Ivan Angelozzi

IN PURSUIT OF LEPTON FLAVOUR VIOLATION

A search for the $\tau \rightarrow \mu\gamma\gamma$ decay with ATLAS at $\sqrt{s}$ = 8 TeV

IN PURSUIT OF LEPTON FLAVOUR VIOLATION

A SEARCH FOR THE $\tau \to \mu\gamma\gamma$ DECAY WITH ATLAS AT $\sqrt{s} = 8$ TEV

IN PURSUIT OF LEPTON FLAVOUR VIOLATION

A SEARCH FOR THE $\tau \to \mu\gamma\gamma$ DECAY WITH ATLAS AT $\sqrt{s} = 8$ TEV

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 28 november 2017, te 12:00 uur

door

IVAN ANGELOZZI

geboren te Rome, Italië

Promotiecommissie:

PROMOTOR	prof. dr. O.B. Igonkina (Radboud Universiteit Nijmegen)
PROMOTOR	prof. dr. ir. P.J. de Jong (Universiteit van Amsterdam)
OVERIGE LEDEN	prof. dr. H.G. Raven (Vrije Universiteit Amsterdam)
	prof. dr. E.L.M.P. Laenen (Universiteit van Amsterdam)
	prof. dr. W. Verkerke (Universiteit van Amsterdam)
	prof. dr. A.P. Colijn (Universiteit van Amsterdam and Universiteit Utrecht)
	dr. R. Bruijn (Universiteit van Amsterdam)
	dr. ir. C.J.G. Onderwater (Rijksuniversiteit Groningen)
	dr. S. Dhaliwal (Brandeis University)

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

ISBN 9789462337718

Cover design by Buro Brouns
Printed by Gildeprint - The Netherlands

This work is part of the research program of the *Stichting voor Fundamenteel Onderzoek der Materie (FOM)*, which is part of the *Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)*. It was carried out at the *Nationaal Instituut voor Subatomaire Fysica (Nikhef)* in Amsterdam, the Netherlands. The research was financially supported by a FOM-Projectruimte grant titled "Search for tau decays to a muon and a photon to understand the lack of anti-matter in the universe" of prof. dr. O.B. Igonkina.

Contents

Introduction

High energy physics is the branch of science that studies the nature of the fundamental building blocks of the universe. An accurate description of the subatomic particles and the laws which govern their interactions is provided by the Standard Model. It was developed in the second half of the twentieth century and its validity has been strengthened during the years by the discovery of the predicted particles. In 2012, the Standard Model picture was completed with the discovery of the last missing piece, the Higgs boson. However, our current understanding of the universe is not complete. For example, the Standard Model does not describe gravity, which is one of the four fundamental forces. Another shortcoming is represented by the fact that only about 5% of the content of the universe consists of the Standard Model particles.

Moreover, the original formulation of the Standard Model does not include masses for the neutrinos, which turned out to be actually massive after the observation of neutrino oscillations. These phenomena indicate the existence of mixing between generations of neutrinos, which can change from one flavour (electron, muon or tau neutrino) to another. The question arises: does flavour violation occur also in the sector of the charged leptons? Theoretical and experimental efforts have been made to answer this question throughout several decades. Nevertheless, no evidence has been found so far. This kind of observation would be an unambiguous sign of physics beyond the Standard Model. The search for lepton flavour violation (LFV) among charged leptons is the main topic of this thesis.

In high energy physics, particle accelerators relying on cutting-edge technologies are employed to perform experimental searches. The Large Hadron Collider (LHC), located at CERN, is currently the most powerful accelerator and it is designed to collide beams of high energetic protons. The ATLAS detector is one of the four detectors which record the collisions delivered by the LHC. It is used to measure the properties of the Standard Model and search for direct evidence of physics beyond the Standard Model.

This thesis presents a search for the lepton flavour violating $\tau \to \mu\gamma\gamma$ decay performed using LHC proton-proton collisions collected by the ATLAS detector at a centre-of-mass energy of $\sqrt{s} = 8$ TeV in 2012. Tau lepton decays offer an interesting opportunity to research LFV since several theories that go beyond the Standard Model predict observable decay rates for such processes. The $\tau \to \mu\gamma\gamma$ channel investigated here adds a new contribution to this scenario, considering that no experimental limits have been set on this decay so far.

Outline of the thesis

This thesis is structured as described below. My original contributions are also indicated.

- **Chapter 1** provides the motivations for the search presented in this thesis. After an introduction to the Standard Model, charged lepton flavour violation is discussed giving an overview of theoretical models and experimental searches.

- **Chapter 2** presents an overview of the LHC and the ATLAS experiment. In Section 2.3, the triggers employed in the LFV $\tau \to \mu\gamma\gamma$ analysis are introduced.

- **Chapter 3** describes the methods used for the reconstruction and identification of particles in ATLAS. Moreover, a dedicated object identification for the $\tau \to \mu\gamma\gamma$ decay, which represents my original contribution, can be found in Section 3.8.

- **Chapter 4** provides a detailed description of the search for the LFV $\tau \to \mu\gamma\gamma$ decay performed at $\sqrt{s} = 8$ TeV with the ATLAS detector. The analysis strategy, the steps of its implementation and the results obtained are presented. I personally developed this analysis during my PhD project.

The manuscript ends with conclusions, outlook and summary.

chapter one

Motivations

The Standard Model gives an accurate representation of the world of particle physics through the description of the subatomic particles and their interactions. The predictive power of the Standard Model has been highlighted by the discovery of particles like the W boson [1, 2], the Z boson [3, 4] and the top quark [5, 6]. The latest success was reached on the 4^{th} of July 2012 when the ATLAS [7] and CMS [8] experiments at the Large Hadron Collider (LHC) announced the discovery of the Higgs boson [9, 10].

The development of the Standard Model in the second half of the twentieth century allowed to describe three interactions: electromagnetic, weak and strong forces. The unification between the weak and the electromagnetic interactions was achieved thanks to the work of Glashow, Weinberg and Salam between 1961 and 1968 [11–13]. The Higgs mechanism, able to explain the mass generation of the elementary particles, was introduced in the theory by Higgs, Brout, Englert, Kibble, Guralnik and Hagen in 1964 [14–16]. The proof of the renormalizability of the model was provided by Veltman and 't Hooft in 1972 [17]. These are some of the milestones in the construction of the Standard Model.

Despite its success, the Standard Model is an incomplete theory. There are still several phenomenological and theoretical issues to address. The asymmetry between matter and antimatter in the universe, the large fraction of "unknown" universe with only 4.9% made of visible matter [18] and the presence of a fourth force, the gravity, which is not included in the theory, are some of the big topics to deal with. These shortcomings leave room for theories that go beyond the Standard Model.

Furthermore, the observation of neutrino oscillations [19, 20] requires an extension of the original formulation of the Standard Model to include neutrino masses, showing a clear evidence of lepton flavour violation (LFV) in the neutrino sector. This scenario implies the possibility of a similar violation, which has not been observed so far, among the charged leptons.

However, even within this extended version of the Standard Model, the expected probability for charged LFV processes is negligible [21, 22]. Nevertheless, the experimental searches in this sector are well motivated by several theories beyond the Standard Model that predict observable decay rates, see e.g. [23, 24]. In this context, an interesting case is represented by LFV tau decays since these leptons are expected to have a strong coupling with physics beyond the Standard Model.

An overview of some theoretical models concerning the charged LFV processes is presented later in this chapter. A particular attention is dedicated to the $\tau \to \mu\gamma$ channel, which is one of the most sensitive probes of LFV in tau decays, and to the $\tau \to \mu\gamma\gamma$ channel, the subject of the search described in this thesis.

1.1 Ingredients of the Standard Model

The Standard Model picture has been completed with the discovery of the Higgs boson in 2012. The Standard Model consists of fermions with semi-integer spin and gauge bosons with integer spin in addition to the Higgs boson, assumed to be a scalar (spin 0).

Fermions

The fermions can be split into quarks and leptons divided into three generations differing only by their mass. For the leptons, the generations are formed by a charged lepton and a chargeless neutrino. The first generation is composed by the electron (e^-) and the electron neutrino (ν_e); the second generation consists of the muon (μ^-) and its neutrino (ν_μ) and the third of the tau (τ^-) and the corresponding neutrino (ν_τ). Leptons can interact via the electromagnetic force (only in the case of charged leptons) and via the weak force (both charged leptons and neutrinos).

For the quarks, up (u) and down (d) quarks constitute the first and lightest generation; the second includes the quarks charm (c) and strange (s) and the third the quarks top (t) and bottom (b). Quarks have an extra quantum number with respect to leptons, the colour charge, which can be of three different types denoted as red, green and blue. Carrying the colour charge, they can interact via the strong force, in addition to the weak and the electromagnetic forces. Quarks are confined in colour neutral bound states called hadrons, which consist of three quarks (baryons) or a quark and an antiquark (mesons).

All the fermions have a corresponding antiparticle with the same mass and the same spin but opposite electric charge.

Gauge bosons

The gauge bosons, with spin 1, are the mediators of the three fundamental forces described in the Standard Model. The photon (γ) is the mediator of the electromagnetic force and it is massless. The massive $W^\pm$ and Z bosons mediate the weak interaction. The eight gluons (g), massless like the photon, mediate the strong interaction carrying the colour charge.

Higgs boson

To complete the picture, we have to take into account the Higgs boson (H), a fundamental component of the Standard Model. Theorized for the first time in 1964 [14], the Higgs is the only scalar boson. It is the particle associated with the Higgs field, which plays a key role in the model. The Higgs mechanism, indeed, illustrates that the particles gain their masses through the interaction with this field. The table of the Standard Model particles is shown in Figure 1.1.

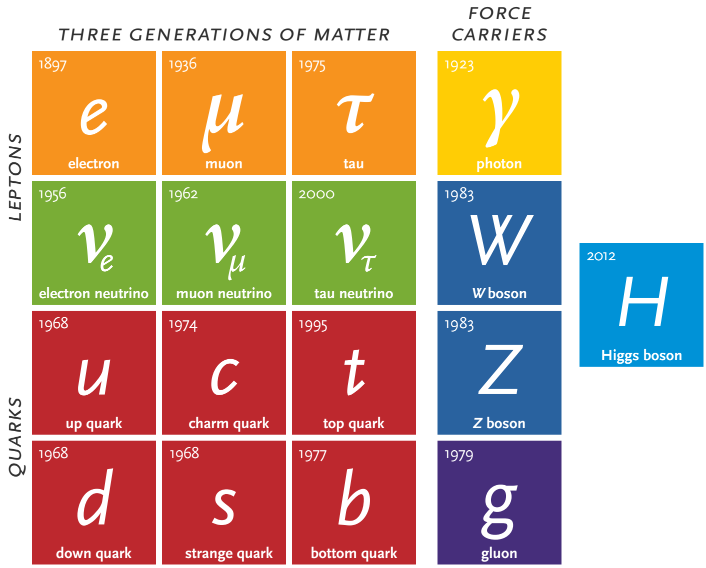


Fig. 1.1 · Schematic view of the Standard Model with the three generations of fermions, the gauge bosons and the Higgs boson. The year of their discovery is also indicated. Figure taken from [25].

1.1.1 Lagrangian of the Standard Model

The Standard Model, from the mathematical point of view, is a quantum field theory based on local gauge invariance. The model is renormalizable [17] and describes three of the four natural forces and the dynamics of the subatomic particles. The Standard Model is invariant under the symmetry group

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

The $SU(3)_C$ group refers to the strong interaction where C stands for the colour charge and the mediators are the eight gluons. $SU(2)_L$ describes the weak interaction where L stands for left-handed[1] since only the left-handed fermions take part in the interaction. Finally, $U(1)_Y$ characterizes the interaction where Y is the hypercharge. The mediators of the electroweak interaction, $SU(2)_L \otimes U(1)_Y$, are the gauge bosons W^1_μ, W^2_μ, W^3_μ and B_μ.

Section 1.2 briefly describes the electroweak theory and the introduction of the Higgs mechanism that explains the origin of the mass of the particles. The discovery of the W and Z bosons and lately of the Higgs boson provided experimental confirmation of this model.

[1] The terms left-handed and right-handed indicate chirality states. Left-handed fermions are arranged in weak isospin doublets while right-handed fermions appear as singlets.

1.2 Electroweak theory and Higgs mechanism

Building blocks of the Standard Model are presented in this section.[1] An exhaustive treatment of these topics is given e.g. in [26, 27].

Quantum Electro-Dynamics (QED) describes the interactions between fermions by means of exchange of photons. The dynamics of a fermion can be derived from the Lagrangian

$$\mathcal{L} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{1.2}$$

where ψ is a fermion field, γ^{μ} are Dirac matrices, $\partial_{\mu} \equiv \partial/\partial x^{\mu}$ is the derivative operator and m represents the mass of the particle. It is possible to impose the invariance of this Lagrangian under simple $U(1)$ local gauge transformations

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

where $\alpha(x)$ is a parameter which depends arbitrarily on the space-time coordinate x. The Lagrangian in equation 1.2 is not invariant under these transformations. To achieve this goal, a new vector field (A_{μ}) is introduced, replacing in addition the derivative ∂_{μ} with the *covariant derivative*, D_{μ}. The latter is defined as

$$D_{\mu} \equiv \partial_{\mu} - ieA_{\mu}, \tag{1.4}$$

where A_{μ} transforms as $A_{\mu} \rightarrow A_{\mu} + \frac{1}{e}\partial_{\mu}\,\alpha$.

In order to consider this field as the physical photon field, a term for its kinetic energy must be added, leading to the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{1.5}$$

where $F_{\mu\nu}$ is the electromagnetic field tensor defined as $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$.

1.2.1 Electroweak unification

To include the weak interactions in the theory, the symmetry group $SU(2)_L$ is introduced. The left-handed fermions form weak isospin doublets while the right-handed fermions form singlets. For instance, the first generation of leptons can be written as

$$\psi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad ; \quad \psi_R = e^-_R. \tag{1.6}$$

The $SU(2)$ transformations act on the left-handed doublets while the right-handed singlets do not interact under $SU(2)$. In this formulation, neutrinos are assumed to be only left-handed and massless.

[1] A description of quantum chromodynamics (QCD), the theory of the strong interactions, is beyond the scope of this thesis and is not given here.

The transformations under the symmetry group $SU(2)_L$ can be expressed as

$$\psi \to \psi' = e^{(i T_a \alpha^a(x))} \psi, \tag{1.7}$$

where T_a ($a = 1, 2, 3$) are the three generators of the $SU(2)_L$ group, 2×2 matrices related to the Pauli matrices, $T_a = \frac{1}{2}\tau_a$.

To realize the unification between weak and electromagnetic interactions, the Lagrangian must be invariant under $SU(2)_L \otimes U(1)_Y$ transformations. Four gauge bosons are associated with this symmetry: W^1_μ, W^2_μ, W^3_μ for the group $SU(2)_L$ and B_μ for the group $U(1)_Y$. The covariant derivative of the electroweak theory, required to preserve the local gauge invariance, is defined as

$$D_\mu = \partial_\mu + igT_a W^a_\mu + ig'\frac{Y}{2}B_\mu, \tag{1.8}$$

where g and g' are the coupling constants of the electroweak interaction. The electroweak Lagrangian can be expressed as the sum of two contributions:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{gauge}}. \tag{1.9}$$

The fermionic component of the Lagrangian is given by

$$\mathcal{L}_{\text{fermion}} = \sum_\psi \bar{\psi}\gamma^\mu (i\partial_\mu - g\frac{1}{2}\tau_a W^a_\mu - g'\frac{1}{2}YB_\mu)\psi. \tag{1.10}$$

The contribution to the electroweak Lagrangian that describes the dynamics of the four gauge bosons is given by

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.11}$$

where $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors for the fields W^a_μ and B_μ, respectively.

The electroweak theory presented so far is successful in the description of the interactions but leaves the gauge bosons and the fermions massless. Furthermore, the introduction of mass terms by hand is forbidden by the gauge invariance. The Higgs mechanism represents the right solution to this issue, allowing a gauge invariant mass generation for the $W^\pm$ and Z bosons and the fermions.

1.2.2 Higgs mechanism

The Higgs mechanism plays a key role in the formulation of the Standard Model. The basic idea is to introduce a complex $SU(2)_L$ doublet with hypercharge $Y = +1$

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.12}$$

written in terms of four scalar fields ϕ_1, ϕ_2, ϕ_3 and ϕ_4. The doublet enters into the Higgs Lagrangian, which is added to the electroweak one and defined as

$$\mathcal{L}_H = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \tag{1.13}$$

where $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ is the Higgs potential and D_μ is the covariant derivative defined in equation 1.8. The parameter λ must be positive in order to have a potential with a bound state of minimal energy (vacuum). Depending on the sign of μ^2, there are two possible cases:

- $\mu^2 > 0$: the potential has only one minimum in $|\phi| = 0$
- $\mu^2 < 0$: there is a manifold of minima for the potential given by

$$|\phi|^2 = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda}. \tag{1.14}$$

The interesting scenario is represented by $\mu^2 < 0$. The vacuum state can be chosen between the infinite number of degenerate states of minimum energy. A suitable candidate is

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} \text{ with } v \equiv \sqrt{\frac{-\mu^2}{\lambda}}, \tag{1.15}$$

where v indicates a non-zero vacuum expectation value. As a consequence, the symmetry $SU(2)_L \otimes U(1)_Y$ gets spontaneously broken to the electromagnetic subgroup $U(1)_{\text{EM}}$. Considering fluctuations around the vacuum, the Higgs doublet can always be generalized as

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

where there is only one real scalar field, $h(x)$, which corresponds to the Higgs boson. This particular choice is denoted as *unitary gauge*.

Mass generation

The Higgs mechanism explains how the elementary particles acquire mass. For the gauge bosons, masses arise from the kinetic term of the Higgs Lagrangian defined in equation 1.13, by substituting the constant term of the Higgs doublet (equation 1.16) in the expression

$$\begin{aligned}(D_\mu\phi)^\dagger(D^\mu\phi) &= \left|\left(-i\frac{g}{2}\tau_a W_\mu^a - i\frac{g'}{2}B_\mu\right)\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}\right|^2 \\ &= \left(\frac{1}{2}vg\right)^2 W_\mu^+ W^{-\mu} + \frac{1}{8}v^2\left(W_\mu^3, B_\mu\right)\begin{pmatrix}g^2 & -gg'\\ -gg' & g'^2\end{pmatrix}\begin{pmatrix}W^{3\mu}\\ B^\mu\end{pmatrix}.\end{aligned} \tag{1.17}$$

The physical bosons $W^{\pm}$ which appear in the equation are obtained by a linear combination of the fields W^1_μ and W^2_μ,

$$W^{\pm}_\mu = \frac{W^1_\mu \mp iW^2_\mu}{\sqrt{2}}, \tag{1.18}$$

and their mass is given by $m_W = \frac{1}{2}gv$. The physical fields corresponding to the photon and the Z boson, A_μ and Z_μ, are derived from a linear combination of the two remaining gauge fields, W^3_μ and B_μ:

$$\begin{aligned} A_\mu &= \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \quad \text{with}\, m_A = 0, \\ Z_\mu &= -\sin\theta_W B_\mu + \cos\theta_W W^3_\mu \quad \text{with}\, m_Z = \frac{gv}{2\cos\theta_W}. \end{aligned} \tag{1.19}$$

The fields A_μ and Z_μ are expressed in terms of the weak mixing angle θ_W with $\sin\theta_W = g'/\sqrt{g^2+g'^2}$ and $\cos\theta_W = g/\sqrt{g^2+g'^2}$. The values of the masses for the photon and the Z boson are obtained calculating the eigenvalues of the 2×2 matrix in equation 1.17. The photon remains massless and the relationship between the $W^{\pm}$ and Z boson masses is given by

$$\frac{m_W}{m_Z} = cos\,\theta_W. \tag{1.20}$$

Furthermore, the spontaneous symmetry breaking introduces a massive scalar particle, the Higgs boson, as shown in equation 1.16. The mass of this particle depends on the parameters of the Higgs potential and it is defined as $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda v^2}$. Since λ is a free parameter in the Standard Model, the value of the mass of the Higgs boson is not predicted.

The same Higgs doublet that generates masses for gauge bosons is able to provide mass terms for fermions through the *Yukawa* interaction. For example, considering the case of the electron, a gauge invariant term can be added to the electroweak Lagrangian:

$$\mathcal{L}_e = -\lambda_e \left[(\bar{\nu}_e, \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R(\phi^-, \bar{\phi}^0) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right]. \tag{1.21}$$

After spontaneous symmetry breaking, implementing the expression in equation 1.16, the Lagrangian term becomes

$$\mathcal{L}_e = -\frac{\lambda_e}{\sqrt{2}} v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{\lambda_e}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)h = -m_e \bar{e}e - \frac{m_e}{v}\bar{e}eh. \tag{1.22}$$

The first term in equation 1.22 can be interpreted as the mass term for the electron with $m_e = \lambda_e v/\sqrt{2}$, where λ_e represents the Yukawa coupling. The second term, proportional to m_e/v, describes the interaction between the Higgs boson and the electron.

The quark masses are generated in a similar way. Unlike the leptons, also the upper components of the $SU(2)_L$ doublets are massive in this case. A detailed explanation can be found in [26].

1.3 Neutrinos: mixing and mass

In the original formulation of the Standard Model previously presented, the neutrinos are massless and right-handed neutrinos are not included. The evidence that these particles have a non-zero mass was provided by several experiments observing neutrino oscillations whereby a neutrino produced with a specific flavour at the source can later be detected with another flavour. The 2015 Nobel Prize in Physics was awarded for this discovery made by two different experiments [19, 20].

Oscillations have been observed using neutrinos produced by four different sources: sun, atmosphere, nuclear reactors and accelerators. In solar neutrino experiments, the disappearance of the electron neutrinos ν_e is observed e.g. in the SNO experiment [28]. The KamLAND measurements of reactor neutrinos show instead the disappearance of the electron antineutrinos $\bar{\nu}_e$ [29]. In the case of atmospheric neutrinos, the Super-Kamiokande experiment observed the disappearance of muon neutrinos ν_μ and antineutrinos $\bar{\nu}_\mu$ [30, 31]. The ν_μ disappearance is confirmed by the long-baseline accelerator experiment K2K [32]. Strong indications for $\nu_\mu \to \nu_e$ oscillation were also given in recent times by the accelerator neutrino experiments T2K [33] and MINOS [34]. Evidences for ν_τ appearance due to $\nu_\mu \to \nu_\tau$ oscillations have been reported by Super-Kamiokande [35] and OPERA [36] collaborations.

These experiments have furnished a compelling evidence for the existence of three-flavour neutrino mixing defined as

$$\nu_l = \sum_{j=1}^{3} U_{lj}\,\nu_j, \tag{1.23}$$

where the neutrino flavour eigenstates, ν_l ($l = e, \mu, \tau$), are a quantum superposition of the mass eigenstates ν_j. The transition between them is realized through the mixing matrix U, called Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. Equation 1.23 implies the non-conservation of the individual lepton numbers L_e, L_μ and L_τ.

The observation of neutrino oscillations proves that neutrinos have mass, but the experimental data provide information only on their squared mass differences. Stringent upper bounds on the absolute values can be set by direct neutrino mass experiments. The Troitsk experiment set the upper limit on the combination of mass eigenstates that makes up the electron antineutrino to 2.2 eV at 95% confidence level [37]. Moreover, the Planck Collaboration reported an upper limit on the sum of the neutrino masses obtained using cosmic microwave background (CMB) and astrophysical data: $\sum_j m_j < 0.23$ eV at 95% confidence level [18].

At the present time, the real nature of the neutrinos is not yet determined. Being electrically neutral, massive neutrinos can be Dirac or Majorana particles. In the first case, the neutrino and its antiparticle, the antineutrino, are distinct and the total lepton number L is conserved in the interactions. The Dirac neutrino masses can be introduced in the Standard Model by including right-handed neutrinos. These are coupled to the left-handed lepton doublet and the conjugate Higgs doublet, which is defined as

$$\phi_c = -i\tau_2\phi^* = \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} \xrightarrow{\textit{symmetry breaking}} \frac{1}{\sqrt{2}} \begin{pmatrix} v+h \\ 0 \end{pmatrix} \tag{1.24}$$

and transforms identically to ϕ (equation 1.12), but has opposite hypercharge, $Y = -1$. The Lagrangian term, for a single generation, can be written as

$$\mathcal{L}_\nu = -\lambda_\nu(\bar{\nu}_l, \bar{l})_L \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} \nu_{lR} + \text{hermitian conjugate.} \tag{1.25}$$

After symmetry breaking, one obtains a Dirac mass term for the neutrinos, described by the following expression

$$\mathcal{L}_D = -m_D(\bar{\nu}_L\nu_R + \bar{\nu}_R\nu_L), \tag{1.26}$$

where $m_D = \lambda_\nu v/\sqrt{2}$. In this scenario, neutrinos acquire mass in an analogous way to quarks and charged leptons. However, the tiny masses of the neutrinos imply that λ_ν is unnaturally small, on the order of 10^{-12}.

As already mentioned, neutrinos could be Majorana particles. In this case, the neutrino field is equal to its charge conjugate and the neutrino and the corresponding antineutrino are the same particle ($\nu = \nu^c = C\bar{\nu}^T$, where C is the charge conjugation operator.) The Majorana mass term can be written as

$$\mathcal{L}_M = -\frac{m}{2}(\bar{\nu}_L\nu_R^c + \bar{\nu}_R^c\nu_L) = -\frac{m}{2}(\bar{\nu}_L C\bar{\nu}_L^T + \text{h.c.}) = -\frac{m}{2}\bar{\nu}\nu, \tag{1.27}$$

where $\nu = \nu_L + \nu_R^c$. In addition, the lepton number can be violated by two units.

An interplay between Dirac and Majorana neutrino mass terms is also possible. This scenario is realized in the *seesaw mechanism*, described in Section 1.4.1, which provides a natural explanation of the lightness of the neutrinos in comparison with the other fermions.

The nature of the neutrinos also affects the PMNS matrix previously introduced. If the massive neutrinos are Dirac particles, there is a single physical phase responsible for CP violation in the lepton sector, in analogy with the single phase in the CKM quark mixing matrix [38]. If the neutrinos are Majorana fermions, the

PMNS matrix contains two additional CP violation phases. It can be parameterized as

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \text{diag}\left(1, e^{i\alpha}, e^{i\beta}\right), \tag{1.28}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, θ_{ij} being the three mixing angles, δ is the Dirac CP violation phase and α and β are the two Majorana CP violation phases.

The three-neutrino oscillation analysis of the experimental data involves six parameters: two neutrino mass differences, $\Delta m^2_{ij} \equiv (m^2_i - m^2_j)$, defined as Δm^2_{21} and Δm^2_{31}, the three mixing angles and the Dirac phase δ. The two Majorana phases do not affect the oscillations. While the smaller of the neutrino mass differences, Δm^2_{21}, can be always assumed to be positive by convention, the sign of Δm^2_{31} cannot be determined from the data available. Therefore, there are two possible hierarchies for the neutrino masses: $m_1 < m_2 < m_3$, defined as *normal* ordering, and $m_3 < m_1 < m_2$, defined as *inverted* ordering. The results of a global analysis performed using data from neutrino oscillation experiments can be found in [39]. The best fit values obtained for the six parameters are given in Table 1.1.

