are virtually aligned, such that the coherent bias in their impact parameter translates directly into a shift along ϕ of the SV's position, and thus into a shift in a_0^{xy}. The different topology of the two vertices is displayed by their *centrality*, shown in Figure 4.9.
The bias' effect on the significance of a_0^{xy} is evaluated performing a Gaussian fit to the distributions of this variable in $\tau \to 3\mu$ simulation without and with applied misalignment as described, with a magnitude is indicative of the worst occurrence in the indicated data periods. The fitted distributions are shown in Figure 4.11, displaying as well the fit parameters. This endeavour shows that to a good approximation the misalignment effect can be represented by a shift of the significance of a_0^{xy}. For this reason the named "weak mode" misalignment is addressed, in this analysis, correcting the value of the significance of a_0^{xy} applying an offset obtained through a Gaussian fit to the same distributions of data events. The correction is determined for each data taking period (each period constituting of a number of consecutive runs) and is shown in Figure 4.12.



(a) (b)

Figure 4.11: Distributions of simulated $\tau \to 3\mu$ events as a function of the triplet's a_0^{xy} significance. On the left the events are produced for nominal alignment (a), while the events shown to the right have an applied misalignment with magnitude comparable to the worst case scenario estimated in actual Run-2 data. Overlaid are two Gaussian fits, quantifying the misalignment impact.

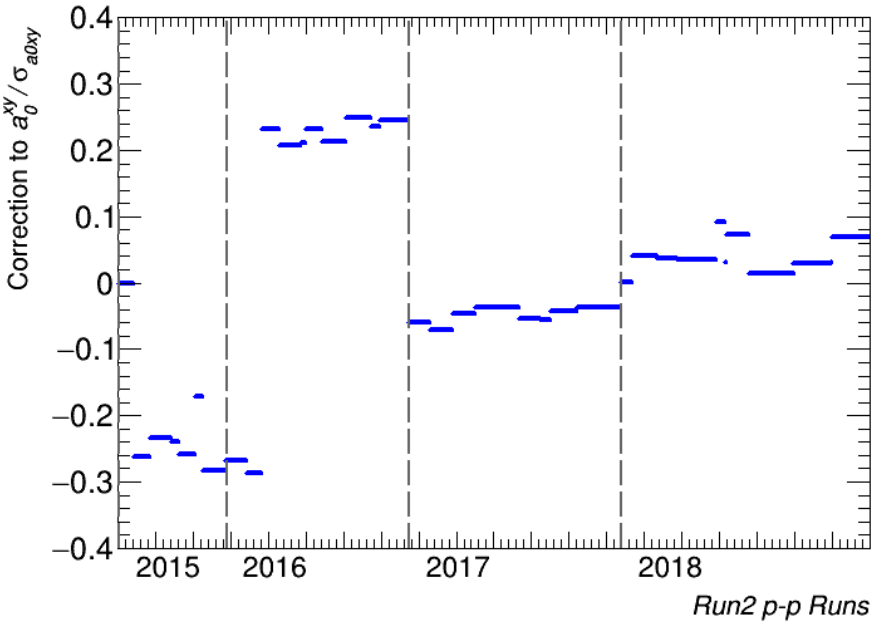


Figure 4.12: Offset values determined for the variable a_0^{xy}, per data taking period. Each bin represents a data run, thus longer segments indicate periods with more runs in them (and not per se more luminosity, as the luminosities collected in each run differ widely). The values of the offsets are determined applying gaussian fits to the a_0^{xy} distributions of data events in each period, similarly to what is shown for simulated $\tau \to 3\mu$ events in Figure 4.11. The dashed lines separate the different years of data taking.

4.3 Muons

4.3.1 Muon Reconstruction

At ATLAS muons are primarily recognized as charged particles traversing the calorimeters and leaving consistent tracks in the ID and in the MS, because of their minimum ionizing nature.
Muons are first reconstructed separately as tracks in the ID and in the MS; in a second step the information of the various subdetectors is combined to determine the momentum with the best possible resolution and perform identification.
In the ID muon tracks are reconstructed precisely like any other track, as described in section 4.1.1.
In the MS, reconstruction begins with the formation of linear *segments* from the hit patterns within each chamber. In this fit, segments are mainly constrained by the multiple hits in MDTs, whereas RPC and TGC hits add spatial resolution in the direction orthogonal to the tubes. For the construction of segments in the CSC, the algorithm initially fits hits in η and ϕ planes, independently.
From segments, MS tracks are created joining segments on the basis of their relative position and angles. This procedure begins using as seeds segments from chambers in the central layer, as they feature the highest number of trigger hits (and thus display the best angular resolution), but eventually proceeds to using un-matched segments from other layers, as seeds.
While segments may initially be part of multiple tracks, a final overlap removal is performed amongst MS tracks, in favour of high-quality tracks. The quality of a MS track is determined on the basis of hit multiplicity and goodness of fit.

The combination of ID and MS tracks into muons follows four different approaches, effective in different detector regions, reflecting the different coverage ranges of the various sub-detectors. In general the reconstruction attempts to use any available detector information to characterise muon candidates.
The four approaches generate the following four muon *types*:

- **Combined (CB) muons** originated from the combination of an ID track and a MS track. A global fit produces the combined track, possibly removing or adding a few hits in the MS to improve the fit and thus the estimate of the track parameters.

- **Segment-tagged muons** these muons are reconstructed when an ID track is extrapolated to the MS, and is there matched to an individual segment.

- **Calorimeter-tagged muons** these are obtained when an ID track is extrapolated to an energy deposition in the calorimeter, compatible with a minimum ionizing particle.

- **Extrapolated muons** in the absence of an ID track, an MS track is extrapolated back to the interaction point, where the muon parameters are then defined.

While CB muons are by far the most reliable muon candidates, segment-tagged muons help reconstruct muons that traverse a single layer of the MS. Calorimeter-tagged muons have the lowest purity amongst muons, but they allow to recover efficiency at $\eta \simeq 0$, where there is a gap in the MS; similarly extrapolated muons permit to reconstruct muons at $\eta > 2.5$, beyond the ID's extension.
For the estimate of muon's energy, the energy loss in the material between ID and MS is accounted for. This is achieved through an analytical parametrization of the average energy loss employing precise information on the detector's geometry, then combined with the actual energy measured in the calorimeter.
The resulting energy resolution for muons is of the order of 1.7% in the central region of the detector, and about 2.3% in the end-caps, in the case of muons up to ∼50 GeV (muons coming from the expected $\tau \to 3\mu$ signal are within this range) [122]. Figure 4.13 shows the di-muon mass resolution, which is used as a proxy for the evaluation of the muon energy resolution. The figure is also a comparison between the resolution in data and simulation.

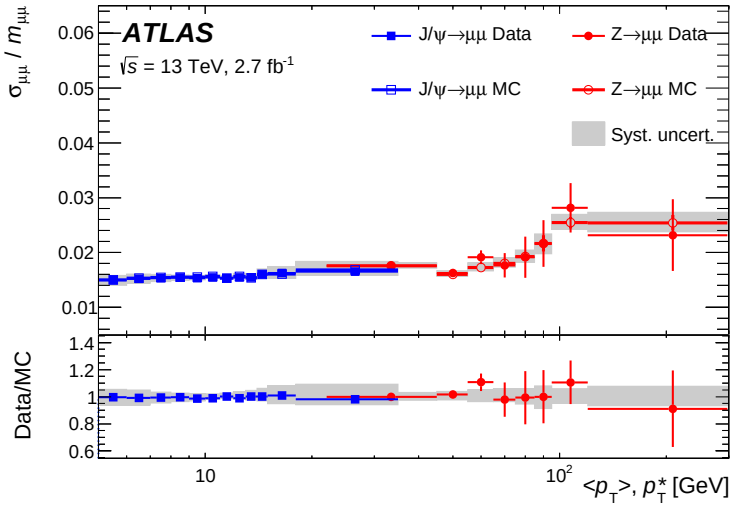


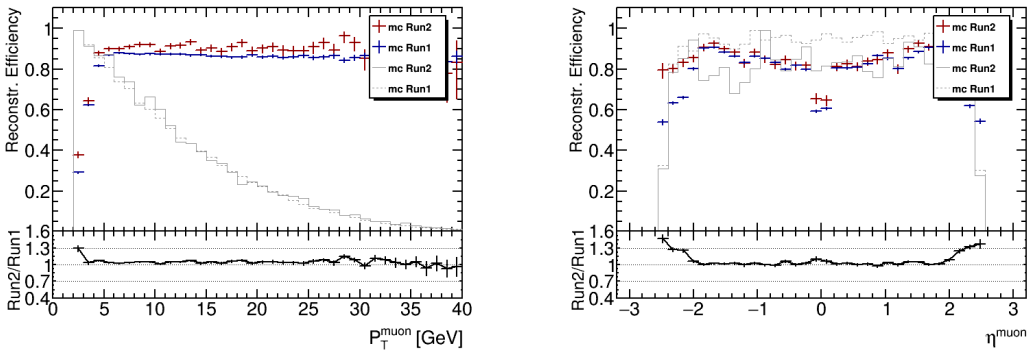
Figure 4.13: Relative resolution of the di-muon invariant mass measured from J/ψ and $Z \to \mu\mu$ events, with respect to average muon momentum variables. This quantity is a measure of the reconstruction's momentum resolution due to a direct proportionality between the two. The blue markers represent the relative J/ψ mass resolution versus the average p_T of the involved muons ($\langle p_T \rangle$). The red markers represent the relative Z mass resolution versus the effective average transverse momentum variable $p_T{}^*$ [3], which avoids the correlation between mass and p_T^μ, through the exclusive use of angular variables. Figure taken from [122].

[3] $p_T{}^* = m_Z \sqrt{\frac{\sin\theta_1 \sin\theta_2}{2(1-\cos\alpha_{12})}}$ where α_{12} is the angle separating the two muons and θ_i the polar angle of each of them. m_Z is the nominal Z mass (91.19 GeV).

4.3.2 Overlap Removal Improvement

It has been mentioned before that from studies related to the Run-1 $\tau \to 3\mu$ analysis it was shown that a considerable number of $\tau \to 3\mu$ candidates had been failing muon reconstruction due to MS track overlap rejection affecting the reconstruction of the close system of three muons. In preparation for this Run-2 analysis the reconstruction efficiency of simulated $W \to \tau(3\mu)\nu$ events with the original (preliminary) Run-2 reconstruction software has been studied and a correction to the muon "overlap removal" algorithm was introduced. This section describes the underlying flaw in preliminary software, the validation of the corrected version and its performance on $W \to \tau(3\mu)\nu$ events.

Of the generated $W \to \tau(3\mu)\nu$ events ∼40% have all muons within acceptance of the ID ($|\eta| < 2.5$) and $p_T > 2.5$ GeV. Reconstructing $W \to \tau(3\mu)\nu$ events with the original software, about 55% of these have their muon triplets reconstructed as three muons, while about 25% have only two muons reconstructed and the third reconstructed merely as an ID track (in the remaining 20%, of the events in acceptance, one or less muons are reconstructed, making them useless for analysis). As a fact, muons with p_T between 2 and 4 GeV are not guaranteed to reach the MS, but the effect is not sufficient to explain the observed inefficiency. This means that a significant fraction of signal events were lost because of the muon reconstruction. Figures 4.14 and 4.15 show the comparison of reconstruction efficiency for simulated signal events, within detector acceptance, between Run-1 and preliminary Run-2 software.



(a) (b)

Figure 4.14: Muon reconstruction efficiency estimates for Run-1 atlas software and preliminary Run-2 software. The efficiencies are shown as a function of muon p_T (a) and pseudorapidity (b). The distributions show the fraction of muons reconstructed as CB muons, matched to generated muons in simulated $\tau \to 3\mu$ events. The baseline is given by events in which all muons are within acceptance, i.e. have $|\eta| < 2.5$ and $p_T > 2$ GeV. Arbitrarily normalized distributions of events generated within acceptance are shown in gray. The simulations are produced using Pythia8 with AU2 set of tuned parameters [149] and MSTW2008LO p-p tune [150]; the reconstruction software used is Athena 17.2.1.4.9 for Run-1 and Athena 20.1.4.13 for Run-2.

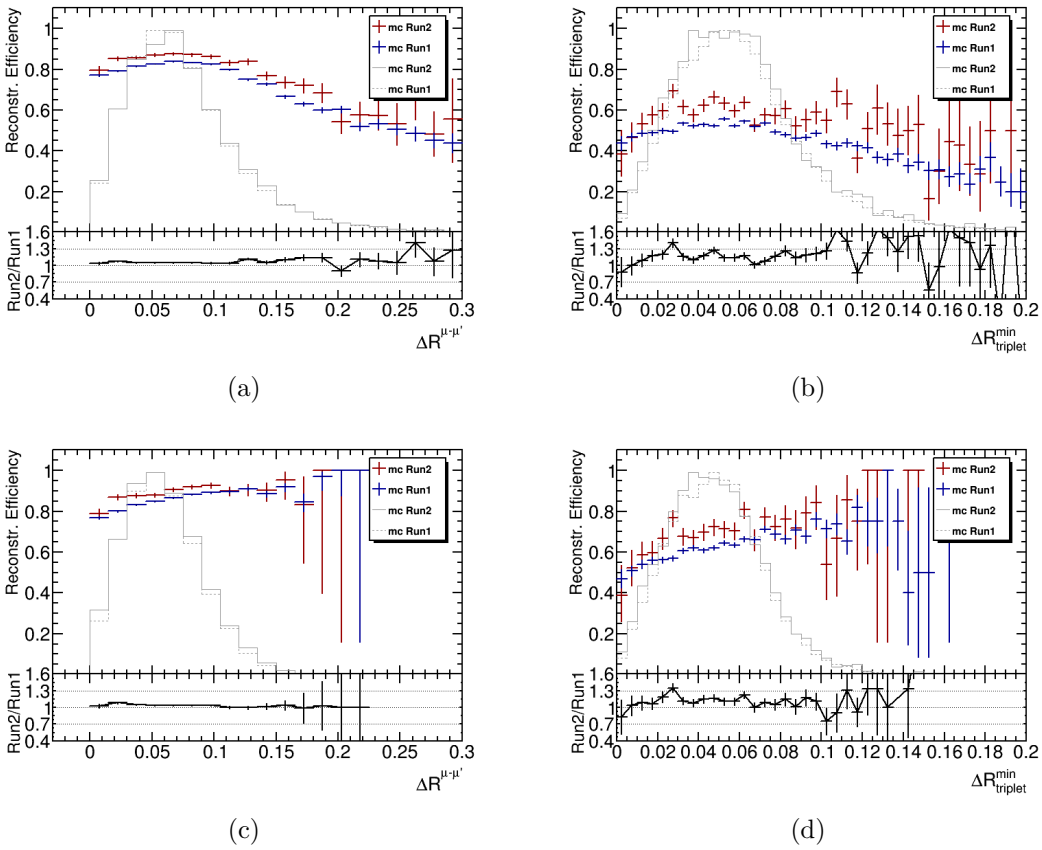


Figure 4.15: CB muon reconstruction efficiency as a function of the ΔR opening to closest muon (a) and (c), and muon-triplet reconstruction efficiency as a function of minimum ΔR between any two muons (b) and (d). Samples and layout are the same as in Figure 4.14. The acceptance selections in (c) and (d) are though tightened to $|\eta| < 2.5$ and $p_T > 5$ GeV, to remove the effect of the drop in reconstruction efficiency at low momenta visible in 4.14a. (a) and (b) are shown as reference; in both distributions (c) and (d) the inefficiency due to the closeness of muons and the consequent overlap of their MS tracks is visible.

In Figure 4.14 one can appreciate the improvements in combined muon reconstruction efficiency for low-momentum muons, as well for muons in the forward region $|\eta| > 1.7$. Figure 4.15 focuses on the performance of the reconstruction of muons as a function of the minimum ΔR between muons. The reconstruction efficiency is higher for Run-2 with respect to Run-1, nevertheless, a persistent trend of degradation of the reconstruction efficiency for closer muons is seen.
Upon closer inspection a flaw was discovered in the exclusion criteria adopted for the muon overlap removal, which can be summarized as follows: the overlap removal step is rejecting one of two MS tracks if they have an overlap of more than 10% of their hits. Exempt from this rejection are tracks that have no shared hits in any precision layer, the

overlapping hits are in these cases coming from the rougher granularity in trigger layers. A criterion which was added in the original Run-2 algorithm allows an overlap up to 25% if tracks cross three precision layers (or more) and share no hits in one of them. This rule was specifically intended to increase the reconstruction efficiency of close-by muons, the flaw is that it penalizes tracks which are independent in more than one precision layer, which should be expected to be less prone to be a fake.
To study the case of extending the overlap removal exceptions with such tracks, various variations were tested on $\tau \to 3\mu$ simulation and a $t\bar{t}$ simulation sample with leptonic final states, where most multi-muon candidates are fakes.
Because the case of tracks independent in (at least) two precision layers is entirely separate from the case of track pairs with one independent layer, it was important to determine the ideal percentage of overlapping hits to be tolerated.
Figure 4.16a shows the reconstruction efficiency for $\tau \to 3\mu$ signal muon triplets, as well as the fake muon fractions for $t\bar{t}$, for variations of the muon reconstruction software allowing muons with more than 10% overlapping hits and independent in two out of three (or more) precision layers ($two\text{-}layer^{\text{indep.}}$ muons), as well as the original configuration and a configuration in which the overlap removal step is removed. The plots show the importance of the overlap removal to contrast fakes, and that introducing $two\text{-}layer^{\text{indep.}}$ has little impact on the fake fraction, for virtually any allowed number of overlapping hits. Figure 4.16b visualizes the per-mille impact of larger overlap acceptance for fake muons in $t\bar{t}$. An ideal maximum overlap limit was chosen to maximise reconstruction efficiency for close-by muons, while keeping the increase in fakes to a minimum. This limit is set at 35%, where the triplet reconstruction efficiency crosses the value attained without overlap removal. Table 4.1 sums up the new criteria adopted by the muon reconstruction to determine whether a muon is rejected due to excessive overlap with another.

The lower quality muon of two muons is rejected if:	they share more than 10% of hits	
Exceptions:	they share no hits in precision layers	
	>2 precision layers; no hit shared in 1 prec. layer;	overlapping hits < 25%
	>2 precision layers; no hit shared in >1 prec. layer;	overlapping hits < 35%

Table 4.1: Criteria followed by muon reconstruction to determine whether two muons are overlapping and therefore the one with lower score needs to be removed. The last line has been added to correct for the inefficiency described in the text.

It must be noticed that, because the overlap removal step causes muons to be refitted with more or fewer hits, the reconstructed muons change between different configurations, such that the reconstruction efficiency with a well-chosen overlap removal selection outperforms the configuration which skips overlap removal all together.
The validity of this new reconstruction configuration was confirmed cross-testing it upon simulated $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events, that were tested against individual disappearances of formerly reconstructed muons. As the result of all tests was deemed satisfactory, the change was applied to Run-2 ATLAS reconstruction software (Athena release 21), establishing a $\sim$20% increase for $\tau \to 3\mu$ signal reconstruction efficiency (Figure 4.16c).

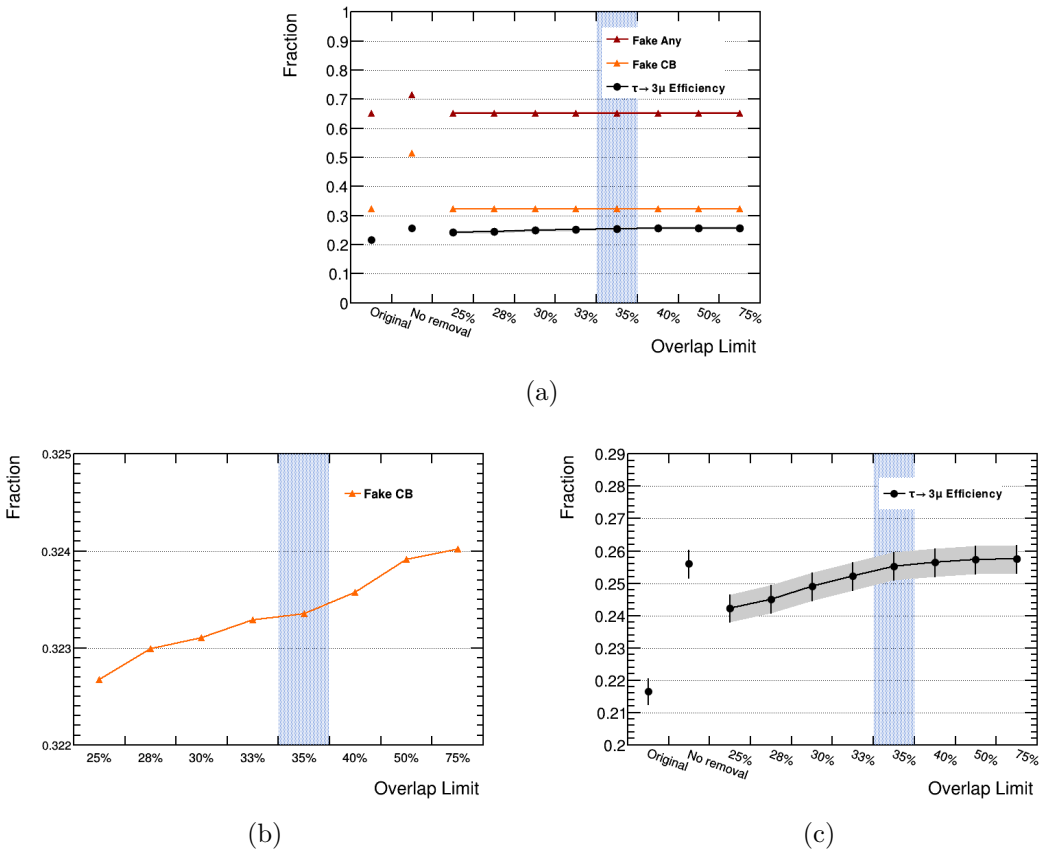


Figure 4.16: $\tau \to 3\mu$ reconstruction efficiency and $t\bar{t}$ fake fraction for variations of the overlap removal criterion for muons. Figure (a) shows the fake fraction in $t\bar{t}$ for any type of muons and for combined muons only. The first of these is largely dominated by calorimeter-tagged muons. The fake fraction is determined by the proportion of reconstructed muons that are not matched to any generated muon. In the same figure the CB reconstruction efficiency for $\tau \to 3\mu$ is shown. This refers to the fraction of events in which all three muons are reconstructed as CB muons. All these values are represented first in the original (Run-2 preliminary) definition of overlap removal (1^{st} bin), then removing the muon removal entirely, and the remaining points show the effect of adding *two-layer*$^{\text{indep.}}$ with a hit overlap cap ranging between 25% and 75%. Figure (b) shows a close up of the $t\bar{t}$ fake fractions, illustrating the weak trend of increasing fakes for larger overlaps. The statistical uncertainty on these values is 4×10^{-3} (not plotted), larger than the variation, however the variation remains significant, because all points refer to the same generated events. Figure (c) shows the same distribution of $\tau \to 3\mu$ reconstruction efficiency shown in (a). The different scale makes it clear that the chosen working point (indicated by a blue band) fully restores the inefficiency caused by the overlap removal in the original configuration (compare 2^{nd} and 1^{st} bin in (c)). The $\tau \to 3\mu$ simulation sample is the same as in Figure 4.14 and 4.15 up to the mentioned variations in the muon reconstruction. The $t\bar{t}$ simulation is generated using Powheg with CT10 [151] tune and uses Pythia6 with the Perugia 2012 [152] tune for particle showering.

With the new reconstruction configuration, of the ∼40% $W \to \tau(3\mu)\nu$ events within ID acceptance, just above 70% are reconstructed as three muons, 16.5% have two muons reconstructed, and only 13.5% have fewer.
This update to the muon reconstruction ultimately permitted for the Run-2 $\tau \to 3\mu$ analysis to avoid the use of a secondary category of triplet candidates formed with one ID track matched to a MS track, as was done in Run-1.

4.3.3 Muon Identification and Isolation

For this analysis the "Medium" quality selection for muons is applied [122]. This is the default selection for ATLAS, requiring muons to be CB or extrapolated. Extrapolated muons are excluded from this analysis because they have been found not to significantly improve the signal acceptance while introducing considerable backgrounds.
The criteria applied to muons for "Medium" selection are summarised in Table 4.2. This selection corresponds to a ∼95% identification efficiency for single muons above 4 GeV. The acceptance of the "Medium" selection is extended with the use of an experimental "Low-p_T" working point for muons with $p_T <$ 5 GeV [153]. This selection extends the ∼95% reconstruction efficiency to $p_T >$ 3 GeV, modifying the criteria for muons in the barrel region, where the energy loss in the calorimeter is most problematic. The working point loosens selections on the MS track, focussing more on inconsistencies in the ID track.
On top of these standard selections, a few quality-related analysis specific selections are applied, these are also listed in Table 4.2.

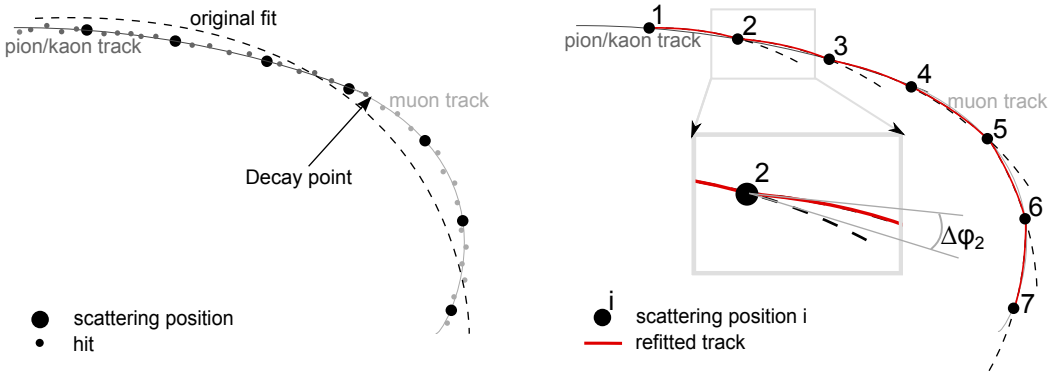
Medium:	Combined muon
($p_T \geq$ 5 GeV)	MS precision hits > 3
	MS precision layers > 1 (>0 if $\lvert\eta\rvert < 0.1$ & MS holes < 2)
	$Sig^{\text{ID-MS}} < 7$
Low-p_T:	Combined muon
($p_T <$ 5 GeV)	MS precision layers > 0
	Out-of-bounds precision hits $= 0$
	$\lvert Sig^{\text{ID-MS}}\rvert$, $\lvert Sig^{\text{IDcurv}}\rvert$, $\lvert Sig^{\text{IDneighbour}}\rvert < 3$
	If $\lvert\eta\rvert > 1.3$, must pass Medium
$\tau \to 3\mu$ specific selection:	$\lvert Sig^{\text{ID-MS}}\rvert < 4$
	$\lvert Sig^{\text{IDcurv}}\rvert < 3$
	MS precision hits > 9

Table 4.2: Muon identification selections. MS precision layers or hits refers to the high granularity MDT/CSC detectors. *Out-of-bounds* precision hits denotes hits in precision layers that are attributed to a track albeit being inconsistent with it. $Sig^{\text{IDneighbour}}$ indicates the scattering neighbouring significance: a quantity similar to Sig^{IDcurv} (see text), but based on the comparison of the two semi-tracks built from all hits preceding and following a pair of consecutive (neighbouring) hits.

Two quantities, essential for the determination of muon quality, are the so-called momentum-balance significance ($Sig^{\text{ID-MS}}$), and scattering-curvature significance (Sig^{IDcurv}) described as follows.

- $\boldsymbol{Sig^{\text{ID-MS}}}$ is the significance of the muon's "momentum imbalance" $\frac{\Delta p}{p} = \frac{p_{\text{ID}} - p_{\text{MS}}}{p_{\text{ID}}}$. where p_{ID} is the momentum of the muon's track in the ID, while p_{MS} is its momentum determined by extrapolating backwards the muons' MS track, and correcting this value by the estimated muon energy loss in the calorimeters [154]. This quantity allows to distinguish genuine muon tracks from candidates arising from the intermediate decay of a light meson (pion or kaon), where the meson decays in the calorimeters.