Parameter	NO best fit $\pm 1\sigma$	IO best fit $\pm 1\sigma$
$\sin^2\theta_{12}$	$0.304^{+0.013}_{-0.012}$	$0.304^{+0.013}_{-0.012}$
$\sin^2\theta_{23}$	$0.452^{+0.052}_{-0.028}$	$0.579^{+0.025}_{-0.037}$
$\sin^2\theta_{13}$	$0.0218^{+0.0010}_{-0.0010}$	$0.0219^{+0.0011}_{-0.0010}$
$\delta/^\circ$	306^{+39}_{-70}	254^{+63}_{-62}
$\Delta m^2_{21}[10^{-5}\,\text{eV}^2]$	$7.50^{+0.19}_{-0.17}$	$7.50^{+0.19}_{-0.17}$
$\Delta m^2_{3l}[10^{-3}\,\text{eV}^2]$	$2.457^{+0.047}_{-0.047}$	$-2.449^{+0.048}_{-0.047}$

Table 1.1 · Best fit values for the three-neutrino oscillation parameters extracted from [39]. The values are obtained assuming to know a priori the neutrino mass ordering: normal (NO) in the first column and inverted (IO) in the second column. Note that $\Delta m^2_{3l} \equiv \Delta m^2_{31} > 0$ for NO and $\Delta m^2_{3l} \equiv \Delta m^2_{32} < 0$ for IO.

1.4 Charged lepton flavour violation

The non-conservation of the individual lepton numbers is established by the neutrino mixing, as mentioned in the previous section. As a consequence, LFV decays in the charged lepton sector, such as $\tau \to 3\mu$, $\tau \to \mu\gamma$, $\mu \to 3e$ and $\mu \to e\gamma$, are allowed in the Standard Model extension that includes massive neutrinos. These processes arise exclusively at loop level. Figure 1.2 shows a one-loop Feynman diagram of the $\tau \to \mu\gamma$ decay mediated by neutrino oscillation.

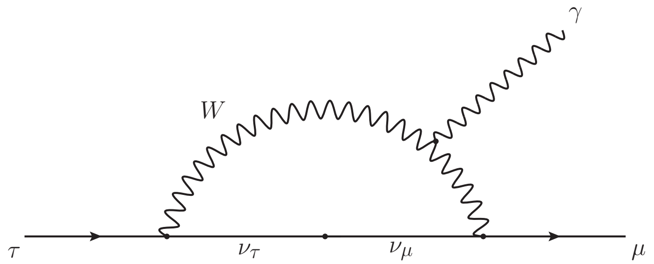


Fig. 1.2 · One-loop Feynman diagram of the LFV $\tau \to \mu\gamma$ process occurring through the $\nu_\tau \to \nu_\mu$ oscillation.

However, the rates of these processes are severely suppressed. They are indeed proportional to a factor $(\Delta m_{ij}/M_W)^4$, where Δm_{ij} represents the mass difference between the neutrinos and M_W is the W boson mass. The branching fractions for $l_i \to l_j\gamma$ decays are on the order of 10^{-40} or smaller [22, 23].

This value is far beyond the range of any kind of experimental search. Any observation of these processes would then be a clear indication of physics beyond the Standard Model. Strongly enhanced branching fractions for these decays, within the reach of present (or future) experiments, are predicted by theories such as the seesaw mechanism and supersymmetry.

1.4.1 The seesaw mechanism

The observation of neutrino oscillations requires an extension of the Standard Model. The seesaw mechanism is a model able to justify the lightness of the neutrinos with the introduction of new heavy particles.

There are three ways to extend the Standard Model through the seesaw mechanism. Neutrino masses can be generated by the exchange of three different types of heavy particles: fermion singlets (type I) [40, 41], scalar triplets (type II) [42, 43] and fermion triplets (type III) [44]. In the following, the type I seesaw model is described.

This is the simplest and most popular possibility, realized by introducing heavy right-handed neutrinos in the Standard Model, in addition to the left-handed ones. The neutrino mass terms can be written as

$$\mathcal{L}_Y^\nu = -\frac{1}{2}\begin{pmatrix}\bar{\nu}_L & \bar{\nu}_R^c\end{pmatrix} M \begin{pmatrix}\nu_L^c \\ \nu_R\end{pmatrix} + \text{hermitian conjugate}, \tag{1.29}$$

containing a 6×6 neutrino-mass matrix (considering three left-handed and three right-handed neutrinos), denoted as M, of the form

$$M = \begin{pmatrix} 0 & m_D \\ m_D^T & M_R \end{pmatrix}. \tag{1.30}$$

The matrix M can always be diagonalized by a unitary matrix U of the same dimension, i.e. $UMU^T = \hat{M}$, to obtain the mass eigenstates for the neutrinos. The terms m_D and M_R indicate the Dirac and Majorana mass matrices, respectively.

The interesting scenario is realized for $M_R \gg m_D$. In this case, right-handed neutrinos have mass $M_N \simeq M_R$ and light left-handed neutrinos have mass defined by the seesaw formula [40]:

$$m_\nu = -m_D M_R^{-1} m_D^T. \tag{1.31}$$

In this formulation, both light and heavy neutrinos are Majorana particles. The possibility to explain the tiny masses of the neutrinos is one of the attractive properties of the seesaw mechanism. The term "seesaw" indicates that the heavier are the heavy neutrinos, the lighter will be the light ones, as results from equation 1.31. Assuming a value of m_ν on the order of 0.1 eV, in agreement with the experimental limits, and the scale of m_D as the electroweak scale (around 246 GeV, the Higgs doublet vacuum expectation value), the mass scale of the heavy neutrinos is $M_N \sim 10^{14}$-10^{15} GeV. The seesaw mechanism generates in this way two very different mass scales for the neutrinos.

In the type I seesaw model, LFV processes can be induced at one-loop through the introduction of Majorana neutrinos. An example is shown in Figure 1.3 for $l_i \to l_j\gamma$ decays.

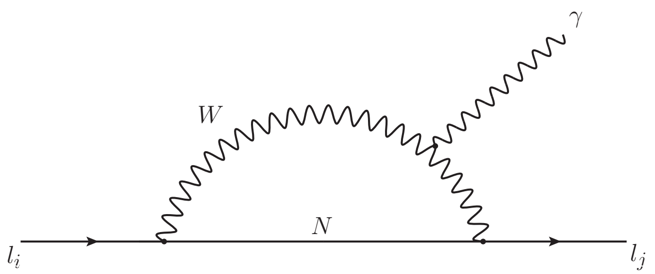


Fig. 1.3 · Feynman diagram for $l_i \to l_j\gamma$ decay with a heavy Majorana neutrino (N) in the loop.

In the seesaw model previously described, the range of predicted rates for LFV processes is broad. For the $\tau \to \mu\gamma$ decay, the largest possible branching fraction, which is reached only in a restricted region of parameter space, is given by $\mathrm{Br}(\tau \to \mu\gamma) \sim 10^{-9}$ [24]. This value is close to the current experimental upper limit, $\text{Br}(\tau \to \mu\gamma) < 4.4 \times 10^{-8}$ at 90% confidence level [45], and can be accessed by planned experiments [46].

There are possible extensions of the seesaw mechanism. An interesting example is represented by a model which introduces the same number of left-handed and right-handed neutral singlets in addition to the left-handed neutrinos [47]. Also in this case, the predicted rates for different charged LFV processes can be in the range of the experiments. For $\tau \to l\gamma$ decays ($l = e, \mu$), the maximal branching fraction is on the order of 10^{-8} [24].

Several models bring right-handed neutrino masses down to the TeV scale (*low-scale* seesaw), due to a weaker hierarchy between the neutrino mass scales. Sizeable branching fractions can be obtained for LFV decays, e.g. for $\tau \to \mu\gamma$ and $\tau \to e\gamma$, they can reach order of 10^{-7} in the study presented in [48].

1.4.2 LFV in supersymmetry

Supersymmetry (SUSY) is one of the most attractive extensions of the Standard Model. This theory implies a symmetry between fermions and bosons. For each Standard Model particle, there is a supersymmetric partner which only differs by the spin of a half-integer. Since these *superpartners* have not been observed so far, they must have a higher mass than the corresponding Standard Model particles. As a consequence, SUSY must be a broken symmetry. This theory could solve Standard Model shortcomings, such as the *hierarchy problem*, and provide a candidate for *dark matter*. A pedagogical introduction to SUSY can be found in [49].

The hypothetical SUSY particles with large mass could represent another source of lepton flavour violation. The *sleptons*, which are the supersymmetric partners of the leptons and carry the same flavour quantum numbers, are expected to play a key role. A large flavour mixing in the slepton sector, indeed, would enhance the branching fractions for rare LFV processes, such as $\tau \to \mu\gamma$ and $\mu \to e\gamma$ decays. Off-diagonal elements in the slepton mass matrices may bring LFV decay rates within the reach of the experiments.

The type I seesaw mechanism can also be implemented in supersymmetric models. In this scenario, LFV in right-handed neutrino Yukawa couplings induces LFV in slepton masses at low energies through renormalization-group effects [50]. This leads to unsuppressed LFV processes at loop level. Rates for $l_i \to l_j\gamma$ decays are enhanced for large values of $\tan\beta$, a SUSY parameter indicating the ratio of the vacuum expectation values of the two Higgs doublets, and for relatively light masses of the superparticles. In particular, for the $\tau \to \mu\gamma$ decay, the branching fraction is predicted to be as large as 10^{-7} [51].

The SUSY framework can also be used to investigate $l_i \to l_j\gamma\gamma$ decays. A study has been realized within the R-parity[1] violating (RPV) supersymmetric model [52]. The $l_i \to l_j\gamma\gamma$ decays can be generated through the diagram shown in Figure 1.4, where a heavy lepton decays into a lighter one and a virtual *sneutrino*, the SUSY partner of the neutrino, which decays into two photons via a triangle loop of

[1] The R-parity is a quantum number defined as $R = (-1)^{3(B-L)+2s}$, where B is the baryon number, L is the lepton number and s is the spin of a particle.

fermions. This possibility could be viable even in absence of observable $l_i \to l_j\gamma$ decays, representing an independent LFV probe.

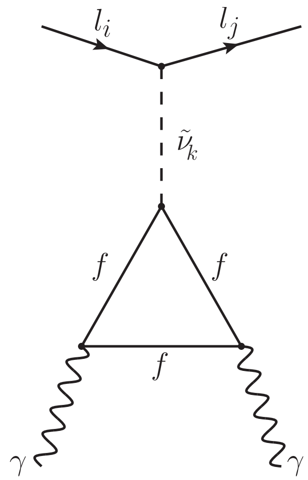


Fig. 1.4 · Topology of RPV SUSY $l_i \to l_j\gamma\gamma$ decay with a sneutrino that decays into two photons via a triangle fermionic loop.

The largest branching fraction for the $\tau \to \mu\gamma\gamma$ channel is obtained with a s-quark in the loop and corresponds to $\text{Br}(\tau \to \mu\gamma\gamma) \approx 10^{-10}$ [52]. No experimental limits have been set on this channel so far.

1.4.3 LFV via the Higgs sector

Another scenario of physics beyond the Standard Model predicts observable rates for charged LFV processes. Flavour violation can be probed also in the Higgs sector. Following the discovery of the Higgs boson, the ATLAS and CMS experiments are investigating its possible LFV decays (see Section 1.5.2).

LFV transitions mediated by Higgs bosons are allowed in extensions of the Standard Model. Two-Higgs-doublet models (2HDMs) may introduce LFV couplings between Higgs bosons and fermions. These models consist of five physical scalar particles: two CP-even neutral Higgs bosons (h and H), a CP-odd pseudoscalar (A) and a pair of charged Higgs bosons ($H^{\pm}$).

Higgs-mediated LFV effects can have a significant impact in rare lepton decays. The $\tau \to \mu\gamma$ process receives contributions from one-loop diagrams, as shown in Figure 1.5, and two-loop Barr-Zee type diagrams [53].

The flavour violating couplings involved in the Higgs-mediated $\tau \to \mu\gamma$ decay are constrained by the experimental bounds on this channel. The current limit on the $\tau \to \mu\gamma$ branching fraction translates into a constraint on the $\mu\tau$ Yukawa couplings: $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2} < 1.6 \cdot 10^{-2}$ [53]. As a result, the branching fraction for the LFV $H \to \tau\mu$ decay can reach up to 10%. More stringent constraints are obtained by direct searches performed in this channel, as described in Section 1.5.2.

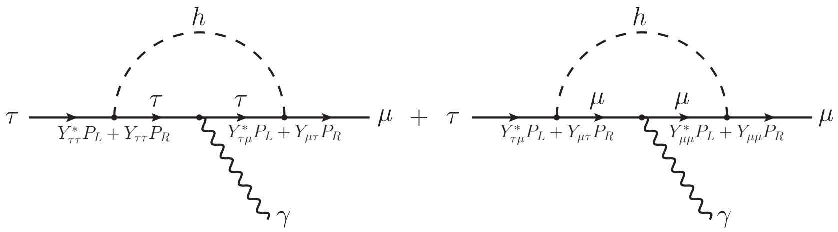


Fig. 1.5 · One-loop diagrams contributing to the LFV $\tau \to \mu\gamma$ decay mediated by a Higgs boson with flavour violating Yukawa couplings.

A prediction of the branching fractions for Higgs-mediated LFV $\tau \to \mu\gamma$ and $\tau \to \mu\gamma\gamma$ decays has been obtained in the study presented in [54]. The calculation concerns decays induced at one-loop level by a $H\tau\mu$ vertex generated by a Yukawa sector extended with operators of dimension higher than four. This results in upper bounds on the branching fractions for $\tau \to \mu\gamma$ and $\tau \to \mu\gamma\gamma$ decays on the order of 10^{-10} and 10^{-12}, respectively.

1.5 Charged LFV searches

The quest for flavour violation involving charged leptons can be carried out studying several decay modes. There is no experimental evidence for charged LFV so far, but limits are constantly improving in recent years, constraining the theoretical models. Several experiments include charged LFV searches in their physics program.

1.5.1 LFV tau decays

The tau leptons provide a large spectrum of LFV decays with different final states. An overview of the experimental upper limits on the branching fraction of the LFV tau decays is shown in Figure 1.6.

The LFV searches in this sector are dominated by the B factory experiments, Belle [56] and BaBar [57]. The large samples of tau leptons available were produced in e^+e^- storage ring colliders via the process $e^+e^- \to \tau^+\tau^-$ operating at the centre-of-mass-energy of the $\Upsilon(4S)$ resonance (10.58 GeV).

At the LHC, the proton-proton collisions produce a huge amount of tau leptons, offering the possibility to study their LFV decays. In this case, the major experimental challenge is represented by the several sources of background which affect the searches. The LHCb [58] and ATLAS experiments have performed analyses in this sector already during the first operational run (Run-1) of the LHC.

$\tau \to \mu\gamma$ decay

Among the LFV tau decays, the final state with a lighter lepton and a photon generates a great interest, motivated by several scenarios (beyond the Standard Model) that predict rates in the experimental range. In particular, we focus our attention on the $\tau \to \mu\gamma$ decay. The signal is characterized by a μ-γ pair with an invariant mass close to the nominal tau mass. The B factory experiments, taking

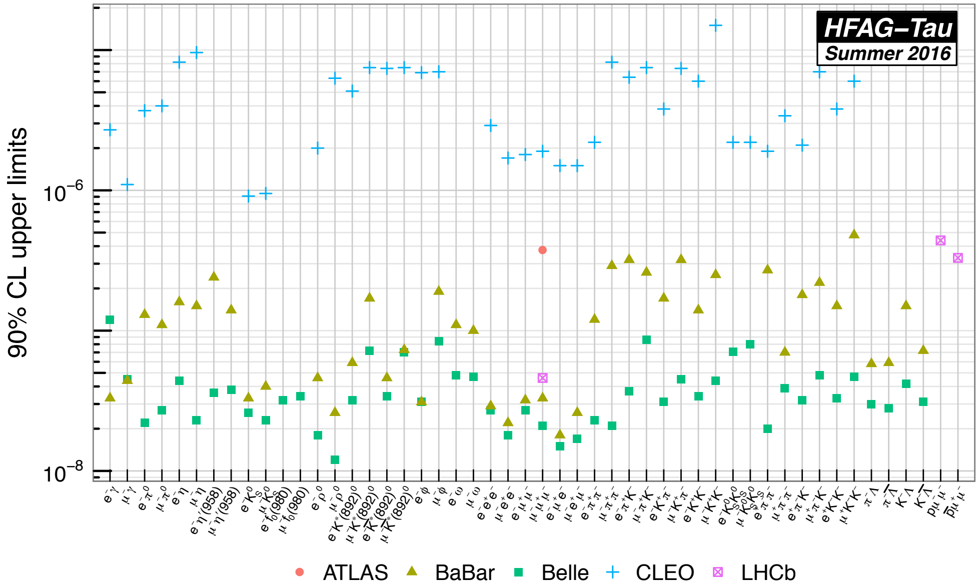


Fig. 1.6 · Summary plot for 90% confidence level upper limits on LFV tau branching fractions. Taken from [55].

advantage of the low-background environment, maintain a high signal efficiency. Currently, the world record upper limit on the $\tau \to \mu\gamma$ channel has been set by the BaBar experiment. As already mentioned, the limit corresponds to $\text{Br}(\tau \to \mu\gamma) < 4.4 \times 10^{-8}$ at 90% confidence level [45]. An analogous result was obtained by the Belle experiment, with a $\tau \to \mu\gamma$ branching fraction upper limit of 4.5×10^{-8} at 90% confidence level [59].

The ATLAS experiment started to search for LFV tau decays during the LHC Run-1. A first analysis of the $\tau \to \mu\gamma$ channel has been performed using the 2012 proton-proton collision data at the centre-of-mass energy of 8 TeV [60]. The tau leptons are produced from $W \to \tau\nu$ and $Z \to \tau\tau$ decays. Unlike the B factory experiments, the search is extremely challenging due to substantial background processes, such as heavy-flavour quark decays and $W \to \mu\nu$ decays. A preliminary result gives an expected upper limit on the $\tau \to \mu\gamma$ branching fraction on the order of 10^{-6} at 90% confidence level.

$\tau \to \mu\gamma\gamma$ decay

Flavour violation in the charged lepton sector can be studied in other tau lepton channels. The $\tau \to \mu\gamma\gamma$ decay may provide an independent LFV probe. Up to now, little attention has been devoted to this three-body decay by assuming a suppression of its rate with respect to $\tau \to \mu\gamma$. Nevertheless, as already described in Section 1.4, some theoretical models do not exclude the possibility of an observation. No experimental limits have been ever set on the LFV $\tau \to \mu\gamma\gamma$ decay. The first search for this channel, performed using the ATLAS detector, is presented in this thesis.

Other LFV tau decays
The branch of $\tau \to 3l$ decays (where l indicates an electron or a muon in any possible combination) is another appealing LFV source. In particular, several experiments performed analyses of the $\tau \to 3\mu$ channel. For the Belle and BaBar experiments, this decay mode has a very clean signature since the backgrounds can be reduced to a negligible level. The most stringent limit on the $\tau \to 3\mu$ channel has been set by Belle and corresponds to 2.1×10^{-8} at 90% confidence level [61].

The $\tau \to 3\mu$ decay is not only tested at the B factories. The LHCb detector, relying on good muon identification and powerful signal-background discrimination, is suitable for this search. Using Run-1 data at 7 and 8 TeV, the LHCb limit is 4.6×10^{-8} at 90% confidence level [62]. Finally, the $\tau \to 3\mu$ channel represents the first LFV tau decay on which the ATLAS collaboration set an upper limit. The analysis, performed using 2012 data at $\sqrt{s} = 8$ TeV, established an upper limit on the $\tau \to 3\mu$ branching fraction of 3.76×10^{-7} at 90% confidence level [63]. The result is not yet competitive with the other experiments, but shows the potential of the ATLAS detector as a probe of LFV in tau lepton decays.

The LFV tau decays include also final states where a lighter lepton is accompanied by one or more hadrons. The Belle and BaBar experiments performed several searches in this sector. The most stringent upper limit has been obtained by Belle in the $\tau \to \mu\rho^0$ decay: 1.8×10^{-8} at 90% confidence level [64].

1.5.2 Alternative LFV decays

The discovery of a single LFV channel would be an unambiguous sign of physics beyond the Standard Model, but would not be sufficient to reveal the underlying theoretical framework. In addition to tau lepton decays, LFV can be explored in muon decays and in the production of leptons in Z, Higgs and heavy neutral particles decays.

The abundant production of muons at high-intensity proton accelerators and the simple final states of their LFV decays allow to set stringent experimental bounds. The $\mu^+ \to e^+\gamma$ decay is characterized by a positron and a photon emitted back-to-back in the rest frame of the decaying muon, each with an energy equal to half the muon mass. The MEG experiment [65], at the PSI in Switzerland, set the lowest upper limit on the $\mu^+ \to e^+\gamma$ branching fraction so far, corresponding to 4.2×10^{-13} at 90% confidence level [66]. The signature of the $\mu \to 3e$ decay consists of three charged particles in the final state providing an invariant mass consistent with the muon mass. The current upper limit on this channel, 1.0×10^{-12} at 90% confidence level, was obtained by the SINDRUM experiment [67]. The Mu3e experiment, in preparation at the PSI at the time of writing, is designed to improve dramatically this result, reaching a sensitivity of 10^{-16} [68].

Furthermore, the muon to electron conversion $\mu^- + N \to e^- + N$ represents a LFV reaction. It consists of a muon that converts into an electron without accompanying neutrinos by exchanging a virtual photon with the capture nucleus (N). The present experimental limit has been obtained by the SINDRUM II experiment using a gold nucleus: $\text{Br}(\mu^- + Au \to e^- + Au) < 7 \times 10^{-13}$ at 90% confidence level [69].

In addition to charged lepton decays, LFV can be probed at higher mass. There are three LFV Z decays to investigate: $Z \to e\mu$, $Z \to \mu\tau$ and $Z \to e\tau$. Several experiments performed direct searches for these decays. At the LEP collider, the OPAL experiment set the stringent upper limits of $\text{Br}(Z \to e\tau) < 9.8 \times 10^{-6}$ and $\text{Br}(Z \to e\mu) < 1.7 \times 10^{-6}$ [70], while the DELPHI experiment placed the upper limit of $\text{Br}(Z \to \mu\tau) < 1.2 \times 10^{-5}$ [71], all with a confidence level of 95%. The LHC experiments exploit the large Z boson production cross-section with the drawback of a less clean environment with respect to LEP. ATLAS and CMS performed analyses of the $Z \to e\mu$ channel at $\sqrt{s} = 8$ TeV, setting the 95% confidence level upper limits at 7.5×10^{-7} [72] and 7.3×10^{-7} [73], respectively, improving the previous LEP limit. A Run-1 search has been performed also for $Z \to \mu\tau$ and $Z \to e\tau$ decays with the ATLAS detector. The 95% confidence level upper limits, $\text{Br}(Z \to \mu\tau) < 1.7 \times 10^{-5}$ [74, 75] and $\text{Br}(Z \to e\tau) < 4.7 \times 10^{-5}$ [75], have not yet improved the limits set by the LEP experiments previously mentioned.

Lepton flavour violating decays of the Higgs boson are also investigated. As already mentioned in Section 1.4.3, indirect constraints can be derived on the branching fractions of $H \to e\mu$, $H \to e\tau$ and $H \to \mu\tau$ decays. The experimental bound on the $\mu \to e\gamma$ channel places a constraint on the order of 10^{-8} on $\text{Br}(H \to e\mu)$ [53, 76]. Searches for $\tau \to \mu\gamma$, $\tau \to e\gamma$ and other decays result in much less stringent constraints, of order 10%, on $\text{Br}(H \to \mu\tau)$ and $\text{Br}(H \to e\tau)$ [53, 76]. Sensitive direct searches for LFV Higgs decays are carried out at the LHC. The ATLAS and CMS experiments performed $H \to \mu\tau$ analyses at $\sqrt{s} = 8$ TeV. In the CMS search, a slight excess of events with a significance of $2.4\,\sigma$ was observed, corresponding to a best fit value of $\text{Br}(H \to \mu\tau) = (0.84^{+0.39}_{-0.37})\%$. An upper limit was set at 1.51% with a confidence level of 95% [77]. A small excess of $1\,\sigma$, which corresponds to a best fit branching fraction of $\text{Br}(H \to \mu\tau) = (0.53 \pm 0.51)\%$, has been observed by ATLAS. The 95% confidence level upper limit is 1.43% [74]. Recently, CMS performed a search for this channel at $\sqrt{s} = 13$ TeV (LHC Run-2) using 2016 data. No significant excess is observed and a 95% confidence level upper limit is set at 0.25% [78]. This result excludes the branching fraction corresponding to the best fit for the $2.4\,\sigma$ excess observed in the 8 TeV analysis. The limit is used to constrain the $\mu\tau$ Yukawa couplings, $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2} < 1.43 \cdot 10^{-3}$ at 95% confidence level. The $H \to e\tau$ channel has also been studied in this search. No evidence is found for this process and the 95% confidence level upper limit is set at 0.61% [78], improving the previous limit obtained using 8 TeV data [79]. Moreover, CMS performed a search for the $H \to e\mu$ channel at $\sqrt{s} = 8$ TeV. The resulting upper limit is found to be $\text{Br}(H \to e\mu) < 0.035\%$ at 95% confidence level [79].

To conclude this overview, we briefly introduce searches for LFV decays of heavy neutral particles, such as Z' bosons and sneutrinos. The Z' boson is a massive hypothetical particle of spin 1, predicted by several Standard Model extensions [80]. A search for a Z' decaying into $e\mu$, $e\tau$ and $\mu\tau$ has been carried out with the ATLAS detector using data collected at $\sqrt{s} = 8$ TeV. The theoretical framework considered is the Sequential Standard Model (SSM) [80], where the Z' is assumed to have the same couplings as the Standard Model Z boson. No significant excess above the Standard Model expectation is observed in any of the three decay modes. Upper limits are placed on the production cross-section times branching fraction, $\sigma(pp \to Z') \times \text{Br}(Z' \to l_i l_j)$, as a function of the Z' mass. For a mass of 1 TeV, the 95% confidence level limits are 1.0, 4.0 and 9.9 fb for $e\mu$, $e\tau$ and $\mu\tau$ channels, respectively. In addition, lower limits are set on the Z' mass, which correspond to 2.5 TeV for the $e\mu$ channel and 2.2 TeV for $e\tau$ and $\mu\tau$ channels [81]. The search was also interpreted in terms of LFV decays of a tau sneutrino ($\tilde{\nu}_\tau$) in a RPV SUSY model. For a $\tilde{\nu}_\tau$ mass of 1 TeV, the 95% confidence level upper limits on $\sigma(pp \to \tilde{\nu}_\tau) \times \text{Br}(\tilde{\nu}_\tau \to l_i l_j)$ are 0.5, 2.7 and 9.1 fb for $e\mu$, $e\tau$ and $\mu\tau$ channels, respectively [81]. A new search performed at $\sqrt{s} = 13$ TeV with the ATLAS detector improved the previous limit on the mass of a SSM Z' decaying to $e\mu$. A lower limit of 3.01 TeV is set at 95% confidence level [82], as shown in Figure 1.7.

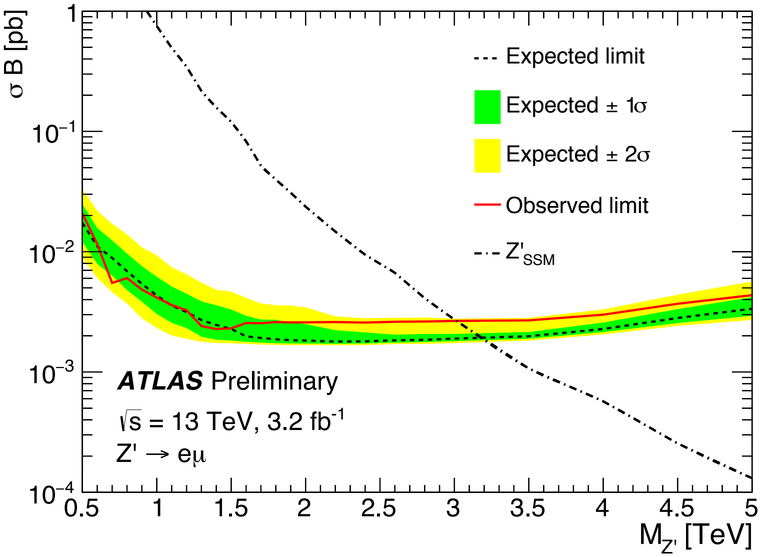


Fig. 1.7 · Upper limits on the Z' production cross-section in decay to an $e\mu$ final state as a function of the Z' mass at 95% confidence level. The expected limit is shown with the ±1 (green) and ±2 (yellow) standard deviation uncertainty bands. Taken from [82].

In summary, the upper limits on the branching fractions of the LFV decays described in Section 1.5 are reported in Table 1.2. The long quest for a LFV discovery in the charged lepton sector continues.

Sector	Search	Limit	Confidence Level	Experiment
Tau decays	$\tau \to \mu\gamma$	4.4×10^{-8} [45]	90%	BaBar
	$\tau \to 3\mu$	2.1×10^{-8} [61]	90%	Belle
		3.8×10^{-7} [63]	90%	ATLAS
	$\tau \to \mu\rho^0$	1.8×10^{-8} [64]	90%	Belle
Muon decays	$\mu \to e\gamma$	4.2×10^{-13} [66]	90%	MEG
	$\mu \to 3e$	1.0×10^{-12} [67]	90%	SINDRUM
$\mu \to e$ conversion	$\mu N \to eN$	7.0×10^{-13} [69]	90%	SINDRUM II
Z decays	$Z \to e\mu$	7.3×10^{-7} [73]	95%	CMS
		7.5×10^{-7} [72]	95%	ATLAS
	$Z \to e\tau$	9.8×10^{-6} [70]	95%	OPAL
		4.7×10^{-5} [75]	95%	ATLAS
	$Z \to \mu\tau$	1.2×10^{-5} [71]	95%	DELPHI
		1.7×10^{-5} [74]	95%	ATLAS
Higgs decays	$H \to e\mu$	0.036% [79]	95%	CMS
	$H \to e\tau$	0.61% [78]	95%	CMS
	$H \to \mu\tau$	0.25% [78]	95%	CMS (13 TeV)
		1.43% [74]	95%	ATLAS (8 TeV)

Table 1.2 · Overview of charged LFV searches for tau, muon, Z and Higgs decays. For each search, upper limit on the branching fraction, confidence level and name of the experiment are indicated.

chapter two

The LHC and the ATLAS detector

Progress and new discoveries in high energy physics need powerful machines exploiting the most innovative technologies. Particle accelerators, undergoing continuous evolution, play a key role to achieve these goals.

This chapter gives a description of the Large Hadron Collider (LHC), a particle collider located at CERN, at the border between France and Switzerland. An overview of the ATLAS detector, which was employed to collect the data used in this thesis, is also presented. In Section 2.3, the triggers used in the LFV $\tau \to \mu\gamma\gamma$ analysis are introduced.

2.1 The Large Hadron Collider

The largest and most powerful accelerator in the world today is the Large Hadron Collider [83]. The first big step for the LHC construction dates back to December 1994 when the CERN council approved the project. The LHC was built in the same tunnel which previously hosted the LEP collider [84], between 1998 and 2008.

The LHC is a two-ring superconducting accelerator situated near Geneva, in a tunnel of almost 27 km circumference, 100 m on average below the ground (between the Jura mountains and the Lac Léman). It is designed to collide beams of protons at the centre-of-mass energy of $\sqrt{s}$ = 14 TeV. In addition, it can also collide heavy (Pb) ions and protons with heavy ions. During the first operational run (Run-1), from 2010 to 2012, the actual centre-of-mass energy for proton-proton (p-p) collisions was $\sqrt{s}$ = 7 TeV in 2010 and 2011 and $\sqrt{s}$ = 8 TeV in 2012.

The LHC aims to answer unsolved physics questions. In Run-1, the discovery of the Higgs boson represented an outstanding success, completing the Standard Model picture. In the next years, the properties of this particle will be further investigated, also through the possible observation of new decay channels. Another main goal is the discovery of new particles at higher energies, predicted by theories beyond the Standard Model. Furthermore, the LHC could be important to understand some of the mysteries of our universe, such as the nature of dark matter and the asymmetry between matter and antimatter.

To pursue these ambitious purposes, two beams collide inside the LHC in four different points, where four detectors are located. ATLAS [7] and CMS [8] are two general purpose experiments, LHCb [58] investigates the physics of B-hadrons and their CP violation properties and ALICE [85] studies mainly the quark-gluon plasma in heavy ion collisions. Three smaller experiments complete the LHC environment: MoEDAL [86] searches for magnetic monopoles; TOTEM [87] studies elastic scattering and the total p-p cross-section; LHCf [88] analyzes the forward production of neutral particles.

2.1.1 Accelerator complex

The high energetic proton beams colliding at the LHC are obtained through a series of intermediate steps based on several accelerators. An electric field is used to strip electrons from hydrogen atoms to get protons, which are accelerated to 50 MeV in the first accelerator, Linac2. Protons are then injected into the Proton Synchrotron Booster (PSB), the first of three circular pre-accelerators, able to increase the energy up to 1.4 GeV. At this stage, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) raise the energy to 25 GeV and 450 GeV, respectively.

The beams are finally injected into the two beam pipes of the LHC. The two proton beams are arranged into bunches, each one containing about 10^{11} protons. They circulate in opposite directions and are accelerated to reach their maximum energy (corresponding to 4 TeV in 2012, for a centre-of-mass energy of $\sqrt{s}$ = 8 TeV). Protons are boosted to the required energies using 16 superconducting radiofrequency (RF) cavities, 8 per beam. They operate at 400 MHz, each delivering a voltage of 2 MV, which results in an accelerating field of 5 MV/m. A scheme of the accelerator complex is shown in Figure 2.1.

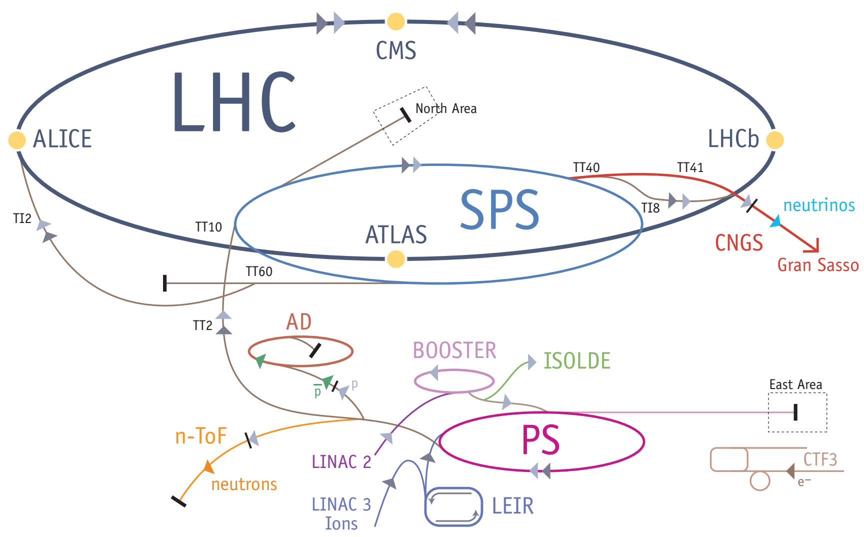


Fig. 2.1 · Schematic representation of the LHC and the injection chain. The four major experiments along the LHC ring are also indicated. Figure taken from [89].

Superconducting magnets are needed to direct the beams inside the LHC. These include 1232 superconducting dipole magnets which keep the beams in their nearly circular orbit, providing a maximum magnetic field of 8.33 T. To obtain this high magnetic field, they are cooled with superfluid helium and operate at a temperature of 1.9 K. In addition, 392 superconducting quadrupole magnets are used to focus the beams.

2.1.2 The LHC parameters

One of the principal goals of the LHC is the discovery of rare processes. The possibility to achieve this objective depends on two key parameters of the collider: the instantaneous luminosity of the beams, L, and the centre-of-mass energy ($\sqrt{s}$). The production rate, R, of a process with cross-section σ, indeed, can be expressed as

$$R = L\,\sigma(\sqrt{s}), \tag{2.1}$$

where σ depends on the centre-of-mass energy. The parameter $\sqrt{s}$ expresses the energy available in the collision, given by the sum of the energies of the two beams. As already mentioned, the LHC operated at a centre-of-mass energy of 7 and 8 TeV during Run-1. The instantaneous luminosity depends only on the beam parameters. For a Gaussian beam distribution, it is defined as

$$L \propto \frac{N_b^2 n_b f_{\text{rev}}}{4\pi\sigma_x\sigma_y}, \tag{2.2}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency and σ_x, σ_y the transverse dimensions of the beam at the interaction point.

Under nominal condition, each beam is made of 2808 bunches and two consecutive bunches are separated by 25 ns. In 2012, the LHC operated using 1380 bunches per beam separated by 50 ns.

The LHC is designed to achieve an instantaneous luminosity of $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. During Run-1, the highest peak luminosity was reached in 2012 with a value of $7.73 \cdot 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$. The total amount of data delivered by the LHC is described by the *integrated* luminosity

$$\mathcal{L} = \int L\,dt, \tag{2.3}$$

measured in inverse barns ($1\text{b} = 10^{-28}\,\text{m}^2$). Figure 2.2a shows the peak instantaneous luminosity per day in 2012; Figure 2.2b displays the integrated luminosity delivered by the LHC and recorded by ATLAS in 2011 and 2012.

During a bunch crossing, many *p*-*p* interactions can occur at the same time. This phenomenon, known as *pile-up*, makes the identification of the particles com-

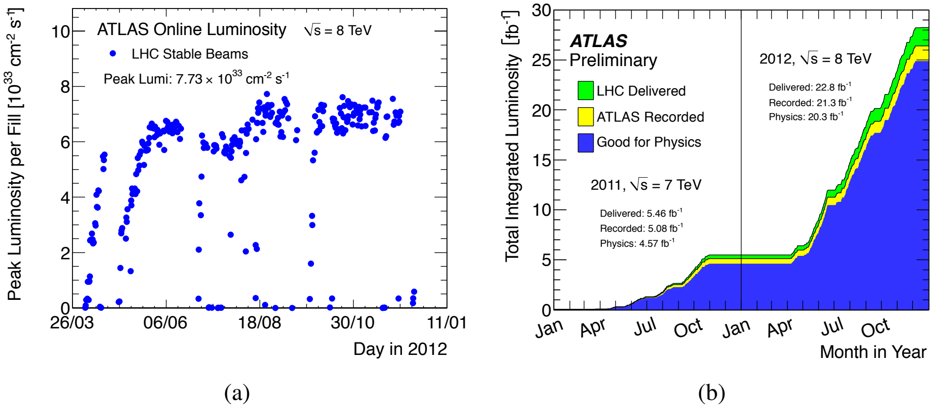


Fig. 2.2 · (a) Maximum instantaneous luminosity versus day delivered to ATLAS during stable beam periods in 2012. (b) Cumulative luminosity versus time delivered by the LHC (green), recorded by ATLAS (yellow) and certified to be good quality data with fully operational sub-detectors (blue) for p-p collisions at 7 and 8 TeV centre-of-mass energy in 2011 and 2012 [90].

ing from the different interactions more challenging. The mean number of interactions per bunch crossing $\langle\mu\rangle$ is defined as

$$\langle\mu\rangle = \frac{L\,\sigma_{\text{inel}}}{n_b f_{\text{rev}}}, \tag{2.4}$$

where σ_{inel} is the total inelastic p-p cross-section (estimated to be about 75 mb at $\sqrt{s} = 8$ TeV [91]). The value of $\langle\mu\rangle$ increased from 9.1 in 2011 to 20.7 in 2012, as shown in Figure 2.3.

The results presented in this thesis are based on the LHC Run-1 data collected during 2012. Table 2.1 summarizes the main LHC parameters, comparing the 2012 performance with their design values.

After a shutdown (LS1) of two years, necessary to upgrade the machine, the LHC started its Run-2 in 2015. The second run is characterized by the increase of the centre-of-mass energy from 8 to 13 TeV, which allows to improve the results obtained during Run-1 and explore new regions of phase-space. The performances of the LHC Run-2 are not discussed in this thesis since the LFV $\tau \to \mu\gamma\gamma$ analysis described here makes use of Run-1 (2012) data.

2.2 The ATLAS detector

The ATLAS (A Toroidal LHC Apparatus) detector [7] is one of the two general purpose detectors at the LHC, designed to detect the particles emerging from the p-p collisions. Its physics program includes several goals, such as:

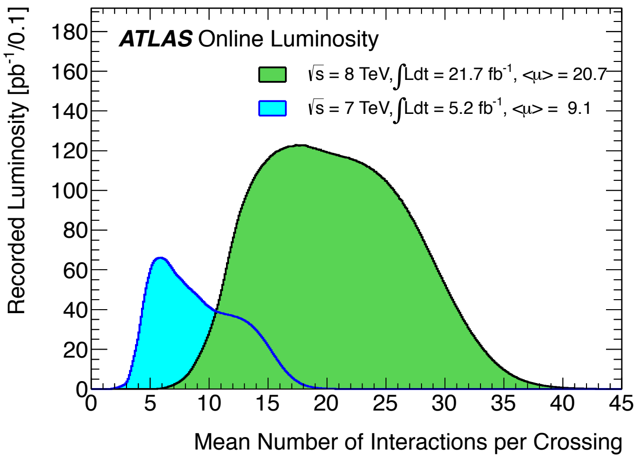


Fig. 2.3 · Luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 (blue) and 2012 (green) data [90].

Parameter	2012	Design
Centre-of-mass energy $\sqrt{s}$ [TeV]	8	14
Peak instantaneous luminosity [10^{33} cm^{-2}s^{-1}]	7.73	10
Delivered integrated luminosity [fb^{-1}/year]	22.8	80-120
Maximum number of colliding bunches	1380	2808
Minimum bunch spacing [ns]	50	25
Average number of interactions $\langle\mu\rangle$	20.7	19.02

Table 2.1 · The parameters deriving from the performances of the LHC during the p-p collisions in 2012 are extracted from [92] and compared to the design parameters taken from [93].

- the search for the Higgs boson;
- precision measurements of Standard Model observables;
- searches for direct evidence of physics beyond the Standard Model in the form of new particles;
- searches for indirect evidence of physics beyond the Standard Model, including lepton flavour violation and CP violation beyond the CKM formalism.

To explore this broad range of physics topics, the ATLAS detector consists of several sub-detectors, each with specific functions. They are installed in concentric layers around the interaction point. Moving outwards from the beam pipe, the sub-detectors are: the *Inner Detector* (ID), the *electromagnetic* (EM) *calorimeter*, the *hadronic calorimeter* and the *Muon Spectrometer* (MS). The magnet system is based on a superconducting solenoid surrounding the ID and three superconducting toroids placed around the calorimeters. The overall layout of the ATLAS detector is illustrated in Figure 2.4. A description of the detector and its components used during Run-1 is given in the following.

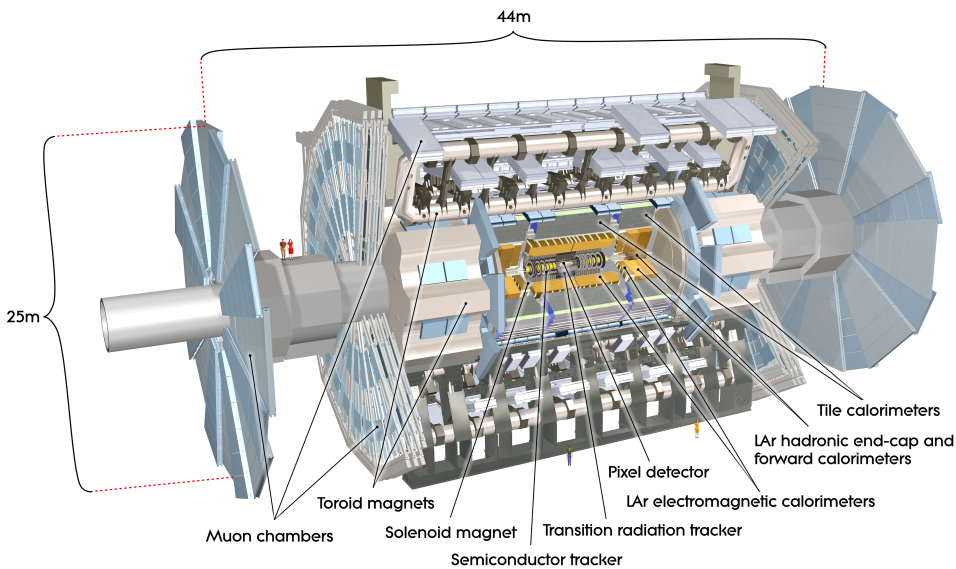


Fig. 2.4 · Cut-away view of the ATLAS detector [94]. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight is approximately 7000 tonnes.

2.2.1 ATLAS coordinate system

The nominal interaction point is defined as the origin of the right-handed coordinate system. The z-axis corresponds to the beam direction, the x-axis points towards the centre of the LHC ring and the y-axis is defined as pointing upwards. Cylindrical coordinates are commonly used: the azimuth angle $\phi \in [-\pi, \pi]$ is measured around the beam axis and the polar angle $\theta \in [0, \pi]$ is the angle with respect to the beam axis. The latter is usually expressed in terms of *pseudorapidity*, defined as $\eta = -\ln \tan(\theta/2)$, since the difference $\Delta\eta$ is invariant under boosts along the beam direction. Distances in the $\eta - \phi$ space are defined by the quantity $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Several parameters are calculated in the x-y plane, such as the transverse momentum, $p_T = p \sin\theta$, the transverse energy, $E_T = E \sin\theta$, and the missing transverse momentum E_T^{miss}.

2.2.2 Magnets

The ATLAS detector features a system of four superconducting magnets [95]. The central solenoid is aligned on the beam axis and provides a magnetic field of 2 T for the ID, bending the trajectories of charged particles. It is 5.8 m long and has a diameter of 2.5 m. The layout has been designed to minimize the amount of material in front of the calorimeters, resulting in a thickness of only 0.66 radiation lengths.[1]

Three large air-core toroids are employed to provide the magnetic field for the MS, a barrel toroid and two end-cap toroids. Each of the three toroids consists of eight coils, assembled symmetrically and radially around the beam axis. The coils of the end-cap toroids are rotated by 22.5° compared to the barrel toroid, in order to optimize the bending power at the interface between the two systems. Both solenoid and toroids operate in a superconductive regime, cooled down at the temperature of 4.5 K using liquid helium. The magnet system of the ATLAS detector is shown in Figure 2.5.

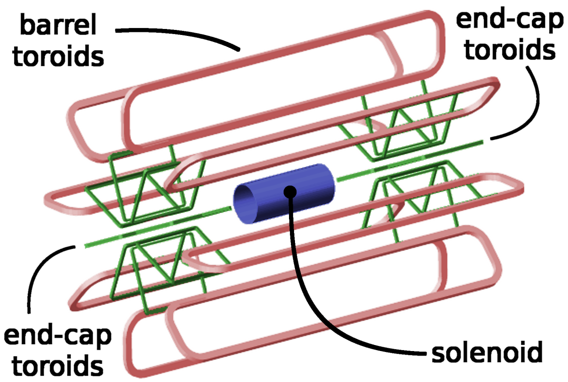


Fig. 2.5 · Schematic view of the ATLAS magnet system which consists of a central solenoid and three toroids (one barrel and two end-caps) [95].

2.2.3 The Inner Detector

The Inner Detector is the closest sub-detector to the beam line. It is designed to provide tracking information, measuring the momentum and impact parameters of charged particles and identifying primary and secondary vertices. It covers the pseudorapidity range $|\eta| < 2.5$. The system is designed to cope with a high-radiation environment. The ID is composed of three complementary sub-detectors: the Pixel Detector, the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT), as shown in Figure 2.6.

[1] The radiation length is defined as the mean distance over which an electron loses all but $1/e$ of its original energy.

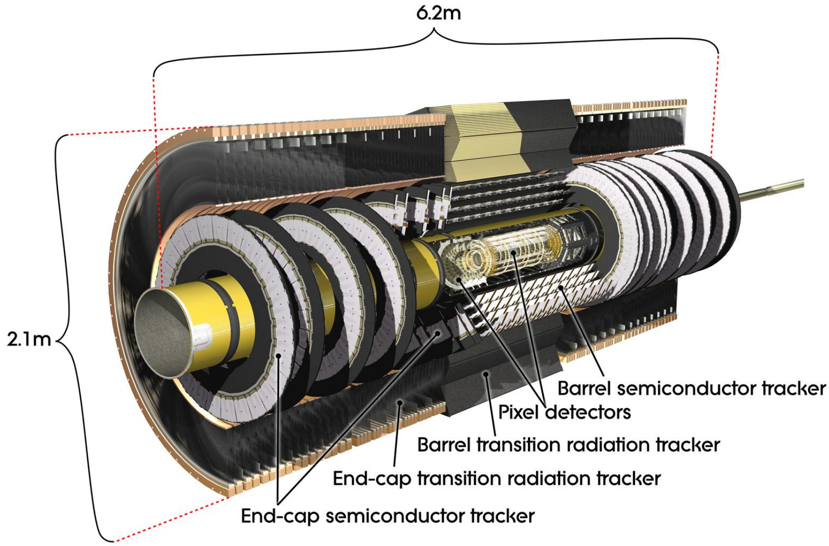


Fig. 2.6 · Cut-away view of the Inner Detector [96].

Pixel Detector

The Pixel Detector is made of three concentric layers around the beam axis in the barrel region and three disks perpendicular to the direction of the beam in the end-cap regions. It consists of 1744 silicon pixel sensors with a nominal pixel size of $50 \times 400\,\mu\text{m}^2$. In total, the Pixel Detector has approximately 80.4 million readout channels. The intrinsic resolution is $10\,\mu$m in the transverse direction (R-ϕ) and $115\,\mu$m in the longitudinal direction (z in the barrel and R in the end-caps). Generally, a charged particle crosses three pixel layers travelling through the detector.

SemiConductor Tracker

The SCT uses silicon microstrips rather than pixels. It consists of four cylindrical layers in the barrel region and nine disks on both sides of the end-cap region. It is made of 4088 two-sided modules for a total number of approximately 6.3 million readout channels. The strip pitch, defined as the distance between the centres of two neighbouring strips, is $80\,\mu$m both for barrel and end-caps. The intrinsic resolution corresponds to $17\,\mu$m in the transverse direction (R-ϕ) and $580\,\mu$m in the longitudinal direction (z in the barrel and R in the end-caps). The SCT is expected to provide eight strip measurements corresponding to four space points for each track.

Transition Radiation Tracker

The TRT is the outermost part of the ID. It is composed of about 350,000 straw tubes with a diameter of 4 mm. Each tube is filled with a gas mixture of 70% Xe, 27% CO_2 and 3% O_2. In the barrel region, the straws are 144 cm long and parallel to the beam axis. In the end-cap region, the straws are 37 cm long and arranged radially in wheels. Each tube acts as a cathode and contains a 31 μm diameter gold-plated tungsten wire, the anode. The TRT provides a large number of hits (typically 36) per track, carrying only R-ϕ information, with an intrinsic accuracy of 130 μm per straw.

The TRT plays also a role in the particle identification. The straws are interleaved with polypropylene fibres (barrel) and foils (end-cap). Charged particles crossing this radiator material emit transition radiation photons, which are absorbed in the Xe-based gas mixture. The intensity of the radiation is proportional to the Lorentz factor (E/m) of the charged particle, which is higher for an electron than for heavier particles, like hadrons, at a given particle energy. Working with two independent thresholds, the TRT can distinguish between electrons and charged hadrons, especially pions.

2.2.4 Calorimeters

The calorimeter system provides precise energy measurements for photons, electrons and hadrons, as well as an estimate of the missing transverse momentum. The calorimeters are sampling devices, made of alternating layers of an absorber and an active medium. Particles lose their energy interacting with the absorber, creating a *shower* of particles. The active medium measures the energy of the electromagnetic and hadronic showers. The system is divided into the EM calorimeter and the hadronic calorimeter, which also includes the forward calorimeter, covering a pseudorapidity range up to $|\eta| < 4.9$. Calorimeters are designed to ensure a good containment for electromagnetic and hadronic showers, stopping most of their particles. A view of the ATLAS calorimeter system is shown in Figure 2.7.

Electromagnetic calorimeter

The EM calorimeter [98] is used to detect and measure the energy of electrons and photons. It consists of a barrel ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$). Lead is used as absorbing material, while liquid argon (LAr) is chosen as active material for its linear response and radiation hardness.

The accordion geometry of the detector provides a full coverage in ϕ without any cracks. The barrel calorimeter consists of two identical halves, separated by a gap of 4 mm at $z = 0$. The end-caps are divided into two coaxial wheels, which are further segmented into eight wedge-shaped modules. The total thickness of the EM calorimeter is greater than 22 radiation lengths, X_0, in the barrel and greater than 24 X_0 in the end-caps.

In the region $|\eta| < 2.5$, the calorimeter is divided in three layers. The inner one has the finest granularity, $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$; the second layer collects the largest fraction of the shower energy; the third and outermost layer collects the tail of the electromagnetic shower and has the coarsest granularity. A presampler, which consists of an active LAr layer, is also installed in the region $|\eta| < 1.8$.

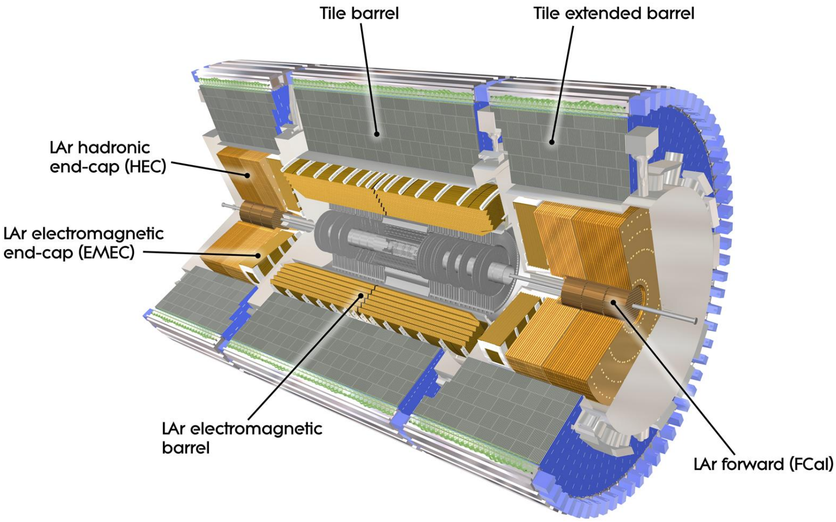


Fig. 2.7 · Cut-away view of the ATLAS calorimeter system [97].

Hadronic calorimeter

The hadronic calorimeter is designed to measure the energy of the hadrons and consists of three sub-systems: the tile calorimeter [99], the LAr hadronic end-cap calorimeter and the LAr forward calorimeter. The tile calorimeter uses steel as the absorbing material and scintillating tiles as the active medium. It is subdivided into a central barrel ($|\eta| < 1.0$) and two extended barrels ($0.8 < |\eta| < 1.7$). Central and extended barrels are divided azimuthally into 64 modules and segmented in depth in three layers.