- $\boldsymbol{Sig^{\text{IDcurv.}}}$ determines the probability of a momentum discontinuity within the ID track of a muon. It is calculated from a series of scatter positions, effectively subdividing the track into multiple segments. For each of these positions the track's residual (s_i) is calculated (equation (4.5)) as the weighted azimuthal deviation of a segment from the continuation of the fit to the previous segment (in the fits the track's momentum is kept constant). The quantity's sign is flipped on the basis of the track's reconstructed charge and weighed by an expected average multiple scattering deviation angle, dependent on the material density in the relevant volume. At this point the track's scattering significance is determined choosing the



(a) (b)

Figure 4.17: Schematic view of the ingredients for the calculation of *ScatteringCurvatureSig*. Figure (a) shows the distribution of hits as well as scattering positions. The fitted track is superimposed as a dashed line. Figure (b) displays the procedure to determine the track residuals for each scattering point: for each segment a track is fitted on the basis of the momentum of the original track, and its extrapolation towards the next segment is compared with the same fit performed there. The difference in the two curve's azimuthal angles at the scattering point is denoted $\Delta\phi_i$. The figure shows a π/K decay to muon half way along the track, leading to positive values of $\Delta\phi_i$ up to the decay point and negative values beyond it. This effect is captured by *ScatteringCurvatureSig*. Figure taken from [154].

scattering position where the local scattering significance (S_P) is maximised, possibly denoting a trend inversion in the residuals caused by a sudden energy loss in its proximity [154]. Figure 4.17 visualizes the procedure displaying the case of a meson decay half-way along a track.
Due to its sensitivity to sudden momentum gradients within the ID Sig^{IDcurv} is complementary to $Sig^{\text{ID-MS}}$ (sensitive to momentum variations beyond the ID). in its function as a discriminant for muon candidates originating from light meson decays.

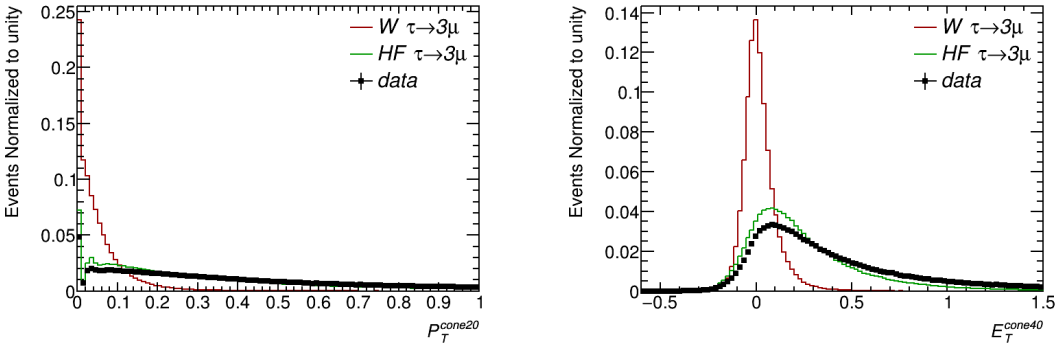
$$s_i = q\frac{\Delta\phi_i}{\phi_i^{\text{msc}}} \qquad S_P(k) = \frac{1}{\sqrt{n}}(\sum_{i=1}^{k} s_i - \sum_{j=k+1}^{n} s_j) \tag{4.5}$$

For the study of muons originating from W decays, identification can be further improved by *isolation* selection. This relies on the observation that the leptonic products of vector boson decays are often isolated from other particles. This especially distinguishes them from non-prompt leptons originating in semi-leptonic decays, which are embedded in hadronic jets. This analysis builds on the isolation criteria derived for muons, adapting them for the use of a $\tau \to 3\mu$ decay.
Isolation is quantified in two ways: summing the momentum of tracks surrounding the candidate (Track Isolation) and adding the energy depositions in the calorimeter, likewise around the candidate's direction of propagation (Calorimeter Isolation).

- **Track Isolation** Track isolation is determined performing a scalar sum of the transverse momenta of all tracks within a ΔR cone of the triplet candidate (the opening is calculated from the triplet momentum $\boldsymbol{P}_{triplet}$ described in 4.2.3). In this calculation ID tracks with $p_T > 500$ MeV are used, and the ID tracks associated to the triplet muons are excluded. To account for the fact that the three muons may be more or less collimated, the acceptance angle depends on the maximum opening angle between the triplet's ID tracks. This is to ensure that isolation for all triplet muons is determined.
  The isolation parameter used in the analysis is P_T^{cone20}, where the scalar sum of the transverse momenta of all tracks within a cone of $\Delta R = \Delta R_{triplet}^{\text{Max}} + 0.2$ is divided by $P_T^{triplet}$.

- **Calorimeter Isolation**: this calorimeter-based isolation E_T^{cone40} is determined summing the transverse energy of topo-clusters (see section 4.4) whose pseudorapidity-azimuthal position falls within ΔR =0.4 of $\boldsymbol{P}_{triplet}$. In this calculation the energy deposition of the muons themselves is removed (only when they do reach the calorimeter within the cone), and the energy contribution of pile-up and the underlying event are subtracted for each event, the same way this is done for the standard muon isolation of ATLAS. This quantity is as well divided by the transverse momentum of the triplet.

The two isolation variables are shown in Figure 4.18, both for simulated W-channel $\tau \to 3\mu$ signal events and loosely selected 2015-2016 data, representative of the backgrounds.



(a) (b)

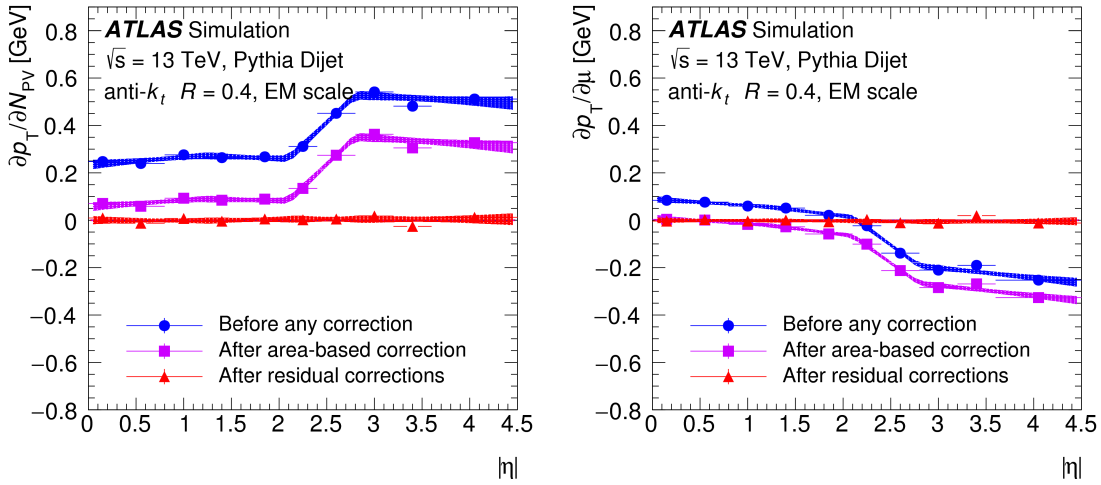
Figure 4.18: Distribution of $\tau \to 3\mu$ simulation events for HF channel (green), W channel (red) and Run-2 data (representative of the backgrounds) as a function of the triplet isolation variables P_T^{cone20} (a) and E_T^{cone40} (b). Events are shown after the loose selection applied for the analysis (Table 5.6).

4.4 Hadronic Jets

Due to the process of color confinement, when coloured particles are produced in a hard interaction, they *hadronize*, producing a shower of color-neutral hadrons. These shower particles are mostly contained within a narrow cone around the momentum of the original particle. These so-called hadronic jets are reconstructed from the ID tracks and calorimeter energy deposition of the shower particles, and constitute the only possible observation of energetic quarks and gluons produced in the hard p-p scattering.
In this analysis *topocluster* jets are employed [155]: jets are reconstructed from clusters of neighbouring calorimeter cells representing "blobs" of deposited energy, clustering is conducted on the basis of topological considerations [156]. For the clustering an anti-k$_t$ algorithm is employed [157], with a radius parameter of 0.4. The major difficulty for jet reconstruction is given by the energy depositions due to pile-up, both in-time and out-of-time. Part of the pile-up influence is countered adjusting to the pile-up level of each run the energy threshold used for the formation of clusters, thus reducing the influence of pile-up upon cluster formation.
A second contribution to pile-up rejection is given by Jet-Vertex Tagging (JVT), an estimate of the probability of a jet having originated from the primary vertex, based on the compatibility of associated ID tracks with the PV itself. JVT permits to recognize topoclusters that are caused by in-time pile-up interactions as well as by residual energy deposited in previous bunch crossings. The latter is ascribed to the ID's response being fast enough to distinguish between bunch crossings. The JVT score for a jet is determined as follows:

$$JVT(jet_i, vertex_j) = \frac{\sum_k P_T(trk_k^{jet_i}, vertex_j)}{\sum_k \sum_\ell P_T(trk_k^{jet_i}, vertex_\ell)} \tag{4.6}$$

Even after a topocluster is identified as a jet, the determination of the jet energy is highly non-trivial [158]. The calculation accounts for the calorimeter's weaker response to energy deposited by hadrons, compared to electro-magnetic interactions. It accounts, as well, for energy depositions which are missed because located outside the boundaries determined by the clustering or in the dead material in front of the calorimeters. Last but not least the contribution of the pile-up energy associated with the underlying event is removed. Figure 4.19 illustrates how the dependence of jet energy p_T on the in-time and out-of-time pileup is countered: firstly subtracting the area-based energy deposits, based on the expected number of pile-up particles associated with a jet, on the basis of its radial extension, and then correcting for the residual dependence (varying as a function of η).

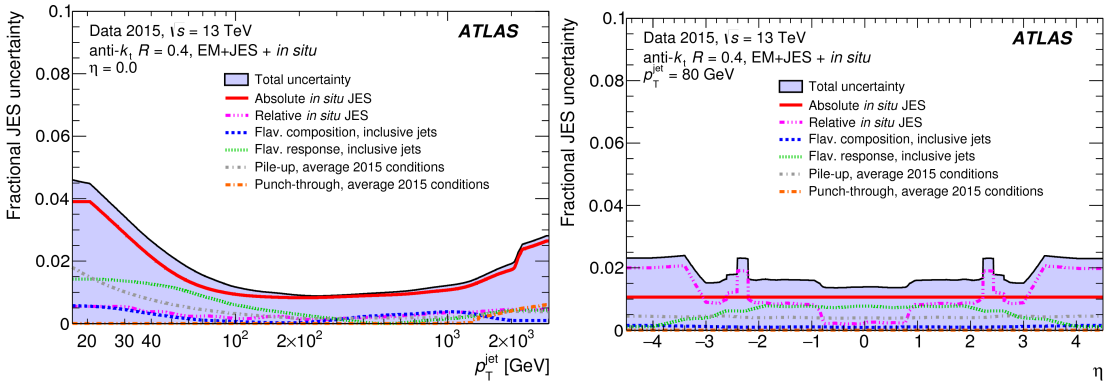


(a) In-time pileup

(b) Out-of-time pileup

Figure 4.19: In-time (a) and out-of-time pileup (b) dependence of the jet p_T of anti-k$_t$ calorimeter jets. The energies are not yet corrected to account for the different impact of hadronic particles with respect to EM, as this has no importance for the subtraction of pile-up energy depositions. The vertical axis represents the derivative of p_T with respect to the number of PVs (averaged over out-of-time pileup level) in (a) and the p_T's derivative with respect to the out-of-time pileup (averaging over number of PVs). The meaning of area-based and residual corrections are explained in the text. Figures taken from [158].

The jet energies are calibrated studying the response on simulated events. Through the study of jets recoiling from Z/γ selected in data, so-called Jet Energy Scale (JES) factors are derived for simulated jets and the systematic uncertainty on the resulting energy scales is determined. Figure 4.20 shows the estimated systematic uncertainty as well as its dependence on jet momentum and pseudorapidity, calculated for Run-2.
While the intrinsic relative energy resolution of the various calorimeters ranges from $10\%/\sqrt{E}$ in the EMCalo (E being divided by GeV), to $50\%/\sqrt{E}$ in the Tile calorimeter



(a) (b)

Figure 4.20: Uncertainty on the jet energy scale after full calibration, determined for 2015 data. (a) shows the energy uncertainty as a function of jet p_T at $\eta = 0$, while (b) shows the uncertainty at $p_T = 80$ GeV as a function of pseudorapidity. The *in situ* jet energy scale uncertainties are the ones derived from the data-driven study of Z/γ+jet and multi-jet balance. They are divided in *absolute* component (independent of η) and *relative*, referring to the correction against η-dependent effects. Figures taken from [158].

and hadronic end-cap up to $100\%/\sqrt{E}$ in the forward calorimeter, the energy scale resolution of reconstructed jets is virtually constant over η. It starts at 4.5% for jets of 20 GeV and progressively improves to about 1% at 200 GeV [158]. The jets used for this analysis are calibrated and *cleaned* (selected on the basis of their quality) as per standard recommendation of the JetEtMiss ATLAS performance group [159]. If jets are within the momentum range 20 GeV $< p_T <$ 60 GeV and are located in the central region ($|\eta < 2.4|$), they are required to pass "Medium" JVT selection (corresponding to a JVT score higher than 0.59). Further a selection of $p_T > 30$ GeV is applied to the jets used in the analysis.

4.5 Missing Transverse Energy

At ATLAS the calculation of Missing Transverse Energy E_T^{miss} [160] is used to gather information on particles that do not leave any signal in the detector. This counts SM neutrinos and hypothetical long-lived neutral states for exotic searches. The principle of E_T^{miss} is that the vector sum of the momenta of all particles emitted in any hard interaction must have a vanishing transverse component, because of momentum conservation with respect to the original colliding partons. The calculation of E_T^{miss} is formed by the negative vector sum of the transverse component of all objects produced in a hard interaction.

$$\boldsymbol{E}_T^{miss} = -\sum \boldsymbol{P}_T^{e,\gamma,jet,\mu} - \sum \boldsymbol{P}_T^{soft} \tag{4.7}$$

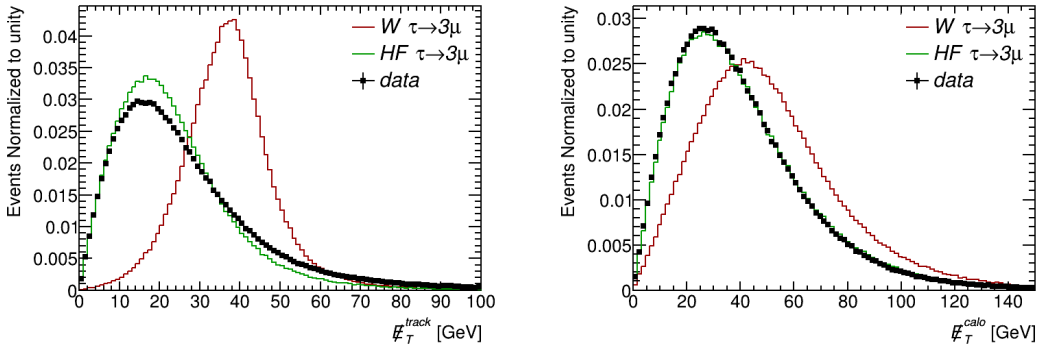
The first term on the right-hand side of (4.7) is the *hard* term, which contains contributions of all calibrated hard (i.e. high-p_T) objects (electrons, photons, hadronic jets

and muons). The second term is the *soft* contribution, which it addresses all the energy, originated in a the hard interaction of interest, that has not been reconstructed as e, γ, jet or μ. Two values for E_T^{miss} can be calculated, based on the two possible choices to produce the soft term: it can be constructed using ID tracks associated to the PV of interest, but not responsible for any of the reconstructed objects ($\not{E}_T^{track}$), or employing all unused calorimeter clusters in the event ($\not{E}_T^{calo}$).
The two definitions are complementary to each other, because of the different information they rely upon. The distribution of $\tau \to 3\mu$ signal events and backgrounds are shown in Figure 4.21c, for both channels. It is though known that $\not{E}_T^{track}$ has the tendency of being more accurate, due to its intrinsic association with the PV, while METcalo is affected more strongly by the level of pileup in the events.
For the $\tau \to 3\mu$ analysis E_T^{miss} is a quantity especially important for the selection of events in the W-channel.
Because in $\tau \to 3\mu$ events the associated PV is not always coinciding with the commonly used highest $\sum p_T$ PV, implicitly used for $\not{E}_T^{track}$ calculation, this analysis makes use of a simpler E_T^{miss}-related quantity for signal-like events. The quantity is referred to as transverse sum of tracks (T_T) and is shown in Figure 4.21. T_T is the magnitude of the transverse component of the negative vector sum of all tracks attributed to a muon triplet's refitted PV with $p_T > 1$ GeV. Adding $\boldsymbol{P}_T^{triplet}$ to it would give a direct estimate for the neutrino p_T in $W \to \tau(3\mu)\nu$ events; not adding this step leaves T_T fairly uncorrelated with $\not{E}_T^{track}$ and $\boldsymbol{P}_T^{triplet}$. The variable is expected to be a more accurate choice than $\not{E}_T^{track}$, to convey information about the signal candidate, albeit being less precise, due to the use of ID track information, instead of calibrated jet parameters, even in the presence of a recoiling jet.
The selection of objects used for the calculation of $\not{E}_T^{calo}$and $\not{E}_T^{track}$, which are not otherwise used in the analysis follows the default quality selection determined by the ATLAS jet and E_T^{miss} combined performance group; the remaining objects respect the same selections listed so far.
The specific selections applied to the inputs of the E_T^{miss} calculation are as follows:

- **electrons** pass ID "Tight", have $p_T > 10$ GeV, and $|\eta| < 2.47$
- **photons** pass ID "Medium", have $p_T > 10$ GeV, and $|\eta| < 2.47$ (objects which are ambiguously reconstructed as photon or electron are used if passing either of the two particle's selection)
- **jets** undergo the same cleaning/JVT selection reported in 4.4, but no additional momentum selection ($p_T > 20$ GeV)
- **muons** undergo the same selection shown in Table 4.2, and are required to have $p_T > 2.5$ GeV

The resolution of E_T^{miss} depends roughly linearly on the average number of interactions, ranging from ~ 12 GeV at $\langle\mu\rangle = 5$ up to 20-25 GeV at $\langle\mu\rangle = 40$ [161] (the resolution of E_T^{miss} is taken as the RMS of the difference between the components of the transverse

momentum p_x and p_y of energetic muons or electrons with the components of E_T^{miss} obtained removing the particle under observation from its calculation).



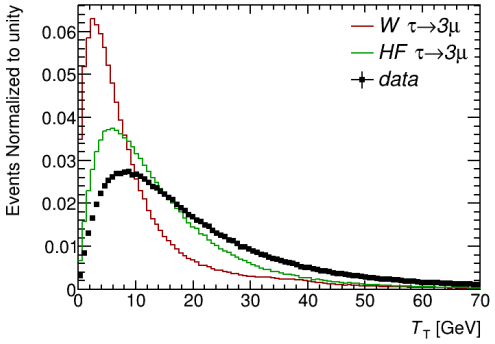


Figure 4.21: Distributions of $\tau \to 3\mu$ events and Run-2 data events representative of backgrounds as a function of $\not{E}_T^{track}$ (a), $\not{E}_T^{calo}$ (b) and T_T (c). The distributions are normalized to unity integral and show events passing the loose selection adopted for the HF analysis (Table 5.6).

Chapter 5

$D_s \to \tau(3\mu)\nu$ Analysis

"There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable.
There is another theory which states that this has already happened."

— Douglas Adams, *The Restaurant at the End of the Universe*

This chapter describes the search for $\tau \to 3\mu$ candidates amongst τ leptons produced in heavy-flavour jets. The dominant production channel for such τ involves the decay of a D_s meson. Events in this channel can be further categorized on the basis of whether this meson is the first long-lived particle coming from the p-p interaction, or whether it is already the product of decay of a b-flavoured hadron. Inclusive decays of B mesons ($B \to \tau X$) constitute a secondary production channel and its contibution is about a third of the D_s one. The estimated production cross section of $D_s^{\pm} \to \tau(3\mu)\nu$ is augmented to account for the extra taus produced (this builds on the strong similarity of the kinematics in $D_s^{\pm} \to \tau(3\mu)\nu$ and $B \to \tau X$). Background events passing the selections are dominated by di-muon resonances combined with unrelated muons or wrongly identified hadrons ($\pi^{\pm}$, $K^{\pm}$); the rejection technique for these events is developed directly on actual collision events. The simulation of the $D_s^{\pm} \to \tau(3\mu)\nu$ signal, which is used to evaluate the trigger and event selection efficiency, is described in section 5.1, followed by the trigger and event selection in sections 5.2 and 5.3 respectively. The D_s production cross-section is normalized using a measurement of the $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ cross section, which is described in Appendix A. The background is estimated on data from the side-bands of the three-muon invariant mass distribution. These side-bands are also used as background sample together with signal Monte Carlo to train a multivariate classifier to increase the signal sensitivity (5.4). The final fit and the resulting limit are presented in 5.5 and 5.6.

5.1 Simulation and Data Sample

For the preliminary analysis presented here, the full Run-2 dataset collected by ATLAS is used ($\mathcal{L}$ =139.6 fb^{-1}). Simulated $D_s^{\pm} \to \tau(3\mu)\nu$ signal events are produced with Pythia8 with A14 tune [162] and NNPDF23LO pdf set [163] (HF related processes are handled using the PythiaB wrapper [164]).
The two-level filter applied during $D_s^{\pm} \to \tau(3\mu)\nu$ event generation, is reported in Table 5.1. The first level requires generated D_s mesons to have transverse momentum larger than 10 GeV and loose pseudorapidity acceptance. The second level mimics the momentum selections that are applied by the analysis triggers: a milder momentum requirement for all three muons ensures events that will pass "3mu4" triggers, and a "staggered" momentum selection accepts events that will pass "mu11_mu6" triggers (which set two different energy thresholds for two independent muons). Further information on this can be found in Appendix A.
Events are simulated separately for prompt and non-prompt D_s mesons (corresponding to the first and third entry in Table 3.1, respectively), then weighted on the basis of the fractions generated by an encompassing simulation. The relative fraction of non-prompt events is found to be 26.6% at *MCDsFilter* level. An ongoing development of the D_s cross-section measurement, presented in Appendix A, aims at measuring the differential fraction of the production cross-section of prompt and non-prompt D_s at ATLAS. This value will eventually be used to normalize the two contributions to the $D_s^{\pm} \to \tau(3\mu)\nu$ signal.

MCDsFilter:	$P_T^{D_S} > 10$ GeV $\lvert\eta^{D_S}\rvert < 3$
MCMixedFilter:	$P_T^{\mu 1}, P_T^{\mu 2}, P_T^{\mu 3} > 3.5$ GeV OR $P_T^{\mu 1} > 10.5$ GeV, $P_T^{\mu 2} > 5.5$ GeV, $P_T^{\mu 3} > 2$ GeV $\eta^{\mu 1}, \eta^{\mu 2}, \eta^{\mu 3} < 3$

Table 5.1: Filtering selections applied to simulated $D_s^{\pm} \to \tau(3\mu)\nu$ events at their generation.

5.2 Triggers

The production of taus from heavy flavour decays is at the edge of the kinematic acceptance of the ATLAS trigger. This makes it necessary to use a combination of triggers to achieve a reasonable efficiency. Because of yearly changes in the beam conditions and trigger operation, the triggers used in this analysis vary from year to year, as shown in Table 5.2. These triggers are selected because virtually unprescaled[1] (each having missed less than 1% of luminosity due to prescaling).
The trigger names are fully descriptive, although not too simple to interpret. The meaning of the short-forms is as follows:

[1]At ATLAS, to maintain under control the rate of events written to disk, triggers may be prescaled, which means that not all events they trigger on are stored. This permits to employ triggers which

- **HLT_**X**mu**Y refers to HLT triggers requiring X muons to be identified with a transverse momentum $p_T > Y$.
- **_nomucomb** means that a trigger does not perform the *muComb* step to match MS tracks with ID tracks.
- **_msonly** denotes triggers that are exclusively based on MS tracks.
- **noL1_nscan**X Narrow scan: the extra muon(s) are sought through full reconstruction of MS tracks within a ΔR cone of X of the single L1 seed.
- **_l2idonly** this trigger applies the p_T requirements on the trigger-level ID tracks. There is no intermediate matching to MS segments, before the full resolution muon reconstruction and selection. Albeit efficient for close-by muons, such triggers were CPU intensive, and thus only used for a limited time.
- **_bTau** a dedicated $\tau \to 3\mu$ trigger implying *nomucomb* and an invariant mass below 2.7 GeV for either two or three muons, as specified in the full name.
- **_bNocut** trigger implying *nomucomb* (plus weak p_T cut at 3.75 GeV and requirment of two muons).
- **_L1MU**Y**_**X**MU**Y' this is a trailing reference to the L1 triggers employed. Similarly to the HLT selection, X represents the multiplicity of muons and $Y^{(')}$ is the p_T selection. The two entries after "L1" mean that the trigger requires events to pass either one of the two L1 triggers (MUY or XMUY').

otherwise would need to be excluded due to their throughput. A trigger's prescale in a specific run, or period, corresponds to the inverse of the fraction of events that are written to disk having passed the trigger in question (thus a trigger active throughout a period would have prescale 1, while a prescale of 2 would indicate that only half of the triggered events were written to disk).

Trigger	Eff.	ΔEff.	$\mathcal{L}$ [pb]	$\mathcal{L} \times$ ΔEff. [pb]
2015				
HLT_mu11_2mu4noL1_nscan03_L1MU11_2MU6	0.106	0.106	3925.3	417.9
HLT_mu20_msonly_mu6noL1_msonly_nscan05	0.081	0.051	3932.9	200.4
HLT_3mu6_msonly	0.061	0.037	3931.8	144.4
HLT_2mu10	0.054	0.016	3933.5	61.3
2016-early				
HLT_mu6_nomucomb_2mu4_nomucomb_bTau_L1MU6_3MU4	0.129	0.129	7090.2	916.8
HLT_mu11_nomucomb_2mu4noL1_nscan03_L1MU11_2MU6	0.115	0.07	7090.2	497.1
HLT_3mu6_msonly	0.061	0.016	7090.2	111.7
HLT_mu20_l2idonly_mu6noL1_nscan03	0.048	0.025	7090.2	178.1
2016-late				
HLT_mu6_nomucomb_2mu4_nomucomb_bTau_L1MU6_3MU4	0.129	0.129	29697	3840.1
HLT_mu11_nomucomb_2mu4noL1_nscan03_L1MU11_2MU6	0.115	0.07	29894.7	2096.1
HLT_3mu6_msonly	0.061	0.016	29894.7	470.8
HLT_mu20_nomucomb_mu6noL1_nscan03	0.046	0.023	29894.7	693.8
2017				
HLT_3mu4_bTau	0.191	0.191	48921	9335.4
HLT_mu11_2mu4noL1_bNocut_L1MU11_2MU6	0.126	0.074	48921	3612.7
HLT_3mu6_msonly	0.066	0.008	48921	379.2
HLT_mu20_mu6noL1_bTau	0.041	0.018	49348.6	901.9
2018				
HLT_3mu4_bTau	0.191	0.191	62943.6	12011.3
HLT_mu11_mu6_bTau	0.159	0.106	62945.4	6694.2
HLT_mu11_2mu4noL1_bNocut_L1MU11_2MU6	0.126	0.02	62943.6	1258.3
HLT_3mu6_msonly	0.066	0.006	62943.6	400.1
HLT_mu20_mu6noL1_bTau	0.041	0.012	63010.2	771.3

Table 5.2: Triggers employed for the HF-channel analysis. The four columns next to each trigger's name represent the signal efficiency for the trigger (determined on MC), the increase in signal efficiency with respect to previously listed triggers, the recorded luminosity (a measure of the amount of data "seen" by the trigger, accounting for the periods in which the trigger was in use, and its prescale) and last the product of recorded luminosity and the relative increase in signal efficiency (a measure of the unique contribution of a trigger). The reported recorded luminosity does not account for losses related to the application of good-run lists (see 5.3.2) and therefore deviates slightly from the values relevant for the analysis. The efficiency of each trigger is calculated on $D_s^{\pm} \to \tau(3\mu)\nu$ simulation, starting with events filtered at generator-level (described in Table 5.1) and having at least three reconstructed CB muons, respectively above the transverse momentum threshold values 5.5 GeV, 3.5 GeV and 2.5 GeV (these mirror the analysis selection in section 5.3.3). For each year only a few of the most relevant triggers are used. The data of 2016 is subdivided into two periods: before the introduction of the B-Physics Delayed data stream and after. The period 2016-late began with run number 302956, which started a few minutes before midnight of June 28th 2016. Triggers are listed in order of decreasing signal efficiency.