The hadronic end-cap calorimeter employs copper plates as absorbers and LAr as active medium. Both end-caps consist of two independent wheels, built from 32 wedge-shaped modules. This calorimeter covers the range $1.5 < |\eta| < 3.2$. The depth of approximately 7.4 interaction lengths,[1] λ, in the barrel and 10 λ in the end-caps is able to minimize the *punch-through*, represented by the hadrons reaching the MS.

The forward calorimeter covers the range $3.1 < |\eta| < 4.9$. It consists of three modules on each side of the detector and uses LAr as active material. The innermost module is optimized for electromagnetic showers and copper is used as absorbing material; the other two measure the hadronic showers, employing tungsten as absorber. The depth of the forward calorimeter is approximately 10 λ.

[1] The interaction length is defined as the mean distance a hadron travels before undergoing a nuclear interaction.

2.2.5 Muon Spectrometer

Muons are minimum ionizing particles for most momenta relevant to ATLAS and are not stopped by the calorimeters. The Muon Spectrometer is the outermost part of the ATLAS detector and is designed to detect these particles and measure their position and momentum. It is based on the magnetic deflection of the muon tracks provided by the system of air-core toroid magnets. The MS measures the momentum of the muons between about 3 GeV (an energy sufficient to escape from the calorimeters) and 3 TeV in the pseudorapidity range $|\eta| < 2.7$. The resolution is approximately 10% for muons with p_T of 1 TeV. In addition to tracking, the MS is also designed to trigger on muons in the region $|\eta| < 2.4$. Therefore, precision-tracking and trigger chambers are incorporated in the system. In the barrel region, the chambers are placed in three cylindrical layers around the beam axis. In the end-caps, the chambers are arranged in four large wheels perpendicular to the beam direction. The tracking chambers consist of *Monitored Drift Tubes* (MDT) and *Cathode Strip Chambers* (CSC); also the trigger chambers are of two types, *Resistive Plate Chambers* (RPC) in the barrel and *Thin Gap Chambers* (TGC) in the end-caps. The layout of the MS is shown in Figure 2.8.

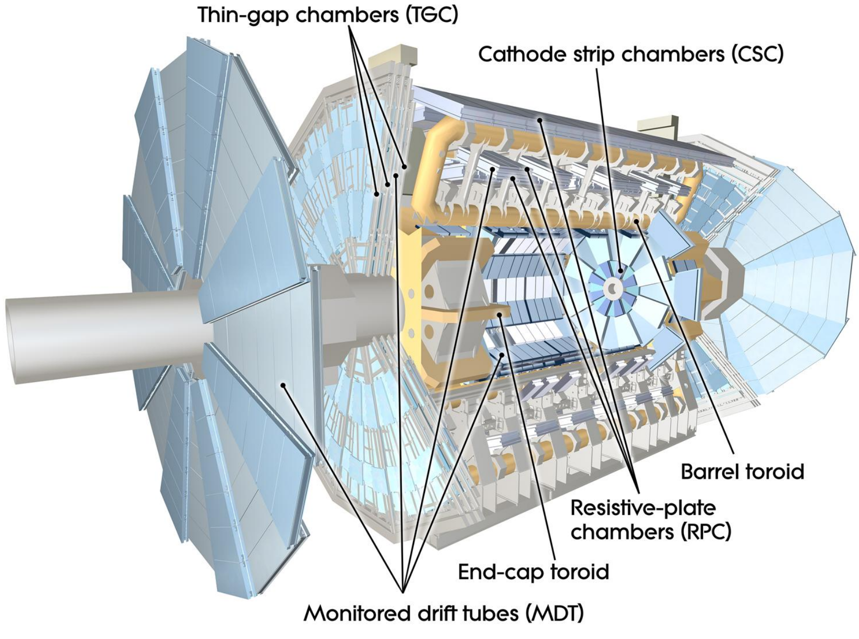


Fig. 2.8 · Cut-away view of the ATLAS Muon Spectrometer [100].

Tracking chambers

The MDT chambers perform precision momentum measurements, covering the range $|\eta| < 2.7$ (except in the innermost layer where the coverage is limited to $|\eta| < 2.0$). They consist of up to eight layers of aluminium drift tubes with a diameter of 30 mm, filled with a gas mixture of 93% Ar and 7% CO_2. Crossing the tube, a muon ionizes few atoms of the gas and the freed electrons are collected at the anode, the central tungsten-rhenium wire with a diameter of 50 μm. The MDTs provide precision measurements in the bending plane, with an average resolution of 80 μm per tube and 35 μm per chamber.

The CSC chambers are used in the innermost layer of the end-caps, between $2 < |\eta| < 2.7$, for their high rate capability. The CSCs are multiwire proportional chambers with cathodes segmented into strips in orthogonal directions. They are filled with a mixture of gas of 80% Ar and 20% CO_2. For each chamber, the resolution is 40 μm in the bending plane and 5 mm in the transverse plane.

Trigger chambers

The trigger chambers are designed to achieve three goals: identify the bunch-crossing, trigger on muons and measure the second coordinate of the muon track, in the direction orthogonal to that determined by the MDT chambers.

The RPC chambers cover the barrel region ($|\eta| < 1.05$) and consist of two parallel resistive plates with a 2 mm gap between them, which is filled with a gas mixture. Each of the three RPC layers (referred to as the three trigger stations) is made of two independent detector sub-layers, which provide measurements of the η coordinate and the ϕ coordinate, respectively. The RPCs have a time resolution of 1.5 ns and a spatial resolution of about 10 mm.

The TGC chambers are multiwire proportional chambers used in the end-cap region ($1.05 < |\eta| < 2.4$). The design of these chambers is characterized by a distance between wire and cathode of only 1.4 mm and a wire spacing of 1.8 mm, which ensures the drift time is shorter than the design bunch spacing of 25 ns. For the TGCs, the time resolution is 4 ns and the spatial resolution is 2-7 mm.

2.2.6 Trigger and data acquisition system

The LHC bunches are designed to collide every 25 ns (during Run-1 the bunch spacing was 50 ns) corresponding to an event rate of 40 MHz, producing a volume of data of approximately 60 TB per second. Such a value is far beyond the maximum rate that can be read out and stored. To reduce the amount of data and select only the fraction of interesting events for physics analyses, ATLAS employs the trigger system. It consists of three levels of event selection: Level-1 (L1), Level-2 (L2) and Event Filter (EF), with the last two forming the High-Level Trigger (HLT). The data acquisition system (DAQ) is used to buffer and transport event data to the permanent storage. A schematic view of the trigger and DAQ system is shown in Figure 2.9.

The L1 trigger is a hardware-based system and performs the initial event selection relying on reduced-granularity information from the calorimeters and the muon detectors. The L1 calorimeter trigger (L1Calo) identifies high energetic photons, electrons, jets and hadronically decaying taus, as well as events with large E_T^{miss}. The L1 muon trigger searches for high p_T muons using the RPC and TGC trigger chambers. The L1 trigger reduces the data rate to about 75 kHz, with a latency time lower than 2.5 μs. The L1 trigger defines the Regions of Interest (RoIs), the

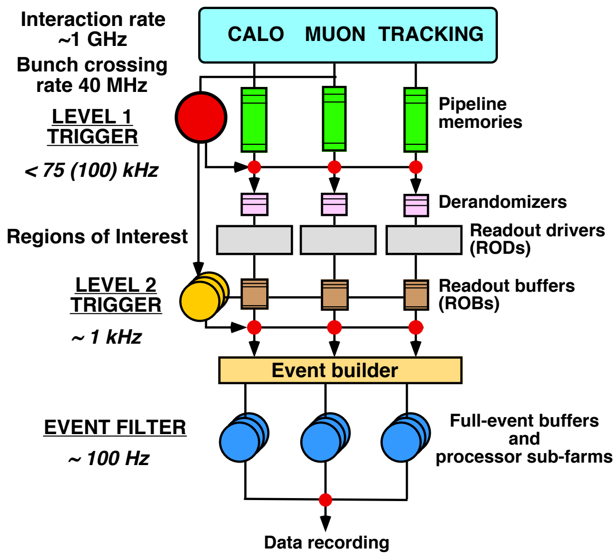


Fig. 2.9 · Diagram of the ATLAS trigger and data acquisition system [101].

coordinates in η and ϕ of those regions of the detector where possible trigger objects are identified in the event. The information from calorimeter and muon triggers is combined by the central trigger processor (CTP), which makes the final L1 accept decision based on lists of selection criteria, implemented as a trigger menu. The DAQ buffers the event data from the detector-specific readout electronics at the L1 trigger rate.

The L2 trigger is a software-based system, seeded by the RoIs provided by the L1 trigger. Within these regions, the L2 trigger performs a fast event reconstruction using the full detector granularity. It is designed to reduce the event rate by an additional factor 30, with an event processing time of about 40 ms.

The EF represents the final stage of the trigger selection. The events passing the L2 requirements are processed using the full information of the ATLAS detector and employing offline analysis procedures. The EF reduces the event rate to around 400 Hz, with an event processing time of approximately 4 s. The events are also separated in distinct physics *streams*, according to their prevalent trigger objects. The selected events are moved to permanent storage to be used in the offline analyses.

2.2.7 ATLAS detector performance

The LFV $\tau \to \mu\gamma\gamma$ analysis presented in this thesis is performed using the 2012 p-p collision data at $\sqrt{s} = 8$ TeV recorded by the ATLAS detector. During that year, as well as throughout Run-1, the detector achieved a successful operational performance. More than 93% of the luminosity delivered by the LHC ($\mathcal{L} = 22.8$ fb^{-1}) was recorded by the ATLAS detector ($\mathcal{L} = 21.3$ fb^{-1}). Each sub-detector sys-

Inner Detector			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.9	99.1	99.8	99.1	99.6	99.6	99.8	100.	99.6	99.8	99.5
All good for physics 95.5										

Table 2.2 · Luminosity weighted relative detector uptime and good quality data delivery (in %) during 2012 stable beams in p-p collisions at $\sqrt{s} = 8$ TeV, corresponding to 21.3 fb^{-1} of recorded data. Values taken from [102].

tem recorded data over 99% of the time while the beams were in stable conditions, totalling a fraction of good quality data above 95%. A summary of the performance of the various sub-detectors during 2012 data-taking is given in Table 2.2.

2.3 Triggers for LFV $\tau \to \mu\gamma\gamma$ analysis

To conclude this chapter, the collection of triggers used in the LFV $\tau \to \mu\gamma\gamma$ analysis is introduced in this section. A key role is played by a set of triggers originally designed for the $\tau \to \mu\gamma$ decay [103]. For this signal, the low energy of the final state objects does not allow an efficient triggering on the single particles. Muon and photon share the momentum of the decaying tau lepton; therefore, a muon with higher p_T is commonly accompanied by a photon with lower p_T and vice versa.

To trigger on these pairs, a set of *muon+photon* triggers characterized by different p_T thresholds is employed. Since the reconstructed invariant mass of the final state objects is usually consistent with the nominal value of the tau mass, a requirement is also applied on the $\mu\gamma$ invariant mass, $m(\mu\gamma)$. For this reason, the denomination *TauMass* is used to identify these triggers. In addition, a requirement is placed on the angular distance between the two particles, which are typically produced with small separation. At the final stage of the trigger selection (EF), the trigger level μ-γ candidates have to satisfy the conditions

$$1\text{ GeV} < m(\mu\gamma) < 2.5\text{ GeV} \quad \text{and} \quad \Delta R(\mu, \gamma) < 0.6, \tag{2.5}$$

which are also useful to reduce the EF rates. The *TauMass* trigger set consists of the following five triggers:

- EF_mu10i_g10_medium_TauMass
- EF_mu10i_loose_g12Tvh_medium_TauMass
- EF_mu20i_tight_g5_loose_TauMass
- EF_mu20i_tight_g5_medium_TauMass
- EF_mu4Ti_g20Tvh_medium_TauMass

In the list above, *mu* and *g* refer to muon and photon, respectively, while the numbers represent the p_T thresholds. Identification criteria are denoted as *loose*, *medium* and *tight*; i indicates isolation, *Tvh* refers to high L1 p_T threshold and veto on the hadronic energy deposition of the photon. The five triggers cover three steps in the plane spanned by muon p_T and photon p_T (muon $p_T \gg$ photon p_T, muon $p_T \approx$ photon p_T and muon $p_T \ll$ photon p_T). For each trigger, the events pass the selection if the energies of a photon and a muon are higher than the given thresholds and the requirements on the μ-γ pair, invariant mass and ΔR, are satisfied.

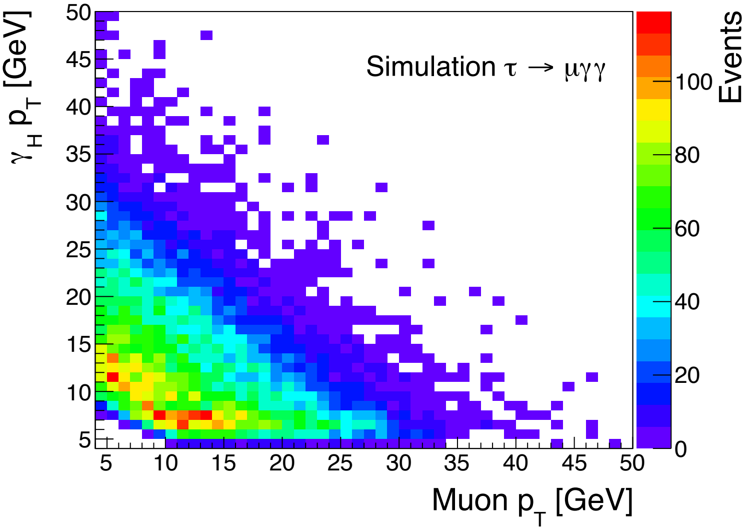
The *TauMass* triggers are also crucial for the LFV $\tau \rightarrow \mu\gamma\gamma$ analysis presented in the thesis. For this three-body decay, the events are selected if the muon and at least one of the two photons in the event satisfy the conditions previously mentioned. In the trigger selection applied to this analysis (see Section 4.3), the set of five *TauMass* triggers is used in combination with two *single muon* triggers and a *tau+muon* trigger to increase the signal efficiency and target other regions of phase-space.

The *single muon* triggers employed are EF_mu24i_tight and EF_mu36_tight. The first trigger, which is used in many physics analyses, is designed to collect muons with $p_T > 24$ GeV. The EF output rate is controlled by applying isolation criteria. In 2012, the ratio between the p_T sum of the tracks within a cone $\Delta R < 0.2$ around the muon and the p_T of the muon itself was required to be less than 0.12 [104]. The L1 p_T threshold is set to 15 GeV. The second trigger, EF_mu36_tight, is designed to select high energetic muons ($p_T > 36$ GeV). The L1 muon p_T threshold is 15 GeV and no isolation requirements are applied.

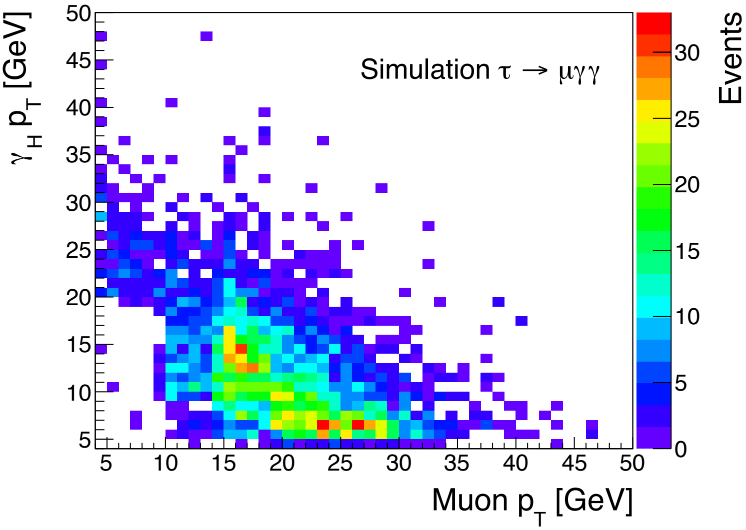
The collection of triggers used in the $\tau \rightarrow \mu\gamma\gamma$ analysis is completed by a *tau+muon* trigger, namely EF_tau20_medium1_mu15. Hadronic tau triggers are used in several analyses and generally require high p_T thresholds in order to maintain under control the trigger rates. For the trigger employed here, the terms *tau20* and *medium1* imply a 20 GeV requirement on the transverse energy, medium selections on shower shape variables and a number of tracks restricted to three or less for an EF tau candidate. For the muon component of the trigger, the term *mu15* indicates a 15 GeV requirement on the p_T of the muon candidate at the EF.

The trigger selection used in the $\tau \rightarrow \mu\gamma\gamma$ search is obtained by the logical *or* combination of all the eight triggers mentioned in this section. Figure 2.10 illustrates the effect of the selection on the Monte Carlo simulation of the $\tau \rightarrow \mu\gamma\gamma$ sample, showing the p_T of the photon (the more energetic of the pair) versus the p_T of the muon before and after the trigger requirements.

The trigger selection represents one of the steps necessary to perform the LFV $\tau \rightarrow \mu\gamma\gamma$ analysis. The results of the selection, which include the trigger efficiency obtained for this signal, will be further discussed in Section 4.3.



(a)



(b)

Fig. 2.10 · Transverse momentum of the higher energetic photon of the pair, defined as γ_H throughout the thesis, versus transverse momentum of the muon for Monte Carlo $\tau \to \mu\gamma\gamma$ sample shown before (a) and after (b) the trigger selection described in the text.

chapter three

Object reconstruction

For each proton-proton collision, the raw information from the ATLAS sub-detectors needs to be combined and processed in order to *reconstruct* the events and *identify* the physics objects involved. The methods used to define these objects and their performances in Run-1 are presented in this chapter. Muons and photons, contained in the final state of the LFV $\tau \to \mu\gamma\gamma$ channel, are described first. An overview of the reconstruction of electrons, jets, tau leptons and missing transverse momentum is also given. Moreover, the analysis of the LFV $\tau \to \mu\gamma\gamma$ decay requires a dedicated object identification, which is illustrated in Section 3.8. The chapter starts with a brief description of the algorithms used for the reconstruction of tracks and vertices.

3.1 Tracks and vertices

The goal of a tracking system is to reconstruct the trajectory and measure the momentum of charged particles through the curvature of their path in the magnetic field. In ATLAS, tracks are reconstructed in the Inner Detector (ID) using a sequence of algorithms. The *inside-out* algorithm is designed for the track reconstruction of primary charged particles, defined as the particles with a mean lifetime greater than 3×10^{-11} s directly produced in a p-p interaction or originated from decays of particles with a lifetime shorter than the value aforementioned.

The inside-out algorithm starts with the identification of seeds formed by triplets of *hits* in the silicon detectors. Moving away from the interaction point, other hits are added using a combinatorial Kalman filter [105]. At this stage, ambiguities in the hit assignment to the track candidates are solved by a scoring algorithm [106], which attributes a different weight to the tracks relying on several parameters, such as the number of hits and *holes*[1] associated with them. Finally, tracks are extended into the TRT. Tracks reconstructed by this algorithm are required to have a $p_T > 400$ MeV.

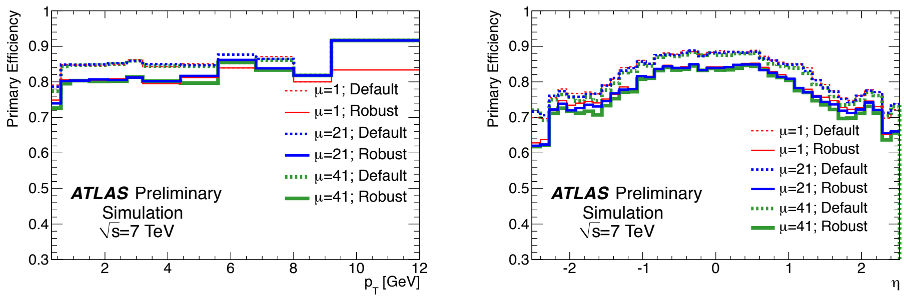
The *outside-in* algorithm is used for the track reconstruction of secondary particles, produced in interactions of primary particles, and starts from TRT segments which are extended towards the interaction point by adding silicon hits in a procedure also known as *back-tracking*.

[1] Missing measurement, although one is expected, given a track trajectory.

For each reconstructed track its *impact parameters* are defined as follows: d_0 represents the distance from the point of closest approach to the z-axis in the transverse plane; z_0 is the z-coordinate at the point of closest approach to the z-axis.

The probability to reconstruct fake tracks, i.e. tracks that cannot be associated with any particle, increases in a high pile-up environment. To reduce this effect, a set of *robust* quality requirements on the reconstructed tracks has been defined. This selection requires tracks with at least 9 hits in the silicon detectors and no holes in the pixel detector. The *default* selection used in 2011 and in the LFV $\tau \to \mu\gamma\gamma$ analysis requires at least 7 silicon hits and at most 2 holes in the pixel detector [107].

The track reconstruction efficiency, determined using Monte Carlo (MC) simulation, is defined as the fraction of particles with $p_T > 400$ MeV and $|\eta| < 2.5$ matched to a reconstructed track. Figure 3.1 shows the track reconstruction efficiency of primary particles comparing the default and robust selections for different pile-up scenarios, as a function of the track p_T and η. The efficiency is insensitive to the increasing pile-up. However, the robust selection lowers the efficiency by 5% due to the tighter quality requirements.



(a) (b)

Fig. 3.1 · Primary track reconstruction efficiency in minimum bias MC samples containing exactly one and on average 21 or 41 interactions as a function of p_T (left) and η (right) [107].

The origin of the reconstructed tracks also needs to be determined. The p-p interaction points, called primary vertices, are reconstructed using an iterative vertex finding algorithm [108], where vertex seeds are found by searching for the global maximum in the z_0 distribution of the tracks. Then, a χ^2-based vertex fitting algorithm [109], which takes as input the seed and nearby tracks, is used to determine the vertex position. Tracks incompatible with the vertex by more than 7σ are used to seed a new vertex and the iterative procedure stops when no additional vertices can be found. The vertex with the highest $\sum p_T^2$ of tracks is labelled as the primary vertex of the hard scattering.

The increasing pile-up affects the vertex reconstruction. Figure 3.2 shows the relation between the average number of reconstructed primary vertices per event and the average number of interactions per bunch crossing for data collected in 2012. At high pile-up, the deviation from the linear dependence is due to effects of *merging* and *shadowing* of the primary vertices [110], which decrease the reconstruction efficiency.

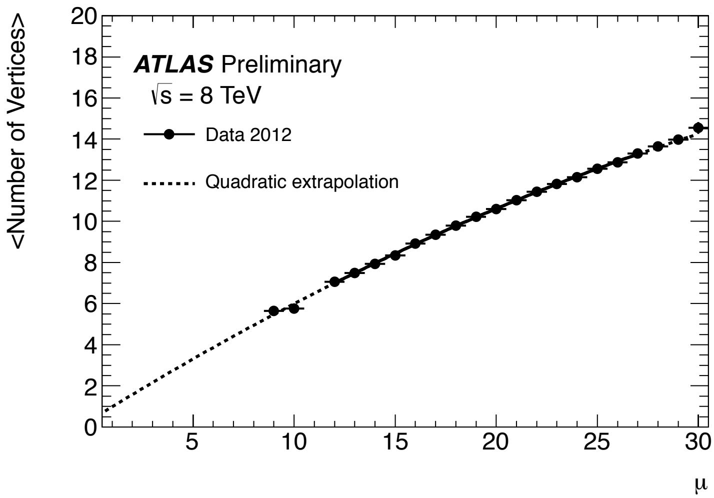


Fig. 3.2 · Average number of reconstructed primary vertices per event as a function of the average number of interactions per bunch crossing in 2012 data. A second order polynomial fit is performed in the upper range of μ [111].

3.2 Muon reconstruction

For the reconstruction and identification of the muons, ATLAS uses information from the ID, the Muon Spectrometer (MS) and, to a lesser degree, from the calorimeter. The combination of measurements made by the ID and MS allows to reject background muons, coming from secondary interactions and π/K decays in flight.

Several reconstruction criteria, based on the information provided by the different sub-detectors, are available for the muon identification. This leads to the definition of several types of muons [112]:

- *Combined* (CB) muons: the track reconstruction is performed independently in the ID and MS. A combined track is formed by the combination of the two measurements. These muons have the highest purity and are the most widely used in ATLAS. The LFV $\tau \to \mu\gamma\gamma$ analysis presented in this thesis uses combined muons.

- *Stand-Alone* (SA) muons: the muon trajectory is reconstructed only in the MS. The track parameters at the interaction point are determined extrapolating the track back to the beam line, including corrections due to energy loss in the calorimeters. SA muons are mainly used in the range $2.5 < |\eta| < 2.7$, not covered by the ID.

- *Segment-tagged* (ST) muons: a track in the ID is identified as a muon if it can be associated with at least one track segment in the precision muon chambers, once extrapolated to the MS. These muons are used to increase the detector acceptance when the muon crosses only one layer of MS chambers (due to its low p_T or regions with a reduced MS acceptance).

- *Calorimeter-tagged* (CaloTag) muons: a track in the ID is classified as a muon if it can be associated with an energy deposit in the calorimeter compatible with a minimum ionizing particle. This type is useful to recover the acceptance in the region $\eta \approx 0$.

The reconstruction of tracks in the spectrometer is affected by acceptance losses in two regions: at $\eta \approx 0$, where the MS is only partially equipped with muon chambers, and in the range $1.1 < \eta < 1.3$, where there are regions with only one layer of MS chambers installed.[1]

The reconstruction of the muons is performed using different algorithms, called *chains*. In Chain 1 [113], a statistical combination of the ID and MS track parameters is performed using the corresponding covariance matrices. Chain 2 [114] performs a global refit of the muon track using the hits from the ID and MS. A new algorithm, Chain 3, has been developed to include the best features of the first two chains and used in parallel to them in 2012. Muons used in the LFV $\tau \to \mu\gamma\gamma$ analysis are reconstructed by the Chain 3 algorithm. The three chains have very similar performances [112].

3.2.1 Efficiency measurements

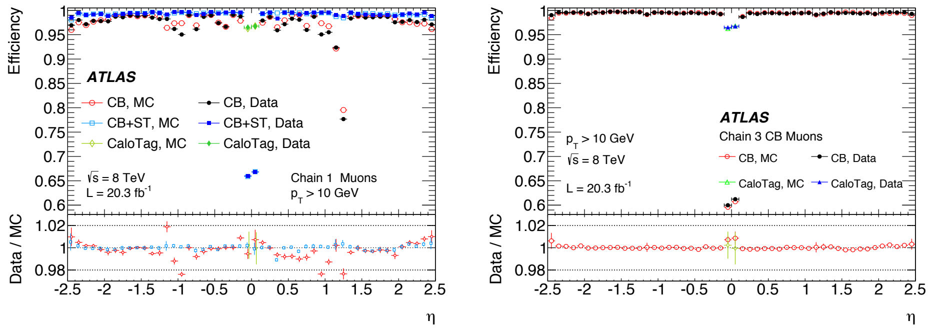
A tag-and-probe method is employed to measure the muon reconstruction efficiency in the region $|\eta| < 2.5$, using $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ (for low p_T muons) samples. One of the two muons, called *tag*, is required to be a CB muon. The other one, called *probe*, reconstructed in either the ID or MS, is used to test the reconstruction efficiency of the other system (MS or ID).

For $Z \to \mu\mu$ decays, events are required to have two opposite-charge isolated muons, with an invariant mass close to the Z mass. A small fraction of the tag-and-probe pairs can originate from other sources, such as $Z \to \tau\tau$ and $t\bar{t}$ decays. To obtain a correct efficiency measurement, these background contributions have to be estimated and subtracted [112].

The efficiency as a function of η for different muon types, reconstructed with the Chain 1 algorithm, is shown in Figure 3.3a. It is about 99% over most of the

[1] The installation of these chambers has been completed during the 2013-2014 LHC shutdown.

regions of the detector. ST and CaloTag muons are used to recover efficiency losses in the region $1.1 < |\eta| < 1.3$ and at $\eta \approx 0$, respectively. Figure 3.3b shows a similar efficiency for muons reconstructed with the Chain 3 algorithm, which are used in this thesis.



(a) (b)

Fig. 3.3 · Muon reconstruction efficiency as a function of η, measured in $Z \to \mu\mu$ events for different muon types, using the Chain 1 (left) and the Chain 3 (right) algorithms [112].

To measure the reconstruction efficiency as a function of p_T, a sample of $J/\psi \to \mu\mu$ events is used for low p_T muons, in addition to the $Z \to \mu\mu$ sample. The performance for CB+ST muons is shown in Figure 3.4. The efficiency is almost independent of the p_T for values above 20 GeV.

The efficiencies measured in the MC simulation and in data are in good agreement, typically within 1%. This small discrepancy is corrected by the scale factor applied to the MC simulation, defined as

$$\mathrm{SF} = \frac{\epsilon^{\mathrm{Data}}}{\epsilon^{\mathrm{MC}}}. \tag{3.1}$$

3.3 Photon reconstruction and identification

The reconstruction of the photons starts from the energy deposits in the EM calorimeter, called clusters. The $\eta - \phi$ space of the EM calorimeter is divided into a grid of towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The energy of the cells inside these components is summed to get the tower energy. The reconstruction is seeded by clusters with transverse energy above 2.5 GeV formed by a sliding-window algorithm [115, 116] with a window size of 3×5 towers.

Then, clusters are matched with tracks reconstructed in the ID, relying on the distance between the cluster barycenter and the extrapolated track intersection point with the second sampling of the calorimeter. If the match is not successful, clusters are classified as *unconverted* photon candidates. For clusters with matched tracks,

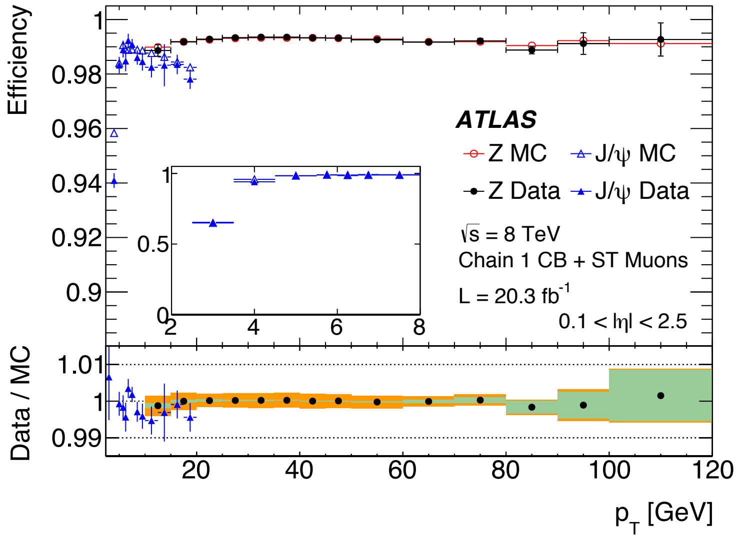


Fig. 3.4 · Reconstruction efficiency for CB+ST muons as a function of the transverse momentum obtained using $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ samples [112].

the candidates can be either electrons[1] (see Section 3.4) or *converted* photons, which convert into a e^+e^- pair when they cross the material in front of the EM calorimeter.

The reconstruction of converted photons includes the reconstruction by the ID of the conversion vertices, within a distance of 800 mm from the primary interaction point. Depending on the number of electron tracks assigned to them, they can be classified as single-track[2] or double-track conversion vertex candidates.

Clusters matched to a conversion vertex are considered converted photon candidates. When these are also reconstructed as electrons, there are two possible outcomes: if the electron track is identified as a track originating from the conversion vertex matched to the same cluster, the candidate is retained as a converted photon, otherwise it is kept as an electron. Further details about the reconstruction procedure can be found in [117].

3.3.1 Photon identification

Prompt photons need to be separated from background photons originating from hadronic jets. In ATLAS, the photon identification is based on a set of rectangular cuts applied on calorimetric variables, which define the shapes of the electromagnetic showers and the leakage energy fraction in the hadronic calorimeter, giving a good separation between prompt photons and jets. A detailed description of these variables is given in [118].

---

1 Positrons are also included in this definition.

2 In this case one of the two electron tracks fails to be reconstructed either because it is very soft or when the two tracks cannot be separated because they are too close to each other.

Two reference sets of cuts, *loose* and *tight*, are defined. The *loose* selection is based on shower shapes in the middle layer of the EM calorimeter and on the hadronic leakage. It is used by the photon triggers and harmonized with the corresponding one for the electrons. The *tight* selection adds information from the strip layer of the calorimeter, which provides a good rejection of photons coming from π^0 decays, exploiting its fine granularity. The *tight* criteria are separately optimized for converted and unconverted photons to obtain an identification efficiency of about 85% for prompt photons with $E_T > 40$ GeV [119]. These identification criteria have been reoptimized to cope with the increasing pile-up in the 2012 data compared to the 2011 data.

An isolation requirement is applied on the photon candidates. The calorimeter isolation transverse energy E_T^{iso} is computed in a cone of size $\Delta R = 0.4$ around the photon candidate and has to be smaller than 4 GeV, excluding the contribution of the photon itself.

The shower shapes of the photons described by the ATLAS MC simulation show a non negligible difference with the ones observed in data, especially in the average values of the distributions of the discriminating variables. To correct this behaviour, the discrepancies between data and simulation are parameterized as simple correction factors and applied to the MC simulation to match the data distributions, using a sample of $Z \to ll\gamma$ events. The correction factors are calculated as the values which minimize the χ^2 between the data and the MC distributions of the photon candidates. This correction has also an impact on the photon efficiency and the MC simulation used in the analysis presented in this thesis.

3.3.2 Photon identification efficiency measurements

Data-driven techniques can be used to measure the photon identification efficiency [118], defined as the ratio of the number of isolated photons which pass the tight selection to the total number of isolated photons.

In a low photon E_T region, between 10 and 80 GeV, $Z \to ll\gamma$ decays provide a pure photon sample to measure the efficiency. These events are selected requiring two leptons, electrons or muons, with opposite charge and $p_T > 15$ GeV. Photon candidates are isolated and with $E_T > 10$ GeV. To avoid a bias in the standard photon identification efficiency due to the proximity of a lepton, a cut of $\Delta R > 0.2\,(0.4)$ is applied between the photon and each of the muons (electrons).

Z+jets events, where a jet can fake a photon, represent an important source of background for a sample of $Z \to ll\gamma$ events with initial state radiation (ISR) photons. To reduce this background, and select final state radiation (FSR) photons, the three-body invariant mass, $m_{ll\gamma}$, and the di-lepton invariant mass, m_{ll}, are required to be between 80 and 96 GeV and between 40 and 83 GeV, respectively.

Two other data-driven methods measure the efficiency at higher energies: the first one uses a pure sample of electrons from $Z \to ee$ decays [118]; the second, known as *matrix method* [118], is based on the track isolation as a discriminating variable between prompt and background photon candidates. The explanation of these methods is not given here, since such energetic photons are not used in this thesis.

The photon identification efficiency can also be predicted in the MC simulation, before (nominal MC) and after (corrected MC) the correction on the photon shower shapes, previously introduced in Section 3.3.1, and can be compared with the data-driven measurements. Figure 3.5 and Figure 3.6 show this comparison in four different regions of pseudorapidity for unconverted and converted photons, respectively.

The EM showers tend to be narrower in the simulation than in data. For this reason, the nominal MC simulation overestimates the efficiency up to $10 - 15\%$. Using the corrected MC simulation, the disagreement with the data is reduced to a few percent. For $E_T > 60$ GeV, there is no significant difference between the data-driven measurements and both simulation predictions.

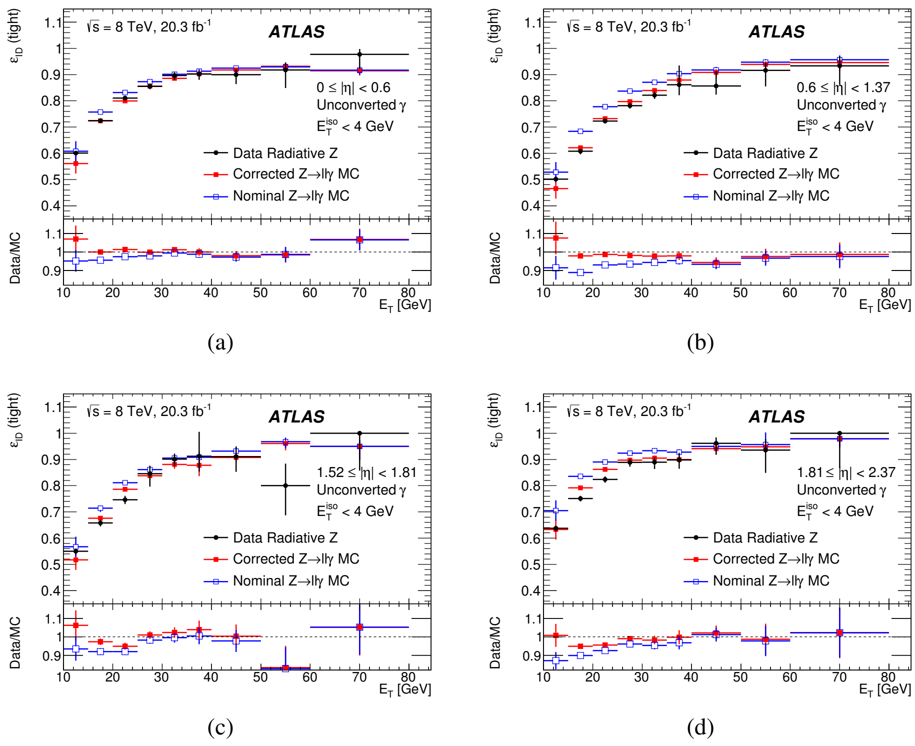


Fig. 3.5 · Comparison of the data-driven identification efficiency measurements of unconverted photons to the nominal and corrected MC predictions as a function of E_T for four different η regions [118].

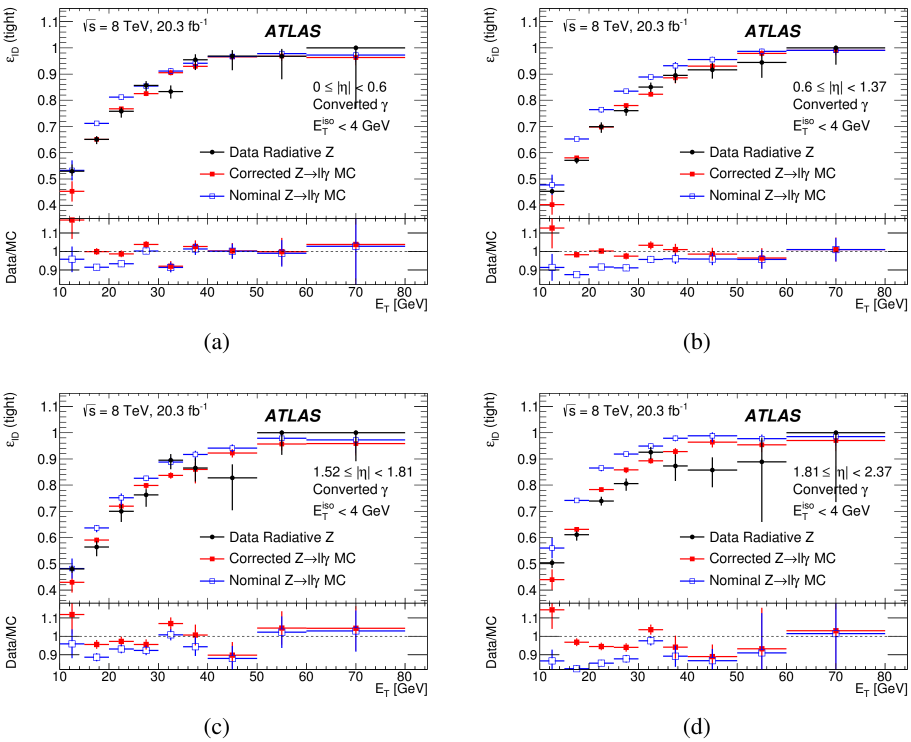


Fig. 3.6 · Comparison of the data-driven identification efficiency measurements of converted photons to the nominal and corrected MC predictions as a function of E_T for four different η regions [118].

3.4 Electron reconstruction and identification

The first step of the electron reconstruction is based on the reconstruction of the EM clusters and follows the same procedure already described for the photons in Section 3.3.

Afterwards, clusters are associated with tracks reconstructed in the ID and extrapolated to the calorimeter, within a specified distance in η and ϕ. The parameters of these electron-track candidates are accurately re-estimated using an optimized electron track fitter, the Gaussian Sum Filter (GSF) [120], which takes into account non-linear energy losses due to bremsstrahlung effects.

An electron candidate is reconstructed if at least one track is successfully matched to the EM cluster. In case of more tracks matched to the same cluster, tracks with silicon hits are preferred, and the one with the smallest ΔR distance is chosen. Electron clusters are then rebuilt, optimizing their size to take into account the different overall energy distributions in the barrel and endcap regions of the calorimeter [121].

Among the objects reconstructed in this way, identification criteria are used to distinguish real prompt electrons from background sources, such as hadronic jets and electrons from photon conversions. The identification is based on a set of discriminating variables combined into different selection criteria. These variables describe the shape of the electromagnetic showers, the hadronic leakage in the calorimeter, the quality of the tracks and track-cluster matching. Both sequential cuts and a multivariate likelihood-based (LH) method are employed.

Selection criteria with an increasing background rejection are designed for cut-based (*loose*, *medium* and *tight*) and LH-based (*looseLH*, *mediumLH* and *very tightLH*) identification. The tighter selections represent a subset of the looser ones, obtained adding new discriminating variables and applying tighter requirements on the variables already used. The *multilepton*[1] selection has been added in 2012 to the cut-based identification [122]. The LH-based identification uses signal and background probability density functions (PDFs) of the discriminating variables. Its three working points are designed to have a better background rejection than the cut-based selections.

The reconstruction and identification efficiency is measured using the tag-and-probe method on clean samples of electrons, such as $Z \to ee$ and $J/\psi \to ee$ decays. One of the two electrons (*tag*) is selected applying strict selection criteria, while the other (*probe*) is used to measure the efficiency. Further details about this tag-and-probe procedure can be found in [122].

Efficiencies measured in MC samples are slightly different from the ones obtained on data. This effect is corrected applying a scale factor to the MC simulation, defined as the ratio of the efficiency measured in data to the one in simulation. The combined reconstruction and identification efficiency for cut-based and LH-based selections is shown in Figure 3.7, as a function of E_T and η.

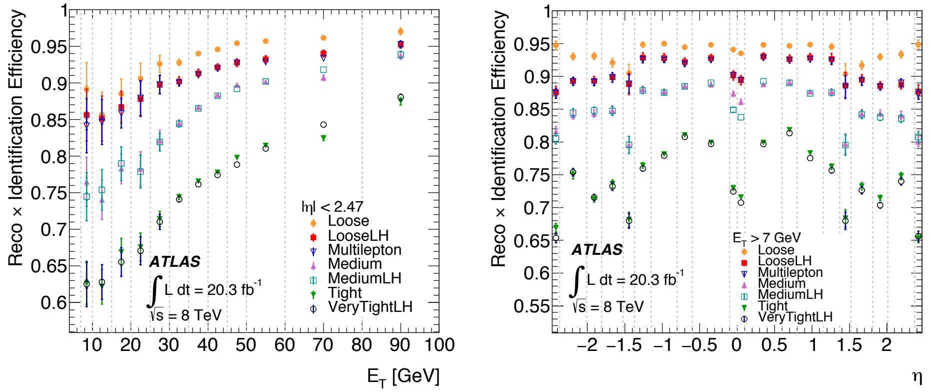
3.5 Jets

Due to colour confinement, quarks and gluons formed in p-p collisions *hadronize*, producing a narrow cone of colourless hadrons, called *jet*, around the direction of the original object.

The reconstruction of jets is based on calorimeter information. Energy deposits in the calorimeter system are grouped together by a topological clustering algorithm [123]. The formation of topological clusters starts from a *seed* cell, whose signal-to-noise ratio is over a certain threshold. Neighbouring cells, with a significant signal compared to noise, are added iteratively to form the topo-clusters.

As a next step, jets can be identified using jet finding algorithms. These algorithms are required to be both collinear and infrared safe, so that the emission of collinear or soft gluons does not affect the definition of a jet. The algorithm used in ATLAS is the anti-k_t algorithm [124], which belongs to the collection of sequential recombination algorithms.

[1] This selection is optimized for low energy electrons in the $H \to 4l$ analysis.



(a) (b)

Fig. 3.7 · Combined electron reconstruction and identification efficiency for cut-based and LH-based selections as a function of E_T (left) and η (right) in 2012. Data efficiency is derived from the measured data-to-MC efficiency ratios applied to MC $Z \rightarrow ee$ events [122].

In these algorithms the distance between two objects, d_{ij}, and the distance between an object and the beam line, d_{iB}, are defined as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^{2p}, \tag{3.2}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i, ϕ_i represent the transverse momentum, the rapidity and the azimuth angle of the object i, respectively. The parameters p and R depend on the specific algorithm. For the anti-k_t algorithm, the parameter p is set to -1 to ensure that soft particles tend to cluster with the hard ones before they cluster among themselves. The parameter R defines the radius of the jet and is equal to 0.4 for the jets used in this thesis. The two distances d_{ij} and d_{iB} defined above are compared. If $d_{ij} < d_{iB}$, the objects i and j are recombined into a new object. On the other hand, if $d_{iB} < d_{ij}$, the object i is defined as a jet and removed from the list. Then, the procedure is repeated to search for the next jet until no objects are left.

The energy of the jets formed from topological clusters needs to be calibrated. Topo-clusters are initially reconstructed at the EM scale, which measures correctly the energy deposited in the calorimeter by electrons and photons. In a second step, topo-clusters are calibrated using the *local cluster weighting* (LCW) method [125], in order to obtain a correct response of the calorimeter to hadrons. The LCW method classifies the topo-clusters as either electromagnetic or hadronic, depending on the shower properties. According to this classification, energy corrections are determined from single pion MC simulations. Dedicated corrections take into account effects of calorimeter non-compensation and energy lost in non-instrumented regions [126].

To relate the reconstructed energy of a jet to that of the corresponding jet at truth-level, *jet energy scale* (JES) corrections are applied. The LCW + JES calibration, which is also employed for the jets used in this thesis, consists of several steps. First, the direction of the jet is corrected to point to the primary vertex instead of the nominal center of the ATLAS detector. Next, another correction is applied to remove the effects of pile-up [127]. The calibration of the energy and direction of a reconstructed jet is then derived comparing its kinematic observables with the ones of the truth jet in MC simulation. Finally, a residual correction obtained from in-situ techniques, which compare the p_T of a jet to the one of a reference object, is applied to jets reconstructed in data.

The total jet energy scale uncertainty is derived adding in quadrature sources of systematic uncertainties arising from in-situ calibration methods, pile-up corrections and other contributions [128]. Figure 3.8 shows the total uncertainty for jets calibrated using the LCW + JES scheme as a function of the jet p_T for central jets and as a function of η.

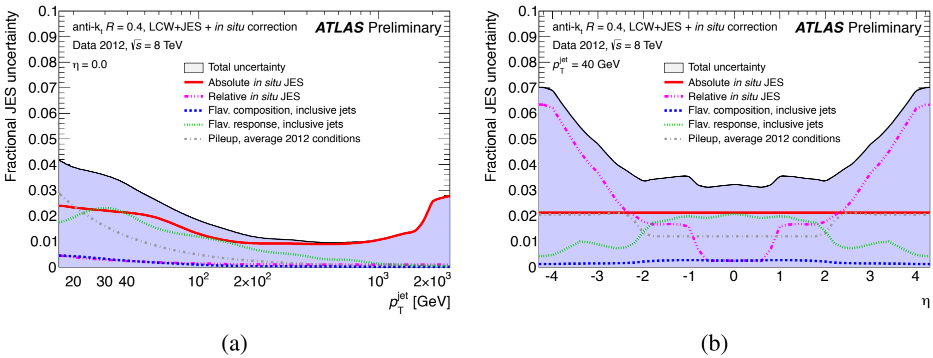


Fig. 3.8 · Fractional jet energy scale systematic uncertainty in 2012, as a function of the jet p_T (left) and η (right) [129].

3.6 Tau reconstruction and identification

Tau leptons decay into either lighter leptons or hadrons in 35% and 65% of the cases respectively [130]. Typically, they decay before reaching the innermost sensitive elements of the ATLAS detector, due to their short lifetime. Therefore, decay products have to be used for the identification. However, the leptonically decaying tau, τ_{lep}, produces (according to the Standard Model) an electron or a muon, in association with neutrinos, which are not easily distinguishable from prompt electrons and muons. The hadronically decaying tau, τ_{had}, produces one or three charged pions, in association with a tau neutrino, in 72% and 22% of the cases, respectively.[1] In 78% of all hadronic decays, associated neutral pions are also produced. Since the neutrino is undetected, only the visible part of the hadronic tau decay is used for reconstruction and identification.

The τ_{had} reconstruction starts from energy deposits in the calorimeter which have been reconstructed as a jet using the anti-k_t algorithm ($R = 0.4$). Jets are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. To reduce the effects of pile-up, a tau vertex association algorithm [131] is used to identify the tau production vertex among the primary vertex candidates in the event. The selected vertex is the one with the highest fraction of the p_T sum of tau candidate tracks, which are in a cone $\Delta R < 0.2$ around the jet seed direction.

Tracks are associated with the τ_{had} if they are in the *core region* $\Delta R < 0.2$ around its direction and satisfy a set of quality criteria [132]. The number of these tracks classifies the τ_{had} candidate as 1-prong or 3-prong. Tracks in the *isolation region* ($0.2 < \Delta R < 0.4$) are used for the calculation of identification variables. Furthermore, a π^0 reconstruction algorithm [133] has been developed to obtain the number of reconstructed neutral pions in the core region.

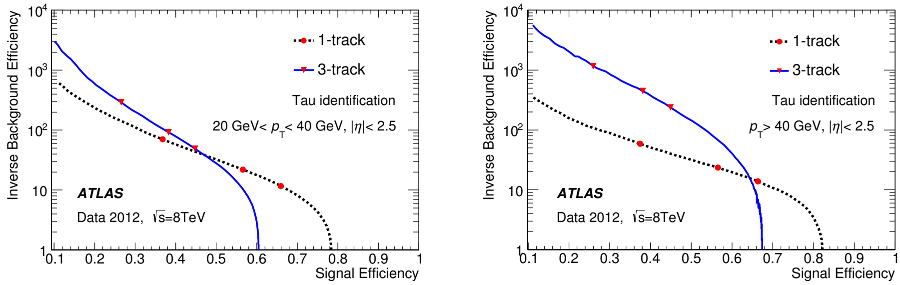
The reconstruction of τ_{had} candidates provides low rejection against quark and gluon jets. A set of dedicated variables is used as input for two Boosted Decision Trees (BDT) algorithms [134], trained separately for 1-prong and 3-prong cases, in order to increase the discrimination power against jets. MC simulation and multijet data samples are used for signal and background, respectively. The variables are based on tracks, shower shape in the calorimeter and π^0 content of the hadronic tau decays. A complete description of these variables is given in [135].

Three working points, corresponding to different identification efficiency values, are defined and labelled as *loose*, *medium* and *tight*. Figure 3.9 shows the inverse background efficiency as a function of the signal efficiency for different p_T ranges.

Measurements of the τ_{had} identification efficiency are performed using a tag-and-probe method on a sample of $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ events, selected requiring the presence of an electron or a muon (*tag*), and then searching for the τ_{had} candidate (*probe*). The small difference observed between the efficiencies measured in data and in simulation is corrected applying scale factors, defined as the ratio of the identification efficiency in data to the one in simulation [135].

The signature of 1-prong τ_{had} can be mimicked by electrons. Therefore, another BDT is used to reject this background contribution, exploiting transition radiation and calorimeter shower shape information. Three working points, labelled as *loose*, *medium* and *tight*, correspond to 95%, 85% and 75% signal efficiency, respectively. Furthermore, a simple cut-based selection is applied to reject the small fraction of muons which deposit enough energy in the calorimeter to be misidentified as τ_{had} candidates. The resulting efficiency is over 96%, with a background rejection of about 40%.

[1] Kaons are produced in the majority of the remaining hadronic decays.



(a) (b)

Fig. 3.9 · Inverse background efficiency versus signal efficiency for the tau identification, for low p_T (left) and high p_T (right) τ_{had} range. The red markers indicate the three working points mentioned in the text [135].

Finally, the τ_{had} energy scale needs to be calibrated. The description of this procedure and the associated systematic uncertainties is beyond the scope of this thesis, but a detailed explanation can be found in [136].

3.7 Missing transverse momentum

Some particles cannot be detected by the ATLAS detector, such as neutrinos and weakly-interacting particles predicted by theories that go beyond the Standard Model. However, their presence can be revealed by the momentum imbalance in the transverse plane, where momentum conservation is expected. The *missing transverse momentum*, $\mathbf{E}_T^{\text{miss}}$, is defined as the negative vector sum of the transverse momenta of all the reconstructed objects [137, 138].

The calorimeter-based $\mathbf{E}_T^{\text{miss}}$ reconstruction, denoted here by $\mathbf{E}_T^{\text{miss,calo}}$, uses energy deposits in the calorimeters and muons reconstructed in the MS or in the ID (for muons in regions not covered by the MS). The energy deposits are associated with high p_T objects in a specific order to avoid double counting in the $\mathbf{E}_T^{\text{miss,calo}}$ calculation:[1]

$$E_{x(y)}^{\text{miss,calo}} = E_{x(y)}^{\text{miss},e} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss,soft}} + E_{x(y)}^{\text{miss},\mu}, \quad (3.3)$$

where the *soft* term is added to the expression to calculate the remaining energy deposits from topological clusters which are not associated with any high p_T object and each term represents the negative sum of the momenta of the reconstructed and calibrated objects.

Figure 3.10 shows the $E_T^{\text{miss,calo}}$ distribution for $Z \to \mu\mu$ events. In this sample, where no real E_T^{miss} due to invisible particles is expected, contributions from im-

1 $E_T^{\text{miss,calo}} = \sqrt{(E_x^{\text{miss,calo}})^2 + (E_y^{\text{miss,calo}})^2}$

perfections in the object reconstruction and detector inefficiencies can be evaluated. A good agreement is observed between data and MC simulation distributions. Reconstruction performances for events with real E_T^{miss} can be studied using $W \to l\nu$ samples, due to the presence of the neutrino.