5.3 Event Selection

The structure of this section mirrors the sequence of selections in the analysis. Subsection 5.3.1 describes secondary vertex building from two muons and a track and relative selections (procedurally an independent selection step named *derivation* [165]). This covers not only the $\tau \to 3\mu$ signal, but also $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ states which are used in validation. After this come event-level selections (5.3.2), followed by a mainly muon-directed quality selection (5.3.3). At this point Loose selections (5.3.4) are applied to generate training samples for the multivariate classifier. After the classifier training events are further skimmed by the Tight selection (5.3.5) previous to the application of a cut to the classifier and the extrapolation of an expected limit on the branching fraction of $\tau \to 3\mu$ (treated in dedicated sections).

5.3.1 Secondary Vertex Building

The selection of candidate triplets begins identifying the triggers that are of interest to the analysis (see 3.4). For all events which fire one or more of our triggers, an attempt is made to find vertices from which 2 muons originate, as described in section 4.2.1. Such di-muon candidates are further selected to have at least 1 hit or hole in the pixel detector and 2 in the SCT, for a minimum total of 3 hits or holes in the associated ID tracks. A loose selection cut is set on the quality of the vertex, and another on the invariant mass of the two muons. The latter is aimed at excluding J/ψ decays from the resulting set of candidates; failing to do so would result in an output containing almost exclusively $J/\psi \to 2\mu$ due to its large cross-section.
Once the di-muon vertices are formed they are extended into three-track vertices, by adding an ulterior ID track compatible with the secondary vertex. These three-track SVs are henceforth referred to as "triplet candidate vertex" (SV^{triplet}). The full set of selections is reproduced in the following table:

ID Tracks selection:	pixel hits/holes >= 1
	SCT hits/holes >= 2
Di-muon selection:	Vertex χ^2 < 110
	Inv. Mass < 2.9 GeV
Triplet selection:	All track's p_T > 1 GeV
	p_T of triplet candidate > 1 GeV
	Mass of the triplet < 5 GeV
	χ^2 of the triplet vertex < 150

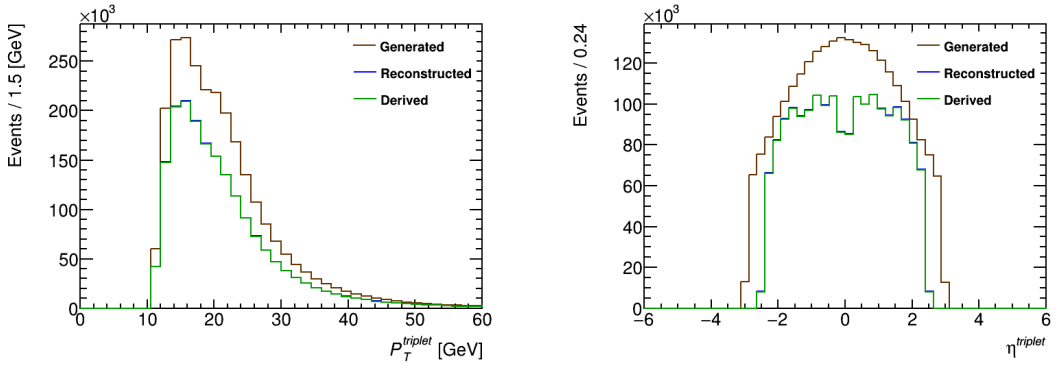
Table 5.3: Selections associated to the construction of triplet candidates.

After rejecting events with no such vertex, besides full information of the newly formed vertices, jets and ID tracks are retained when satisfying the following criteria:

Tracks:	is ID track of a muon is part of a selected triplet candidate is associated to one of the PV candidates matched to any of the triplet candidates [2]
Jets:	$p_T > 20$ GeV

Table 5.4: Skimming selection for tracks and jets.

These selections do select not only potential $\tau \to 3\mu$ candidates, but also "2μ + 1track" vertices, amongst which $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ and $D^{\pm} \to \phi(2\mu)\pi^{\pm}$, which are used in this analysis for normalization and for the validation of developed selections.
The selections described here happen alongside the computation of vertex-related quantities (see 4.2.3).
Figure 5.1 displays the performance of these selections. The distribution of *generated* events are not representative of the inclusive D_s production at ATLAS, but are loosely reduced to the trigger acceptance and pseudorapidity acceptance of ATLAS, by the generator-level filter applied during simulation (see 5.1). Figure 5.1, shows the distributions of generated, reconstructed and *derived* muon triplets (i.e. successfully passing secondary vertex building): the reconstruction efficiency, for events passing the generator filter, is 72.5%, while vertex building efficiency with respect to reconstruction (corresponding to the selections in Table 5.3) is about 99.8%. The <1% events that are lost in this step do not to impact the analysis as they fail successive selections, specifically "MCP ID" (Table 5.5) or $P - value^{\text{triplet}}$ (Table 5.6) selection.



(a) (b)

Figure 5.1: Distributions showing the performance of secondary vertex building in relation to the totality of simulated $D_s^{\pm} \to \tau(3\mu)\nu$ events. In brown the distribution of events obeying generator level filtering (described in 5.1). In blue are all events for which reconstructed muons are matched to muons originating from the decay of the simulated $\tau \to 3\mu$ chain. In green are the events in which the secondary vertices corresponding to the $\tau \to 3\mu$ decays are correctly re-built. The minimal difference between this distribution and the blue one is a measure of the good performance of the *derivation* procedure.

[2] See 4.2.3

5.3.2 Event Level Selection

After the triplet candidates are built, events are selected to be "good for physics analysis" (included in the Good-Run list), i.e. recorded when all parts of the detector were fully functional and the proton beam was stable [166] (as shown in Figure 2.2a, about 93% of the recorded events are "good for physics").
Events are then checked against detector-flags denoting that some subdetectors were experiencing performance-impairing problems [167] (for example noise bursts in the LAr calorimeter).
At this point events are required to have triggered any of the triggers chosen for the corresponding data period as indicated in section 3.4. The average estimated trigger efficiency for simulated $\tau \to 3\mu$ events successfully passing the secondary vertex building is 17% for the 2015 triggers and for both sets of 2016 triggers, 21.6% for 2017 triggers and 26% for 2018 triggers.

5.3.3 Muon Quality Selection

After the building of muon triplets and the event-level selection, further selections are applied to retain only good quality muon candidates (Table 5.5). The first cut assesses the quality of the Inner-Detector track associated with the muons. This selection is moderately tighter than the one reported in Table 5.3, and is an ATLAS standard (defined in the Muon Combined Performance (MCP) group's Muon Selection Tool [168]). Two kinematic cuts follow, ensuring that the ID tracks are within the acceptance of the Inner Detector and that the muon momenta are higher than a set of threshold values.
The next set of selections address the track muons leave in the Muon Spectrometer, and their compatibility with the relative Inner Detector tracks. The muons are required to be "combined" (CB) and pass the MCP "Medium" or "Low-p_T" quality selection, as described exhaustively in section 4.3.3. On top of these some custom selections are applied on the muon-track quality quantities $Sig^{\text{ID-MS}}$ and Sig^{IDcurv} (see 4.3.3), and on the number of precision MS hits associated to the muon tracks; these contribute to the separation of $\tau \to 3\mu$ events from backgrounds, leading to a $\sim$7% improvement on the sensitivity to $\mathcal{B}(\tau \to 3\mu)$.

ID Track cuts:	All tracks pass MCP ID cut $\lvert\eta_{\mu 1}^{\text{Ref}}\rvert, \lvert\eta_{\mu 2}^{\text{Ref}}\rvert, \lvert\eta_{\mu 3}^{\text{Ref}}\rvert < 2.5$ ${p_T}_{\mu 1}^{\text{Ref}} > 5.5$ GeV, ${p_T}_{\mu 2}^{\text{Ref}} > 3.5$ GeV, ${p_T}_{\mu 3}^{\text{Ref}} > 2.5$ GeV
Muon cuts:	CB muons only Muons working point = Medium OR Low-p_T $\lvert Sig^{\text{ID-MS}}\rvert < 2.2$ $\lvert Sig^{\text{IDcurv}}\rvert < 2.5$ n° precision hits > 9

Table 5.5: Track and Muon quality cuts adopted for the HF-channel. The selections are described in the text. The label "Ref" denotes components of the refitted track momenta (as described in 4.2.3).

5.3.4 Loose Selection

Events passing all previous selections are subject to a so-called Loose selection, developed to provide the optimal training sample for a BDT. The Loose selections have been specifically optimized to ensure the best possible performance of the multivariate classifier and are shown in Table 5.7. The first selection ensures that one of the three muons carry opposite charge with respect to the others (thus respecting charge conservation for a τ decay). The second selection is mainly a protection against triplet candidates built from photon-conversion muon pairs, whose invariant mass results being $2m_\mu \simeq 212MeV$. The selection is kept slightly tighter, to account for resolution effects on muon energies, whilst not impacting significantly the expected signal distribution.
The next selection loosely cuts on the quality of the secondary vertex, while the last selection is chosen to cancel a bias introduced by the "di-muon" selection adopted in the vertex building (Table 5.3): that selection, while being meant to reject J/ψ candidates, actually allows a few of them to pass, due to the sequential selection of the third muon.

Loose Selection:	$\lvert\sum_{i=1}^{3} q_i\rvert = 1$
	$M_{os1}, M_{os2}, M_{ss} > 325$ MeV
	$P\text{-}value^{\text{triplet}} > 0.1$
	$M_{os1}, M_{os2} < 2.9$ GeV

Table 5.6: Loose selection used for the HF-channel. q_i is the charge associated to the reconstructed muons, M_{os1} is the invariant mass calculated from the refitted tracks of the opposite-sign muon pair with highest transverse-momentum, M_{os2} is the invariant mass calculated from the other opposite-sign muon combination, M_{ss} is the invariant mass for the same-sign muon pair. $P - value^{\text{triplet}}$ is the p-value associated to the secondary vertex fit.

5.3.5 Tight Selection

A Tight selection is applied after the training of the multi-variate classifier. It has been verified that maintaining the looser selection for the training sample leads to a better background separation by the BDT. The Tight selection is summarised in Table 5.7, it rejects the narrow resonances $\rho^0 \to \mu\mu$ and $\phi \to \mu\mu$. An ulterior cut on the significance of a_0^{xy} is included as it is found to add separation power to a purely BDT based selection on this variable.
Rejection of the resonances can not be handled by the BDT, because including di-muon masses in the training leads to a strong correlation between the classifier and the three-muon mass $M^{triplet}$, making it impossible to use side-bands in this variable to characterise the backgrounds (as described in 5.4.3).

Tight Selection:	$\lvert M_{os1}, M_{os2} - 775 \text{ MeV}\rvert > 50 \text{ MeV}$
	$\lvert M_{os1}, M_{os2} - 1020 \text{ MeV}\rvert > 60 \text{ MeV}$
	$a_0^{xy}/\sigma_{a0xy} < 5.$

Table 5.7: Tight selection used for the HF-channel.

5.4 BDT Analysis

Characteristic of the HF-channel is a pronounced similarity between the sought $D_s^{\pm} \to \tau(3\mu)\nu$ signature and the multi-muon backgrounds. This makes it even more important than for the W channel to rely on a multivariate classification to successfully distinguish the two.
The multivariate classifier employed is a Boosted Decision Tree (BDT) method of the TMVA toolkit [169], part of the ROOT framework, developed at CERN for the handling of petabyte sized data [170] [171]. The training described in this thesis is the baseline built for the HF channel of the $\tau \to 3\mu$ analysis. Beside it, a parallel attempt to employ a Neural Network for the final signal and background separation is being pursued.

5.4.1 BDT Input Variables

The set of variables that is used to train the BDT is outlined in Table 5.8: Two ID-track related quantities are introduced: $\prod_{P\text{-}value}^{tracks}$ and $\Delta\chi^2_{4track}$. The first is chosen as a global handle to the goodness of fit of the three muons' ID tracks. $\Delta\chi^2_{4track}$ is the minimum variation in χ^2 of the secondary vertex, when adding a fourth track to the fit; it serves as a handle for background events in which the triplet is (partially) built from muons belonging to independent vertices from which further charged particles originate.
Besides the previously introduced isolation variables, a further variable is added to handle hadronic activity in the underlying event: H_T^{1jet} constitutes a measure of the counterbalance of the muon triplet candidate by intense hadronic activity. It is defined as $H_T^{1jet} = \|\boldsymbol{P}_T^{triplet} + \boldsymbol{p}_T^{\text{lead}jet30}\|$, i.e. the variable is constructed using only the leading jet in each event, and this only when its transverse momentum exceeds 30 GeV. In events where no such hadronic jet is present, the variable coincides with $P_T^{triplet}$, indicating large momentum imbalance. This alternative of the, more common, vector sum of jets avoids having to deal with the low reliability of the modelling of underlying moderate-momentum jets.
A set of four E_T^{miss} related variables is used. Besides the previously described T_T variable, a $\not{E}_T^{track}$ based transverse mass is added[3], as well as the angles between the muon triplet and the two standard estimates of E_T^{miss} in ATLAS (see chapter 4.5).
Two slightly more elaborate estimators are employed to gather information about the

[3] The transverse mass calculated from $\not{E}_T^{track}$ and the triplet candidate is defined as follows:

$$M_T^{track} = \sqrt{2 \cdot \not{E}_T^{track} p_T^{triplet} (1 - \cos\Delta\phi_{\not{E}_T^{track} - triplet})} \tag{5.1}$$

Variable	Description	Where defined
$P_T^{triplet}$	Transverse momentum of the muon triplet candidate	(4.2.3)
η^{triplet}	Pseudorapidity of the muon triplet candidate	
a_0^{xy}/σ_{a0xy}	Significance of muon triplet transverse impact parameter w.r.t. PV	(4.2.3)
L^{xy}/σ_{Lxy}	Significance of muon triplet longitudinal impact parameter w.r.t. PV	(4.2.3)
$\lambda^{\text{triplet}}/\sigma_\lambda^{\text{triplet}}$	Significane of the triplet candidate's lifetime	(4.2.3)
χ^2_{triplet}	Secondary vertex fit χ^2	(4.2.3)
$P\text{-}value^{\text{triplet}}$	The P-value associated to the triplet vertex	
$\prod_{P\text{-}value}^{tracks}$	Product of P-values of each of the three muon's ID tracks	
$\Delta\chi^2_{4track}$	Impact of fourth track on SV χ^2	
E_T^{cone20}	Calorimeter isolation ($\Delta R = 0.2$)	(4.3.3)
E_T^{cone40}	Calorimeter isolation ($\Delta R = 0.4$)	(4.3.3)
P_T^{cone20}	Track isolation ($\Delta R = 0.2$)	(4.3.3)
H_T^{1jet}	Single-jet hadronic imbalance	
T_T	Transverse sum of track momenta	(4.5)
M_T^{track}	Transverse mass calculated with $\not{E}_T^{track}$	
$\Delta\phi_{\not{E}_T^{track}-triplet}$	Azimuthal angle between triplet and $\not{E}_T^{track}$	
$\Delta\phi_{\not{E}_T^{calo}-triplet}$	Azimuthal angle between triplet and $\not{E}_T^{calo}$	
$\Delta M_{corr}^{\text{triplet}}$	Variation in candidate invariant mass adding a collinear neutrino	
$M_{miss}^{\text{triplet}}$	Missing mass estimate	

Table 5.8: List of variables adopted for the training of a Boosted Decision Tree multivariate classifier. Detailed definition of the variables is found in text, or where indicated in the third column, in the reported sections.

presence of the neutrino in the $D_s^\pm \to \tau(3\mu)\nu$ decay chain, on the basis of the kinematic properties of the reconstructed process: $M_{miss}^{\text{triplet}}$ and $\Delta M_{corr}^{\text{triplet}}$ (both values are closer to zero for signal than for backgrounds).

$M_{miss}^{\text{triplet}}$ is an estimate of the "neutrino mass" based on the muon triplet four-momentum and a four-momentum constructed for the initial decaying particle (here labelled as D_s)

$$\begin{aligned} M_{miss}^{\text{triplet}} &= \sqrt{m_{D_s}^2 + M^{triplet\ 2} - 2E_{D_s}E_{triplet} + 2P_{D_s}P_{triplet}\cos\Delta\alpha_{S-P}} \\ m_{D_s} &= M^{triplet} + \Delta M^{pdg} = M^{triplet} + 192\ \text{MeV} \\ \boldsymbol{P}_{D_s} &= \frac{m_{D_s}}{M^{triplet}} P_{triplet}\hat{\boldsymbol{S}} \end{aligned} \tag{5.2}$$

the four-momentum of the "D_s" is constructed as follows: its mass is obtained adding the difference between the nominal masses of Ds and τ (ΔM^{pdg}) to M^{triplet}. The D_s momentum is given by the triplet momentum scaled by the factor $m_{D_s}/M^{\text{triplet}}$ oriented along the vector connecting PV and SV ($\boldsymbol{S}$, see Figure 4.3). α_{S-P} represents the angle

between the triplet momentum and the vector $\boldsymbol{S}$. The second variable is defined as

$$\Delta M_{corr}^{\text{triplet}} = \sqrt{2(P_T^{triplet} \sin \Delta\alpha_{S-P})^2 + M^{triplet\ 2}} - M^{triplet} \tag{5.3}$$

starting from the premise of a two body decay, this is a measure of the maximum mass that is available to the "neutrino". The first term after the equal sign is a measure of the mass of the initial decaying particle (say D_s), adding to the triplet mass the momentum component of the triplet (and the recoiling "neutrino") in the direction orthogonal to the initial particle's propagation, given by the vector connecting PV and SV). The $M^{triplet}$ value is then subtracted from the result to minimise an otherwise strong correlation with the muon triplet mass itself (important because the classifier should not interfere with the final $M^{triplet}$ selection). Four of the input variables are displayed in Figure 5.2, for simulated signal events and data. Corresponding distributions for the full set of variables are reported in Appendix B. The distribution of data events can be safely used to represent background events, as the number of signal events allowed for by the current limit on the $\tau \to 3\mu$ branching fraction is still negligible after the used selections. The distributions mark the similarity present between signal and the backgrounds, showing why a multivariate classifier is employed to extract what little separation power each of the variables provides. A further observation that can be made is that the variables do not display any specific distinction between narrow and broad di-muon resonances (narrow ones being removed by the tight selection), further justifying the training of the multivariate classifier after the loose selection alone, to provide for extra statistics, beneficial for the process.

To validate the correct modelling of the variables used as input for the BDT training, the simulation's performance has been evaluated on a sample of $D_s^{\pm} \to \phi\pi^{\pm}$ events. D_s has about 4.5% chance to decay into a $\phi\pi$ and phi has a branching fraction into two muons of $2.86 \pm 0.19 \times 10^{-4}$. This may sound like a small amount, but in view of the large production cross section of D_s at the LHC, this channel provides enough statistics to permit the evaluation of spectral distributions of events (the relevant fiducial production cross-section is $\mathcal{O}(10^9$ fb), see (A.3)).

The presented variables are deemed to be well-reproduced in simulated events, within uncertainties, through the comparison of simulated $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events with the bin-wise count of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events performed on real data. Thanks to the strong similarity between the channels $D_s^{\pm} \to \tau(3\mu)\nu$ and $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$, this validation extends to simulated $\tau \to 3\mu$ events. Four of the relevant validation plots are reported in Figure 5.3 for illustration. Detailed description of this validation procedure is part of an extensive document which is planned for publication in 2019, and is not the focus of the present thesis. A "mismodeling" systematic uncertainty is introduced to account for minor mismodeling effects of the BDT input variables in simulation (section 5.5.3).

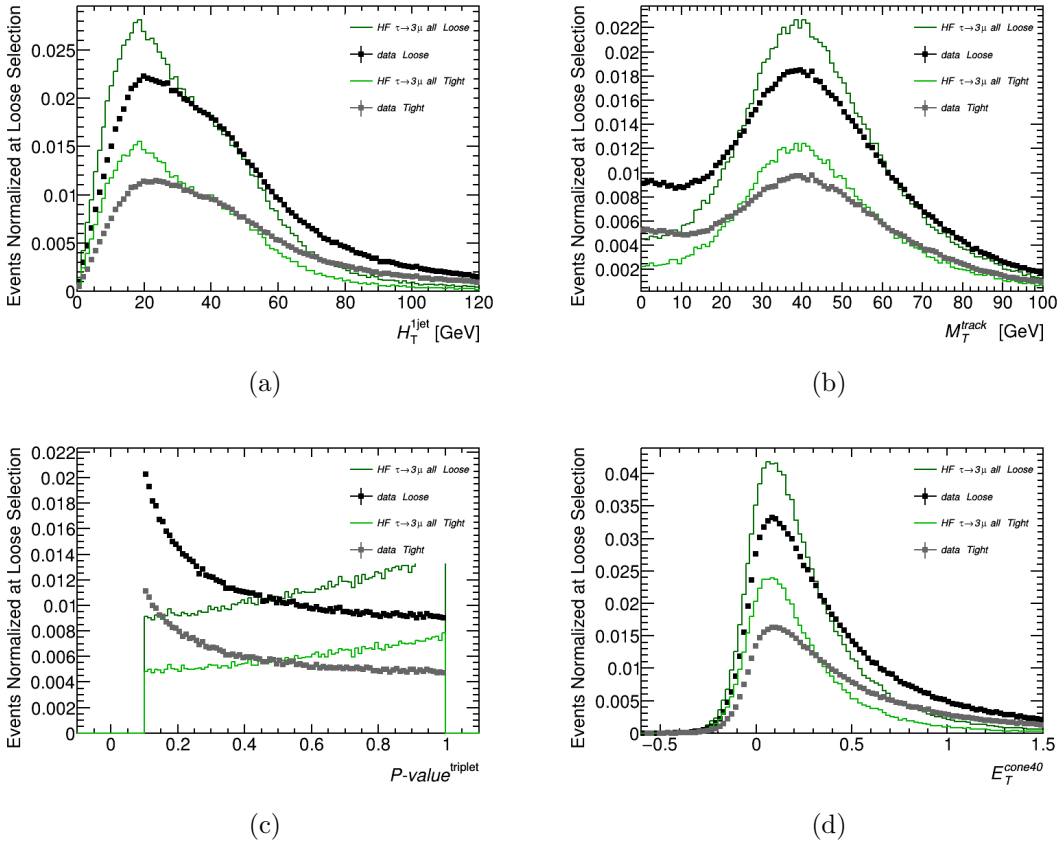


Figure 5.2: Illustration of four variables used for to the BDT training. The figures show simulated $D_s^{\pm} \to \tau(3\mu)\nu$ events and data collected in Run-2, representative of the backgrounds in the analysis. The distributions are normalized to unity integral after the application of the loose selection (Table 5.6). The same distributions are shown in light green and gray after the application of the tight selection.

5.4.2 BDT training

As inputs for the training of the multivariate classifier events are taken which pass the loose selection (5.3.4). For background data events are further reduced to a Training Region (TR, see Table 5.10), given by a relatively wide window around the τ mass (though events are excluded when within ∼90 MeV from it to avoid signal contamination, the mass resolution of the signal peak being about 40 MeV). The use of the TR means that training events are not too far from the kinematics of the signal (narrowly peaking at M_τ), supporting an optimal performance of the classifier.
The adopted classifier is, as already mentioned, a Boosted Decision Tree (BDT) method of the TMVA toolkit [169]. Gradient-based boosting is applied for the training, and the settings shown in Table 5.9 are applied. The settings are determined from performance of the resulting classifier amongst a range of choices, their values only minimally differing

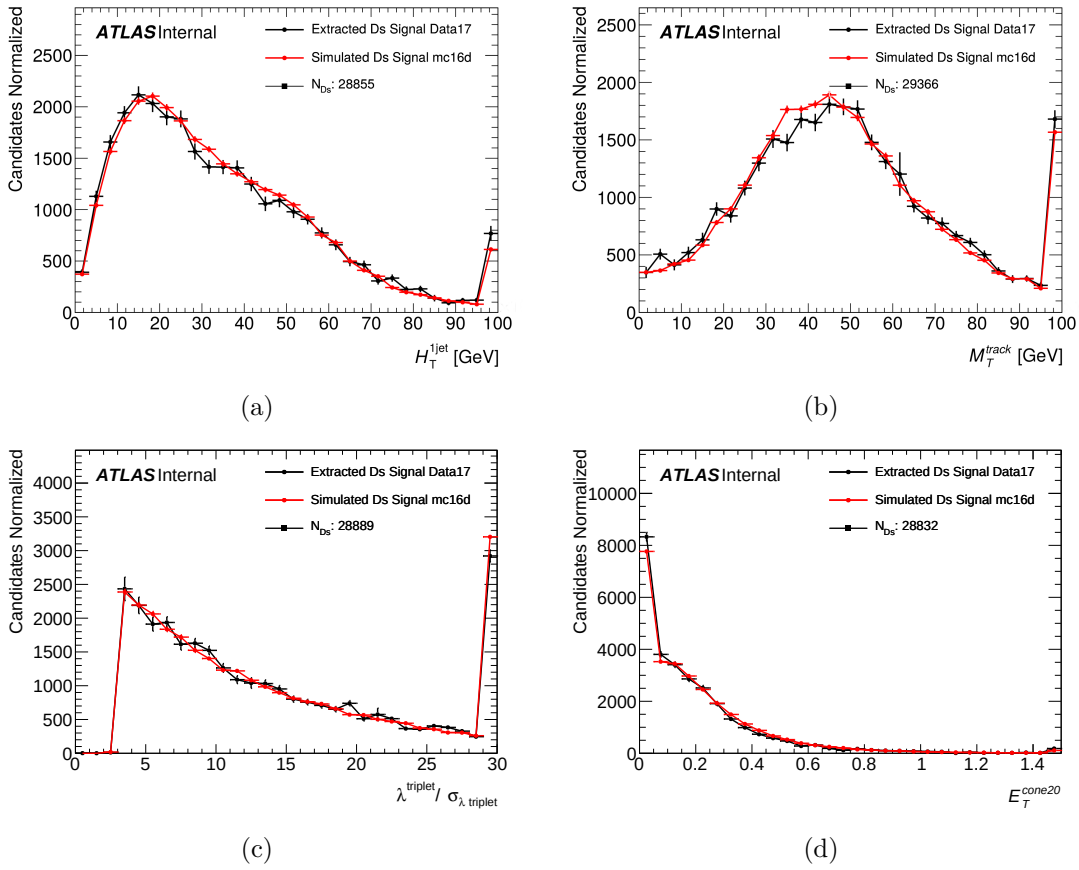


Figure 5.3: Distributions for the validation of BDT input variables, overlapping simulated $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events (red) with counts of the same decays observed in ATLAS data (black). The distributions show H_T^{1jet} (a), M_T^{track} (b), $\lambda^{\text{triplet}}/\sigma_{\lambda}^{\text{triplet}}$ (c) and E_T^{cone20}(d). The last bin of each plot includes events above the plotting range. The simulation of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ is generated as described in Appendix A, and undergoes the same reconstruction as the given $D_s^{\pm} \to \tau(3\mu)\nu$ signal. The number of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events in data is extracted with the use of a simultaneous fit to the distribution of the three particle invariant mass, like the one shown in Figure A.1a. However, the fit is here performed for separate bins of the variable under study. The distributions validate the reasonable reproduction of the shown quantities by the simulation. This allows their use in the multivariate classifier without need of further corrections to the simulation.

from the suggested defaults. The minimum *NodeSize* has been reduced from 5% to 1% and the allowed *MaxDepth* of the trees is increased to 5, to improve performance (the high number of training events, ∼500k for both signal and background, reduces the chance of overtraining).

The BDT classifier is cross-trained on 50% of the available events after loose selection. I.e. both signal and background events are randomly split into two exclusive sets each containing about 50% of the events. For the sake of this explanation they shall be

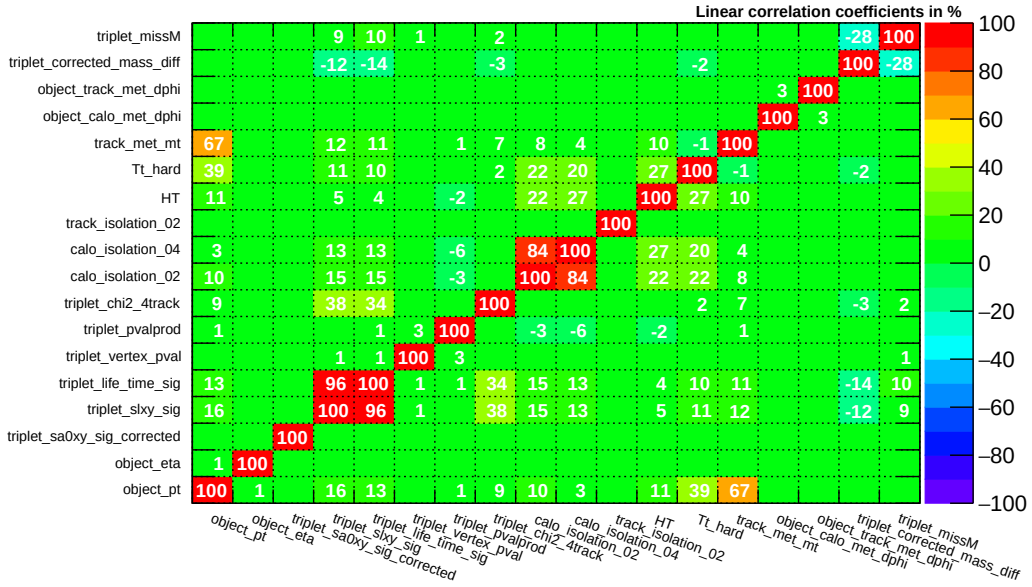
NTrees	=	1000
BoostType	=	Grad
Shrinkage	=	0.10
UseBaggedBoost		
BaggedSampleFraction	=	0.5
nCuts	=	20
NodeSize	=	1%
MaxDepth	=	5

Table 5.9: Configuration parameters employed for the BDT training.

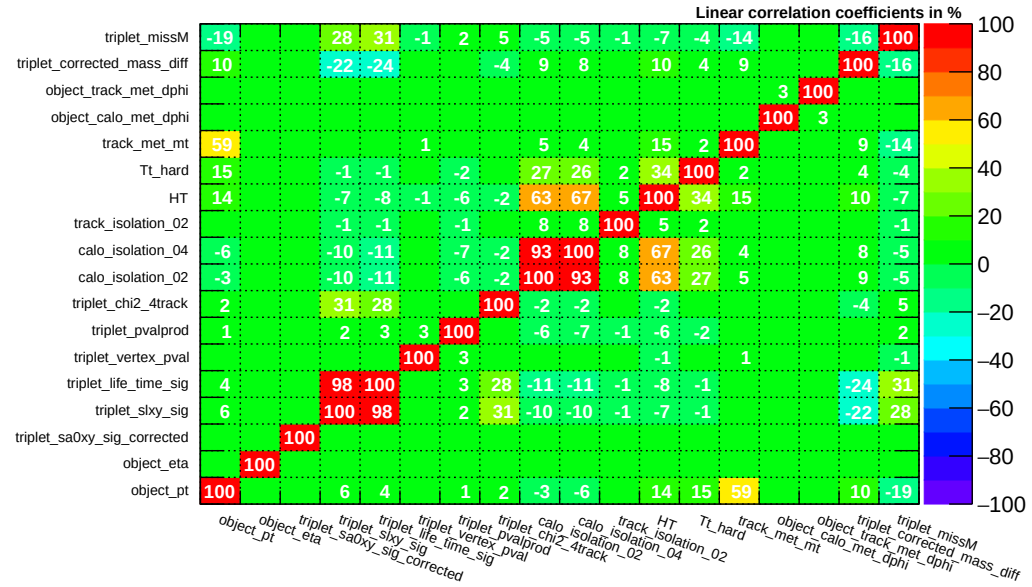
Training Bands (TB)	$M^{\text{triplet}} \epsilon [1400, 1690] \cup [1870, 2200]$ MeV
Side Bands (SB)	$M^{\text{triplet}} \epsilon [1450, 1690] \cup [1870, 2110]$ MeV
Signal Region (SR)	$M^{\text{triplet}} \epsilon [1713, 1841]$ MeV

Table 5.10: Definitions of the various M^{triplet} ranges relevant for the analysis. TB represents the region for background data events to be used in the BDT classifier training. SB are the bands where curves are are fitted to background events to estimate the expected background population in the signal region SR, as illustrated in the next section (5.4.3). The latter is a 64 MeV wide window around the τ mass, it contains most of the expected signal, and defines the "single bin" in which the expected upper limit on $\mathcal{B}(\tau \to 3\mu)$ is determined through a *counting experiment* (Section 5.5.2).

labelled set A and set B. A classifier (BDT_A) is then trained exclusively on the basis of signal and background data events of set A and another classifier (BDT_B) is trained on events of set B. At this point every event receives a classification score from the classifier whose training it has not contributed to. Thus events of set A receive a score from BDT_B, and events of set B receive a score from BDT_A. Because of the random subdivision of events for the training, the two classifiers have very similar performance. This enables the use of all scores as a unique classifier for the final selection (henceforth labelled $Score^{BDT}$). Figure 5.5 shows the habitual overlay between training and testing samples separately for the two BDT trainings. Figure 5.4 shows the linear correlation coefficients for the input variables, a measure of how independent the various inputs are from each other. A small number of high correlations indicate that the set of variables might be reduced without significantly affecting the classifier's separation power, though the overall picture is reasonable.

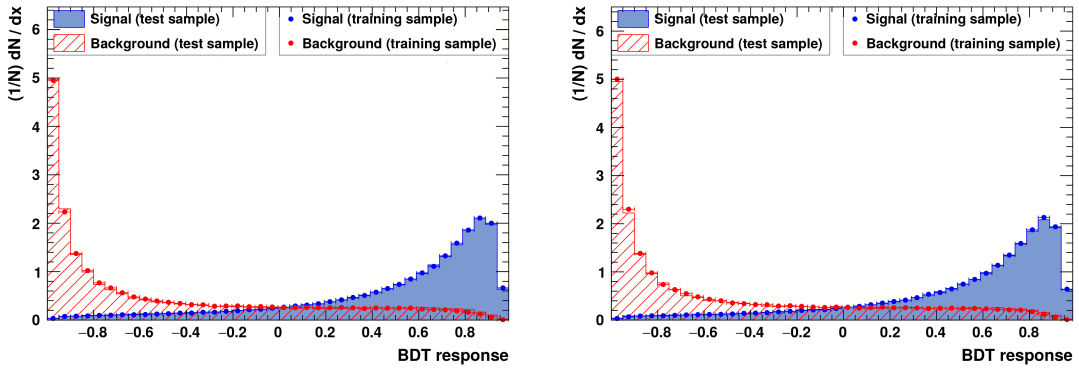


(a)



(b)

Figure 5.4: The images show the correlations between the input variables for HF signal sample(a) and for the background sample (b).



(a) (b)

Figure 5.5: Overlay of the BDT classifier's response for background data events (red) and for MC signal events (blue), both for the half sample employed for its training (dot markers) and the remaining events (filled histograms). Training and evaluation samples are inverted between (a) and (b), as described in the text.

5.4.3 Extrapolation of Expected background

The number of expected background events in the SR is obtained from the number of events found in the SB region, via a transfer factor f_{SB}^{transf}. The factor is determined fitting a 5^{th} degree Chebychev polynomial to the M^{triplet} distribution in the SB and taking the ratio between its integral over the SR and the SB (an unbinned fit is used for this).

$$f_{SB}^{\text{transf}} = \int_{SR} g(x)dx \;/\; \int_{SB} g(x)dx \tag{5.4}$$

The shape of the M^{triplet} distribution varies depending on the cut applied to $Score^{BDT}$, because of the different nature of the backgrounds populating individual parts of the M^{triplet} spectrum. For this reason the M^{triplet} fit is performed after the $Score^{BDT}$ cut that is eventually chosen ($Score^{BDT} > 0.8$). Figure 5.6 shows the fit result, while Figure 5.7 shows the data distributions of M^{triplet} obtained from the selections $Score^{BDT} > -0.8$ and $Score^{BDT} > 0$, to display the variation in shape. Two systematic uncertainties are associated with f_{SB}^{transf}: "fit" and "model". The first propagates the post-fit uncertainty on the polynomial's parameters onto f_{SB}^{transf}. The "model" uncertainty embodies the arbitrarity of the curve choice and is determined using 2^{nd}, 3^{rd} and 4^{th} degree polynomials, instead. Both types of uncertainty are displayed in Figure 5.6. The resulting transfer factor, with its uncertainties, is the following:

$$f_{SB}^{\text{transf}} = 0.249 \pm 0.016^{(fit)}\ {}^{+0.009}_{-0.016}{}^{(model)} \tag{5.5}$$

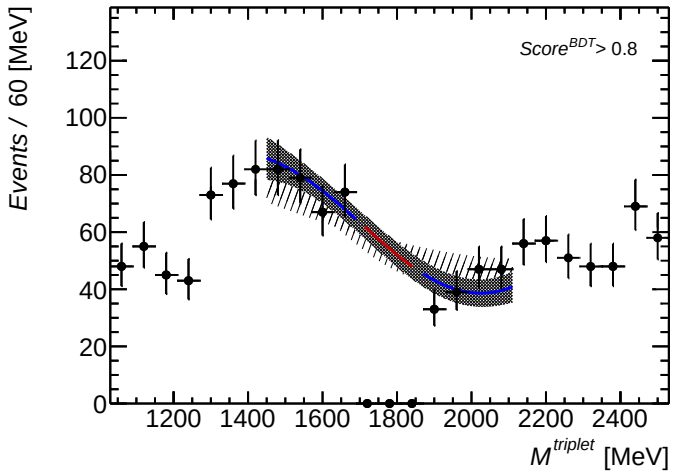
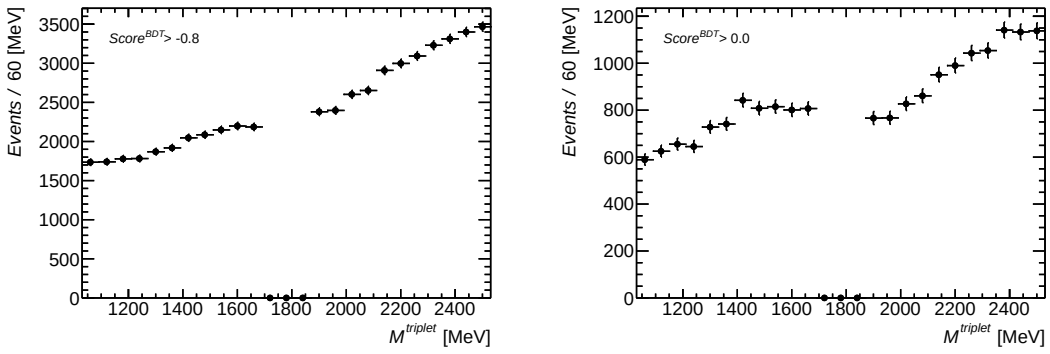


Figure 5.6: Result of a 5^{th} degree Chebychev polynomial fit to the M^{triplet} distribution of Run-2 selected data events passing $Score^{BDT} > 0.8$. Events in the range [1690, 1870] MeV are blinded. The data is subject to the full selection described in 5.3, including *tight* selection. The fitting function is shown in black within the SB region and in red in the SR. The "fit" uncertainty is indicated by a gray band, while the "model" uncertainty is shown by a dashed band.



(a) (b)

Figure 5.7: Distributions of M^{triplet} in full Run-2 data, as reference for its shape dependence on the $Score^{BDT}$ selection. The events are selected as in Figure 5.6, except for the selection $Score^{BDT} > -0.8$ in (a) and $Score^{BDT} > 0$ in (b).

5.5 Upper Limit Calculation

5.5.1 CLs Limit Setting

The measure for the analysis' success is its sensitivity to the observation of a $\tau \to 3\mu$ decay, or, conversely, to set a sufficiently low upper limit on the branching fraction $\mathcal{B}(\tau \to 3\mu)$, to be relevant with respect to the result of previous searches. This is achieved by determining an exclusion limit at 90% Confidence Level (CL) [172] on $\mathcal{B}(\tau \to 3\mu)$. This in turn is virtually identical to determining the maximum number of $\tau \to 3\mu$ events that can exist in the SR whilst leaving the total number of occurring events within 90% of the possible outcomes that would be expected from a population of only background events. For this analysis the limit calculation follows a so-called counting experiment (or one-bin experiment): the evaluation depends on the count of events obeying a specific selection, within the unknowns in the procedure (systematic uncertainties).
The statistical procedure (frequentist) adopted to evaluate the compatibility between a hypothesis or model and the corresponding measurement is based on the *likelihood function*, describing the probability of obtaining a specific measurement, under the assumption that a particular model is true.

$$L(n|s,b) = L(n|\mu,b) \tag{5.6}$$

To lowest order of approximation, the likelihood function for a counting experiment (formula (5.6)) is a Poissonian distribution with expected occurrence set at the sum of expected signal (s) and background counts (b), and the occurrence value set at the number of observed counts (n).
In order to to test the strength of a possible $\tau \to 3\mu$ signal, it is better to formulate the likelihood in terms of the signal strength μ (right-hand side of equation (5.6)), which, for this analysis is chosen to coincide with $\mathcal{B}(\tau \to 3\mu)$. This means that $\mu = 1$ corresponds to testing a model in which τ always decays into three muons, while $\mu = 0$ is what is called the zero-hypothesis: no LFV decay expected, only backgrounds.
The uncertainties encountered in the description of the signal and background affect the precision with which a model can be tested: the likelihood is, in fact, dependent on the value that *nuisance parameters* assume (further parameters on which the models depend, but which are not of interest). Nuisance parameters are expected to follow individual likelihood distributions, so the overall likelihood can be expressed as in equation (5.7).

$$L(n|\mu,b,\boldsymbol{\theta}) = P(n|\lambda(\mu,b,\boldsymbol{\theta})) \times C_{\text{syst}}(\boldsymbol{\theta},\boldsymbol{\theta}_0) \tag{5.7}$$

In the equation P indicates a Poisson distribution with n being the number of observed events, λ the expected occurrence determined by the signal strength μ, the background contribution b and the specific value assumed by all the relevant nuisance parameters $\boldsymbol{\theta}$ (with expectation value $\boldsymbol{\theta}_0$). The rightmost factor then represents the likelihood term of the nuisance parameters (ideally a product of Gaussians, as this can greatly simplify calculations).
The evaluation of the compatibility of the $\mu \neq 0$ hypothesis with data is determined

from the test statistic for positive signal exclusion (5.8) [173].

$$\widetilde{q}_\mu = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\hat{\boldsymbol{\theta}}}(\mu))} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{5.8}$$

This test statistic saturates against non-physical negative signal contributions (first line) and is one-sided (last term), as it is meant to set an upper limit on the signal strength. A single hat on variables indicates their Maximum Likelihood Estimate (MLE), while $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ represents the values of the nuisance parameters that maximise the likelihood with a fixed signal strength μ.
Given the test statistic in (5.8) one can associate a p-value to a possible measurement as:

$$p_\mu = \int_{n^{\text{observ.}}}^{\infty} f(\widetilde{q}_\mu|_\mu) dn \tag{5.9}$$

where $f(\widetilde{q}_\mu|_\mu)$ indicates the probability distribution function associated with $\widetilde{q}_\mu$ for fixed μ, expressed as a function of the observed number of events n. The integral covers the range of observed events between the experimentally observed number of events and infinity (the full range associated to signal strength equal or larger than the MLE value). The p-value p_μ thus represents the expected fraction of experiments that would lead to setting a higher upper limit on μ, than the performed one, and constitutes the confidence level for rejecting the chosen signal strength. To determine an upper limit with a particular CL this reasoning is reversed, finding the signal strength that has the desired p-value, given the observed measurement.
This procedure has one flaw: it might lead to the rejection of a signal even though the experiment would not agree any better with the zero-hypothesis of having only background events. For this reason a modified p-value p' is used in practice (also referred to as CL_s), to protect from this fallacy:

$$p'_\mu = \frac{p_\mu}{1 - p_b} \tag{5.10}$$

where p_b indicates the p-value associated with the background only hypothesis. The quantity $1 - p_b$ is known as the power of the test, and represents the probability to accept the background-only hypothesis, were it true.
Before determining an actual limit, the sensitivity to calculate the measurement is determined through the calculation of an expected limit. This background-only hypothesis is used to represent the median expected measurement with no fluctuations in absence of a signal (Asimov dataset). For this it is possible to calculate value and uncertainty using an asymptotic approximation to represent $f(\widetilde{q}_\mu|_\mu)$ [174] (accurate if the sample size n is large).

5.5.2 Calculation procedure and working point choice for BDT classifier

The value of $Score^{BDT}$ at which the final cut is applied is determined through a scan over the variable's range, in search of sensitivity to the smallest possible branching fraction (in other words the cut leading to the lowest expected limit). To this end the expected limit is estimated using the Asimov method introduced in the previous section. The specific application employed is an integration of the HistFitter framework [175] [176], constructed on RooFit and RooStats classes [177] [178].
Three primary quantities determine the sensitivity associated with any $Score^{BDT}$ cut: the number of expected background events that are expected in the SR "bin", the number of signal events expected to fall within the SR if the signal hypothesis is true with strength 1, and the number of actually observed events in the bin. To know the expected limit reachable with a specific selection, the number of observed events is substituted with the best estimate of the expected events in the background-only hypothesis, given by the extrapolated number of background events itself. The procedure to estimate the number of signal events and background events in the SR is now summarised.
Calculating the number of backgrounds is straightforward:

$$N^{Bkg}_{SR,\text{expected}} = N^{data}_{SB} \times f^{\text{transf}}_{SB} \tag{5.11}$$

i.e. the number of data events passing the full selection chain in the SB (defined in Table 5.10), multiplied by f^{transf}_{SB} (described in 5.5).
The calculation of the number of expected signal events is more articulated:

$$\begin{aligned} N^{\text{Signal}}_{SR,\text{expected}}(\mu) &= \sigma^{\tau(\text{HF})}_{p\text{-}p} \times \mathit{Accept.} \times \epsilon^{\text{Sig.}} \times \mathcal{B}(\tau \to 3\mu) \times \mathcal{L} \\ &= [\sigma^{D_s} \times \mathcal{B}(D_s \to \tau\nu) + \sigma^{b} \times \mathcal{B}(b \to \tau X)] \times \mathit{Accept.} \times \epsilon^{\text{Sig.}} \times \mu \times \mathcal{L} \\ &= \mu \times \epsilon^{\text{Sig.}}_{fiducial} \times [\sigma^{D_s}_{\text{norm}} \times \mathcal{B}(D_s \to \tau\nu) + \sigma^{b}_{\text{norm}} \times \mathcal{B}(b \to \tau X)] \times \mathcal{L} \\ &= \mu \times \frac{N^{\text{Sig.}}_{SR}}{N^{\text{Sig.}}_{generated}} \times [\sigma^{D_s}_{\text{norm}} \times \mathcal{B}(D_s \to \tau\nu) \times f^{B\to\tau X}_{D_s}] \times \mathcal{L} \end{aligned} \tag{5.12}$$

The first line shows the calculation in its most general from: the number of expected events is proportional to $\sigma^{\tau(\text{HF})}_{p\text{-}p}$ (the total production cross section for HF-channel taus at $\sqrt{s} = 13$ TeV), adjusted by $\mathit{Accept.}$ (the detector acceptance), $\epsilon^{\text{Sig.}}$ (the comprehensive trigger, reconstruction and selection efficiency for signal events), scaled by the branching fraction of $Ds \to \tau\nu$ and the one of $\tau \to 3\mu$, which by construction coincides with the signal strength μ, finally multiplied by the corresponding recorded luminosity $\mathcal{L}$. For the signal normalization $\sigma^{D_s}_{norm}$ is used, representing the fiducial production cross section at generator filter level (determined in equation (A.3)). Correspondingly the signal efficiency is expressed with respect to the generator filter ($\epsilon^{\text{Sig.}}_{fiducial}$), the latter is calculated from the number of simulated events that pass full selection and are localized within the SR. The D_s production cross-section σ^{D_s}, or more accurately $\sigma^{D_s}_{\text{norm}}$ is determined

experimentally from the study of the $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ channel. The production cross-section of taus via the secondary channel $b \to \tau X$ ($\sigma^{b\to\tau X}$) is determined partially from PDG values for branching fractions and partially through Pythia8 simulation [179]. To be precise $B^{0/\pm}$ decays to $\tau\nu D^{\pm/0(*)}$ are taken from PDG; remaining contributions are estimated through Pythia. Preliminary studies indicate that this channel is reasonably emulated by the combination of prompt and non-prompt D_s-originated $\tau \to 3\mu$ events. For this reason no ATLAS simulation of $B \to \tau X$ events is used in this analysis, instead the contribution of the $B \to \tau X$ channel is handled scaling the production cross section $\sigma^{D_s}_{\text{norm}}$ by a factor

$$f_{D_s}^{B\to\tau X} = 1.32 \pm 0.08 \tag{5.13}$$

Figure 5.8 displays the expected upper limit on $\mathcal{B}(\tau \to 3\mu)$ for a range of cuts on $Score^{BDT}$. Based on this distribution a cut of $Score^{BDT} > 0.8$ is chosen.

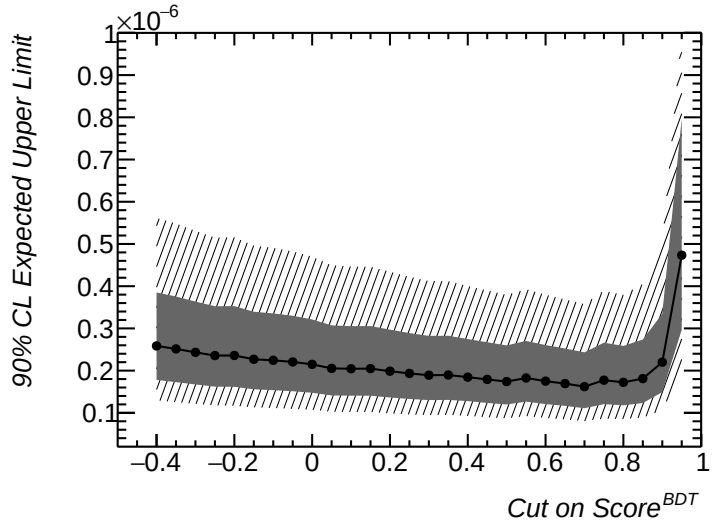


Figure 5.8: Scan of the expected upper limit on $\mathcal{B}(\tau \to 3\mu)$ for various cuts on $Score^{BDT}$. The values are determined through *asymptotic* (Asimov) implementation of HistFitter for a 1-bin experiment.

5.5.3 Systematic Uncertainty

A list of all systematic uncertainties applied for the calculation of the expected upper limit on $\mathcal{B}(\tau \to 3\mu)$ is shown in Table 5.11. The percentages indicate the relative variation of the corresponding sample (signal or background). For the data-driven background estimate the uncertainties are the ones linked to the SB to SR extrapolation, described in section 5.4.3. The signal's uncertainty is treated in two parts: a fully characterized production cross section uncertainty and a selection efficiency uncertainty. The latter accounts for trigger modelling, muon identification uncertainty, jet-energy/E_T^{miss}

uncertainties and the impact of minor mismodelling of BDT input variables on the multivariate classifier's performance. The efficiency uncertainty is conservatively estimated to be 15%, but two further values are used to show how this uncertainty affects the result: 10% for an intermediate scenario and 5% for a more optimistic take. The 15% scenario is close to the uncertainties faced in the Run-1 analysis, dominated by trigger efficiency uncertainty (∼11%) and selection efficiency uncertainty (∼8%); the latter was governed by the non-standard selections that were adopted through the use of muon trackparticles (see 3.3) [132]. The intermediate scenario (10%) reflects the expected improvement in trigger efficiency corrections, benefiting from higher statistics in data, and the avoidance of non-standard muon objects. Finally, the 5% scenario is a template for the best case achievable.

Type	Description	Relative Value
Signal:	Cross-section	±13%
	▷ D_s $\sigma_{\text{fiducial}}^{[15,100]}$	±9.8%
	▷ Normalization	±3.3%
	▷ $\mathcal{B}(D_s \to \tau\nu)$	±4.3%
	▷ $f_{D_s}^{B\to\tau X}$	±6.1%
	Efficiency (indicative)	±15%(±5%)
Background:	Fit	±6.3%
	Model	+3.7% −6.4%
Luminosity:	−	±1.7%

Table 5.11: Systematic uncertainties included in the current upper limit evaluation. The contributions to the signal production cross-section uncertainty are individually listed. The "Normalization" entry refers to the second term of equation (A.7). The uncertainty on the signal "Efficiency" refers to $\epsilon_{fiducial}^{Sig.}$ in equation (5.12) and is described in the text.

5.6 Result

An upper limit for $\mathcal{B}(\tau \to 3\mu)$ is determined with the presented form of the HF-channel analysis at ATLAS, making use of the full 139.0 fb^{-1} of integrated luminosity collected during Run-2.

Table 5.12 gives an overview of the selections used in the analysis and their impact on signal and background events, as well as the final event count in the SR.

The expected limit for the three scenarios introduced in section 5.5.3 are listed in Table 5.13. The values do not vary much between scenarios, indicating that the reach of the analysis is not dominated by the "Efficiency" systematic. The conservative 15% uncertainty scenario is used to calculate the observed limit.

The dataset has been unblinded revealing 120 events in the SR, slightly more than the expected 116.5, but well within the uncertainties (the shape of the unblinded distribution can be seen in Figure 5.9). This led to the observed limit:

$$\mathcal{B}(\tau \to 3\mu) < 1.98 \times 10^{-7} \quad (90\%\ \text{CL}) \tag{5.14}$$

The underlying asymptotic CL scan for the conservative scenario is shown in Figure 5.10.

Further improvement of the signal efficiency uncertainty will likely improve this limit (10% uncertainty being the most probable final value). Amongst the other uncertainties (Table 5.11) there is still room for improvement in the cross-section evaluation: the D_s cross-section measurement is not finalized and it is possible that the measurement will be extended downwards to D_s momenta of 10 GeV, allowing for a more direct derivation of the production cross section of HF-channel taus.
The background extrapolation procedure does not allow for particular gain: the uncertainty deriving from the transfer factor is smaller than the one found in Run-1, but the uncertainty is inversely proportional to the number of events available for the fit (one of the reasons the shape of the limit scan in Figure 5.8 results "flattened out"). What happened in the Run-1 analysis, and is expected to repeat in a Run-2 W-channel analysis (see section 3.2), is that, as only a handful of background events remain after the final selection, the extrapolation procedure becomes more involved, leading to further sources of uncertainty.