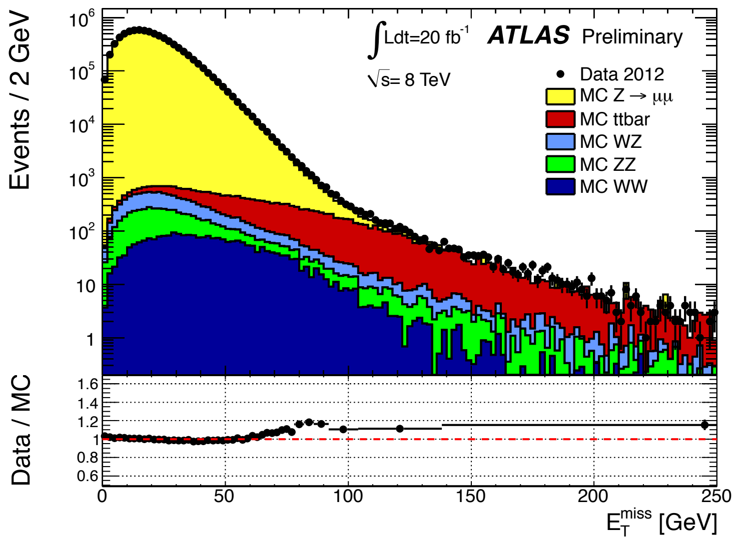


Fig. 3.10 · Distribution of calorimeter-based E_T^{miss} as measured in a data sample of $Z \to \mu\mu$ events, compared to MC simulation [138].

The overall systematic uncertainty on the $E_T^{\text{miss,calo}}$ measurement is given by the combination of the uncertainties on all individual terms, which depend on the event topology. For events with W and Z bosons decaying to leptons, the main contributions originate from uncertainties on the scale and resolution of charged leptons, the jet energy scale and the soft term [138].

A track-based method, denoted here by $\mathbf{E}_T^{\text{miss,trk}}$, can also be applied to reconstruct $\mathbf{E}_T^{\text{miss}}$. Using tracks reconstructed in the ID, which are associated with a primary vertex, this method is useful to remarkably reduce the pile-up dependence, which affects the calorimeter-based reconstruction. The $\mathbf{E}_T^{\text{miss,trk}}$ is calculated as the negative vector sum of the transverse momenta of tracks with $p_T > 500$ MeV and $|\eta| < 2.5$, originating from the primary vertex. Also the track-based reconstruction has its limitations, which are related to the absence of information on the neutral particles, since their momenta are not included, and to the smaller geometrical coverage of the ID with respect to the calorimeters.

In the LFV $\tau \to \mu\gamma\gamma$ analysis, a corrected version of $\mathbf{E}_T^{\text{miss,trk}}$ has been developed. This method, denoted by $\mathbf{E}_T^{\text{miss},\mu\text{corr}}$, removes from the calculation the transverse momentum of the muons and is employed to enhance the discrimination between signal and background sources. The two reconstruction methods $\mathbf{E}_T^{\text{miss,calo}}$ and $\mathbf{E}_T^{\text{miss},\mu\text{corr}}$ are used in the analysis presented in the next chapter.

3.8 Dedicated LFV object identification

The analysis of the LFV $\tau \to \mu\gamma\gamma$ channel, which is the main topic of this thesis and which will be presented in detail in the next chapter, requires a dedicated object identification. This analysis uses the tau leptons produced from the decay of W bosons. Such taus have a transverse momentum between 20 and 50 GeV. Subsequently in this model, the tau decays to a muon and two photons which are usually separated by a small angle, with $\Delta R(\mu, \gamma)$ and $\Delta R(\gamma, \gamma)$ of the order of 0.1. An efficient identification needs to take into account the proximity of the muon to the photons and to be able to recognize the pair of photons. The strategy developed to achieve this result is illustrated in this section.

3.8.1 Muon correction to the photon variables

Due to the small distance between muon and photon (see Figure 3.11), the energy deposition in the calorimeter by the muon modifies the shower shapes of the photons. A muon typically deposits around 300 MeV in the EM calorimeter and 2 GeV in the hadronic one. For this reason, the tight cut applied on the hadronic energy deposit of a photon by the standard photon identification would reject a large fraction of the $\tau \to \mu\gamma\gamma$ events. Loosening the hadronic veto cut in the standard identification, however, is not sufficient to correctly identify these events due to the presence of the two photons which are also close to each other and can overlap.

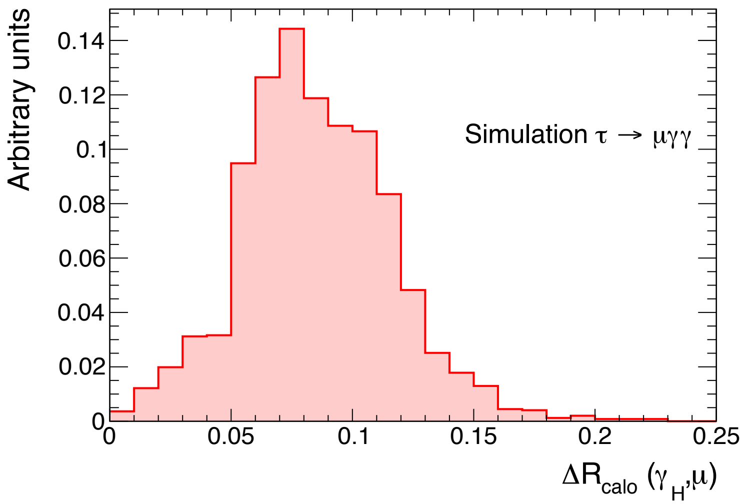


Fig. 3.11 · Angular distance, calculated at the calorimeter radius, between the muon and the higher energetic photon of the pair, defined as γ_H in the text, for MC $\tau \to \mu\gamma\gamma$ sample.

In our analysis, the proximity of the muon to the photons requires a correction to the shower shape variables and the energy deposition of the photons, in addition to the standard MC correction already applied. This correction is based on the distance between the muon and each of the photons calculated at the calorimeter radius (ΔR_{calo}), taken as 1.47 m. In case $\Delta\eta = |\eta^{\gamma} - \eta^{\mu}| < 0.16$ and $\Delta\phi_{\text{calo}} = |\phi^{\gamma}_{\text{EM2}} - \phi^{\mu}_{\text{calo}}| < 0.16$, where the parameters correspond to the center of the cluster in the second layer of the calorimeter, the shower shape variables are fitted as a linear function of $\Delta\phi_{\text{calo}}$ (Figure 3.12).

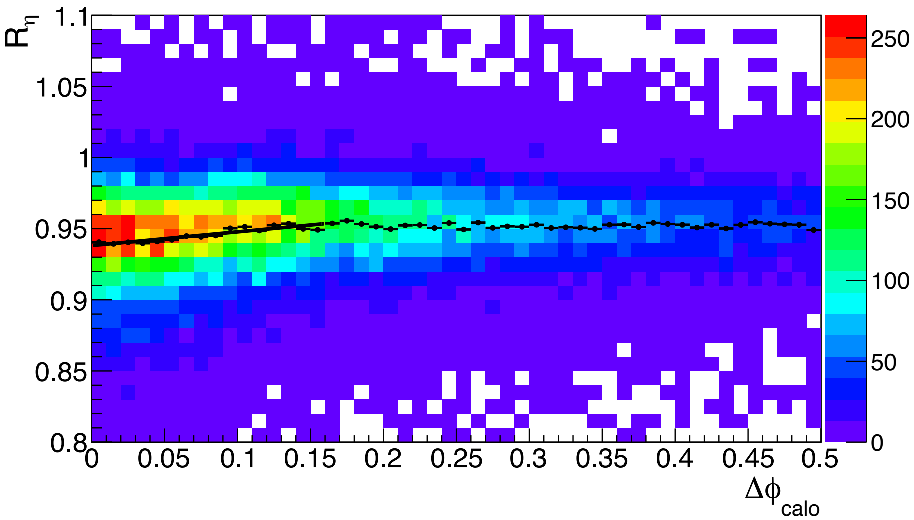


Fig. 3.12 · Dependence of R_η, defined later in Table 3.1, on the distance between photon and muon ($\Delta\phi_{\text{calo}}$), using MC $Z \to \mu\mu\gamma$ events. The default MC correction is already applied. The profile histogram (black) displays the mean value of R_η for each bin of the horizontal axis. The linear fit (black solid line) for photons close to the muon is also shown.

The muon correction, indicated as μC, is applied only on the variables showing a significant deviation from constant as specified below:

- $E^{\mu C} = E - 300\,\text{MeV}$
- $E_T^{\mu C} = E^{\mu C}/\cosh\eta$
- $R_{\text{had}}^{\mu C} = 0.006/(0.074 - 0.363 \times \Delta\phi_{\text{calo}})$
- $R_{\eta}^{\mu C} = 0.948/(0.933 + 0.103 \times \Delta\phi_{\text{calo}})$
- $E_{\text{ratio}}^{\mu C} = 0.914/(0.900 + 0.053 \times \Delta\phi_{\text{calo}})$

A detailed explanation of the definition of the shower shape variables will be given later in this section in Table 3.1.

To extract these corrections a sample of FSR photons from MC simulated $Z \to \mu\mu\gamma$ events is used, similarly to what is done in the standard photon identification, but including events with $\Delta R(\mu, \gamma) < 0.2$. Then, they are verified on a $Z \to \mu\mu\gamma$ sample taken from data. The effect of the correction on one of the shower shape variables, R_η, for photons close to the muon, is illustrated in Figure 3.13.

3.8.2 Di-photon BDT

The identification of the photon pair is crucial in the LFV $\tau \to \mu\gamma\gamma$ analysis. Here the two photons, close to each other, are reconstructed in two separated clusters. The standard photon identification variables, when necessary, are corrected to take into consideration the presence of the muon, as previously described.

To better discriminate the $\tau \to \mu\gamma\gamma$ signal from the background sources, a multivariate approach, based on the BDT method, is used. This is provided by the TMVA toolkit [134], integrated in the ROOT analysis framework [139].

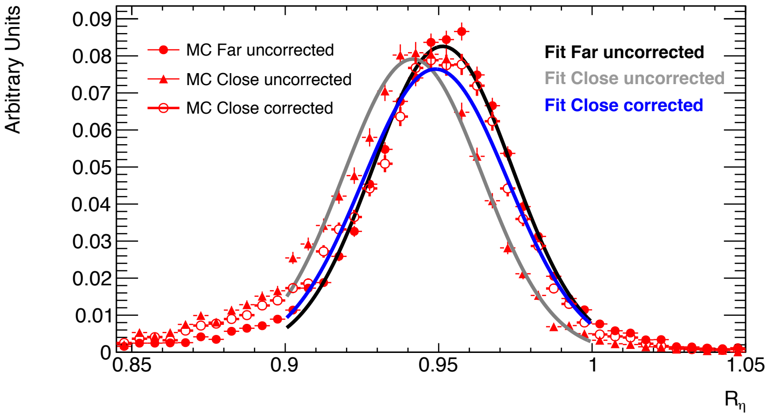
The procedure starts with two samples given as input to the TMVA, one for signal and one for background. The two samples are each split in two parts: the *training* sample is used to train the multivariate BDT method, called classifier, to distinguish between signal and background events; afterwards, the *test* sample is used to validate the performance of the training in the classification of the events.

The BDT training defines a sequence of cuts to classify the events as signal-like or background-like, exploiting a set of discriminating variables. The training starts from the root node. Here the algorithm selects the variable and the corresponding cut value that provide the best separation between signal and background. Following this criterion, the training sample is split in two subsamples. Signal-like events go to a new node to the right, background-like events go to a new node to the left. The process is repeated for the successive nodes. The iterative sequence of these binary splits creates the tree structure. The final nodes are labelled as "signal" or "background" depending on the majority of the events contained.

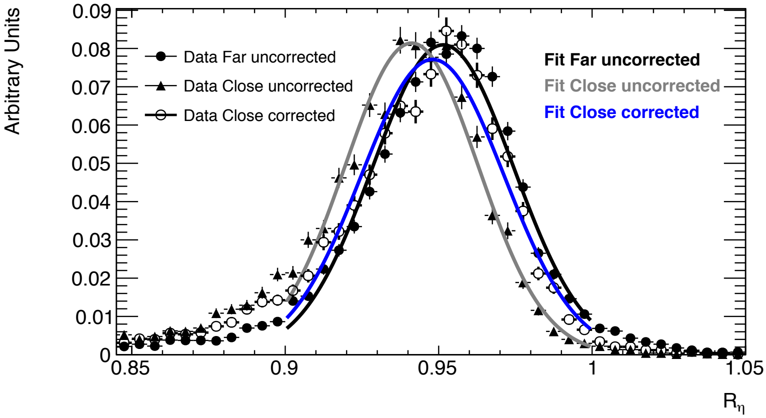
To increase the stability of the training and improve the performance, a "forest" of decision trees derived from the same training sample is built. A larger weight is given to the signal events misidentified as background (and vice versa) after the first tree. The reweighted training sample is used to develop a second decision tree. This procedure is typically repeated hundreds of times in a process called *boosting*. The final output of the BDT assigns to the event a score between -1 (100% background-like) and +1 (100% signal-like).

In our case, the BDT training is developed to separate photons from jets. The $\tau \to \mu\gamma\gamma$ signal, from MC simulation, and dijet background from data (defined in the next paragraph) are the two samples involved in the training. The amount of data corresponds to the period D of the 2012 data-taking, for an integrated luminosity of $3.1\,\text{fb}^{-1}$.

The background sample for the training consists of two back-to-back jets, where one of the two is reconstructed as a $\tau \to \mu\gamma\gamma$ candidate, in a p_T range



(a)



(b)

Fig. 3.13 · R_η distribution for MC (top) and data (bottom) $Z \to \mu\mu\gamma$ samples. The uncorrected distributions for photons in proximity to the muon (triangle), with $\Delta\eta < 0.16$ and $\Delta\phi_{\text{calo}} < 0.16$, are shifted after the muon correction (empty circle) strongly reducing the discrepancy with the distributions for isolated photons, far from the muon (filled circle).

between 10 and 45 GeV. The angular distance between the two objects is required to be $\Delta\phi > 2.5$ (inverting the cut applied to the signal sample). To increase the fraction of signal-like events in this background sample, a requirement on the muon isolation is also applied.

The variables selected for the di-photon BDT training are given in Table 3.1. Several photon shower shape variables are also employed in the standard photon identification described in Section 3.3. Some of them are used only for the more energetic photon of the pair (indicated as γ_H) and others only for the less energetic one (indicated as γ_L). In addition, a set of variables originally developed for standard τ_{had} reconstruction and identification provides further information. They represent hadronic shower shape variables and allow to separate events with one track and two photons (e.g. $\tau \to \rho\nu$ decays) from quark and gluon jets.

The distributions of the variables given as input to the BDT training are shown from Figure 3.14 to Figure 3.17. For each variable three distributions are compared: the two samples used in the training, signal and dijet data ($3.1\,\mathrm{fb}^{-1}$), and a subset of the data sample ($3.3\,\mathrm{fb}^{-1}$) used later in the analysis (see next chapter) passing the preselection described in Section 4.3.

After the training and testing phase, the TMVA toolkit produces important information to evaluate the results.[1] In Figure 3.18a the final BDT output score shows a good separation between signal and background distributions. In Figure 3.18b the ROC (Receiver Operating Characteristic) curve exhibits this discrimination power through the possibility of keeping a high signal efficiency and rejecting a large fraction of dijet background at the same time, selecting an appropriate cut value on the BDT output score. The correlation matrices are useful to avoid high correlation (or anti-correlation) between the variables, which could compromise the performance of the training. These matrices are shown in Figure 3.19, for signal and background samples.

To highlight the important role of the di-photon BDT in the LFV $\tau \to \mu\gamma\gamma$ object identification, we compare the efficiencies obtained using this method with the *tight* selection of the standard photon identification modified to remove the veto on the energy deposition in the hadronic calorimeter. Signal efficiency increases dramatically from a few percent using the standard photon identification to almost 90%, requiring the BDT score to be greater than 0,[2] as shown in Figure 3.20a. Applying this cut, more than 85% of the background is rejected. The comparison with the standard identification of the dijet background efficiency is shown in Figure 3.20b.

Finally, to verify the stability of the results related to the different conditions of the data-taking periods during 2012, the di-photon BDT is retrained on another sample of data. While the signal sample is the same of the previous training, the dijet background sample is obtained from a successive period of data-taking (period H), corresponding to $1.4\,\mathrm{fb}^{-1}$.

The performances of the two different trainings are compared considering the identification efficiencies on signal and dijet background from period D (the sample used in the first training). The efficiencies are calculated requiring the two BDT scores to be greater than 0, as a function of the transverse energy of the tau candidate. The performances of the two trainings are comparable, leading to similar efficiency rates, as shown in Figure 3.21, for both signal and dijet background.

[1] Several options are available to customize the BDT classifier, see [134] for details. The results shown in this section are obtained with this configuration: NTrees=200, MinNodeSize=2.5%, MaxDepth=2, BoostType=AdaBoost:AdaBoostBeta=0.5, SeparationType=GiniIndex, nCuts=20.

[2] The cut on the final BDT score will be optimized for the analysis presented in the next chapter.

Variable name	Description
γ_H and γ_L	
R_ϕ	Ratio of the energy in 3×3 cells over the energy in 3×7 cells
$\mathrm{w}_{\eta 2}$	Lateral width of the shower
E^1_{max}	Energy of strip (in first sampling) with maximal energy deposition
ΔE	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with minimal value between the first and second maximum
E_{ratio}	Ratio of the energy difference associated with the largest and second largest energy deposits over the sum of these energies
only γ_H	
R_η	Ratio of the energy in 3×7 cells over the energy in 7×7 cells
f_1^{core}	Fraction of energy in core of the strip layer over the total energy in the EM calorimeter
only γ_L	
R_{had}	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (over the range $0.8 < \lvert\eta\rvert < 1.37$)
f_1	Fraction of energy in the strip layer over the total energy in the EM calorimeter
$\tau \to \mu\gamma\gamma$ candidate	
$\gamma^{\mathrm{corr}}_{\mathrm{iso}}$	Energy in a cone $\Delta R < 0.4$ around γ_H obtained subtracting γ_L energy and normalizing on the sum of the energies of the two photons
averageIntPerXing	Average number of interactions per bunch crossing
tau	
f_{EM}	Fraction of energy deposited in the EM calorimeter over the energy sum in the EM and hadronic calorimeter
f_{PSS}	Fraction of energy deposited in the presampler and the strip layer over the energy sum in the EM and hadronic calorimeter
S_2^{BDT}	BDT score for the π^0 candidate with the second-highest score
$f^{\mathrm{corr}}_{\mathrm{cent}}$	Pile-up corrected fraction of energy within $\Delta R < 0.1$ over the total tau energy
$\Delta\eta^2_{\mathrm{strip}}$	Uncalibrated E_T weighted width in the strip layer within $\Delta R < 0.4$
caloIso$_{\mathrm{corr}}$	Pile-up corrected sum of E_T of clusters around jet seed within $0.2 < \Delta R < 0.4$
$f^{\mathrm{clus}}_{\mathrm{lead2}}$	Sum of two leading cluster energy over all cluster energy
CalibFactor	Calibration energy factor
$E^{\mathrm{corr}}_{\mathrm{pile\text{-}up}}$	Pile-up correction to the tau energy
E_T^{EM}	E_T from the EM calorimeter at the EM scale
E_T^{IsolHad}	E_T in hadronic calorimeter (at EM scale) within $0.2 < \Delta R < 0.4$

Table 3.1 · Definition of the input variables used for di-photon BDT training.

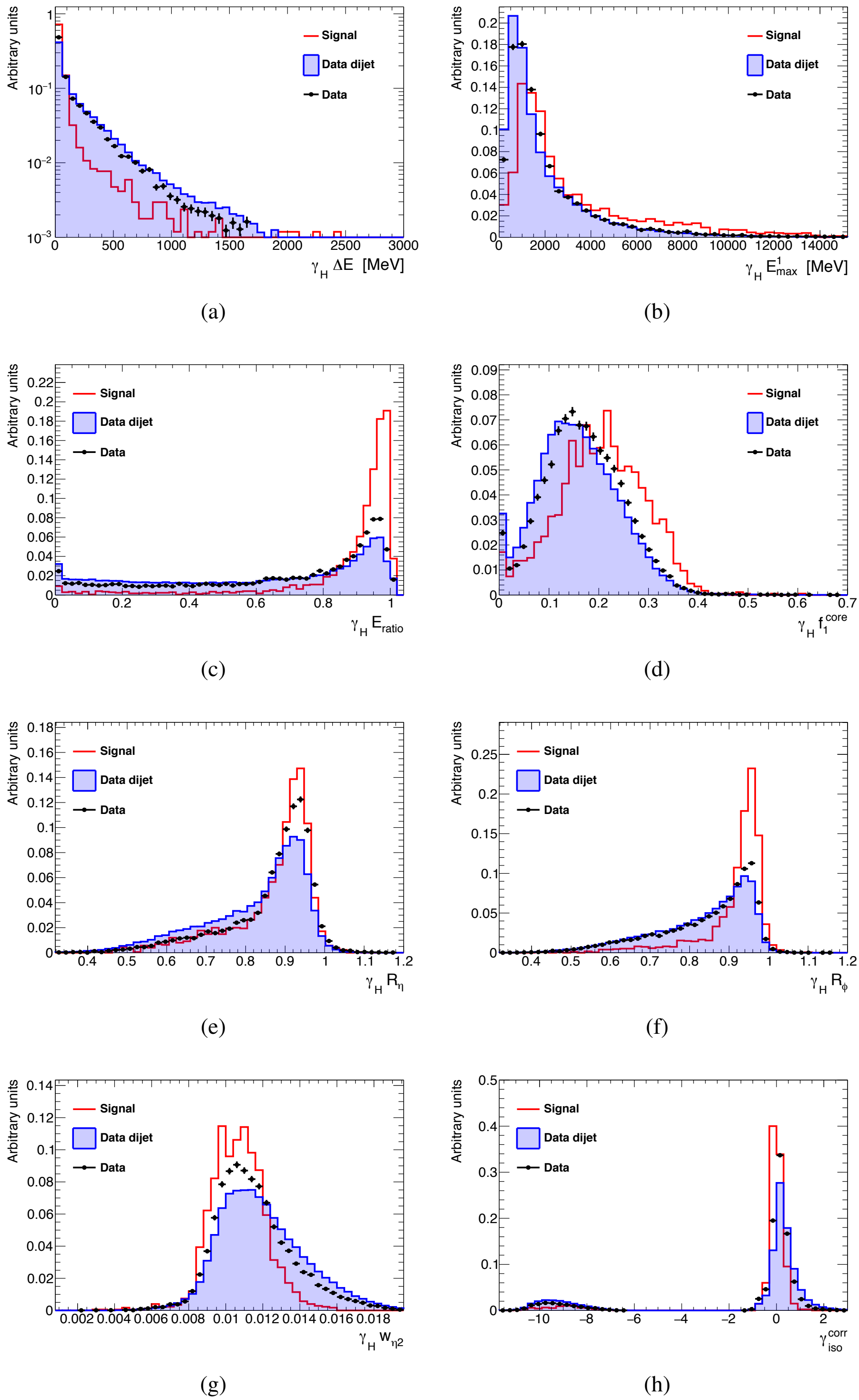


Fig. 3.14 · Input variables for di-photon BDT training. For each variable, the distributions represent signal (in red), dijet background from data (in blue) and a sample of data passing the $\tau \to \mu\gamma\gamma$ preselection described in Section 4.3 (black dots).

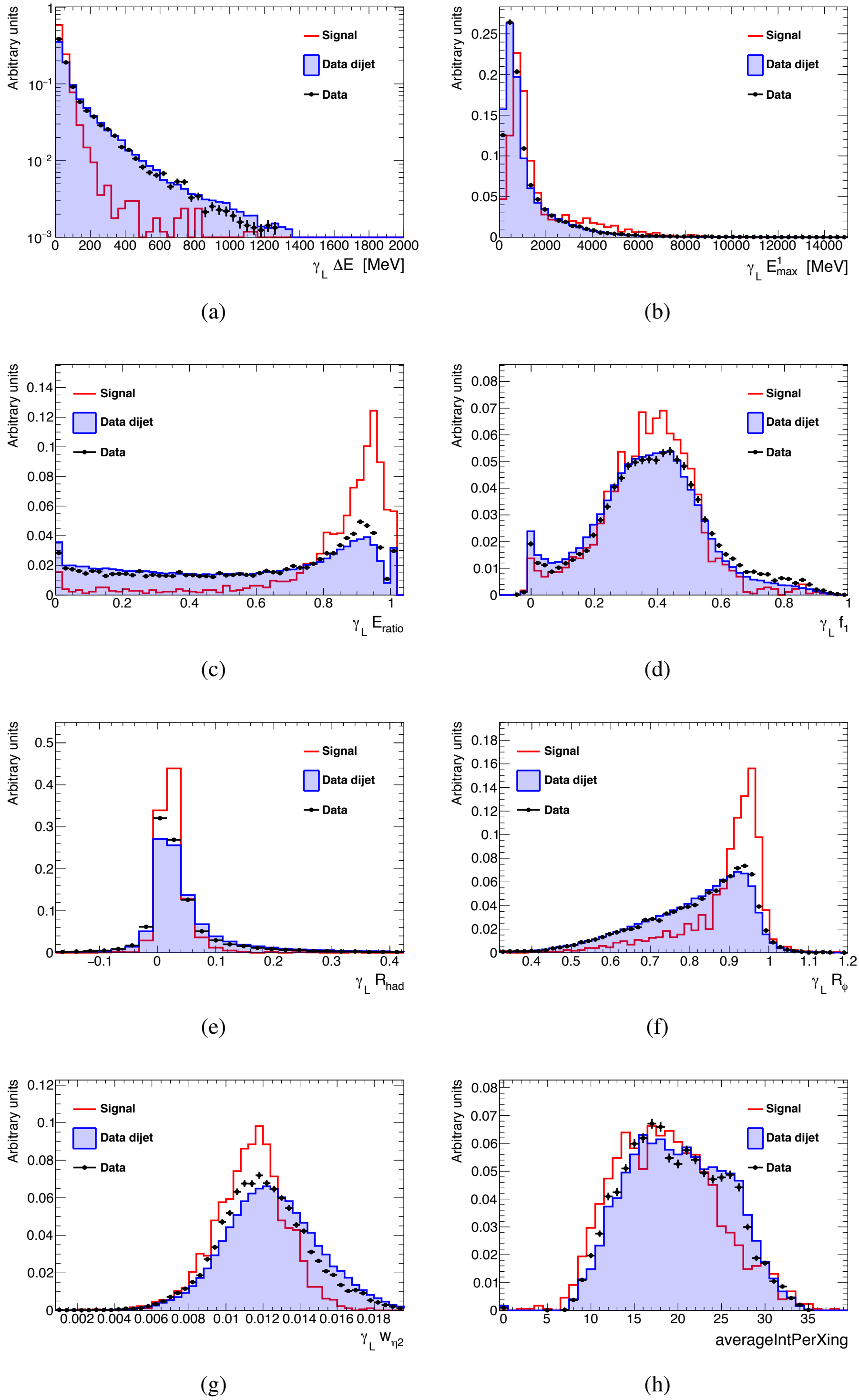


Fig. 3.15 · Input variables for di-photon BDT training. For each variable, the distributions represent signal (in red), dijet background from data (in blue) and a sample of data passing the $\tau \to \mu\gamma\gamma$ preselection described in Section 4.3 (black dots).