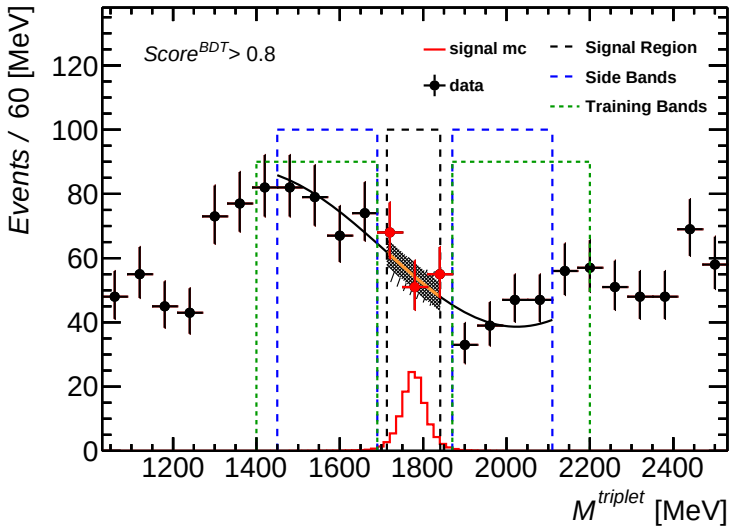


Figure 5.9: The three-muon invariant mass distribution in the range $M^{\text{triplet}} \in [1030, 2530]$ MeV. Signal, shown in the red histogram, and data events, shown as solid dots, pass the *tight* selection (section 5.3) and the additional cut $Score^{BDT}>0.8$ as indicated. Red markers denote data points inside the blinded region, disclosed only at unblinding. The signal histogram is scaled to $\mathcal{B}(\tau \to 3\mu) = 2\times10^{-7}$. The polynomial fit used to extrapolate the expected background into the SR is shown in black, its section representing the expected background in the SR is marked in orange. The fit uncertainty is shown shadowed, while the model uncertainty is dashed. The region covered by the SR is marked by a black-bordered box, while the SB are shown with a blue border. The Training Bands (TB) are similarly marked by a green border.

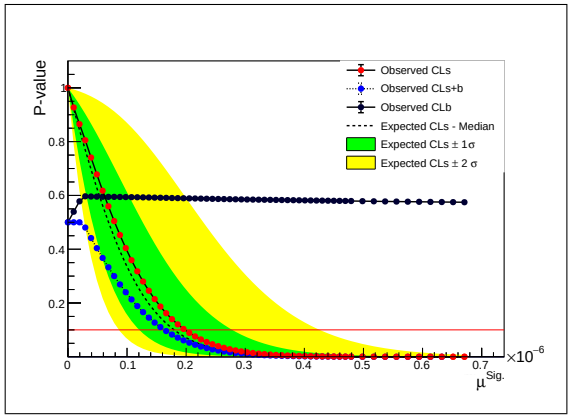


Figure 5.10: CL_s p-values for the range of interest of the $\mathcal{B}(\tau \to 3\mu)$ signal strength. The red points mark the CL_s for the observed events. The dashed line marks the expected CL_s and corresponding 1σ and 2σ bands are shown in green and yellow, respectively. The (expected) upper limit is determined at 90% Confidence Level by the intersection of the relative CL_s curve with the red line at p-value 0.1 ($1 - 90\%$).

Cut-Flow		
Cut	Signal (Weighed)	Data Events
Event-Level sel.	772878.8	3486568
Muon Quality sel.	613137.2	1663249
Loose sel.	499105.3	497807
Tight sel.	261838.8	256208
$Score^{BDT} > 0.8$	82028.0	2437
In Side-Bands (SB)	2278.9	468
In Signal Region (SR)	75912.1	$116.5^{\dagger}$
Observed (SR)	–	120

Table 5.12: Event count after each selection. The numbers show the relative impact of the selections on signal and backgrounds. Data events are subject to blinding ($M^{\text{triplet}} \notin [1690, 1870]$ MeV), except for the observed events listed in the last line. The number of data events in the SR (†) is determined applying the transfer factor. The last two lines show that 120 events have been observed in the SR, while 116.5 were expected, as described in the text. Signal events are not scaled to production cross-section.

The reach of the analysis is limited by the ability to separate signal events from background ones, problematic due to the challenging similarity between the two. Any further contribution to the separation power would directly lower the branching fractions that can be probed. To this end a plan is ongoing to further adjust the multivariate method used in the final selection (including the probable switching to the use of neural network training).
Another source of improvement might be to extend the list of triggers to include *prescaled* ones, increasing the signal throughput by about 10%, though it remains to be verified

that the extra triggers do not negatively impact the trigger efficiency uncertainty.

Scenario	Expected Limit ($\times 10^{-7}$)	1σ up	1σ down
Eff. uncertainty = 5%	1.75	2.63	1.19
Eff. uncertainty = 10%	1.76	2.68	1.20
Eff. uncertainty = 15%	1.80	2.78	1.21

Table 5.13: Expected upper limit on $\mathcal{B}(\tau \to 3\mu)$ and 1σ variations for the three scenarios considered for the signal efficiency uncertainty.

Discussion and Conclusion

"Now I know: don't be scared granny is right, just be prepared. Isn't it nice to know a lot! And a little bit not."

— Stephen Sondheim, *I know things now (lyrics)*

This thesis presents a preliminary version of the search for the LFV decay $\tau \to 3\mu$ with the ATLAS detector, based on the full p-p collision data collected at $\sqrt{s} = 13$ TeV in Run-2 (139 fb^{-1}). The work focuses on taus produced in heavy-flavour hadronic jets (HF-channel) and constitutes an unicum at ATLAS. To search for the signal in this channel secondary vertices are constructed from the ID tracks of three-muon systems and derived vertexing quantities are employed to separate the kinematics of $\tau \to 3\mu$ decays from the abundant background of SM hadronic showers. In the interest of efficiently selecting signal events, dedicated triggers were developed to trigger on events with di-muon systems with transverse momenta as low as 6 (11) GeV for the softer (more energetic) muon and systems of three muons with p_T of just 4 GeV. In addition reconstruction algorithms were improved since Run-1, leading to a ∼20% improvement in the reconstruction of muon triplets, rendering it unnecessary to resort to lower-purity incomplete muon objects (as was done in Run-1).
The expected limit on $\mathcal{B}(\tau \to 3\mu)$ is determined from the count of signal events passing all selections and falling within the Signal Region (SR): $M^{\text{triplet}} \in$ [1713,1841] MeV, as shown in Figure 5.9 (M^{triplet} is the invariant mass of the three muons' refitted ID tracks). The number of background events expected in the SR is determined from the data events falling in two Side Bands (SB: $M^{\text{triplet}} \in [1450, 1690] \cup [1870, 2110]$ MeV) via the unbinned likelihood fit of polynomial curves.
Machine learning is employed, in the form of Boosted Decision Trees (BDTs), to attain separation power from a number of variables, displaying differences between signal and backgrounds (section 5.4). Amongst the quantities in question are characterizations of the displacement of the τ-decay vertex, its goodness of fit, the isolation of the τ candidate from other energetic particles in the event and two mass-estimates of the particle originating the τ which aim to account for the invisible neutrino, pair-produced with the τ.
To model the signal, the process $D_s^{\pm} \to \tau(3\mu)\nu$ is simulated with Pythia8 and scaled to the fiducial production cross-section of D_s with p_T above 10 GeV, measured on 2017 data, through the study of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events (Appendix A) [180]. The secondary

signal contribution from $B \to \tau X$ is handled scaling the $D_s^{\pm} \to \tau(3\mu)\nu$ signal by a factor $f_{D_s}^{B \to \tau X}$ [179]. No separate Monte Carlo modelling of this signal contribution is used.

A measure of the sensitivity of the analysis, in the absence of an observed signal, is expressed by the expected upper limit on the on the branching fraction $\mathcal{B}(\tau \to 3\mu)$, which is evaluated to 1.80 $\times 10^{-7}$ at 90% CL (for the conservative scenario).
The SR has been unblinded and the observed upper limit on $\mathcal{B}(\tau \to 3\mu)$ is of 1.98 $\times 10^{-7}$ at 90% CL (the data points disclosed during unblinding can be seen in Figure 5.9). This observed limit is slightly higher than the expected one, as about four events more than expected were observed, but well within the 1σ band, supporting the validity of the background extrapolation procedure.
The use of the HF-channel is a novel approach at ATLAS and, although the BDT is not able to provide a very strong separation between signal and backgrounds, this analysis already expects to give a better result than the limit set previously by ATLAS.
The BDT training shown here represents a baseline configuration: a Neural Network approach is currently under consideration, aiming for a significant improvement to signal efficiency and background rejection.
In view of the ongoing effort, this search remains blinded and only major sources of systematics are treated. The uncertainties on the signal selection efficiency are replaced with three effective uncertainty scenarios. The accessible limit displays a small (percent level) dependence on this uncertainty.

The presented analysis is limited by the difficulty in separating the topologies of signal and backgrounds. The introduction of the IBL in Run-2 (2.4.1) has improved the resolution on the transverse displacement of the secondary vertex formed by the three signal muons (see a_0^{xy} and L^{xy} in 4.2.3), though the overall resolution of reconstructed tracks is substantially unaltered.
Next to the HF-channel search, a Run-2 W-channel is being pursued. This search has similarities with the Run-1 ATLAS analysis [64], and is expected to have a reach of the order of 4×10^{-8} [181] (the value is determined from the Run-1 result, based on the increase in integrated luminosity, W production cross-section and the improvements in trigger efficiency). Considering the improvement expected for the eventual HF-channel analysis, it is likely that the two channels will contribute about equally to sensitivity of the Run-2 analysis.

Appendix A

Normalization to D_s cross-section

A.1 Introduction

Topic of this appendix is the determination of a normalization factor for the $D_s^{\pm} \to \tau(3\mu)\nu$ signal to the D_s production at ATLAS, used in equation (5.12).
To determine this factor a measurement of the (fiducial) differential production cross section of D_s mesons at $\sqrt{s} = 13$ TeV p-p collisions in ATLAS is used. Section A.2 briefly outlines the measurement of the production cross-section. Due to some variations in the procedure adopted, the fiducial measurement presented here differs from what is going to be reported in [180], an upcoming publication of a Run-2 ATLAS measurement of the (unfolded) production cross section of D_s at $\sqrt{s} = 13$ TeV. The differences are highlighted in the text.
Section A.3 reports the few steps needed to obtain the normalization factor applied to the HF-channel signal simulation and its value.

A.2 Fiducial cross section of D_s

The calculation presented here differs from [180] in the use of a single trigger and 2017 ATLAS data, while the publication is going to rely on several triggers and full Run-2 dataset (thus leading to a reduction in the statistical error on $\sigma_{\text{fiducial}}^{[15,100]}$, reported in equation (A.3)). Other differences include the extension of the measurement to a $p_T^{D_s}$ of 10 GeV and changes to the generation of the $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ sample (to attain reliability at meson momenta $\sim$10 GeV). In the note the generator filter's pseudorapidity acceptance is further changed to $|\eta| < 2.5$ (see MCDsFilter in Table A.1 and A.3), introducing an ulterior difference with respect to the $D_s^{\pm} \to \tau(3\mu)\nu$ signal sample.

A.2.1 Data sample and simulation

The calculation presented here relies exclusively on 2017 data, selected by the single unprescaled trigger: HLT_mu11_mu6_bDimu2700. A corresponding simulation sample of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events is used. This sample is produced with Pythia8 with A14

tune [162] and NNPDF23LO pdf set [163] (the PythiaB wrapper is used to handle the simulation of events involving bottom processes [164]). A two-level filter is applied to the simulated events, as reported in Table A.1. The first requires generated D_s mesons to have transverse momentum larger than 10 GeV, and the second selects on the muons produced in the ϕ decay and the pion. This filter both requires a minimum transverse momentum and rejects events where the products are outside the detector's pseudorapidity coverage. The momenta selected are chosen on the basis of muon trigger acceptance and minimal track reconstruction quality for the pion.
The efficiency of these filters is determined based on a second set of filter-less "truth-only" simulation samples. These avoid any simulation of the ATLAS detector, energy depositions or shower development, while providing full information about the generated events. Events of both (full) simulation and data are subject to the building of secondary vertices described in 5.3.1, prior to further selections.

MCDsFilter:	$P_T^{D_S} > 10$ GeV
	$\lvert\eta^{D_S}\rvert < 3$
MC441Filter:	$P_T^{\mu 1}, P_T^{\mu 2} > 4$GeV
	$P_T^{\pi} > 1$GeV
	$\lvert\eta^{\mu 1}\rvert, \lvert\eta^{\mu 2}\rvert, \lvert\eta^{\pi}\rvert < 3$

Table A.1: Filtering selections applied to simulated $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events at their generation.

A.2.2 Selections

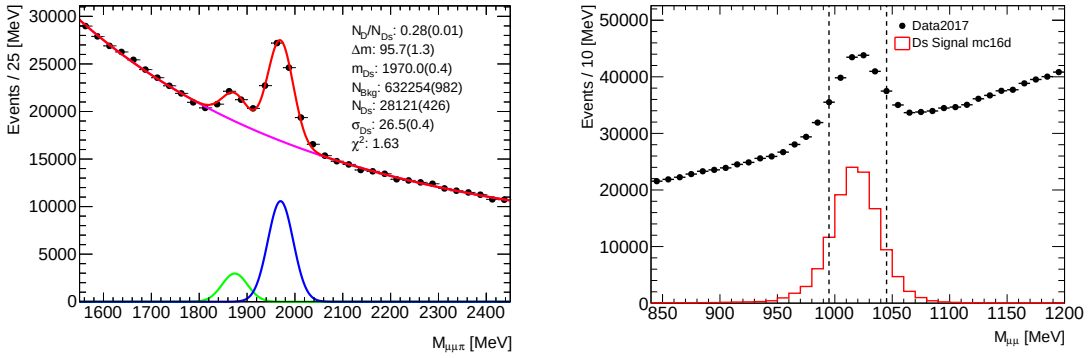
Events are subject to the selections in Table A.2 to reduce the background contamination in data. Particularly effective is the "ϕ selection" (last entry), which selects a 50 MeV window around the ϕ mass for the invariant mass of two muons. Its effect can be appreciated in Figure A.1b, where this variable is shown for data and $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ simulation, prior to this selection.

μ/π selection:	$P_T^{\mu 1}, P_T^{\mu 2} > 4$ GeV	
	$P_T^{\pi} > 1$ GeV	
triplet selection:	$P_T^{\text{triplet}} > 10$ GeV	$\lvert Q^{\text{triplet}}\rvert = 1$
	$\lvert\eta^{\text{triplet}}\rvert < 2.7$	$L^{xy}/\sigma_{L^{xy}} > 1$
	1.5 GeV $< M^{\text{triplet}} <$ 2.5 GeV	$a_0^{xy}/\sigma_{a_0^{xy}} < 4$
	$P\text{-}value^{\text{triplet}}_{\text{Vertex}-\text{fit}} > 10^{-3}$	
	Single vertex per event: maximum $P\text{-}value^{\text{triplet}}_{\text{Vertex}-\text{fit}}$	
ϕ selection:	$\lvert M^{\mu 1,\mu 2} - 1020 MeV\rvert < 25 MeV$	

Table A.2: Selection to isolate $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events.

A.2.3 Procedure

The $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ cross section is determined differentially, within nine bins of the meson's p_T (visible in Figure A.2 and A.3). The width of the transverse momentum bins is of 5, 10 and 20 GeV, larger widths being used at higher momenta, to maintain sufficient statistics for the measurement (in data). Within each p_T bin, the number of reconstructed $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events is determined fitting the D_s resonance in the M^{triplet} distribution. For the fit an exponential curve is used to model the falling backgrounds and two gaussians are used to capture the peak of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ and the partially overlapping $D^{\pm} \to \phi(2\mu)\pi^{\pm}$ (Figure A.1a). The differential cross section is then



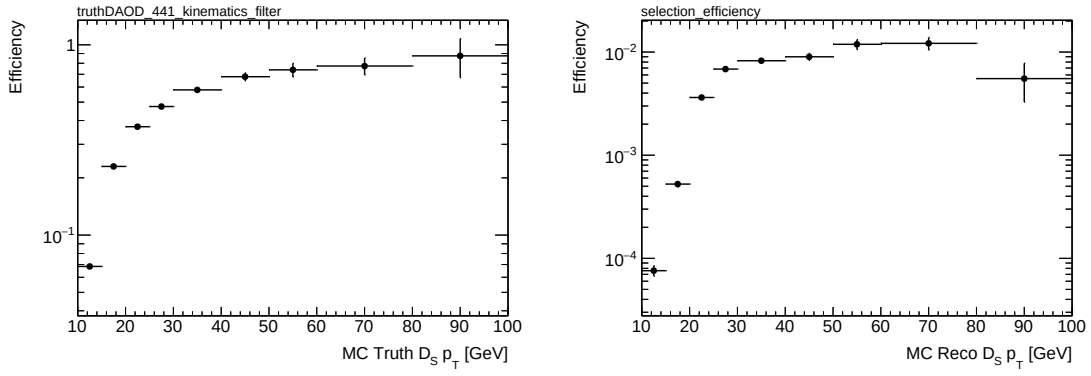
(a) (b)

Figure A.1: Figure (a) illustrates the fit procedure to determine the number of $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events reconstructed at ATLAS. The dots correspond to 2017 data after application of the selections of Table A.2; the curves are the result of a fit involving two Gaussians (in green for $D^{\pm} \to \phi(2\mu)\pi^{\pm}$ and in blue for $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$) and an exponential (in magenta). The red curve is the sum of all three.
Figure (b) illustrates the ϕ mass selection, showing the di-muon invariant mass distribution in data 2017 events (black) and corresponding $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ simulation (red), previous to this cut. The normalization of the two distributions is arbitrary. The cut window is indicated by dashed lines.

determined through the equation:

$$\frac{d\sigma}{dp_{\text{T}}} = \frac{N_{Ds}}{\int Ldt \times \epsilon_{\text{MC441Filter}} \times \epsilon_{\text{MCreco}} \times \mathcal{B}(D_S \to \phi\pi \to 2\mu\pi) \times \text{BinWidth}_{p_T}} \tag{A.1}$$

where $\epsilon_{\text{MC441Filter}}$ denotes the efficiency of the second filter in Table A.1 for the specific momentum bin and is shown in Figure A.2a (determined from "truth-only" samples). The term ϵ_{MCreco} is the "reconstruction" efficiency; it is determined from $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ simulation and incorporates all effects of detector acceptance, particle reconstruction efficiency, trigger efficiency and the selections listed in the previous section. The values of this efficiency are shown in Figure A.2b.
$\mathcal{B}(D_S \to \phi\pi \to 2\mu\pi)$ is the branching fraction of D_s into $\phi\pi$ multiplied by the branching



(a) (b)

Figure A.2: $p_T^{D_S}$ bin-wise efficiency of generator filter *MC441Filter* (a) and of combined detector acceptance, triggering and event reconstruction (b). Both distributions are constructed on the basis of simulated $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events.

fraction of ϕ into 2μ. For the branching fractions the PDG values are used [136]. Its value is of $1.29 \pm 0.14 \times 10^{-5}$, the uncertainty being dominated by $\mathcal{B}(D_S \to \phi\pi)$ and constituting one of the leading sources of uncertainty in the cross-section measurement.

A.2.4 Result

The determined differential cross-section is shown in Figure A.3. From this, the fiducial cross-section values are determined summing over the various bins:

$$\sigma_{\text{fiducial}}^{[\text{Threshold},100]} = \sum_{\text{Threshold}}^{100GeV} \frac{d\sigma}{dp_{\text{T}}} \Delta p_{\text{T}} \tag{A.2}$$

For the normalization of $\tau \to 3\mu$ the cross-section the fiducial region above 15 GeV is used; the first bin ([10,15] GeV) is deemed unreliable at this stage due to unresolved systematic effects related to trigger modelling and incomplete coverage of the simulation. The resulting value is the following:

$$\sigma_{\text{fiducial}}^{[15,100]} = 2.7 \pm 0.3 \times 10^9 \text{fb} \tag{A.3}$$

This value is going to differ from [180] by about 10%, as the note is based on a fiducial volume with $\eta^{D_s} < 2.5$.

[1] GM-VFN stands for general-mass variable-flavour-number scheme.

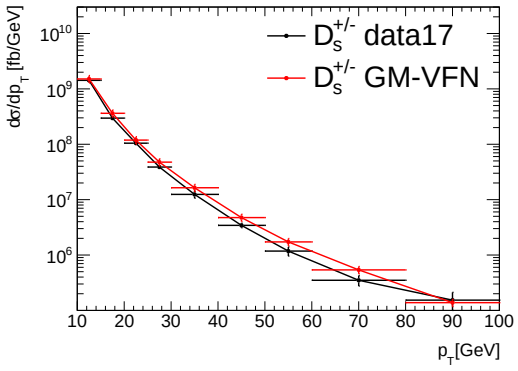


Figure A.3: $p_T^{D_s}$ bin-wise fiducial cross section for the production of energetic D_s mesons at ATLAS. The distribution determined from 2017 p-p collision data is shown in black and the GM-VFN[1] prediction is shown in red. The estimate is not extended beyond 100 GeV due to high uncertainty associated with low statistics.

A.3 $D_s \to \tau(3\mu)\nu$ Normalization

It is not entirely trivial to use the result of equation (A.3) to normalize the $D_s^\pm \to \tau(3\mu)\nu$ signal, due to differences in the generator filtering used for the two signals.
Firstly, the $\tau \to 3\mu$ analysis uses a simulated signal with $p_T^{D_S}$ as low as 10 GeV, while the fiducial cross section given in (A.3) refers to D_s momenta above 15 GeV. Therefore the value of the cross section has to be extrapolated to this value.
A second discrepancy is given by the difference in what can be called the "kinematic filter" applied in the generation of the two samples. $D_s^\pm \to \phi(2\mu)\pi^\pm$ simulation is subject to the filter *MC441Filter* (reported in Table A.1), while $D_s^\pm \to \tau(3\mu)\nu$ simulation is filtered through *MCMixedFilter* (explained in the next section).[2]
The next subsection describing the *MCMixedFilter* and reports its efficiency for $D_s^\pm \to \tau(3\mu)\nu$ signal, this appendix then closes with the factors required to normalize the $\tau \to 3\mu$ signal to the D_s production cross-section (A.3).

A.3.1 Filter efficiency

The two-level filter applied during $D_s^\pm \to \tau(3\mu)\nu$ event generation, is reported in Table A.3. Its particularity is the acceptance of events on the basis of two separate strategies: a milder momentum requirement for all three muons, or a "staggered" momentum distribution. The filter is motivated by the analysis triggers: the flat requirement ensures events that will pass "3mu4" triggers while the staggered selection corresponds to "mu11_mu6" triggers.

[2]It is interesting to notice that even if the simulated events of $D_s^\pm \to \phi(2\mu)\pi^\pm$ and $D_s^\pm \to \tau(3\mu)\nu$ had been using the same filter, a slight difference in their kinematics would unavoidably have led to different filter efficiencies (the difference is due to the μ/π mass gap, the extra neutrino in $D_s^\pm \to \tau(3\mu)\nu$ and the resonant nature of the intermediate ϕ, in the case of $D_s^\pm \to \phi(2\mu)\pi^\pm$).

MCDsFilter:	$P_T^{D_S} > 10$ GeV $\lvert\eta^{D_S}\rvert < 3$
MCMixedFilter:	$P_T^{\mu 1}, P_T^{\mu 2}, P_T^{\mu 3} > 3.5$ GeV OR $P_T^{\mu 1} > 10.5$ GeV, $P_T^{\mu 2} > 5.5$ GeV, $P_T^{\mu 3} > 2$ GeV $\eta^{\mu 1}, \eta^{\mu 2}, \eta^{\mu 3} < 3$

Table A.3: Filtering selections applied to simulated $D_s^{\pm} \to \tau(3\mu)\nu$ events at their generation.

The full efficiency of *MCMixedFilter* is determined from simulation to be:

$$\epsilon_{\text{MCMixedFilter}} = 6.1\% \pm 0.2\% \tag{A.4}$$

Analogously to the procedure used for $\epsilon_{\text{MC441Filter}}$ in A.2.3, the calculation is based on a "truth-only" simulated sample.

A.3.2 Normalization Factor

Because the fiducial cross-section reported in (A.3) refers to $p_T^{D_S}$ above 15 GeV an extrapolation factor is required to extend the value to a threshold of 10 GeV.
The calculation of this factor is straightforward: precisely like it has been done for the *MC*-efficiencies, the number of events passing a $p_T^{D_S} > 15$ GeV cut is divided by the number of events passing the *MCDsFilter* (for simulated $D_s^{\pm} \to \tau(3\mu)\nu$ as $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ alike). The determined efficiency is:

$$\epsilon^{15}_{\text{MCDsFilter}} = 25.2\% \pm 0.4\% \tag{A.5}$$

Writing the expression for the expected number of $D_s^{\pm} \to \tau(3\mu)\nu$ events it simple to introduce a unique normalization factor (or cross-section) $\sigma^{D_S}_{\text{norm}}$, which permits to scale the simulated $D_s^{\pm} \to \tau(3\mu)\nu$ events to the production cross-section of D_s.

$$\begin{aligned} N^{(D_S)\tau\to 3\mu}_{\text{expected}} &= \sigma^{[15,100]}_{\text{fiducial}} \times \frac{\epsilon_{\text{MCMixedFilter}}}{\epsilon^{15}_{\text{MCDsFilter}}} \times \epsilon^{(D_S)\tau\to 3\mu}_{\text{MCreco}} \times \mathcal{B}(D_S \to \tau\nu) \times \mathcal{B}(\tau \to 3\mu) \times \mathcal{L} \\ &= \sigma^{D_S}_{\text{norm}} \times \epsilon^{(D_S)\tau\to 3\mu}_{\text{MCreco}} \times \mathcal{B}(D_S \to \tau\nu) \times \mathcal{B}(\tau \to 3\mu) \times \mathcal{L} \end{aligned} \tag{A.6}$$

Inserting the values described so far:

$$\sigma^{D_S}_{\text{norm}} = \sigma^{[15,100]}_{\text{fiducial}} \times \frac{\epsilon_{\text{MCMixedFilter}}}{\epsilon^{15}_{\text{MCDsFilter}}} = 6.5 \pm 0.6 \times 10^8 \text{fb} \tag{A.7}$$

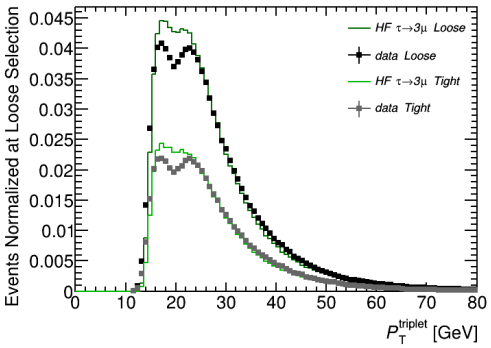
The limitation of the fiducial cross section to $p_T^{D_S} < 100$ GeV is of no consequence, because of the virtual absence of energetic events outside this selection in the finite-statistics simulation of $D_s^{\pm} \to \tau(3\mu)\nu$ events (the impact is estimated to be of the order 0.01%, irrelevant in comparison with the normalization uncertainty).
One might argue that determining $\epsilon^{15}_{\text{MCDsFilter}}$ from simulation may be affected by the

same misrepresentation of the simulation at low D_s momenta which plagued the [10, 15] GeV bin of A.3. While this is indeed true, using this value in equation A.7 cancels this effect out. Furthermore, in the $\tau \to 3\mu$ analysis events with $p_T^{D_S} < 15$ GeV have no impact, as virtually none of them passes the selections (this can be seen for instance in Figure B.1a).

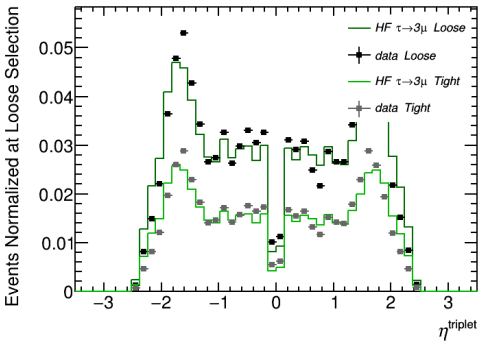
Appendix B

Full Set of BDT Inputs

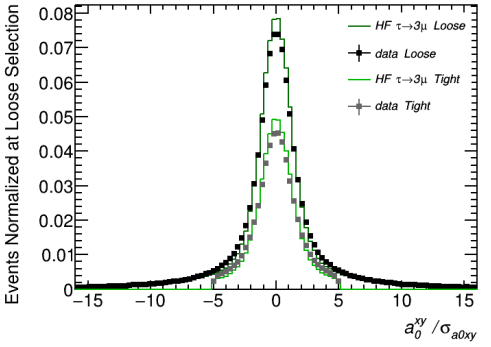
Distributions of all input variables employed for the training of the BDT discriminator used in the HF-analysis. The variables are listed and described in Table 5.8.