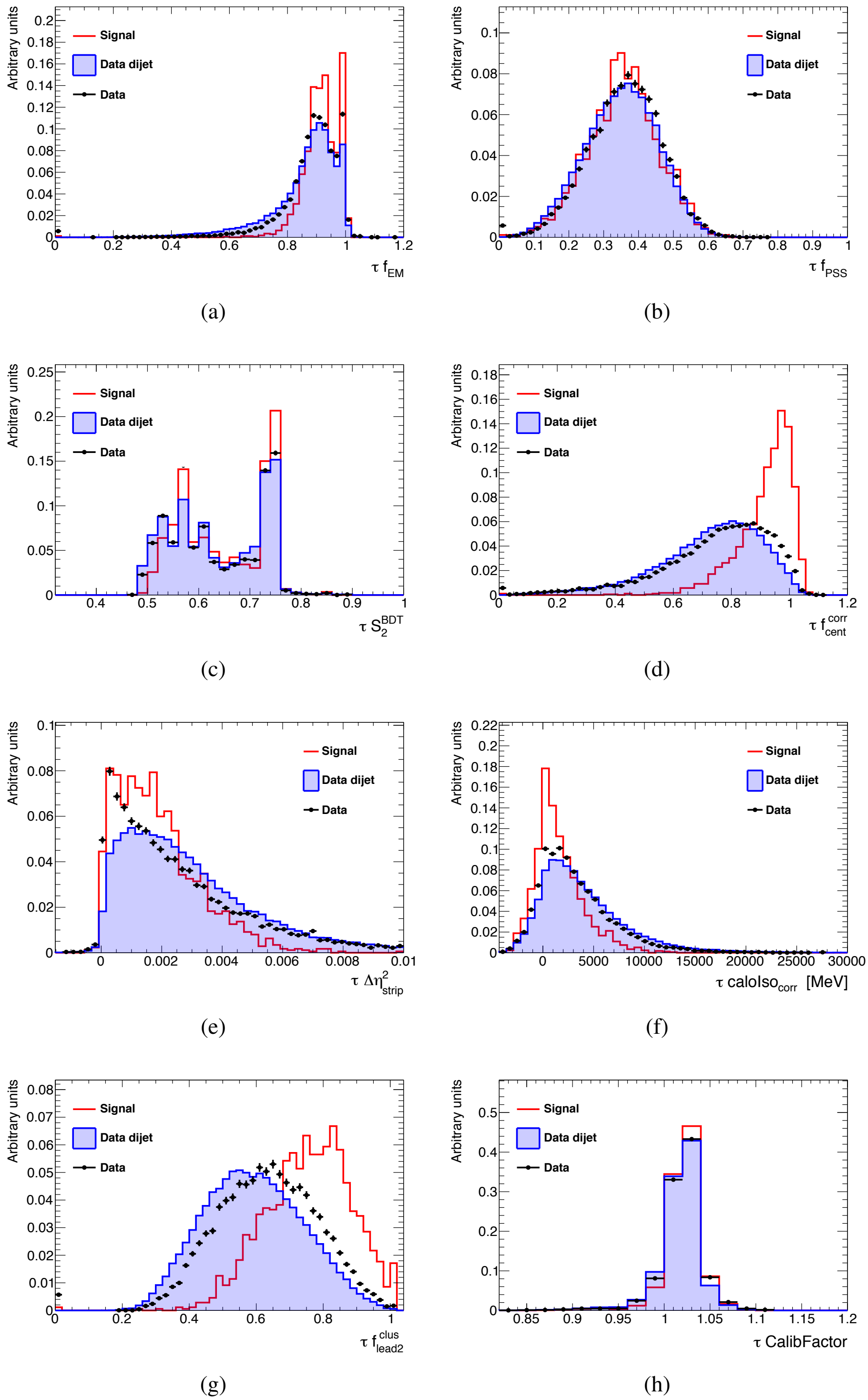


Fig. 3.16 · Input variables for di-photon BDT training. For each variable, the distributions represent signal (in red), dijet background from data (in blue) and a sample of data passing the $\tau \to \mu\gamma\gamma$ preselection described in Section 4.3 (black dots).

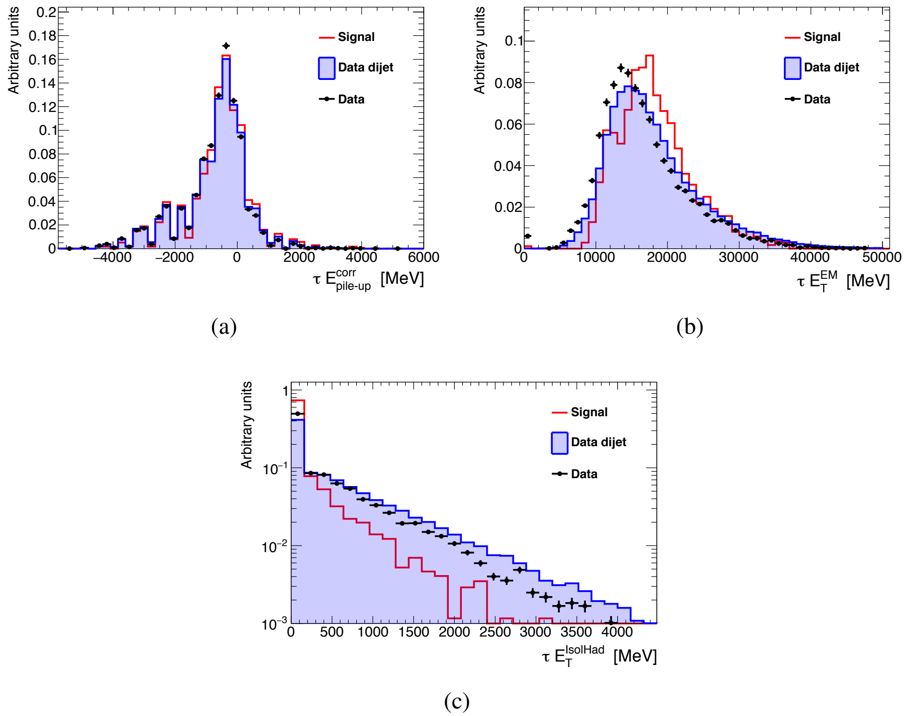
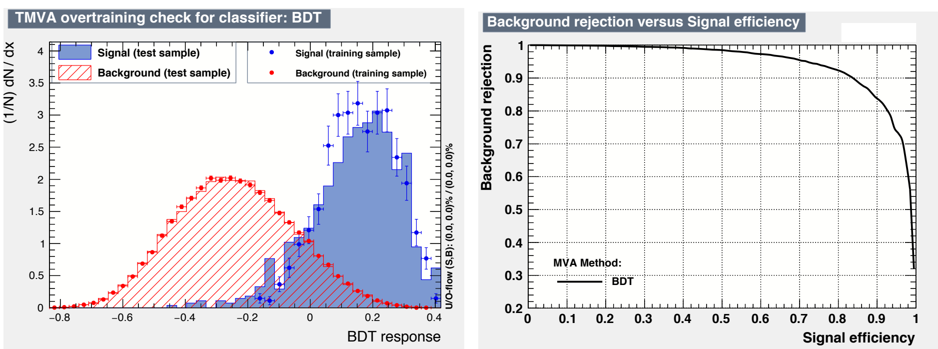
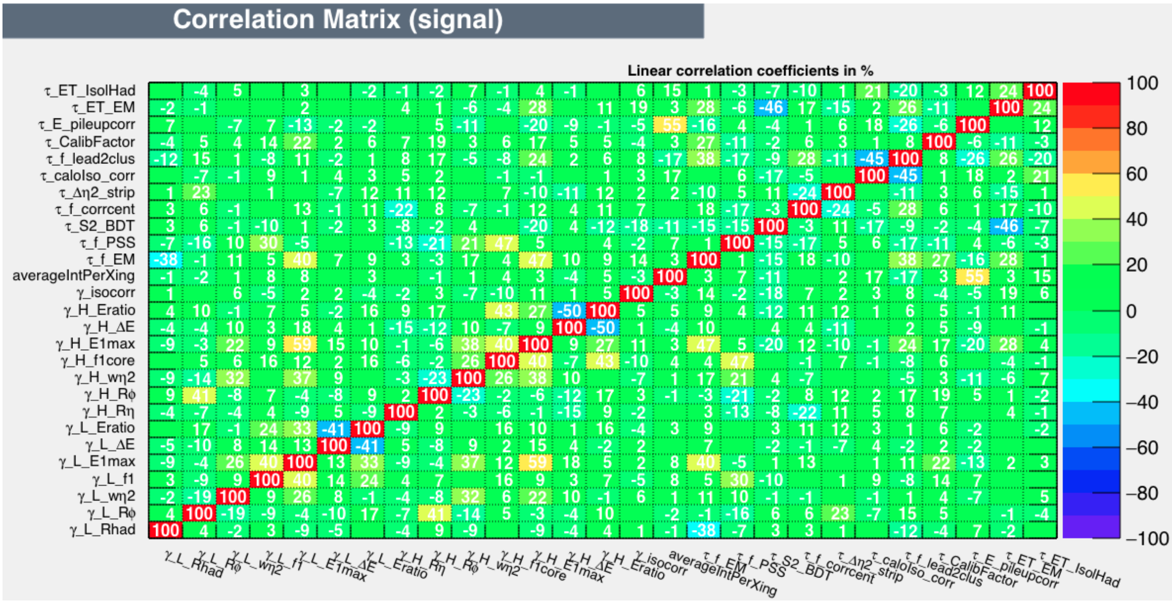


Fig. 3.17 · Input variables for di-photon BDT training. For each variable, the distributions represent signal (in red), dijet background from data (in blue) and a sample of data passing the $\tau \rightarrow \mu\gamma\gamma$ preselection described in Section 4.3 (black dots).

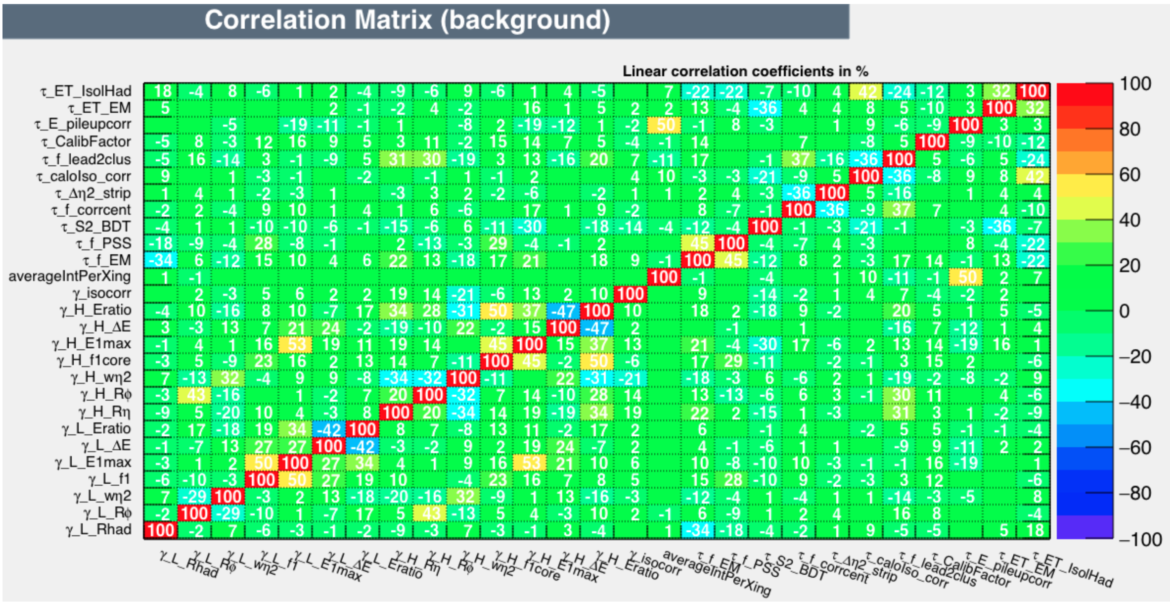


(a) (b)

Fig. 3.18 · Results of the BDT training. (a) Final BDT score of test and training samples for signal (blue) and background (red). (b) Background rejection as a function of signal efficiency (ROC curve).

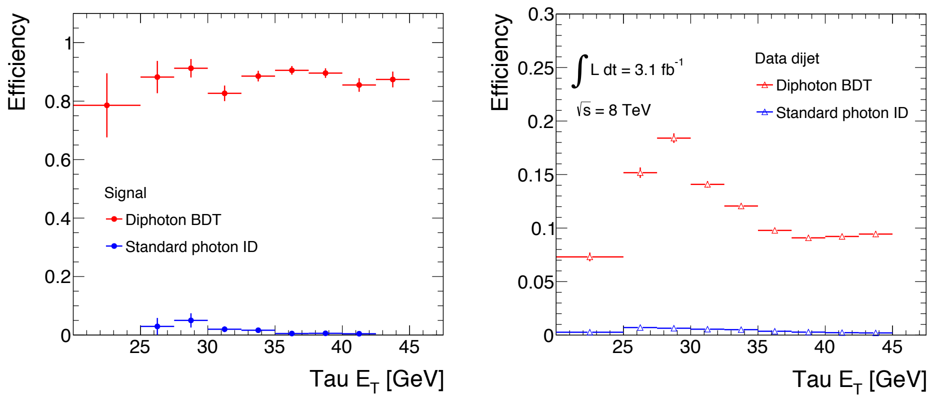


(a)



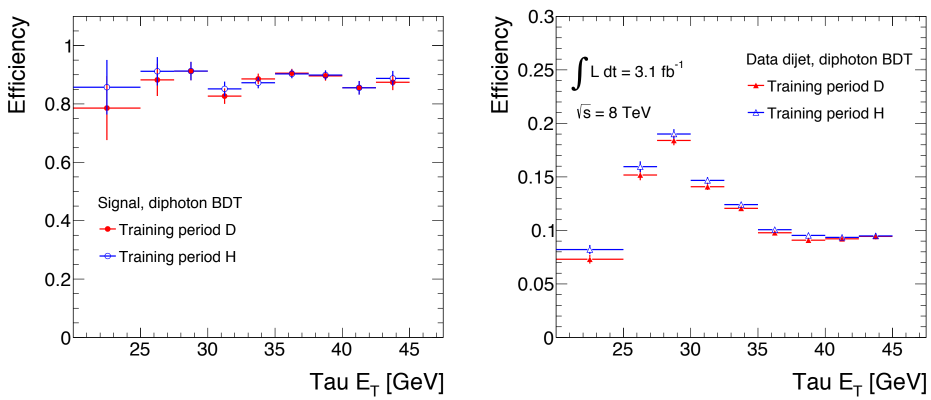
(b)

Fig. 3.19 · Correlation, in %, between the input variables for signal (a) and dijet background (b) samples.



(a) (b)

Fig. 3.20 · Comparison of the di-photon BDT efficiencies (red) to the ones of the standard photon identification (blue), for signal (left) and dijet background (right), as a function of the transverse energy of the tau candidate.



(a) (b)

Fig. 3.21 · Efficiency comparison on signal (left) and period D ($3.1\,\text{fb}^{-1}$) dijet background (right) between the training performed on the same period, in red, and the one performed on period H ($1.4\,\text{fb}^{-1}$), in blue, as a function of the transverse energy of the tau candidate.

chapter four

LFV $\tau \to \mu\gamma\gamma$ analysis

The analysis presented in this chapter describes the search of the charged lepton flavour violating $\tau \to \mu\gamma\gamma$ process. The source of the tau leptons taken in consideration is the $W \to \tau\nu$ decay. The search is performed using the data collected with the ATLAS detector in 2012 at $\sqrt{s} = 8$ TeV.

In ATLAS, the study of the $\tau \to \mu\gamma$ decay (with one photon instead of two) carried out using Run-1 (2012) data [60] resulted to be challenging due to the large amount of background affecting the analysis. As a consequence, the 2012 $\tau \to \mu\gamma$ analysis is not competitive with the limits set by the B factory experiments. The $\tau \to \mu\gamma\gamma$ channel investigated here represents a new intriguing probe of LFV. As previously described, some scenarios of physics beyond the Standard Model predict sizeable rates; in some models these are independent from the $\tau \to \mu\gamma$ case. Moreover, no experimental limits on the $\tau \to \mu\gamma\gamma$ channel have been set up to now.

4.1 Analysis strategy

The LFV $\tau \to \mu\gamma\gamma$ analysis is mostly based on data-driven methods. The background estimation, indeed, is derived directly from data, while the MC simulation is used to model the signal. The MC predictions for the background are only used as a cross-check and for a better understanding of its composition. The data and MC simulation samples will be described in Section 4.2.

Following a loose selection, which reduces the initial amount of data, the analysis relies on two main steps: object identification and event selection. The first one focuses on the examination of the LFV tau decay products, the muon and the pair of photons, for which the di-photon BDT is employed, as previously explained in Section 3.8. The event selection is used to investigate the properties of the $W \to \tau\nu$ decay and to further separate the signal from the background sources. To fully achieve this goal, a multivariate approach, which depends on the performance of a second BDT, called W kinematic BDT, is the strategy adopted once again.

Different kinematical regions are selected using the LFV tau candidate mass, given by the invariant mass of the three tau decay products, $m(\mu\gamma\gamma)$. The *signal region* is defined to maximize the signal sensitivity and reduce the background contamination. The region of the *sidebands* is used to estimate the background in the signal region. More details on the definition of these regions will be given in Section 4.4.1.

After the two BDT trainings are performed, the expected limit on the branching fraction of the LFV $\tau \to \mu\gamma\gamma$ channel is minimized. The procedure is based on the scan of several cut values applied on the two BDT scores, in order to select the combination which gives the lowest upper limit under the assumption that there is no signal excess.

The upper limit on the number of signal events in the signal region is determined at 90% confidence level using HistFitter [140], a software framework for statistical data analysis. Setting the limit on the $\tau \to \mu\gamma\gamma$ branching fraction also involves the estimation of the detector acceptance and efficiency terms for the signal and of the number of tau leptons produced from $W \to \tau\nu$ decays during the 8 TeV data-taking period.

4.2 Data and simulation samples

The data sample available for this analysis, collected during 2012, corresponds to an integrated luminosity of 19.1 fb^{-1}. Data have to fulfill quality requirements: all the relevant components of the detector have to be in good working condition when the events are recorded. Periods from A to L of the 2012 data-taking are used in the analysis. The data format consists of *AODs*, from which the *D3PD* samples used in the analysis are derived.[1]

The MC signal sample is modelled using $W \to \tau\nu$ production, where the tau lepton is forced to decay isotropically into a muon and two photons. The sample is generated by PYTHIA8 [141], using the AU set of tuned parameters [142] and the leading-order parton distribution function set MSTW2008LO [143]. The signal sample contains 100000 events because of enough statistics, after a filter is applied at the generator level. This filter requires at least two photons in the event, each with $p_T > 4$ GeV and $|\eta| < 2.7$; their invariant mass has to be lower than 1.8 GeV at truth-level. The muon is required to have $p_T > 3.8$ GeV and $|\eta| < 2.7$. This selection results in a generator efficiency of 25%. For an integrated luminosity of 19.1 fb^{-1}, a sample of this size corresponds to a $\tau \to \mu\gamma\gamma$ branching fraction of 1.7×10^{-3}.

As already mentioned, the background estimation is obtained directly from the data. Nevertheless, the MC simulation samples provide information on the major backgrounds. The Standard Model W and Z decays, namely $W \to \mu\nu$, $W \to \tau\nu$, $Z \to \mu\mu$ and $Z \to \tau\tau$ are the samples investigated, together with heavy-flavour (HF) and $t\bar{t}$ decays. The W and Z decays are generated using ALPGEN [144] and JIMMY [145]. The $t\bar{t}$ sample is produced by MCAtNLOJimmy [146]. The HF sample consists of $b\bar{b} \to X\mu$ and $c\bar{c} \to X\mu$ (muon $p_T > 15$ GeV) processes and it is generated using PYTHIA8B [141]. The production of this last sample was severely limited by the substantial amount of CPU time required. For a comparison to data (Section 4.6), a weight is applied to scale the MC background samples to

[1] *AOD* and *D3PD* are two data formats used by the ATLAS collaboration. *AODs* are centrally produced at CERN and contain detailed information about physics objects, the derived *D3PDs* contain ROOT [139] ntuples with information relevant for a specific analysis.

the correct data luminosity, using the number of MC generated events and the cross-section of the specific process modeled by the MC sample. A weight of the order 200 is applied to the HF sample. Detailed information on these MC samples for background processes will be provided in Section 4.6.

4.3 Preselection

The first selection of the events in the samples is performed by a set of preliminary cuts, denoted as *preselection*, which reject a significant part of the backgrounds and enhance the signal of this search. In particular, these cuts target the W boson decay and the $\tau \to \mu\gamma\gamma$ candidate in each event.

Object and event preselection

Several cuts are applied on the properties of the objects present in the signal final state, photons[1] and muons. The requirement on the transverse momentum of each of the two photons, γ_H (the more energetic of the pair) and γ_L, corresponds to $p_T > 3.8$ GeV. To further reduce the background contribution from the jets, the p_T sum of the two photons in the event has to be above 9 GeV.[2] Reconstructed photons are restricted to the pseudorapidity range $|\eta| < 2.47$; the transition region between barrel and end-cap calorimeters, $1.37 < |\eta| < 1.52$, is also excluded.

The muons involved in the LFV $\tau \to \mu\gamma\gamma$ analysis are combined muons, reconstructed by the Chain 3 algorithm. To pass the selection, muon candidates are required to have $p_T > 4$ GeV and $|\eta| < 2.4$. The default selection on the track quality in the ID, described in Section 3.1, is applied to these muons.

In addition to the object selection, a set of cuts targets the properties of the whole $W \to \tau\nu$ event. The p_T of the LFV tau candidate is required to be above 20 GeV, a value consistent with the kinematics of a tau lepton produced from the W decay. Another selection is applied on the transverse mass of the W boson, calculated using the calorimeter-based E_T^{miss} reconstruction in this case, given by

$$m_T(\tau, E_T^{\text{miss,calo}}) = \sqrt{2\, E_T^{\text{miss,calo}}\ E_T^{\text{tau}}(1 - \cos(\Delta\phi(\tau, E_T^{\text{miss,calo}})))}. \tag{4.1}$$

The minimum requirement on this quantity is set to 30 GeV, compatible with the kinematics of the signal events. Then, in order to suppress the multi-jet background, events with $\Delta\phi$ larger than 2.5 radians between the LFV tau candidate and any jet with $p_T > 20$ GeV are rejected. Finally, the difference between the $\mu\gamma\gamma$ invariant mass and the nominal value of the tau mass has to be within a range of 2.5 GeV.

---

[1] Three types of photons are considered for this analysis: unconverted, converted and low p_T photons reconstructed from EM topo-seeded clusters.

[2] Starting from this requirement and throughout the analysis, the photon energy is corrected to take into account the proximity of the muon, as explained in Section 3.8.1.

Trigger selection

A fundamental step in the selection of the events is represented by the triggers. The *TauMass* triggers, a set of dedicated *muon+photon* triggers described in Section 2.3, play a crucial role to select the $\mu\gamma\gamma$ combinations which represent the signal final state, sharing the momentum of the tau candidate. The trigger level candidates have to satisfy a requirement on the $\mu\gamma$ invariant mass at the EF, namely 1 GeV $< m(\mu\gamma) <$ 2.5 GeV.

To increase the signal trigger efficiency, two *single muon* triggers are also used in combination with the previous ones: EF_mu24i_tight and EF_mu36_tight. The trigger configuration for this analysis is completed by the *tau+muon* trigger, EF_tau20_medium1_mu15, which allows to almost double up the signal efficiency. The trigger selection is based on the logical operation *or* of the following EF triggers:

- EF_mu10i_g10_medium_TauMass
- EF_mu10i_loose_g12Tvh_medium_TauMass
- EF_mu20i_tight_g5_loose_TauMass
- EF_mu20i_tight_g5_medium_TauMass
- EF_mu4Ti_g20Tvh_medium_TauMass
- EF_mu24i_tight
- EF_mu36_tight
- EF_tau20_medium1_mu15

In this list, *mu*, *tau* and *g* refer to muon, hadronic tau decay and photon, respectively, while the numbers are the p_T thresholds. Identification criteria are denoted as *loose*, *medium* and *tight*; i indicates isolation, *Tvh* refers to high L1 p_T threshold and veto on the hadronic energy deposition of the photon. The trigger efficiency obtained by this combination is about 20% for the events of the signal sample remaining after the object and event preselection, as shown in Figure 4.1. The mass distribution of the LFV tau candidate, for a subset of the data events (3.3 fb^{-1}) which pass the trigger selection (also expressed in its components), is shown in Figure 4.2.

Further requirements

At this stage of the preselection, additional requirements are implemented to reduce the sources of background and select signal-like events. The first cut is applied on the invariant mass of the photon pair in the event, which must satisfy $m(\gamma_H, \gamma_L) > 300$ MeV, to reject the large amount of neutral pions in the low di-photon mass region. Figure 4.3 shows the di-photon invariant mass distribution for signal and a subset of the data sample (3.3 fb^{-1}).

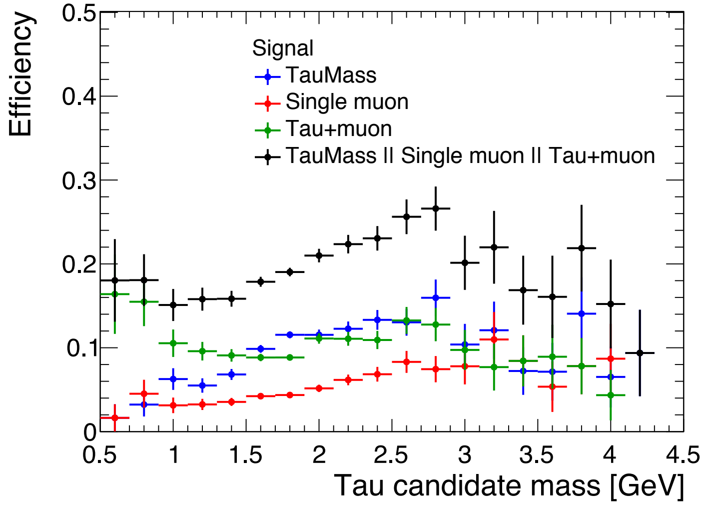


Fig. 4.1 · Signal trigger efficiency as a function of the LFV tau candidate mass, obtained using the set of *TauMass* triggers (blue), the *single muon* triggers (red), the *tau+muon* trigger (green) and their *or* combination (black), which is the configuration used in the analysis.

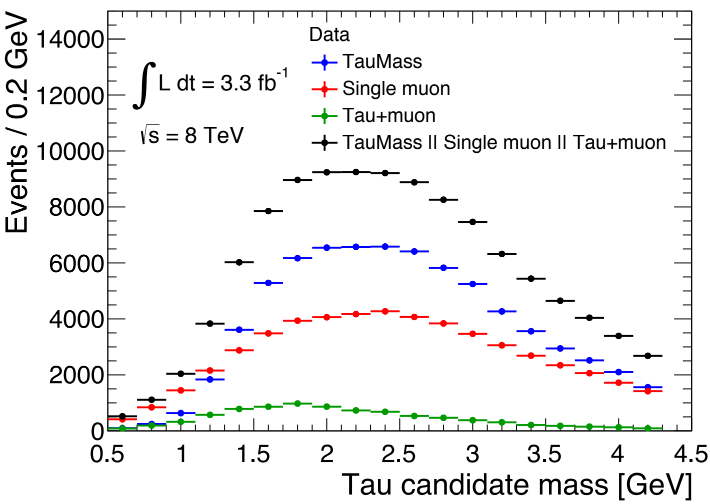


Fig. 4.2 · Mass distribution of the LFV tau candidate for a subset of the data sample after different trigger selections, which consist of: the set of *TauMass* triggers (blue), the *single muon* triggers (red), the *tau+muon* trigger (green) and their *or* combination (black), which is the configuration used in the analysis.