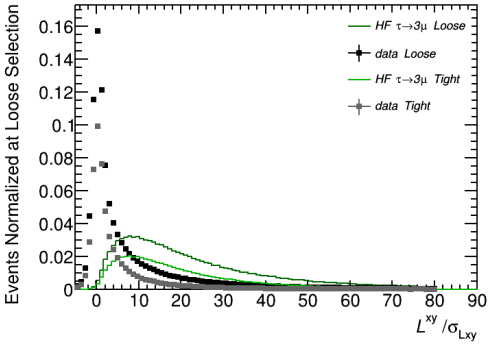
(a)



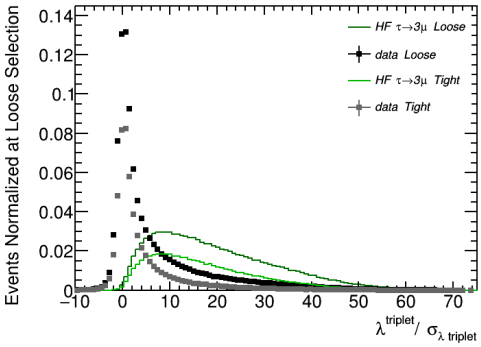
(b)



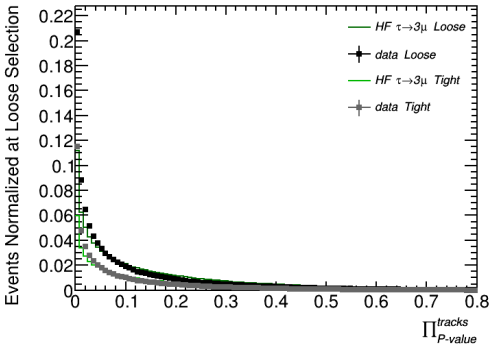
(c)



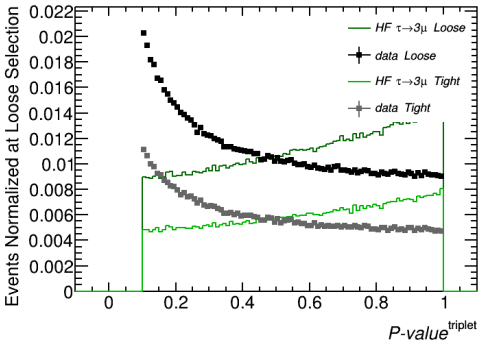
(d)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
λ triplet/ σλ triplet

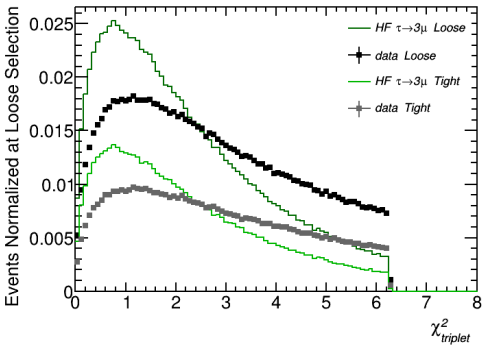
(e)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
Π tracks P-value

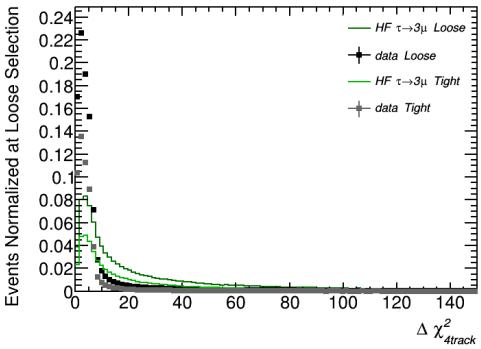
(f)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
P-value triplet

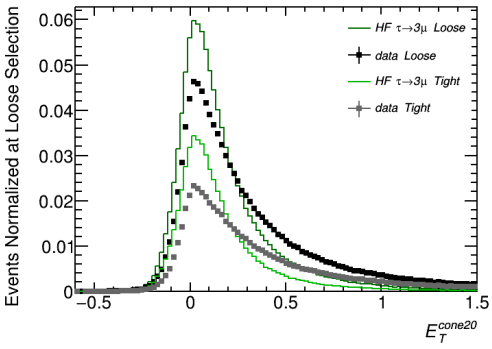
(g)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
χ2 triplet

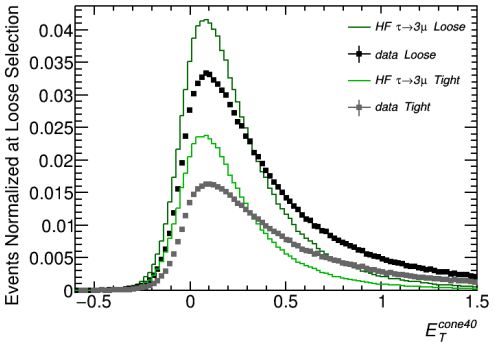
(h)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
Δ χ2 4track

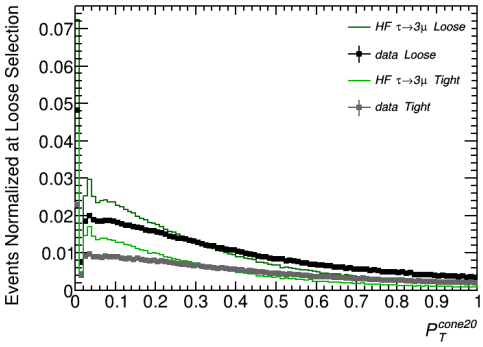
(i)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
ET cone20

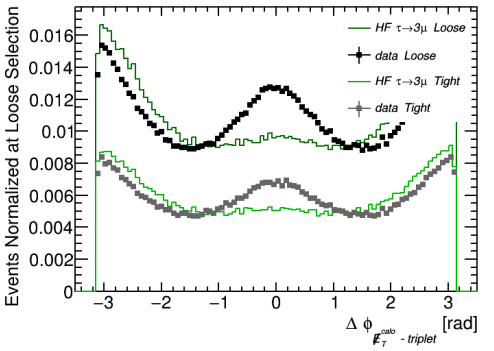
(j)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
E_T^{cone40}

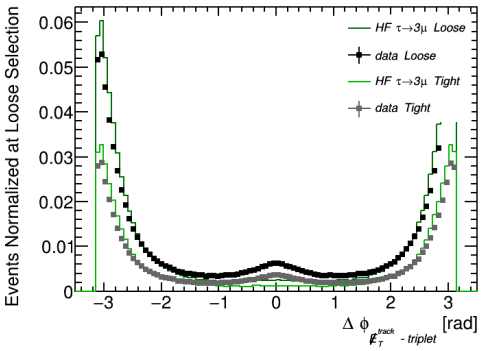
(k)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
P_T^{cone20}

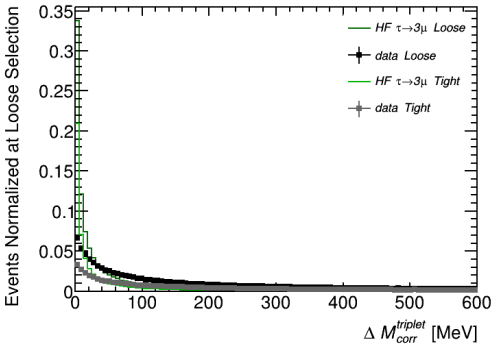
(l)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
Δ φ $\not{E}_T^{calo}$ - triplet [rad]

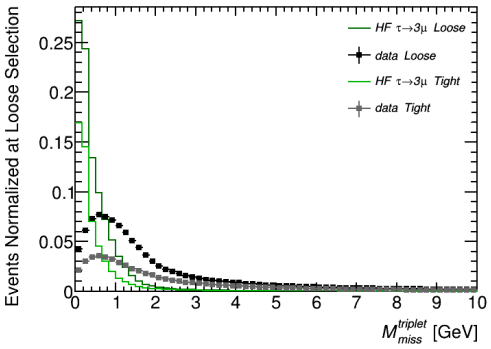
(m)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
Δ φ $\not{E}_T^{track}$ - triplet [rad]

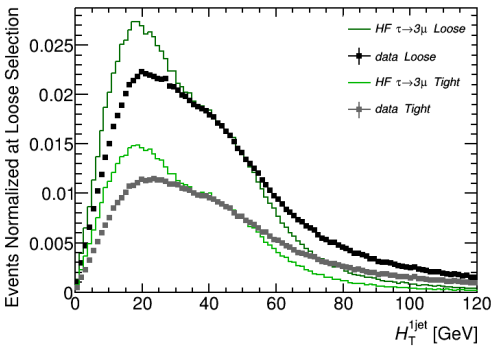
(n)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
Δ $M_{corr}^{triplet}$ [MeV]

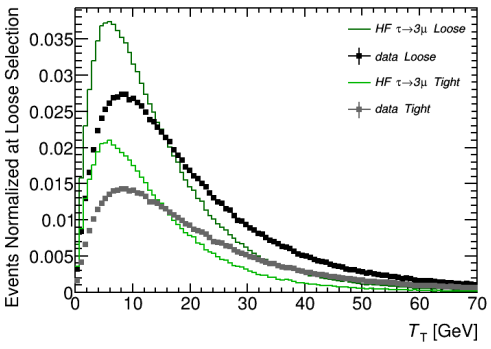
(o)


Events Normalized at Loose Selection
HF τ→3μ Loose
data Loose
HF τ→3μ Tight
data Tight
$M_{miss}^{triplet}$ [GeV]

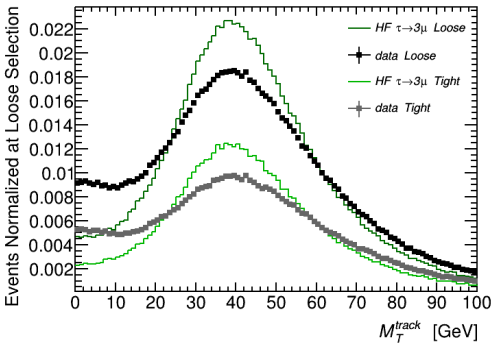
(p)



(q)



(r)



(s)

Figure B.1: Distributions of BDT input variables. The figures show simulated $D_s^{\pm} \to \tau(3\mu)\nu$ events and 139.6 fb^{-1} of data collected in Run-2. The distributions are normalized to unity integral after the application of the loose selection (Table 5.6). The same distributions are shown in light green and gray after the application of the tight selection.

Bibliography

[1] Martin L. Perl. "The Discovery of the tau lepton". In: *The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings, Conference, Stanford, USA, June 24-27, 1992*. 1992, pp. 79–100. URL: http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-PUB-5937.

[2] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[3] R. Brandelik et al. "Measurements of tau Decay Modes and a Precise Determination of the Mass". In: *Phys. Lett.* 73B (1978), pp. 109–114. DOI: 10.1016/0370-2693(78)90183-1.

[4] Jasper Kirkby. "Tau Studies from DELCO". In: *Conf. Proc.* C780428 (1978), p. 631. URL: https://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-PUB-2127.

[5] Geert Jan Besjes, Nicolo de Groot, and Sascha Caron. "Pushing SUSY's boundaries: Searches and prospects for strongly-produced supersymmetry at the LHC with the ATLAS detector". Nov. 2014. URL: http://cds.cern.ch/record/1981755.

[6] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[7] F. Halzen and Alan D. Martin. "Quarks and Leptons: an introductory course in modern particle physics". 1984. ISBN: 0471887412, 9780471887416.

[8] Debaprasad Maity and Pankaj Saha. "Connecting CMB anisotropy and cold dark matter phenomenology via reheating". In: (2018). arXiv: 1801.03059 [hep-ph].

[9] J. Einasto. "Dark matter and large scale structure". In: *ASP Conf. Ser.* 252 (2001), p. 85. arXiv: astro-ph/0012161 [astro-ph].

[10] Albert Bosma. "The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types". PhD thesis. 1978. URL: http://hdl.handle.net/11370/9490a028-214d-4eb3-974d-3cedcfde44f2.

[11] Manfred Lindner, Moritz Platscher, and Farinaldo S. Queiroz. "A Call for New Physics : The Muon Anomalous Magnetic Moment and Lepton Flavor Violation". In: *Phys. Rept.* 731 (2018), pp. 1–82. DOI: 10.1016/j.physrep.2017.12.001. arXiv: 1610.06587 [hep-ph].

[12] Graziano Venanzoni. "The New Muon g–2 experiment at Fermilab". In: *Nucl. Part. Phys. Proc.* 273-275 (2016), pp. 584–588. DOI: 10.1016/j.nuclphysbps.2015.09.087. arXiv: 1411.2555 [physics.ins-det].

[13] Brandon Murakami. "The Impact of lepton flavor violating Z-prime bosons on muon g-2 and other muon observables". In: *Phys. Rev.* D65 (2002), p. 055003. DOI: 10.1103/PhysRevD.65.055003. arXiv: hep-ph/0110095 [hep-ph].

[14] R. Aaij et al. "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays". In: *JHEP* 08 (2017), p. 055. DOI: 10.1007/JHEP08(2017)055. arXiv: 1705.05802 [hep-ex].

[15] Roel Aaij et al. "Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$". In: *Phys. Rev. Lett.* 115.11 (2015), p. 111803. DOI: 10.1103/PhysRevLett.115.159901,10.1103/PhysRevLett.115.111803. arXiv: 1506.08614 [hep-ex].

[16] J. P. Lees, Poireau, et al. "Evidence for an Excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ Decays". In: *Phys. Rev. Lett.* 109 (10 Sept. 2012), p. 101802. DOI: 10.1103/PhysRevLett.109.101802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.109.101802.

[17] M. Huschle et al. "Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle". In: *Phys. Rev.* D92.7 (2015), p. 072014. DOI: 10.1103/PhysRevD.92.072014. arXiv: 1507.03233 [hep-ex].

[18] HFLAV group. "Average of R(D) and R(D*) for Summer 2018". https://hflav-eos.web.cern.ch/hflav-eos/semi/summer18/RDRDs.html. Accessed on May 6, 2019.

[19] Marat Freytsis, Zoltan Ligeti, and Joshua T. Ruderman. "Flavor models for $\bar{B} \to D^{(*)}\tau\bar{\nu}$". In: *Phys. Rev.* D92.5 (2015), p. 054018. DOI: 10.1103/PhysRevD.92.054018. arXiv: 1506.08896 [hep-ph].

[20] Stephen P. Martin. "A Supersymmetry primer". In: (1997), pp. 1–98. DOI: 10.1142/9789812839657_0001,10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[21] W. Grimus. "Introduction to left-right symmetric models". In: *Elementary particle physics. Proceedings, 4th Hellenic School, Corfu, Greece, September 2-20, 1992. 1&2*. 1993, pp. 619–632. URL: http://lss.fnal.gov/archive/other/uwthph-1993-10.pdf.

[22] G. C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Phys. Rept.* 516 (2012), pp. 1–102. DOI: 10.1016/j.physrep.2012.02.002. arXiv: 1106.0034 [hep-ph].

[23] Ben Gripaios. "Lectures on Effective Field Theory". In: (2015). arXiv: 1506.05039 [hep-ph].

[24] Julian Heeck. "Interpretation of Lepton Flavor Violation". In: *Phys. Rev.* D95.1 (2017), p. 015022. DOI: 10.1103/PhysRevD.95.015022. arXiv: 1610.07623 [hep-ph].

[25] Frans R. Klinkhamer. "Sphalerons in the Standard Model". Oct. 2016. URL: https://www.rug.nl/research/vsi/events/groenewold/klinkhamer.pdf.

[26] V A Rubakov and Mikhail E Shaposhnikov. "Electroweak baryon number non-conservation in the early universe and in high energy collisions". In: *Phys. Usp.* 39.hep-ph/9603208. CERN-TH-96-013. INR-913 (Mar. 1996), 461–502. 113 p. URL: http://cds.cern.ch/record/297990.

[27] John Ellis and Kazuki Sakurai. "Search for Sphalerons in Proton-Proton Collisions". In: *JHEP* 04.KCL-PH-TH-2016-03. KCL-PH-TH-2016-03. LCTS-2016-02. CERN-TH-2016-011. IPPP-16-02 (Jan. 2016), 086. 17 p. URL: http://cds.cern.ch/record/2124347.

[28] Johannes Albrecht. "Lepton Flavour Universality tests with B decays at LHCb". In: *53rd Rencontres de Moriond on QCD and High Energy Interactions (Moriond QCD 2018) La Thuile, Italy, March 17-24, 2018*. 2018. arXiv: 1805.06243 [hep-ex].

[29] Super-Kamiokande Collaboration. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[30] SNO Collaboration. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.011301.

[31] S. N. Ahmed et al. "Measurement of the total active B-8 solar neutrino flux at the Sudbury Neutrino Observatory with enhanced neutral current sensitivity". In: *Phys. Rev. Lett.* 92 (2004), p. 181301. DOI: 10.1103/PhysRevLett.92.181301. arXiv: nucl-ex/0309004 [nucl-ex].

[32] T. Araki et al. "Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion". In: *Phys. Rev. Lett.* 94 (2005), p. 081801. DOI: 10.1103/PhysRevLett.94.081801. arXiv: hep-ex/0406035 [hep-ex].

[33] Y. Ashie et al. "Evidence for an oscillatory signature in atmospheric neutrino oscillation". In: *Phys. Rev. Lett.* 93 (2004), p. 101801. DOI: 10.1103/PhysRevLett.93.101801. arXiv: hep-ex/0404034 [hep-ex].

[34] K. Abe et al. "Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam". In: *Phys. Rev. Lett.* 107 (2011), p. 041801. DOI: 10.1103/PhysRevLett.107.041801. arXiv: 1106.2822 [hep-ex].

[35] Abe, K., Hayato, Y., et al. "Evidence for the Appearance of Atmospheric Tau Neutrinos in Super-Kamiokande". In: *Phys. Rev. Lett.* 110 (18 May 2013), p. 181802. DOI: 10.1103/PhysRevLett.110.181802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.181802.

[36] N. Agafonova et al. "Discovery of τ Neutrino Appearance in the CNGS Neutrino Beam with the OPERA Experiment". In: *Phys. Rev. Lett.* 115.12 (2015), p. 121802. DOI: 10.1103/PhysRevLett.115.121802. arXiv: 1507.01417 [hep-ex].

[37] M. C. Gonzalez-Garcia and Yosef Nir. "Neutrino masses and mixing: Evidence and implications". In: *Rev. Mod. Phys.* 75 (2003), pp. 345–402. DOI: 10.1103/RevModPhys.75.345. arXiv: hep-ph/0202058 [hep-ph].

[38] P. F. de Salas et al. "Status of neutrino oscillations 2018: first hint for normal mass ordering and improved CP sensitivity". In: (2017). arXiv: 1708.01186 [hep-ph].

[39] Mattias Blennow et al. "Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering". In: *JHEP* 03 (2014), p. 028. DOI: 10.1007/JHEP03(2014)028. arXiv: 1311.1822 [hep-ph].

[40] Qing-Guo Huang, Ke Wang, and Sai Wang. "Constraints on the neutrino mass and mass hierarchy from cosmological observations". In: *Eur. Phys. J.* C76.9 (2016), p. 489. DOI: 10.1140/epjc/s10052-016-4334-z. arXiv: 1512.05899 [astro-ph.CO].

[41] G. Hernández-Tomé, G. López Castro, and P. Roig. "Flavor violating leptonic decays of τ and μ leptons in the Standard Model with massive neutrinos". In: (2018). arXiv: 1807.06050 [hep-ph].

[42] Frank F. Deppisch, P. S. Bhupal Dev, and Apostolos Pilaftsis. "Neutrinos and Collider Physics". In: *New J. Phys.* 17.7 (2015), p. 075019. DOI: 10.1088/1367-2630/17/7/075019. arXiv: 1502.06541 [hep-ph].

[43] Aneesh V. Manohar. "Introduction to Effective Field Theories". In: *Les Houches summer school: EFT in Particle Physics and Cosmology Les Houches, Chamonix Valley, France, July 3-28, 2017*. 2018. arXiv: 1804.05863 [hep-ph].

[44] Alexander Merle. "keV Neutrino Model Building". In: *Int. J. Mod. Phys.* D22 (2013), p. 1330020. DOI: 10.1142/S0218271813300206. arXiv: 1302.2625 [hep-ph].

[45] A. Boyarsky et al. "Sterile Neutrino Dark Matter". In: (2018). arXiv: 1807.07938 [hep-ph].

[46] Takehiko Asaka, Mikko Laine, and Mikhail Shaposhnikov. "Lightest sterile neutrino abundance within the nuMSM". In: *JHEP* 01 (2007), p. 091. arXiv: hep-ph/0612182 [hep-ph].

[47] Jackson D. Clarke, Robert Foot, and Raymond R. Volkas. "Electroweak naturalness in the three-flavor type I seesaw model and implications for leptogenesis". In: *Phys. Rev.* D91.7 (2015), p. 073009. DOI: 10.1103/PhysRevD.91.073009. arXiv: 1502.01352 [hep-ph].

[48] Marco Drewes. "The Phenomenology of Right Handed Neutrinos". In: *Int. J. Mod. Phys.* E22 (2013), p. 1330019. DOI: 10.1142/S0218301313300191. arXiv: 1303.6912 [hep-ph].

[49] Morad Aaboud et al. "Search for heavy Majorana or Dirac neutrinos and right-handed W gauge bosons in final states with two charged leptons and two jets at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 01 (2019), p. 016. DOI: 10.1007/JHEP01(2019)016. arXiv: 1809.11105 [hep-ex].

[50] Georges Aad et al. "Search for heavy Majorana neutrinos with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 07 (2015), p. 162. DOI: 10.1007/JHEP07(2015)162. arXiv: 1506.06020 [hep-ex].

[51] CMS Collaboration. "Search for heavy Majorana neutrinos in the same-sign dilepton channel in proton-proton collisions at $\sqrt{s} = 13$ TeV". Tech. rep. CMS-PAS-EXO-17-028. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2308273.

[52] Kyu Jung Bae et al. "Leptogenesis scenarios for natural SUSY with mixed axion-higgsino dark matter". In: *JCAP* 1601 (2016), p. 012. arXiv: 1510.00724 [hep-ph].

[53] Sacha Davidson, Enrico Nardi, and Yosef Nir. "Leptogenesis". In: *Phys. Rept.* 466 (2008), pp. 105–177. DOI: 10.1016/j.physrep.2008.06.002. arXiv: 0802.2962 [hep-ph].

[54] Reyco Henning. "Current status of neutrinoless double-beta decay searches". In: *Reviews in Physics* 1 (2016), pp. 29–35. ISSN: 2405-4283. URL: http://www.sciencedirect.com/science/article/pii/S2405428316000034.

[55] Asmaa Abada et al. "Effective Majorana mass matrix from tau and pseudoscalar meson lepton number violating decays". In: *JHEP* 02 (2018), p. 169. DOI: 10.1007/JHEP02(2018)169. arXiv: 1712.03984 [hep-ph].

[56] Andre de Gouvea. "GeV seesaw, accidentally small neutrino masses, and Higgs decays to neutrinos". In: (2007). arXiv: 0706.1732 [hep-ph].

[57] Takaaki Nomura and Hiroshi Okada. "Neutrino mass generation with large $SU(2)_L$ multiplets under local $U(1)_{L_\mu - L_\tau}$ symmetry". In: (2018). arXiv: 1805.03942 [hep-ph].

[58] Alberto Lusiani. "Search for Lepton-Flavor-Violating Tau Decays at the B-factories". In: *PoS* HQL2010 (2010), p. 054. arXiv: 1012.3733 [hep-ex].

[59] Julian Heeck. "Lepton flavor violation with light vector bosons". In: *Phys. Lett.* B758 (2016), pp. 101–105. arXiv: 1602.03810 [hep-ph].

[60] Wolfgang Altmannshofer, Marcela Carena, and Andreas Crivellin. "$L_\mu - L_\tau$ theory of Higgs flavor violation and $(g-2)_\mu$". In: *Phys. Rev.* D94.9 (2016), p. 095026. arXiv: 1604.08221 [hep-ph].

[61] Andre de Gouvea and Petr Vogel. "Lepton Flavor and Number Conservation, and Physics Beyond the Standard Model". In: *Prog. Part. Nucl. Phys.* 71 (2013), pp. 75–92. DOI: 10.1016/j.ppnp.2013.03.006. arXiv: 1303.4097 [hep-ph].

[62] Robert H. Bernstein and Peter S. Cooper. "Charged Lepton Flavor Violation: An Experimenter's Guide". In: *Phys. Rept.* 532 (2013), pp. 27–64. arXiv: 1307.5787 [hep-ex].

[63] K. Hayasaka et al. "Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced Tau+Tau- Pairs". In: *Phys. Lett.* B687 (2010), pp. 139–143. DOI: 10.1016/j.physletb.2010.03.037. arXiv: 1001.3221 [hep-ex].

[64] Georges Aad et al. "Probing lepton flavour violation via neutrinoless $\tau \longrightarrow 3\mu$ decays with the ATLAS detector". In: *Eur. Phys. J.* C76.5 (2016), p. 232. DOI: 10.1140/epjc/s10052-016-4041-9. arXiv: 1601.03567 [hep-ex].

[65] Bernard Aubert et al. "Searches for Lepton Flavor Violation in the Decays $\tau^{pm} \to e^\pm\gamma$ and $\tau^\pm \to \mu^\pm\gamma$". In: *Phys. Rev. Lett.* 104 (2010), p. 021802. DOI: 10.1103/PhysRevLett.104.021802. arXiv: 0908.2381 [hep-ex].

[66] U. Bellgardt et al. "Search for the decay $\mu^+ \to e^+e^+e^-$". In: *Nuclear Physics B* 299.1 (1988), pp. 1–6. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(88)90462-2.

[67] Bao Baldini et al. "Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment". In: *The European Physical Journal C* 76.8 (Aug. 2016), p. 434. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4271-x.

[68] Wilhelm H. Bertl et al. "A Search for muon to electron conversion in muonic gold". In: *Eur. Phys. J.* C47 (2006), pp. 337–346. DOI: 10.1140/epjc/s2006-02582-x.

[69] "Search for Lepton Flavor Violation in Z decays in pp collisions at $\sqrt{s}$=8 TeV". Tech. rep. CMS-PAS-EXO-13-005. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2019863.

[70] Georges Aad et al. "Search for the lepton flavor violating decay $Z \to e\mu$ in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector". In: *Phys. Rev.* D90.7 (2014), p. 072010. DOI: 10.1103/PhysRevD.90.072010. arXiv: 1408.5774 [hep-ex].

[71] R. Akers et al. "A Search for lepton flavor violating Z0 decays". In: *Z. Phys.* C67 (1995), pp. 555–564. DOI: 10.1007/BF01553981.

[72] Hartger Weits. "Looking for Lepton Flavour Violation with the ATLAS Detector". PhD thesis. Amsterdam U., 2016. URL: https://inspirehep.net/record/1492453/files/thesis_H_Weits.pdf.

[73] P. Abreu et al. "Search for lepton flavor number violating Z0 decays". In: *Z. Phys.* C73 (1997), pp. 243–251. DOI: 10.1007/s002880050313.

[74] Georges Aad et al. "Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector". In: *Eur. Phys. J.* C77.2 (2017), p. 70. arXiv: 1604.07730 [hep-ex].

[75] Vardan Khachatryan et al. "Search for lepton flavour violating decays of the Higgs boson to $e\tau$ and $e\mu$ in proton–proton collisions at $\sqrt{s} = 8$ TeV". In: *Phys. Lett.* B763 (2016), pp. 472–500. arXiv: 1607.03561 [hep-ex].

[76] CMS Collaboration. "Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and $e\tau$ in proton-proton collisions at $\sqrt{s} = 13 \sim$ TeV". In: (2017). URL: https://cds.cern.ch/record/2264540.

[77] Morad Aaboud et al. "Search for lepton-flavor violation in different-flavor, high-mass final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Submitted to: Phys. Rev.* (2018). arXiv: 1807.06573 [hep-ex].

[78] Bernard Aubert et al. "The BaBar detector". In: *Nucl. Instrum. Meth.* A479 (2002), pp. 1–116. DOI: 10.1016/S0168-9002(01)02012-5. arXiv: hep-ex/0105044 [hep-ex].

[79] A. Abashian et al. "The Belle Detector". In: *Nucl. Instrum. Meth.* A479 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[80] Y. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016". In: *Eur. Phys. J.* C77.12 (2017), p. 895. arXiv: 1612.07233 [hep-ex].

[81] Marzia Bordone et al. "Low-energy signatures of the PS^3 model: from B-physics anomalies to LFV". In: (2018). arXiv: 1805.09328 [hep-ph].

[82] Zaineb Calcuttawala et al. "New Physics with Lepton Flavour Violating Decay $\tau \to 3\mu$". In: (2018). arXiv: 1802.09218 [hep-ph].

[83] Thomas Faber et al. "A unified leptoquark model confronted with lepton non-universality in B-meson decays". In: (2018). arXiv: 1808.05511 [hep-ph].

[84] Lorenzo Calibbi, Andreas Crivellin, and Tianjun Li. "A model of vector leptoquarks in view of the B-physics anomalies". In: (2017). arXiv: 1709.00692 [hep-ph].

[85] Debajyoti Choudhury et al. "Minimal unified resolution to $R_{K^{(*)}}$ and $R(D^{(*)})$ anomalies with lepton mixing". In: *Phys. Rev. Lett.* 119.15 (2017), p. 151801. DOI: 10.1103/PhysRevLett.119.151801. arXiv: 1706.08437 [hep-ph].

[86] Wolfgang Altmannshofer et al. "Quark flavor transitions in $L_\mu - L_\tau$ models". In: *Phys. Rev.* D89 (2014), p. 095033. DOI: 10.1103/PhysRevD.89.095033. arXiv: 1403.1269 [hep-ph].

[87] Sheldon L. Glashow, Diego Guadagnoli, and Kenneth Lane. "Lepton Flavor Violation in B Decays?" In: *Phys. Rev. Lett.* 114 (2015), p. 091801. DOI: 10.1103/PhysRevLett.114.091801. arXiv: 1411.0565 [hep-ph].

[88] Ferruccio Feruglio, Paride Paradisi, and Andrea Pattori. "On the Importance of Electroweak Corrections for B Anomalies". In: *JHEP* 09 (2017), p. 061. DOI: 10.1007/JHEP09(2017)061. arXiv: 1705.00929 [hep-ph].

[89] Yuji Kajiyama. "Lepton flavor violating tau decay by R-parity violation in a family symmetry". In: *Summer Institute 2007 (SI2007) Fuji-Yoshida, Japan, August 3-10, 2007.* 2007. arXiv: 0711.0286 [hep-ph].

[90] Avelino Vicente. "Lepton flavor violation beyond the MSSM". In: *Adv. High Energy Phys.* 2015 (2015), p. 686572. DOI: 10.1155/2015/686572. arXiv: 1503.08622 [hep-ph].

[91] Athanasios Dedes, John R. Ellis, and Martti Raidal. "Higgs mediated $B^0_{s,d} \to \mu\tau$, $e\tau$ and $\tau \to 3\mu$, $e\mu\mu$ decays in supersymmetric seesaw models". In: *Phys. Lett.* B549 (2002), pp. 159–169. DOI: 10.1016/S0370-2693(02)02900-3. arXiv: hep-ph/0209207 [hep-ph].

[92] K. S. Babu and Christopher Kolda. "Higgs mediated tau $\longrightarrow$ 3 mu in the supersymmetric seesaw model". In: *Phys. Rev. Lett.* 89 (2002), p. 241802. DOI: 10.1103/PhysRevLett.89.241802. arXiv: hep-ph/0206310 [hep-ph].

[93] Igor P. Ivanov. "Building and testing models with extended Higgs sectors". In: *Prog. Part. Nucl. Phys.* 95 (2017), pp. 160–208. DOI: 10.1016/j.ppnp.2017.03.001. arXiv: 1702.03776 [hep-ph].

[94] Jin-Jun Zhang et al. "Flavor Violating Higgs Couplings in Minimal Flavor Violation". In: (2018). arXiv: 1807.00921 [hep-ph].

[95] S. Myers and E. Picasso. "The design, construction and commissioning of the CERN large Electron–Positron collider". In: *Contemporary Physics* 31.6 (1990), pp. 387–403. DOI: 10.1080/00107519008213789.

[96] Andreas Hoecker. "Physics at the LHC Run-2 and Beyond". In: *2016 European School of High-Energy Physics (ESHEP 2016) Skeikampen, Norway, June 15-28, 2016*. 2016. arXiv: 1611.07864 [hep-ex].

[97] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[98] The ATLAS Collaboration. ""Atlas data summary"". https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/DataSummary/2015/. Accessed on April 10, 2018.

[99] The ATLAS Collaboration. "Stable Beam Runs 2017". https://twiki.cern.ch/twiki/bin/view/Atlas/StableBeamRuns2017. Accessed on April 10, 2018.

[100] V Cindro et al. "The ATLAS Beam Conditions Monitor". In: *Journal of Instrumentation* 3.02 (2008), P02004. URL: http://stacks.iop.org/1748-0221/3/i=02/a=P02004.

[101] F. Lasagni Manghi. "The LUCID detector ATLAS luminosity monitor and its electronic system". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 824 (2016), pp. 311–312. ISSN: 0168-9002. URL: http://www.sciencedirect.com/science/article/pii/S0168900215015272.

[102] Morad Aaboud et al. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[103] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[104] Torbjörn Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[105] The TOTEM Collaboration. In: *Journal of Instrumentation* 3.08 (2008), S08007. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08007.

[106] "TOTEM Upgrade Proposal". Tech. rep. CERN-LHCC-2013-009. LHCC-P-007. Geneva: CERN, June 2013. URL: https://cds.cern.ch/record/1554299.

[107] M. Aaboud et al. "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 117.18 (2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. arXiv: 1606.02625 [hep-ex].

[108] G. Antchev et al. "First measurement of elastic, inelastic and total cross-section at $\sqrt{s} =$ 13 TeV by TOTEM and overview of cross-section data at LHC energies". In: (2017). arXiv: 1712.06153 [hep-ex].

[109] Roel Aaij et al. "Measurement of the inelastic *pp* cross-section at a centre-of-mass energy of 13 TeV". In: (2018). arXiv: 1803.10974 [hep-ex].

[110] ATLAS Collaboration. "ATLAS detector and physics performance: Technical Design Report, 1". Technical Design Report ATLAS. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/391176.

[111] Juerg Beringer et al. "Offline beam spot reconstruction in ATLAS during Run 1". Tech. rep. ATL-COM-DAPR-2014-013. Geneva: CERN, Dec. 2014. URL: https://cds.cern.ch/record/1978679.

[112] S Haywood et al. "ATLAS inner detector: Technical Design Report, 1". Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331063.

[113] S Haywood et al. "ATLAS inner detector: Technical Design Report, 2". Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331064.

[114] M Capeans et al. "ATLAS Insertable B-Layer Technical Design Report". Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. Sept. 2010. URL: https://cds.cern.ch/record/1291633.

[115] "ATLAS Insertable B-Layer Technical Design Report Addendum". Tech. rep. CERN-LHCC-2012-009. ATLAS-TDR-19-ADD-1. May 2012. URL: https://cds.cern.ch/record/1451888.

[116] The ATLAS Collaboration. "ATLAS Run-2 Detector Status". Accessed on October 19, 2019. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector.

[117] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008). URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[118] Tavernier S. "Interactions of Particles in Matter. In: Experimental Techniques in Nuclear and Particle Physics". Berlin, Heidelberg: Springer, 2009. URL: https://doi.org/10.1007/978-3-642-00829-0_2.

[119] "ATLAS liquid argon calorimeter: Technical design report". In: (1996). URL: https://cds.cern.ch/record/331061.

[120] "ATLAS tile calorimeter: Technical Design Report". Technical Design Report ATLAS. Geneva: CERN, 1996. URL: http://cds.cern.ch/record/331062.

[121] J. Colas, M. Pripstein, and W.A. Wenzel. "The electrostatic transformer". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 294.3 (1990), pp. 583–590. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(90)90301-L.

[122] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[123] Johannes Mellenthin. "Measurement of muon misidentification rates in $Z \to \mu\mu$ events for the ATLAS detector". July 2013. URL: https://indico.mpp.mpg.de/event/2410/session/17/contribution/29/material/slides.

[124] E. Diehl. "Calibration and Performance of the ATLAS Muon Spectrometer". In: *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2011, Providence, USA, August 9-13, 2011.* 2011. arXiv: 1109.6933 [physics.ins-det].

[125] The ATLAS TDAQ Collaboration. "The ATLAS Data Acquisition and High Level Trigger system". In: *Journal of Instrumentation* 11.06 (2016), P06008. URL: http://stacks.iop.org/1748-0221/11/i=06/a=P06008.

[126] Georges Aad et al. "Performance of the ATLAS Trigger System in 2010". In: *Eur. Phys. J.* C72 (2012), p. 1849. DOI: 10.1140/epjc/s10052-011-1849-1. arXiv: 1110.1530 [hep-ex].

[127] Aranzazu Ruiz-Martinez and ATLAS Collaboration. "The Run-2 ATLAS Trigger System". Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, Feb. 2016. URL: https://cds.cern.ch/record/2133909.

[128] Morad Aaboud et al. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J.* C77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[129] M Shochet et al. "Fast TracKer (FTK) Technical Design Report". Tech. rep. CERN-LHCC-2013-007. ATLAS-TDR-021. June 2013. URL: http://cds.cern.ch/record/1552953.

[130] Swagato Banerjee et al. "Tau and muon pair production cross-sections in electron-positron annihilations at $\sqrt{s} = 10.58$ GeV". In: *Phys. Rev.* D77 (2008), p. 054012. DOI: 10.1103/PhysRevD.77.054012. arXiv: 0706.3235 [hep-ph].

[131] Paul Seyfert and Ulrich Uwer. "The search for $\tau \to \mu\mu\mu$ at LHCb". Jan. 2015. URL: http://cds.cern.ch/record/2002363.

[132] N Tal Hod et al. "Search for Lepton Flavor Violation in $\tau \longrightarrow 3\mu$ Decays". Tech. rep. ATL-COM-PHYS-2014-1271. Geneva: CERN, Oct. 2014. URL: https://cds.cern.ch/record/1951374.

[133] Hasan Ogul et al. "High Order QCD Predictions for Inclusive Production of W Bosons in pp Collisions at $sqrts = 13$ TeV". In: *Advances in High Energy Physics* 2016.7865689 (2016). URL: https://doi.org/10.1155/2016/7865689.

[134] G. Aad et al. "Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13TeV$ with the ATLAS detector". In: *Physics Letters B* 759 (2016), pp. 601–621. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2016.06.023.

[135] Georges Aad et al. "Measurement of the inclusive $W^{\pm}$ and Z/gamma cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Phys. Rev.* D85 (2012), p. 072004. arXiv: 1109.5141 [hep-ex].

[136] M. Tanabashi and K. et al. Hagiwara. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[137] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

[138] "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2". In: *The European Physical Journal C* 77.10 (Oct. 2017), p. 673. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5225-7.

[139] Georges Aad et al. "A neural network clustering algorithm for the ATLAS silicon pixel detector". In: *JINST* 9 (2014), P09009. arXiv: 1406.7690 [hep-ex].

[140] Pierre Billoir and S. Qian. "Fast vertex fitting with a local parametrization of tracks". In: *Nucl. Instrum. Meth.* A311 (1992), pp. 139–150. DOI: 10.1016/0168-9002(92)90859-3.

[141] Wolfgang Waltenberger, Rudolf Frühwirth, and Pascal Vanlaer. "Adaptive vertex fitting". In: *Journal of Physics G: Nuclear and Particle Physics* 34.12 (2007), N343. URL: http://stacks.iop.org/0954-3899/34/i=12/a=N01.

[142] "Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV". Tech. rep. ATL-PHYS-PUB-2015-026. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037717.

[143] "Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment". Tech. rep. ATLAS-CONF-2010-069. Geneva: CERN, July 2010. URL: https://cds.cern.ch/record/1281344.

[144] "$H \to 2e2\mu$ candidate event with 25 additional reconstructed primary vertices recorded in 2016." In: (Mar. 2018). URL: https://cds.cern.ch/record/2309698.

[145] V. Kostyukhin. "VKalVrt - package for vertex reconstruction in ATLAS." In: (2003).

[146] "ATLAS Inner Detector Alignment Performance with February 2015 Cosmic Rays Data". Tech. rep. ATL-PHYS-PUB-2015-009. Geneva: CERN, Apr. 2015. URL: https://cds.cern.ch/record/2008724.

[147] "Alignment of the ATLAS Inner Detector with the initial LHC data at $\sqrt{s} = 13$ TeV". Tech. rep. ATL-PHYS-PUB-2015-031. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2038139.

[148] "Study of the mechanical stability of the ATLAS Insertable B-Layer". Tech. rep. ATL-INDET-PUB-2015-001. Geneva: CERN, June 2015. URL: https://cds.cern.ch/record/2022587.

[149] "Summary of ATLAS Pythia 8 tunes". Tech. rep. ATL-PHYS-PUB-2012-003. Geneva: CERN, Aug. 2012. URL: http://cds.cern.ch/record/1474107.

[150] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[151] Hung-Liang Lai et al. "New parton distributions for collider physics". In: *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[152] Peter Zeiler Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". In: *Phys. Rev.* D82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[153] Zhi Zheng. "Identification of very-low transverse momentum muons with the ATLAS experiment". Tech. rep. ATL-COM-MUON-2018-064. Geneva: CERN, Nov. 2018. URL: hhttps://cds.cern.ch/record/2647098.

[154] Nicole Manuela Ruckstuhl and D C van den Boom. "Muon signatures in ATLAS A search for new physics in $\mu\acute{s}\mu\acute{s}$ events". 2012. URL: https://cds.cern.ch/record/1461551.

[155] Georges Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. arXiv: 1603.02934 [hep-ex].

[156] W. Lampl et al. "Calorimeter clustering algorithms: Description and performance". In: (2008). URL: https://cds.cern.ch/record/1099735.

[157] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[158] M. Aaboud et al. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[159] "Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector". In: (2015). URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2015-029/.

[160] Morad Aaboud et al. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (2018). arXiv: 1802.08168 [hep-ex].

[161] The ATLAS collaboration. "$E_{\mathrm{T}}^{\mathrm{miss}}$ performance in the ATLAS detector using 2015-2016 LHC p-p collisions". In: (2018). URL: https://cds.cern.ch/record/2625233.

[162] "ATLAS Run 1 Pythia8 tunes". Tech. rep. ATL-PHYS-PUB-2014-021. Geneva: CERN, Nov. 2014. URL: https://cds.cern.ch/record/1966419.

[163] Stefano Carrazza, Stefano Forte, and Juan Rojo. "Parton Distributions and Event Generators". In: *Proceedings, 43rd International Symposium on Multiparticle Dynamics (ISMD 13)*. 2013, pp. 89–96. arXiv: 1311.5887 [hep-ph].

[164] Maria Smizanska. "PythiaB interface to Pythia6 dedicated to simulation of beauty events." https://svnweb.cern.ch/trac/atlasoff/browser/Generators/PythiaB/trunk/doc/PythiaB.pdf. Accessed on January 17, 2019. 2005.

[165] J. Catmore, J. Cranshaw, et al. "A new petabyte-scale data derivation framework for ATLAS". In: *Journal of Physics: Conference Series* 664.7 (2015), p. 072007. URL: http://stacks.iop.org/1742-6596/664/i=7/a=072007.

[166] ATLAS. "Data Quality Information for Data". https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010. Rev: 83. 2017.

[167] ATLAS. "Data Preparation Checklist for Physics Analysis". Rev: 21. 2017. URL: https://twiki.cern.ch/twiki/bin/view/Atlas/DataPreparationCheckListForPhysicsAnalysis.

[168] ATLAS. "Muon Selection Tool". https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool. Rev: 41. 2017.

[169] Andreas Hocker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: (2007). arXiv: physics/0703039 [physics.data-an].

[170] I. Antcheva, M. Ballintijn, et al. "ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization". In: *Computer Physics Communications* 180.12 (2009), pp. 2499–2512. ISSN: 0010-4655. URL: http://www.sciencedirect.com/science/article/pii/S0010465509002550.

[171] ROOT. "ROOT Data Analysis Framework: User's Guide". https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.html. Accessed on September 7, 2018.

[172] Olaf Behnke et al., eds. "Data analysis in high energy physics". Weinheim, Germany: Wiley-VCH, 2013. URL: http://www.wiley-vch.de/publish/dt/books/ISBN3-527-41058-9.

[173] E. Gross. "Practical Statistics for High Energy Physics". In: *CERN Yellow Reports: School Proceedings* 3.0 (2018), p. 199. ISSN: 2519-805X. URL: https://e-publishing.cern.ch/index.php/CYRSP/article/view/755.

[174] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011), p. 1554. arXiv: 1007.1727 [physics.data-an].

[175] J. M. Lorenz et al. "HistFitter - A flexible framework for statistical data analysis". In: *J. Phys. Conf. Ser.* 608 (2015), p. 012049. DOI: 10.1088/1742-6596/608/1/012049.

[176] M. Baak et al. "HistFitter software framework for statistical data analysis". In: *Eur. Phys. J.* C75 (2015), p. 153. DOI: 10.1140/epjc/s10052-015-3327-7. arXiv: 1410.1280 [hep-ex].

[177] Wouter Verkerke and David P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003), MOLT007. arXiv: physics/0306116 [physics].

[178] Christian Gumpert et al. "Software for statistical data analysis used in Higgs searches". In: *J. Phys. Conf. Ser.* 490 (2014), p. 012229. DOI: 10.1088/1742-6596/490/1/012229.

[179] "Search for $\tau \to 3\mu$ decays using τ leptons produced in D and B meson decays". Tech. rep. CMS-PAS-BPH-17-004. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2668282.

[180] M. Bedognetti et al. "Differential cross section measurement of $D^{\pm}$ and $D_s^{\pm}$ meson production with ATLAS". Tech. rep. ANA-BPHY-2018-10-INT1. Amsterdam: Nikhef, July 2019. URL: https://gitlab.cern.ch/atlas-physics-office/BPHY/ANA-BPHY-2018-10/ANA-BPHY-2018-10-INT1/tree/master.

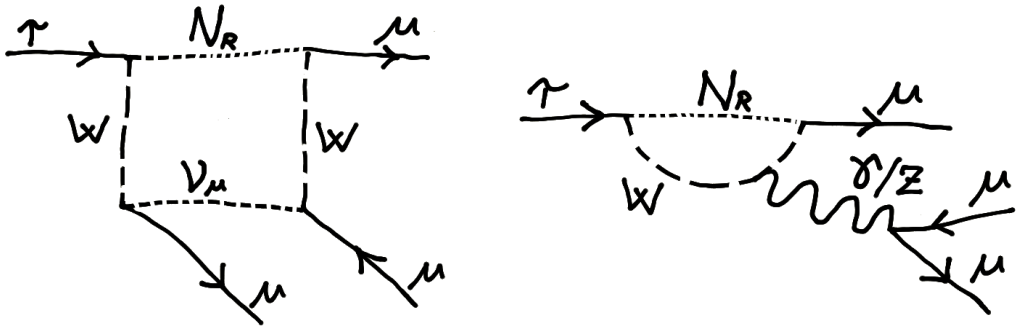
[181] N Tal Hod et al. "Search for Lepton Flavor Violation in $\tau \longrightarrow 3\mu$ Decays". Tech. rep. ATL-COM-PHYS-2014-1271. Geneva: CERN, Oct. 2014. URL: https://cds.cern.ch/record/1951374.

Summary

High-energy physics studies the tiniest knowable components of the physical world. Its mathematical formulation is the Standard Model of particle physics: a theory postulating a number of fundamental particles and forces through which they interact. The Standard Model (SM) agrees to an astonishing degree with experimental observations, however it has a few shortcomings. These include:

- it cannot explain why there is virtually no anti-matter to match the amount of matter in the visible universe
- it does not describe the observed Dark Matter
- its explanation for neutrino masses, albeit mathematically sound, is arbitrary, because their origin from Majorana states is not excluded
- it fails to reproduce the precision measurement of the muon's gyromagnetic moment
- its prediction of relative branching fractions of some semi-leptonic decays may be in disagreement with recent measurements at LHCb.

For these reasons there is the need for a larger model to extend the standard one (so-called BSM).
In the SM lepton flavour is conserved for charged leptons, therefore it is possible to study BSM physics searching for its manifestation in processes that violate the conservation of lepton flavour. Many such processes have been sought, so far without success, and limits have been set on the probability that they may occur. In this thesis a search is presented for the Lepton Flavour Violating (LFV) decay $\tau \to 3\mu$ (a tau lepton decaying into three muons) using data collected by the ATLAS experiment on proton-proton collisions with a centre-of-mass energy of 13 TeV. The decay, should it exist, would entail the existence of yet unobserved interactions between charged leptons of the II and III generation. This may in turn be connected to Majorana states mediating the origin of neutrino masses, through Feynman diagrams like the ones shown in Figure S.1. The possible LFV coupling between tau and muons leading to the decay can have many explanations, e.g. leptoquarks, heavy bosons similar to the mediators of the electroweak interaction, or supersymmetric fields. Failing to observe any such LFV process, one sets limits on their rate of occurrence, permitting to severely constrain the parameter space of the theoretical models in question.
A way to study high-energy physics is observing the controlled collision of particles, such



(a) (b)

Figure S.1: Two Feynman diagrams responsible for the possible contribution of right-handed neutrinos (N_R) to the decay $\tau \to 3\mu$.

as at the Large Hadron Collider (LHC), currently the largest accelerator in the world. The LHC is located just outside the city of Geneva, Switzerland and describes a circle, about 8.6 km in diameter. Between 2015 and 2018 it has collided proton beams at 13 TeV centre-of-mass energy and its four main experiments have recorded the collisions. This thesis makes use of data collected by the ATLAS detector. ATLAS permits to reconstruct the energy and nature of the produced particles to study the properties of the SM to high precision and explore the products of potential BSM interactions.

At ATLAS a trigger system is used to select about 1000 events per second which are then stored to disk for later processing. For the $\tau \to 3\mu$ search events are selected which have three muons emitted close to each other (a feature originating from the smallness of the tau mass with respect to the momenta of candidates within the detector's acceptance). To select these events a combination of several multi-muon triggers is used, including a number of triggers designed specifically for the analysis. The triggers vary year-by-year, in accordance with ever-changing trigger menus employed in the operation of ATLAS.
The events selected by the trigger system are processed by offline reconstruction software, which fully characterizes the particles produced. For this analysis, the muons used are exclusively so-called *combined* muons: muons whose charged track is reconstructed in the Inner Detector tracker (ID) and in the muon spectrometer chambers surrounding the calorimeters.
In view of this analysis a change has been applied to the ATLAS muon reconstruction software to guarantee optimal reconstruction of closely-packed muons. Valuable information about the signal candidates is obtained reconstructing the displaced secondary vertex, at which the paths of the three muons intersect each other (Figure S.2). Having a position for the secondary vertex enables one to estimate the longitudinal and transverse displacement of the three muons' mother particle, and other quantities of interest, and compare these with the same properties of simulated $\tau \to 3\mu$ signal events.
For the transverse displacement (a_0^{xy}) a systematic bias is observed and understood to

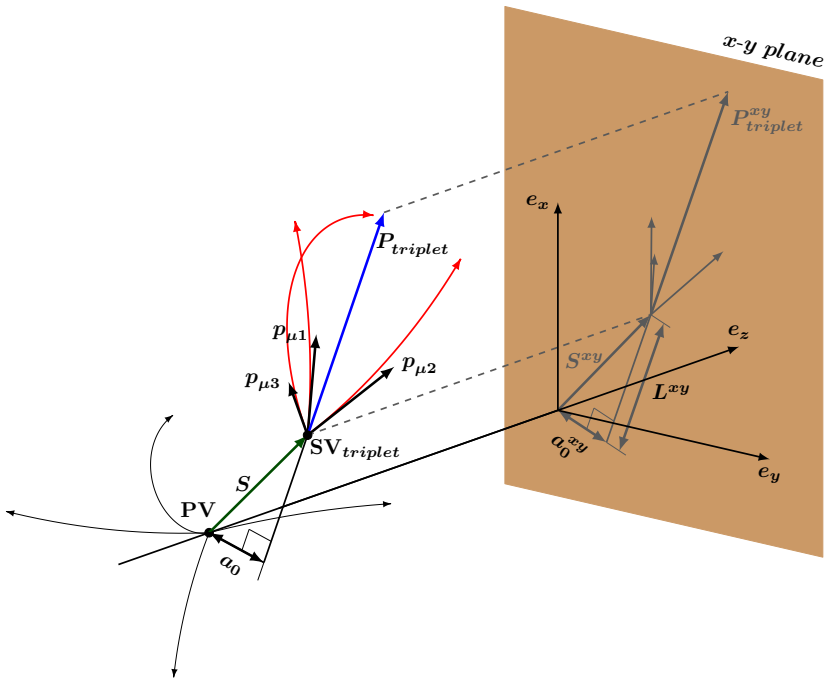


Figure S.2: Graphical representation of a muon-triplet vertex. For clarity the curvature of tracks is exaggerated and projections on the x-y plane are coloured gray. The tracks of the three muons are indicated in red, their refitted momenta are shown in black, and the candidate muon triplet momentum is shown in blue. Several quantities used to train the BDT classifier in the analysis are labelled.

be the effect of residual misalignment of the silicon layers, remaining after the otherwise very successful alignment procedure. The bias is characterized as a function of data taking period, and subtracted from the variable a_0^{xy} in the collected data.
Signal events are further characterised by the isolation of the three-muon candidates from energetic tracks in the ID (P_T^{cone}) and from energy depositions in the calorimeter coming from the underlying collision event (E_T^{cone}). The two quantities are calculated similarly to the customary isolation of individual muons, but correcting for the contemporary presence of three muons. To search for the decay $\tau \to 3\mu$, we focus on the most abundant source of τ leptons: hadronic jets. Accounting for the detector's acceptance, the channel is dominated by $D_s \to \tau\nu$, followed by $B \to \tau X$ (the decay of a B meson to a tau and other particles). The combination of these two channels is here referred to as HF-channel. For the analysis the signal is modelled as a simulation of prompt and non-prompt $D_s^{\pm} \to \tau(3\mu)\nu$ events only, building on the similarity between the secondary channel ($B \to \tau X$) and this one. A previous analysis at ATLAS, performed on Run-1 data, targeted a different channel $W \to \tau\nu$, leading to taus with higher energy, but less abundant than the HF-channel.
The present analysis uses a series of cut-based selections to reduce the background contamination, dominated by events with two muons and an associated π/K meson pro-

duced in jets. These selections primarily target the quality of the muon candidates and accept 65% of reconstructed and triggered signal events, whilst rejecting 86% of background events. After the cut-based selection, the most sensitive variables are fed to a multivariate classifier (BDT). The classifier is trained on simulated signal events against data events, for which the invariant mass of the muon-triplet falls within two "training" bands: $M^{\text{triplet}} \in [1400, 1690] \cup [1870, 2200]$, which excludes the peak described by signal events. The variables used in the BDT training are validated on $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events: simulation is compared bin-wise with signal counts extracted from 2017 data (a fit of the D_s mass peak is used to estimate the events in data). The selection on the BDT discriminant is chosen scanning for the minimum expected limit on the branching fraction of tau into three muons ($\mathcal{B}(\tau \to 3\mu)$), after rejecting the di-muon resonances ρ^0 and ϕ. The BDT selection reduces the backgrounds by a factor 100, while retaining about a third of signal events. To calculate the limit, the signal events are scaled to match a fiducial measurement of the D_s production cross-section at ATLAS for $\sqrt{s} = 13$ TeV p-p collisions, and additionally to a simulation-based estimate of $\sigma^{B\to\pi X}$. A limit

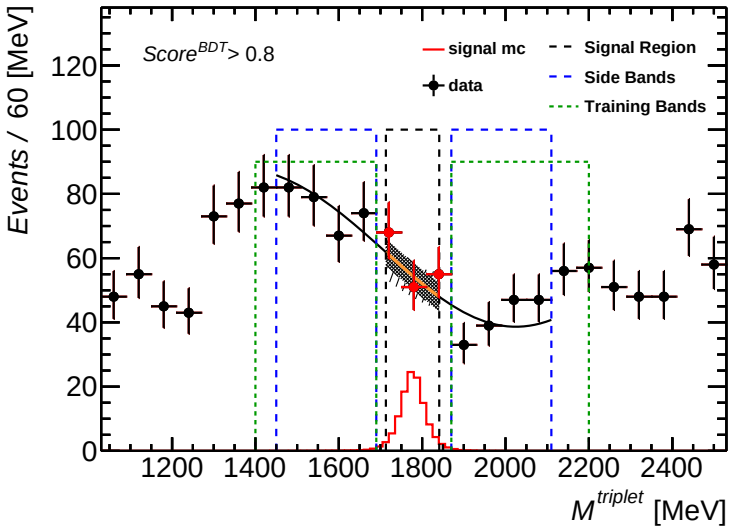


Figure S.3: The three-muon invariant mass distribution in the range $M^{\text{triplet}} \in [1030, 2530]$ MeV. Signal, shown in the red histogram, and data events, shown as solid dots, pass the *tight* selection (section 5.3) and the additional cut $Score^{BDT}$>0.8 as indicated. Red markers denote data points inside the blinded region, disclosed only at unblinding. The signal histogram is scaled to $\mathcal{B}(\tau \to 3\mu) = 2{\times}10^{-7}$. The polynomial fit used to extrapolate the expected background into the SR is shown in black, its section representing the expected background in the SR is marked in orange. The fit uncertainty is shown shadowed, while the model uncertainty is dashed. The region covered by the SR is marked by a black-bordered box, while the SB are shown with a blue border. The Training Bands (TB) are similarly marked by a green border.

on $\mathcal{B}(\tau \to 3\mu)$ is determined from the number of observed data events falling within the SR ($M^{\text{triplet}} \in [1713, 1841]$). The number of expected background events is obtained interpolating the distribution of data events from signal-devoid side-bands, as can be seen in Figure S.3.

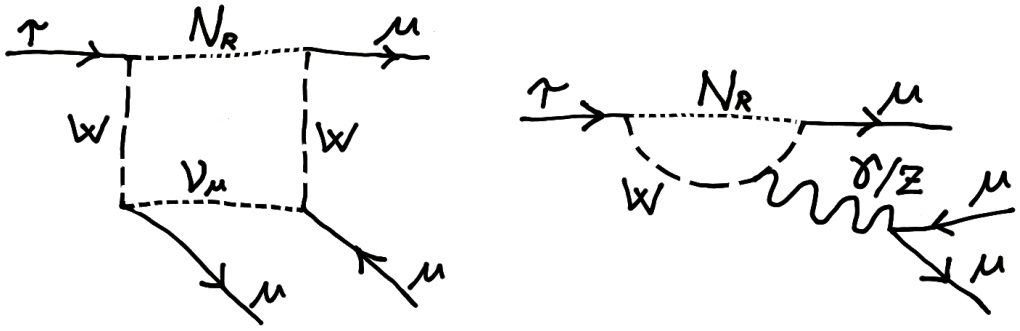
The analysis yields an observed upper limit on $\mathcal{B}(\tau \to 3\mu)$ of 1.98 $\times 10^{-7}$ at 90% CL. This limit is consistent with the expected limit, in absence of signal, of $1.8^{+1.0}_{-0.6} \times 10^{-7}$. The limiting factor for the analysis is the difficulty in separating the topologies of signal and backgrounds. A Neural Network approach is currently under consideration, to improve the separation power. The result may eventually be combined with a Run-2 analysis focussing on the W-channel, bearing similarities with the Run-1 ATLAS analysis [64].

Samenvatting

Hoge energie fysica bestudeert de kleinst kenbare bouwstenen van de fysieke wereld. Haar wiskundige formulering is het Standaardmodel van de deeltjesfysica: een theorie die een aantal fundamentele deeltjes postuleert samen met een stelsel krachten waardoor de deeltjes interacties met elkaar aangaan. Het Standaardmodel (SM) stemt in verbazingwekkende mate overeen met experimentele metingen, maar heeft desondanks enkele tekortkomingen:

- het kan niet verklaren waarom er vrijwel geen anti-materie is in vergelijking met de hoeveelheid materie in het zichtbare universum
- het beschrijft niet de geobserveerde Donkere Materie
- de verklaring voor neutrinomassa's, hoewel wiskundig nauwkeurig, is ongefundeerd, voorzover hun oorsprong uit Majorana-staten niet uitgesloten is
- het slaagt er niet in om de precisiemeting van het gyromagnetische moment van de muon te reproduceren
- de voorspelling van relatieve vervalsfracties van sommige semi-leptonische deeltjesvervallen zijn mogelijk niet in overeenstemming met recente metingen bij LHCb.

Om deze redenen is er behoefte aan een groter model om het standaardmodel uit te breiden (BSM genoemd).
In het SM is lepton-flavour behouden voor geladen leptonen, daarom is het mogelijk om BSM-fysica te bestuderen door het zoeken naar haar manifestatie in processen die het behoud van lepton-flavour schenden. Veel van dergelijke processen zijn beproefd, tot nu toe zonder succes, en vervolgens werden limieten bepaald voor de waarschijnlijkheid dat ze zouden kunnen optreden. In dit proefschrift wordt een zoektocht gepresenteerd voor het Lepton-Flavour schendende (LFV) verval $\tau \to 3\mu$ (de verval van een tau lepton naar drie muonen) met behulp van data verzameld door het ATLAS-experiment over proton-proton botsingen met een massamiddelpuntenergie van 13 TeV. Het verval zou, als het bestaat, het bestaan bevestigen van tot zover niet geobserveerde interacties tussen geladen leptonen van de II en III generatie. Dit kan op zijn beurt in verband staan met Majorana-toestanden die de oorsprong van neutrinomassa's bemiddelen, door Feynman-diagrammen zoals weergegeven in Figuur S.1. De mogelijke LFV-koppeling tussen tauonen en muonen die tot het verval leidt, kan meerdere verklaringen hebben,



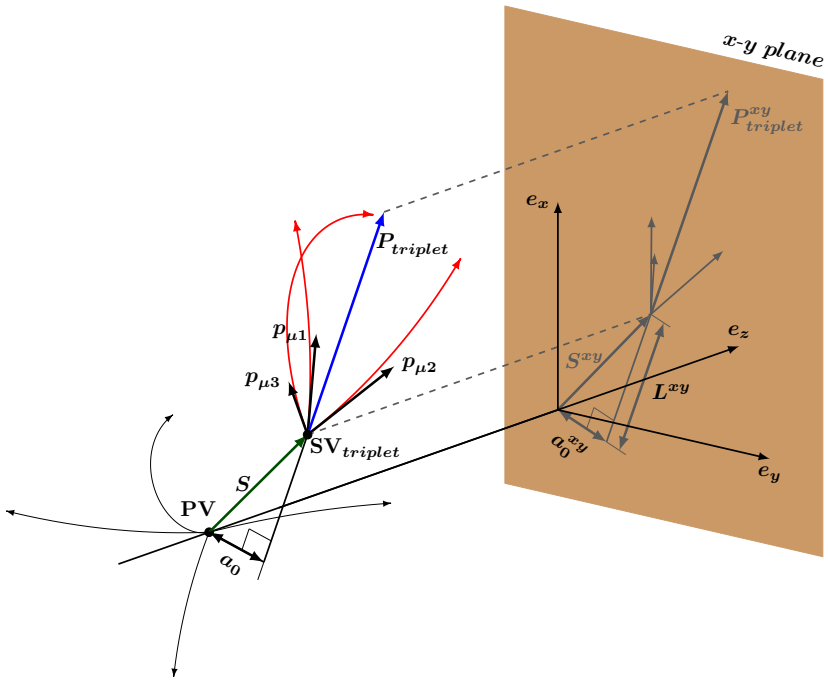
(a) (b)

Figuur S.1: Twee Feynman-diagrammen verantwoordelijk voor de mogelijke bijdrage van rechtshandige neutrino's (N_R) aan het verval $\tau \to 3\mu$.

bijvoorbeeld Leptoquarks, zware bosonen vergelijkbaar met de mediators van de electrozwakke interactie of zelfs supersymmetrische velden. Als een dergelijk LFV-proces niet geobserveerd wordt, worden limieten op hun waarschijnlijkheid van optreden, waardoor de parameterruimte van de theoretische modellen in kwestie sterk kan worden beperkt. Het bestuderen van hoge-energie fysica kan door het observeren van de gecontroleerde deeltjesbotsingen, zoals bij de Large Hadron Collider (LHC), tegenwoordig de grootste deeltjesversneller ter wereld. De LHC ligt net buiten de stad Genève, in Zwitserland, en beschrijft een cirkel met een diameter van ongeveer 8.6 km. Tussen 2015 en 2018 werden er protonbundels met een massamiddelpuntenergie energie van 13 TeV tegen elkaar gebotst terwijl de vier belangrijkste experimenten bij het LHC de botsingen registreerden. Dit proefschrift maakt gebruik van data verzameld door de ATLAS-detector. ATLAS maakt het mogelijk om de energie en de aard van de geproduceerde deeltjes te reconstrueren, om de eigenschappen van het SM nauwkeurig te bestuderen en de producten van potentiële BSM-interacties te onderzoeken.

Bij ATLAS wordt een triggersysteem gebruikt om ongeveer 1000 events per seconde te selecteren die dan op harde schijf opgeslagen worden om later verwerkt te worden. Voor de $\tau \to 3\mu$ analyse zijn botsingsevents geselecteerd met drie dicht op elkaar propagerende muonen (een kenmerk dat ontstaat uit de kleinheid van de tauon-massa met betrekking tot de impuls van de kandidaten binnen de acceptatie van de detector). Om deze events te selecteren, wordt een combinatie van verschillende multi-muon-triggers gebruikt, waaronder een aantal triggers die specifiek voor deze analyse zijn ontworpen. De triggers variëren van jaar tot jaar, in overeenstemming met steeds veranderende trigger-menu's die bij ATLAS werden gebruikt. De events die door het triggersysteem zijn geselecteerd, worden verwerkt met behulp van offline reconstructiesoftware, die de geproduceerde deeltjes volledig karakteriseert. Voor deze analyse zijn de gebruikte muonen uitsluitend zogenaamde *gecombineerde* muonen: muonen waarvan het geladen spoor gereconstrueerd is in de Inner Detector tracker (ID) en in de muonspectrometerkamers

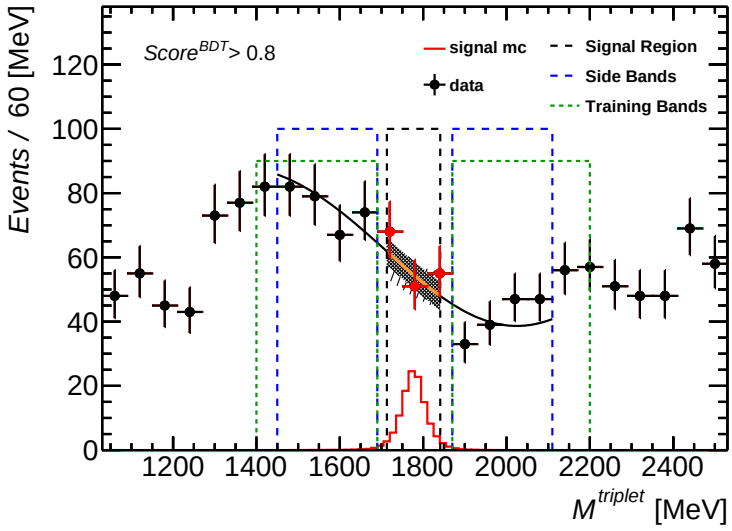
die de calorimeters omringen.
In voorbereiding tot deze analyse is een wijziging aangebracht in de ATLAS-reconstructie software om een optimale reconstructie van dicht opeengepakte muonen te garanderen. Waardevolle informatie over de signaalkandidaten wordt achterhaald door het veschoven



Figuur S.2: Grafische weergave van een muon-driepuntsvertex. Voor de duidelijkheid is de kromming van de sporen overdreven en projecties op het x-y-vlak zijn grijs gekleurd. De sporen van de drie muonen zijn in rood aangegeven, hun herberekende impulsen zijn in zwart weergegeven, en de kandidaat muon-triplet-impuls is als een blauwe pijl weergegeven. Verschillende hoeveelheden die gebruikt worden om de BDT-classificator te trainen voor deze analyse zijn gelabeld.

secundaire hoekpunt te reconstrueren, waar de paden van de drie muonen elkaar kruisen (Figuur S.2). De kennis van de positie van het secundaire hoekpunt maakt het mogelijk om de longitudinale en transversale verplaatsing van het moederdeeltje van de drie muonen te schatten, naast andere hoeveelheden van belang, en deze te vergelijken met dezelfde eigenschappen van gesimuleerde $\tau \to 3\mu$ signaal-events.
Voor de transversale verplaatsing (a_0^{xy}) is een systematische bias geobserveerd en begrepen als het effect van residuele uitlijningsfouten van de siliconendetector-lagen, achterblijvend na de anders uiterst succesvolle uitlijningsprocedure. De bias wordt gekenmerkt als een functie van data-collectie periode en afgetrokken van de variabele a_0^{xy} in de verzamelde data.
Signaal-events worden verder gekenmerkt door de isolatie van de drie muon-kandidaten van energetische sporen in de ID (P_T^{cone}) en van energiedeposities in de calorimeter, afkomstig van de onderliggende botsingsevent (E_T^{cone}). De twee hoeveelheden worden

berekend op een manier vergelijkbaar met de gebruikelijke isolatie van individuele muonen, maar gecorrigeerd voor de gelijktijdige aanwezigheid van drie muonen. Om het $\tau \to 3\mu$ verval te zoeken, richten we ons op de meest overvloedige bron van τ leptonen: hadronische jets. Rekening houdend met de acceptatie van de detector, wordt dit kanaal gedomineerd door $D_s \to \tau\nu$, gevolgd door $B \to \tau X$ (het verval van een B-meson naar een tauon en verdere deeltjes). De combinatie van deze twee kanalen wordt hier HF-channel genoemd. Voor de analyse wordt het signaal gemodelleerd als een simulatie van prompt en niet-prompt $D_s^{\pm} \to \tau(3\mu)\nu$ events alleen, bouwend op de gelijkenis tussen de secundaire kanaal ($B \to \tau X$) en deze. Een eerdere analyse bij ATLAS, uitgevoerd op Run-1 data, was gericht op een ander kanaal $W \to \tau\nu$, wat tot tauonen met hogere energie leidt, maar minder overvloedig dan het HF-channel.
De huidige analyse maakt gebruik van een reeks op sneden gebaseerde selecties om de achtergrond contaminatie te verminderen, deze is gedomineerd door events met twee mu-



Figuur S.3: De drie-muon invariante massadistributie in het bereik $M^{\text{triplet}} \in [1030, 2530]$ MeV. Signaal, weergegeven in het rode histogram en data events, weergegeven als volle stippen, passeren de *tight* selectie (sectie 5.3) en de extra snijd $Score^{BDT}$>0.8 zoals aangegeven. Rode markeringen geven data-punten binnen de verhulde regio aan, alleen onthuld gedurende de *unblinding* phase. Het signaalhistogram wordt geschaald naar $\mathcal{B}(\tau \to 3\mu) = 2{\times}10^{-7}$. De polynoom fit die wordt gebruikt om de verwachte achtergrond te extrapoleren naar de SR (signaalregio) wordt in zwart weergegeven, de sectie ervan die de verwachte achtergrond in de SR weergeeft is oranje gemarkeerd. De fit-onzekerheid wordt als grijze band weergegeven, terwijl de modelonzekerheid is gestreept is. Het oppervlak dat binnen de SR valt, is gemarkeerd door een zwart kader, terwijl de SB met een blauwe rand weergegeven zijn. De trainingsgebieden (TB) worden op dezelfde manier gemarkeerd door een groene rand.

onen en een toegeschreven π/K meson geproduceerd in jets. Deze selecties zijn vooral gericht op de kwaliteit van de muon-kandidaten en accepteren 65% van gereconstrueerde en getriggerde signaal-events, terwijl ze 86% van de achtergrond-events tegenhouden. Na de op sneden gebaseerde selectie worden de meest gevoelige variabelen in een multivariate discriminator (BDT) gespijsd. De classifier is getraind op gesimuleerde signaal-events tegen data-events, waarvoor de invariante massa van het muon-triplet binnen twee trainingsgebieden valt: $M^{\text{triplet}} \in [1400, 1690] \cup [1870, 2200]$, die de piek, beschreven door signaal-events, uitsluiten. De variabelen die worden gebruikt in de BDT-training worden gevalideerd op $D_s^{\pm} \to \phi(2\mu)\pi^{\pm}$ events: simulatie wordt *bin-wise* vergeleken met signaaltellingen geëxtraheerd uit data van 2017 (een fit van de D_s-massapiek wordt gebruikt om de events in data te schatten). De selectie op de BDT-discriminant wordt gekozen door het scannen van het minimum in de verwachte limiet op de vervalsfractie van tauon in drie muonen ($\mathcal{B}(\tau \to 3\mu)$), na het verwerpen van de di-muonresonanties ρ^0 en ϕ. De BDT-selectie vermindert de achtergronden met een factor 100, terwijl ongeveer een derde van de signaal-events behouden blijft. Om de limiet te berekenen, worden de signaal-events geschaald tot een fiduciale meting van de productiedoorsnede van D_s bij ATLAS voor $\sqrt{s} = 13$ TeV p-p botsingen, en daarbovenop tot een op simulatie gebaseerde schatting van $\sigma^{B \to \tau X}$. Een limiet op $\mathcal{B}(\tau \to 3\mu)$ is bepaald op basis van het aantal geobserveerde data-events binnen de Signaalregio (SR) ($M^{\text{triplet}} \in [1713, 1841]$ MeV). Het aantal verwachte achtergrond-events is bepaald door het interpoleren van de distributie van data-events uit signaalvrije zijgebieden, zoals weergegeven in figuur S.3.

De analyse levert een geobserveerd bovenlimiet op $\mathcal{B}(\tau \to 3\mu)$ van 1.98×10^{-7} bij 90% CL. Deze limiet stemt overeen met de verwachte limiet, bij afwezigheid van signaal, van $1.8^{+1.0}_{-0.6} \times 10^{-7}$. De beperkende factor voor de analyse is de moeilijkheid om de topologieën van signaal en achtergronden te scheiden. Momenteel wordt een *Neural Network*-aanpak overwogen om de onderscheidingskracht te verbeteren. Het resultaat zou uiteindelijk gecombineerd kunnen worden met een Run-2 analyse gericht op het W-channel, soortgelijk met de Run-1 ATLAS-analyse [64].

About the Author

Matteo Bedognetti was born on June 27$^{\text{th}}$, 1987, in Chiavenna, a small town enclosed in the Italian Alps. He graduated at the "Liceo Scientifico Leonardo da Vinci" high-school in Chiavenna.
He succesively transferred to Milan, where he enrolled at the "Statale" University to study applied physics. In October 2011 he obtained his Bachelor degree with a thesis on the development of a portable multi-channel readout system for particulate detectors.
In 2012 the author moved to the Netherlands to study particle and astroparticle physics at the University of Amsterdam. He completed his degree in 2014 with a thesis on a study of the $Z \to \mu\mu\gamma$ decay channel as a control sample for the search for the lepton-flavour-violating $\tau \to \mu\gamma$ decay with the ATLAS detector. During the summer of 2013 he spent six weeks in Geneva, participating in the CERN Summer Student Programme.
In September 2014 the author began his PhD at the Nikhef institute in Amsterdam, under affiliation to the Radboud University in Nijmegen, where his supervisor Prof. Dr. Olga Igonkina was appointed Professor in September 2015. His work was initially focused upon the future inner detector of ATLAS, the ITK (to be installed during the long shut-down commencing in 2024); he specifically worked on the set-up of a laboratory at Nikhef for the test readout of the ITK's silicon modules and on robustness studies of a number of design layouts through simulation. He later shifted his efforts to the search for the lepton-flavour violating decay $\tau \to 3\mu$, the results of which are reported in this thesis.
During his years as a PhD student he has been teaching assistant to the UvA/VU course about C++, held by Prof. Dr. Wouter Verkerke, and assisted in the organization of the 15$^{\text{th}}$ international workshop on tau lepton physics held in Amsterdam in 2018. After the tragic passing of Prof. Igonkina in May 2019 Prof. Dr. Nicolo de Groot stepped in as promotor for his thesis.

Acknowledgements

Completing this thesis has been an incredible adventure. Five years are a long time; many great people have come into my life and some have left again. Whilst some have had a direct influence in the completion of this work I must thank each and everyone for the help, support, affection and the fun that have ultimately permitted me to succeed.

First and foremost I thank my promotor and co-promotor Nicolo and Auke-Pieter for their support and patience, for the innumerable comments the text and for pragmatically keeping me on the most direct path towards completion.

A special thanks goes to the *reading committee*: Jörg Hörandel, Wouter Verkerke and Bert Schellekens for agreeing to critically read through the whole text.

A huge thank you to Olya, who has always been there for me throughout the ups and downs of the crazy roller-coaster ride that is a PhD. Your passion for particle physics accompanied me through a fun and exciting time of my Master thesis, and inspired me to pursue a PhD career. It was an honour to be part of your LFV team for all these years. Pier-Olivier, Noam, Sami, Jörn, Ivan, Marcus, Ann-Kathrin, JJ, Pepijn, Terry and Edwin, all amazing people and great colleagues! I will miss you all.
Marcus, you're a force of nature. Remember when you were asking me questions about the analysis even before your postdoc at Nikhef had formally begun? And all the times I would feel guilty asking you for small favours which you would promise to "have a look at" in the evening: how many times have I been baffled by the ungodly hours of the night at which the answers would be delivered. You were an incredible support to the progression of the analysis, despite working together with all of Olya's students at once. I can't but feel that Stefanie needs to be thanked for this too.
Edwin, your hard work and your inputs have been very much appreciated. Keep up the good job, and do not doubt yourself!

For backup with the theory chapter, in particular, I thank Bert, Jan-Willem and Sacha. You were incredibly kind, promptly answering all my emails starting with an awkward "I have a question,".
To the Tokyo group: Dai, Kohei and Masahiro, thank you for your hard work and the collaboration in bridging between the Run-1 analysis and the Run-2 one.

Max, Wolfgang, Peter, Susumu, Susanne and all the others who went out of their way to help me with technical problems and had the patience to have a sit down and answer all my questions, this would not have been possible without you.
Also thanks to all the non-researchers at Nikhef that keep making lives easier for us nerds: in particular Ed, Pieter and Jeff, but also all the secretaries and the helpdesk, always super responsive and incredibly kind. Thank you Stan for the new coffee machines, thus ensuring that blood-caffeine levels would not drop below the safety limit (and for keeping the institute up and running).

Of my many colleagues I want to thank Tim, who began this trip with me; it was immensely helpful to have you to share problems (and solutions) day after day. –Who filled up */project/atlas* this time?– Your level-headedness has helped me overcome many a creeping doubt that I came to struggle with over the years. Also thank you for playing ping-pong and table-football (oh, and escape rooms) with me even though I could not equal your skill, and of course for being the best roomie event after event.
A special *thanks!* to all who have made my permanence at CERN super fun! Here's to all the waiting at the coffee machines to lunch/snack/borrel together. To all the Friday evenings spent chatting over a few pitchers of beer –Just how slow is that tap?– hoping that someone would take pity and volunteer a ride home after the last buses had left. And of course I could not forget the (oh sooo nerdy) pub-quizzes at Charly's. Thank you Antonia, Luca, Remco, Ruchi, Dominik, Xiaoli, Snežana, Sara, Łucasz and Fritz.
Thanks to all the LHCb-ers who allowed me to stay productive through challenging table-football games during breaks: Michele, Elena, Mick, Jacco, Maarten, Katja, Mauricio, Lais. There have been days where smashing that ball around has definitely kept me from smashing my monitor and other electronics across the office.
I want to thank all the *Nikheffer*s. Thank you Pierfrancesco, Lydia, Stefan, Gabriele, Ingrid, Magdalena, Wouter, Koen, Ivan, Rosemarie, Veerle, Siim, Jörn, Priscilla, Geert-Jan, Martijn, Pieter, Rolf, Nika, Vince, Enrico, Vasilis, Panos, Afroditi, Karel, Stergios, Michele, Naghmeh, Andrea, Giuseppe, Elisa, Sabrina, Jacopo F., Jacopo M. and Suvayu for being welcoming and helpful; for all the borrels and nights out. You were simply amazing!
And let's not forget the "new" generations, Fabrizia, Veronica, Jeroen, Giuseppe, Remco and Lucrezia, Broos, Marc, Birgit, Kees, Milo, Anamika, Maria, Anuradha, Archisman, Ash, Brian, Marko, Hannah, Stefano, Carlo, Federica, Rahoul, Alice, some maybe already beyond my time, but still, I like to think that we are all part of a great dynasty of Nikhef PhD students. We all struggled through difficult times and found support in the good heart of our fellow colleagues.
Laura, Rahoul and Tim D., it was a pleasure to run with you guys to ventilate my brain every now and again.
Thanks to Lydia and Lennaert for the evenings spent in Utrecht playing music and composing songs with questionable lyrics.
For tolerating my endless complaints over coffee I especially thank Birgit, Terry, Edwin and Federica. Sometimes a man really just needs to let out steam (definitely true if we

are talking about me).

To the best housemates of Pregnin: Morgane, Harri, Maria and Tommi. You have made me feel welcome right from the start. I loved all the great dinners and barbecues. Thanks for all the times you let me ride with you to CERN, making me avoid to bike in chill, rain or snow.
Now to Laura and Šarunas, we made ourselves a home and elevated it to Klipperweg Empire. There are too many reasons I am thankful to you to fit in a text. One for all let me recall the evenings sitting on the balcony reading in the sunset with Dumbo and hot tea. I love you guys.
Thanks also to the later *Empires*, to Maria, Gerardas, Barbora, Tom, Kyriakos, Korinna and Martin, you were very cool house mates!
Van harte bedankt! tot mijn familieleden in Nederland: Marcel, Hans en André voor de voortdurende steun. Martin, Nicole, Edward, Age, Kees, Jaap en Bea, jullie moet ik danken voor al de leuke eetentjes en de humor en voor de warmste verwelcoming.
To all the friends from all over the world that I met in the Netherlands: Prem for your hospitality, Nina for cutting through the bulls**t (and making delicious Moussaka), Sandra for being part of my childhood, Kandan for the food, Shah for the music, Giuseppe and Leo for all the games played, I feel privileged to know all of you.
Thank all the *Wushuers*, for the evenings sweating and the nights playing games and partying!
To all the (drunken) sailors with whom I had the pleasure of exploring new waters: Justina, Zbyněk, John, Rob, Rasa, Natalija, Prem, Andries, Anuradha, Laura, Giovanni and Šarunas. Thank you for spending with me some of the best holidays I have ever experienced and making me discover the passion for sailing.
Thank you also Jasper, Atie, Cherry, Patras, Genavaite, Ioanna, Dimitra, Christian, Carsten, Maggie, Shadow and Dumbo, you know what you did...
I want to thank Giovanni, with your big heart and hard head you have helped me through quite some tough time. Thank you for always being there for me. I look forward to see what's up for us next.
Last but not least, a big fat thank you to my family and friends who remained back in Italy, and who tolerate my showing up way too rarely. Grazie Antonio e Therese per essere sempre stati dalla mia parte nonostante io abbia la testa tanto dura, mi sia messo a studiare "robe difficili" e a girovagare per il mondo. Siete dei genitori spettacolari. Grazie oneesan Esther per essere semplicemente la migliore sorella che ci sia. Grazie Davide per avermi dato dei nipoti magnifici quanto Sebastian ed Ettore. Grazie Caterina per la copertina sprizzante carattere. Grazie Marco per tutte le chiaccherate e le camminate. Grazie a Roberta e famiglia per tutte le avventure che abbiamo passato insieme.