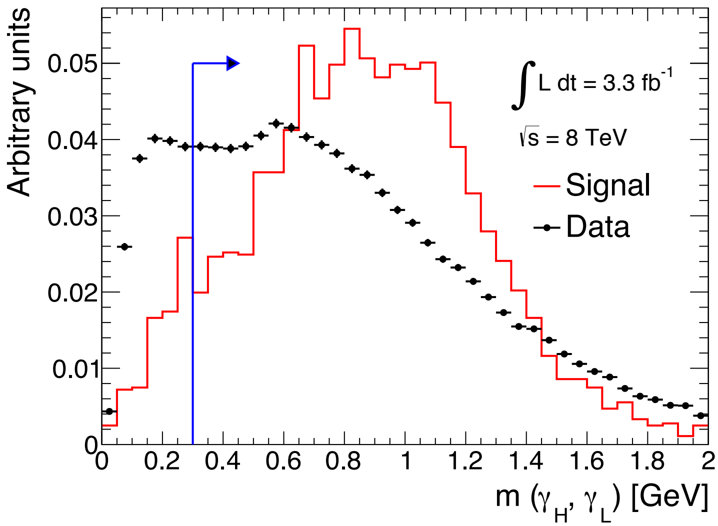


Fig. 4.3 · Di-photon invariant mass distribution for signal (red) and a subset of the data sample (black dots). The cut applied at 300 MeV is indicated by the blue line.

Furthermore, to solve a technical generator issue, a selection which concerns only the signal sample ensures that there is exactly only one LFV tau decay in the event. Subsequently, two further requirements are applied to both samples. The first one is employed to tighten the cut on the p_T sum of the two photons ($\sum p_T > 10$ GeV). The second one is based on the muon isolation, defined as the ratio between the scalar sum of tracks p_T within a cone of $\Delta R = 0.4$ around the muon, where the contribution of the photon conversion tracks is removed, and the p_T of the muon itself. This quantity has to be smaller than 0.5%.

At the end of the preselection, the signal sample consists of 1909 events, less than 2% of the initial one. In the data sample, 138762 events pass the set of cuts described in this section. A tighter event selection and the application of the di-photon BDT training results to the remaining events in the two samples, will be presented in the next section.

4.4 Selection of $W \to \tau\nu$ events

Following the object identification realized mainly through the di-photon BDT training, described in Section 3.8, the next step targets the $W \to \tau\nu$ decay. This further selection allows to reject the fraction of the background which is kinematically inconsistent with the signal. A number of cuts is applied on several variables before performing the second BDT training (W kinematic BDT), which requires a sufficient number of events.

The selection starts tightening the cut already applied in the preselection on the transverse mass, $m_T(\tau, E_T^{\text{miss,calo}})$, requiring a value greater than 40 GeV. About 20% of data (background) is rejected, while a limited loss of signal efficiency occurs. Figure 4.4 shows the distribution of this variable for signal and data samples.

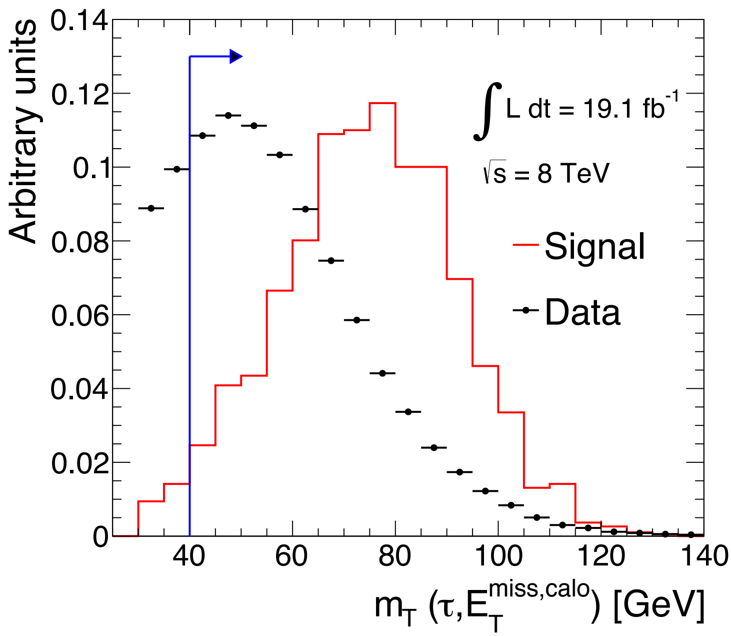


Fig. 4.4 · Transverse mass of the W boson, calculated using $E_T^{\text{miss,calo}}$, for signal (red) and data (black dots) samples. The cut applied on this variable, at 40 GeV, is indicated by the blue line.

In a second step, two cuts involving the muon transverse impact parameter and its significance are defined: $|d_0| < 0.5$ mm and $|d_0|/\sigma(d_0) < 10$, where $\sigma(d_0)$ is the estimated uncertainty on the parameter, obtained from the track fitting. This selection allows to discriminate between signal muons and background muons from π/K decays in flight. The muon d_0 parameter used in the analysis is defined *unbiased*, i.e. taken with respect to the primary vertex refitted without muon track. Figure 4.5 compares signal and data distributions for these two quantities, showing the respective requirements.

In addition, the results of the di-photon BDT training are used to classify the events of the samples under consideration. The di-photon BDT score distribution for the signal and data samples is shown in Figure 4.6. The events are selected requiring a value of the score above 0, a cut that rejects a large fraction of background-like events, reducing the data sample by 80%. The events passing all the cuts previously applied are used for the W kinematic BDT training. The effect of the cuts mentioned in this section on the number of signal and data events is shown in Table 4.1.

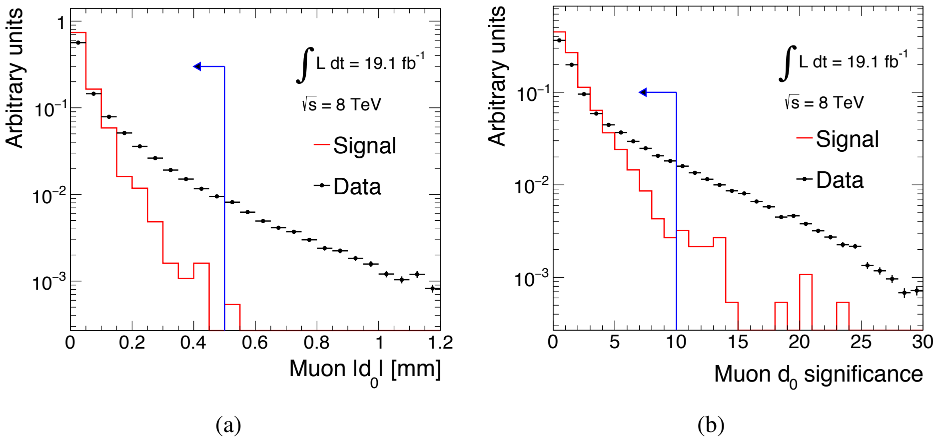


Fig. 4.5 · Distributions of muon $|d_0|$ (left) and d_0 significance (right) for signal (red) and data (black dots) samples. The cuts applied on these variables are indicated by the blue lines.

Cuts	Signal events	Data events
Preselection	1909	138762
$m_T(\tau, E_T^{\text{miss,calo}}) > 40$ GeV	1864 (97.6%)	112680 (81.2%)
Muon $\lvert d_0 \rvert < 0.5$ mm	1862 (99.9%)	107104 (95.1%)
Muon d_0 significance < 10	1838 (98.7%)	95404 (89.1%)
Di-photon BDT score > 0	1611 (87.6%)	19111 (20.0%)

Table 4.1 · List of the requirements applied in the selection that follows the preselection and the di-photon BDT training. For each cut, the number of events passing the selection and the relative efficiency (in brackets) are shown for the signal sample, initially containing 100000 events, and the data sample.

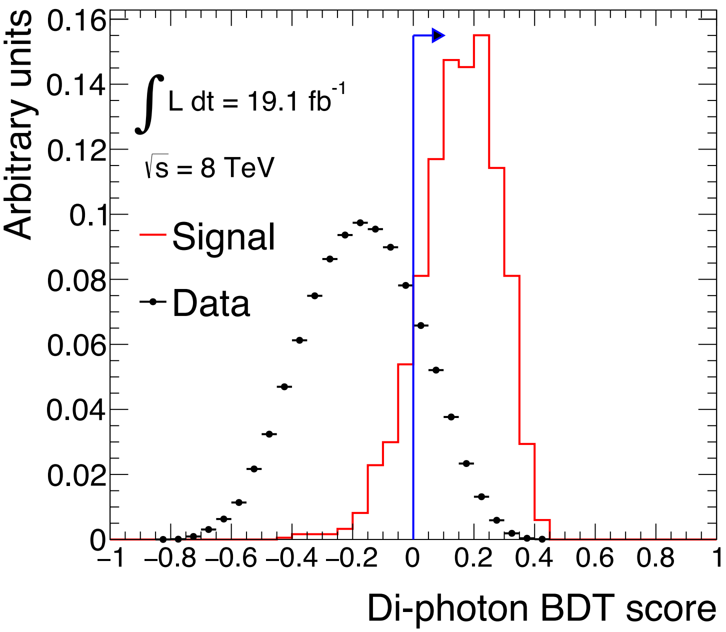


Fig. 4.6 · Di-photon BDT score for signal (red) and data (black dots) samples. A value greater than 0 (blue line) is required to pass the selection.

4.4.1 Signal region and sidebands

A signal region (SR) is selected relying on the distribution of the LFV tau candidate mass for the signal sample. As already explained, the tau mass is estimated as the invariant mass of the three decay products. The SR is defined as an interval around the nominal value of the tau mass, namely $1.6 \leq m(\mu\gamma\gamma) \leq 2.0$ GeV. An excess of signal events is expected in this range, which includes the majority of the signal sample (67%) at this stage of the selection. A slightly wider region than the SR itself ($1.5 \leq m(\mu\gamma\gamma) \leq 2.1$ GeV) is excluded from the data sample during the optimization, to avoid possible biasing (*blind analysis*). The spectrum of the tau candidate mass, for both signal and data, is shown in Figure 4.7.

Another region of the tau mass distribution is represented by the sidebands (SB), specifically defined for the data sample. This region includes two ranges on the two sides of the SR: $1.1 < m(\mu\gamma\gamma) < 1.5$ GeV (left sideband) and $2.1 < m(\mu\gamma\gamma) < 2.9$ GeV (right sideband). The choice of these ranges allows to pick a fraction of background which is consistent with the one in the SR and match the trigger level selection of the $\mu\gamma\gamma$ combinations. In the BDT training presented in the next section, only the data events in the SB will be used. Then, in the following phase, the amount of background in the SR will be estimated from the number of remaining events in the SB at the end of the selection.

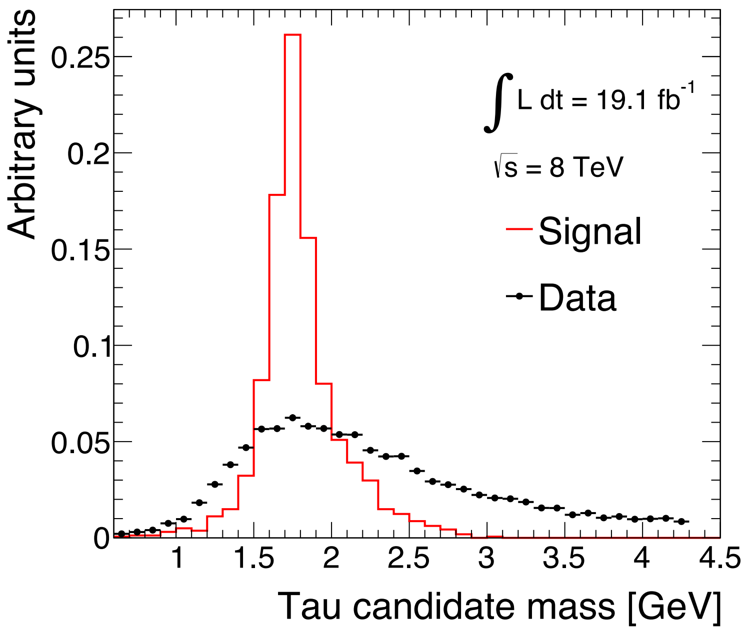


Fig. 4.7 · Mass distribution of the LFV tau candidate for the signal (red) and data (black dots) samples involved in the analysis.

4.5 W kinematic BDT

At this stage of the selection, the W kinematic BDT training is performed to target the residual background, where heavy-flavour (HF) and $W \to \mu\nu$ decays are the main contributions. The multivariate approach enhances the discrimination between signal and background, even if the phase space differs very little at this level. Two samples are given as input to the TMVA toolkit: the signal sample, MC $\tau \to \mu\gamma\gamma$, and the SB of the data (background sample). Both samples are split into halves, which form the training and test samples. The number of training (test) events involved in the W kinematic BDT is equal to 804 (807) for the signal and 4193 (4056) for the background.

A set of discriminating variables is employed in the BDT to target the kinematic and topological differences between signal and background events. The two E_T^{miss} reconstruction methods, $E_T^{\text{miss,calo}}$ and $E_T^{\text{miss},\mu\text{corr}}$, presented in Section 3.7, are used as input together with variables which involve these quantities, such as angular distances and the transverse mass of the W boson. The muon variables, e.g. the *momentum balance significance*, are mainly useful to reject background muons from π/K decays in flight. Variables describing the kinematic properties of the LFV tau candidate and the contribution of the jets in the event are also employed in the training. The complete list of the input variables is given in Table 4.2.

The distributions of these discriminating variables are shown from Figure 4.8 to Figure 4.10. For each variable, in addition to the distributions of the two samples used in the BDT training (signal and the SB of data), the data sample in the SR is also represented. The variables are shown after the requirement on the di-photon BDT score is applied.

As already explained during the description of the previous BDT training (Section 3.8.2), the TMVA toolkit offers the possibility to display and evaluate the results of the training for a given configuration.[1] In Figure 4.11a, the BDT output score for test and training samples shows the separation power between signal and background achieved with this method. The performance of the training is also evaluated through the ROC curve, which shows the background rejection versus the signal efficiency (Figure 4.11b). Finally, the matrices which indicate the correlation coefficients between the input variables are shown in Figure 4.12, for signal and background samples.

In the optimization phase, described later in this chapter, the cut value on the W kinematic BDT score, together with the one on the di-photon BDT score, will be determined in order to minimize the upper limit on the $\tau \to \mu\gamma\gamma$ branching fraction.

[1] Summary of the BDT settings used in this configuration: NTrees=200, MinNodeSize=5%, MaxDepth=2, BoostType=AdaBoost:AdaBoostBeta=0.5, SeparationType=GiniIndex, nCuts=20. A detailed explanation of these options is provided in [134].

Variable name	Description
Tau E_T	Transverse energy of the LFV tau candidate
Tau η	Pseudorapidity of the LFV tau candidate
$m_T(\tau, E_T^{\text{miss},\mu\text{corr}})$	Transverse mass of the W boson, calculated using $E_T^{\text{miss},\mu\text{corr}}$
$\Delta\phi(\tau, E_T^{\text{miss,calo}})$	Difference between azimuth angles of the LFV tau candidate and $E_T^{\text{miss,calo}}$
$\Delta\phi(\tau, E_T^{\text{miss},\mu\text{corr}})$	Difference between azimuth angles of the LFV tau candidate and $E_T^{\text{miss},\mu\text{corr}}$
$\Delta\phi(\tau, E_T^{\text{miss,calo}})_{\text{W restframe}}$	Difference between azimuth angles of the LFV tau candidate and $E_T^{\text{miss,calo}}$ in W boson restframe
Muon ScatCurvSignif	Deflection in muon track trajectory compared to the expected deviation from multiple scattering
Muon ScatNeighSignif	Significance of a change in muon track trajectory (kink) expected in presence of $\pi/K \to \mu$ decay
Muon MomBalSignif	Difference between the muon momentum measured by the ID and the MS divided by the uncertainty on the energy loss measured by the calorimeters
Muon d_0 significance	Ratio between the absolute value of the transverse impact parameter of a muon and its associated uncertainty
$E_T^{\text{miss,calo}}$	Calorimeter-based E_T^{miss}
$H_T^{\text{leading jet}}$	Scalar p_T sum of the LFV tau candidate and the leading jet in the event with $p_T > 20$ GeV and $\Delta R > 0.2$ of the tau candidate
H_T^{jets}	Scalar p_T sum of the LFV tau candidate and the jets in the event with $p_T > 20$ GeV and $\Delta R > 0.2$ of the tau candidate
$E_T^{\text{miss},\mu\text{corr}}$	Track-based E_T^{miss}, corrected to remove the transverse momentum of the muons

Table 4.2 · Definition of the input variables used for W kinematic BDT training.

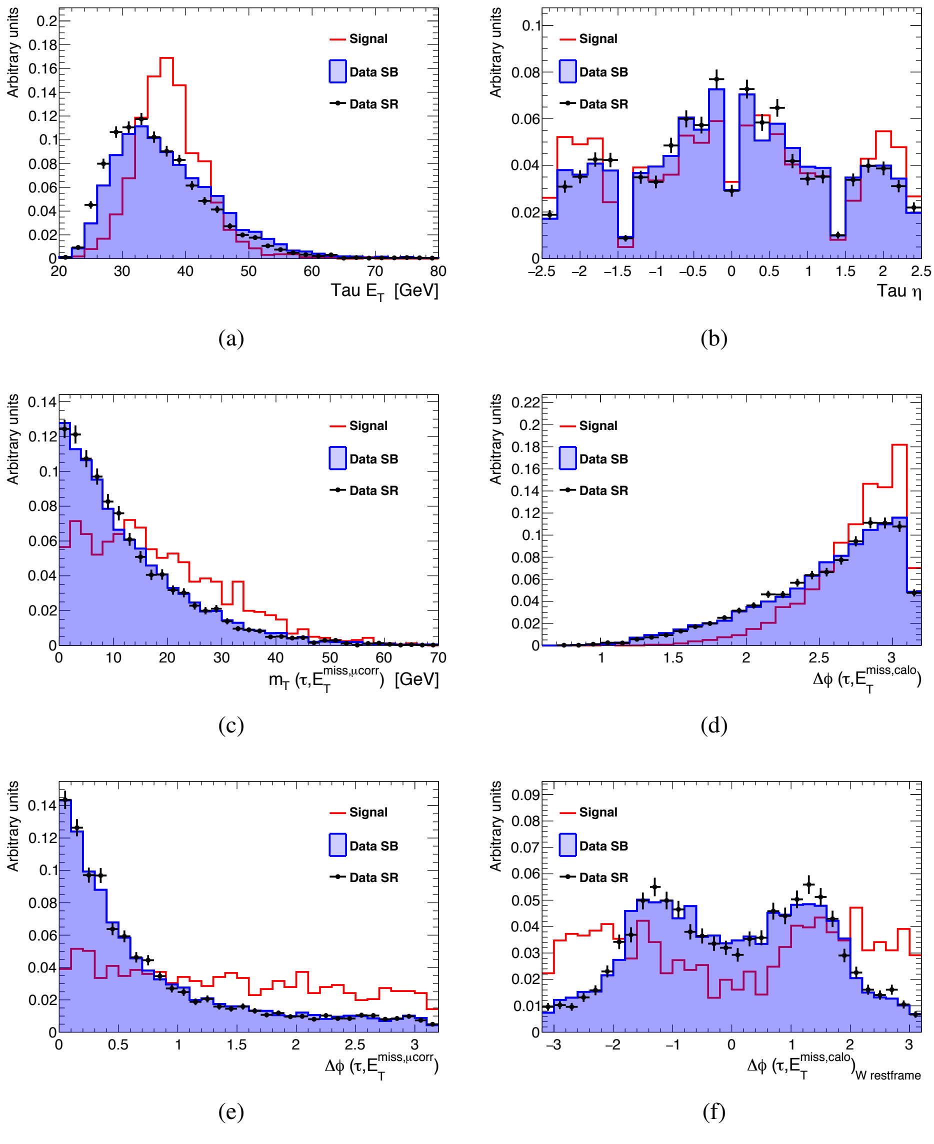


Fig. 4.8 · Input variables for W kinematic BDT training. For each variable, the distributions represent signal (in red), data in the SB (in blue) and data in the SR (black dots). They are shown using the full statistics available (test + training samples) after the di-photon BDT score > 0 requirement.

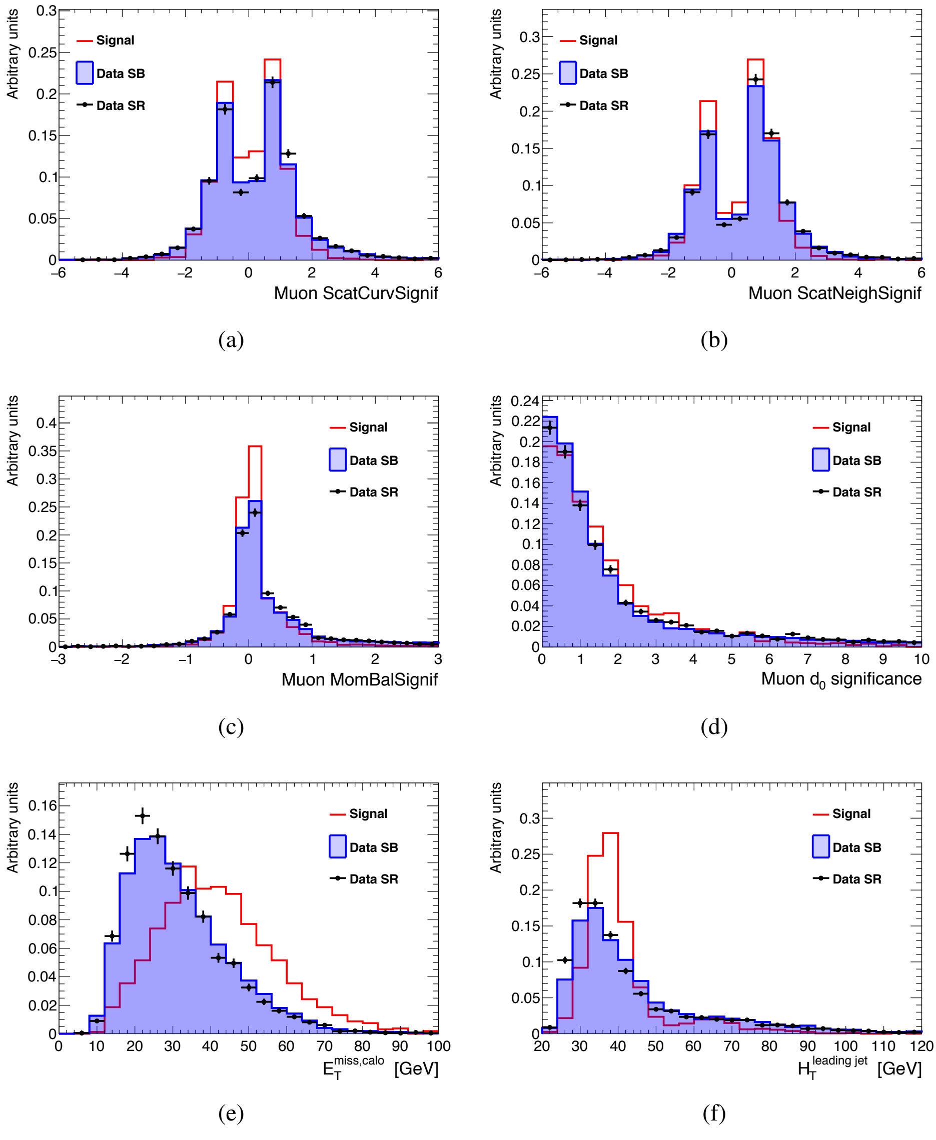
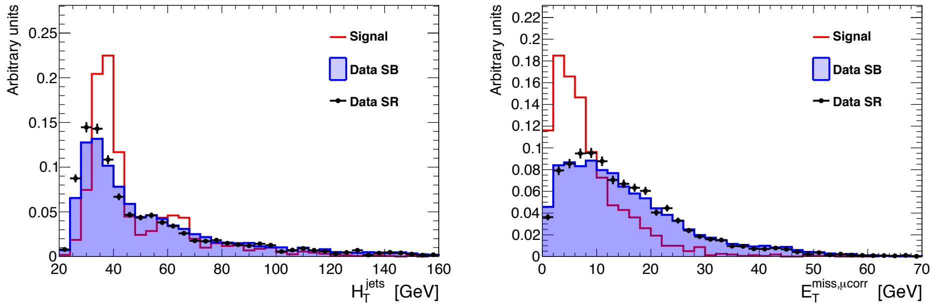
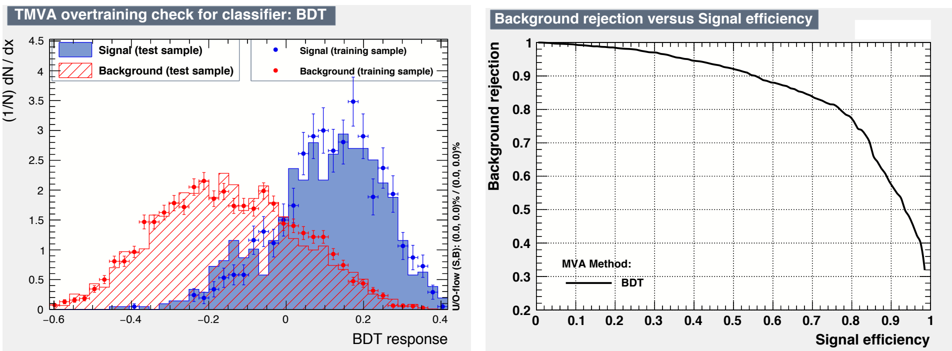


Fig. 4.9 · Input variables for W kinematic BDT training. For each variable, the distributions represent signal (in red), data in the SB (in blue) and data in the SR (black dots). They are shown using the full statistics available (test + training samples) after the di-photon BDT score > 0 requirement.



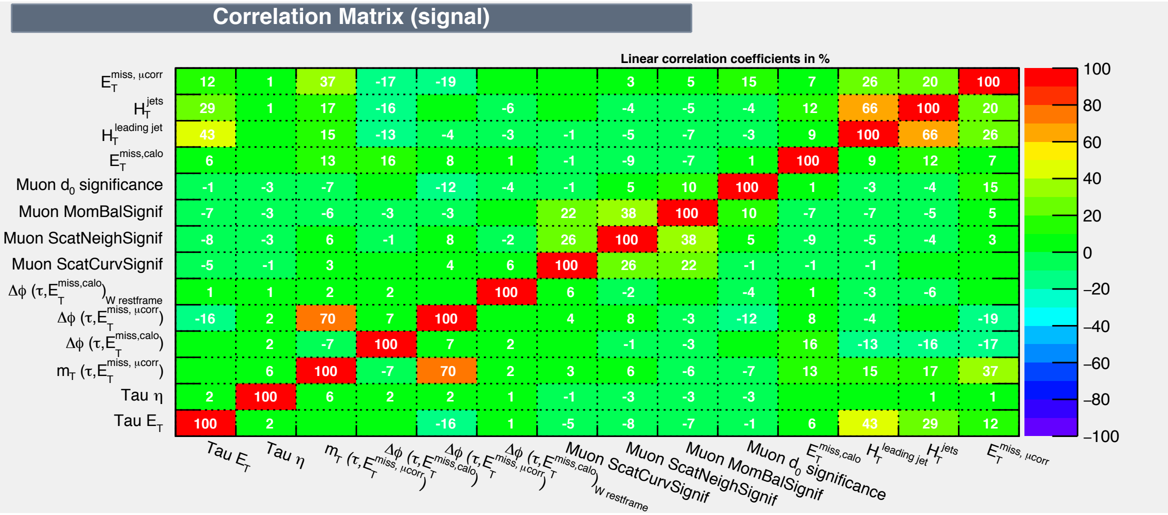
(a) (b)

Fig. 4.10 · Input variables for W kinematic BDT training. For each variable, the distributions represent signal (in red), data in the SB (in blue) and data in the SR (black dots). They are shown using the full statistics available (test + training samples) after the di-photon BDT score > 0 requirement.

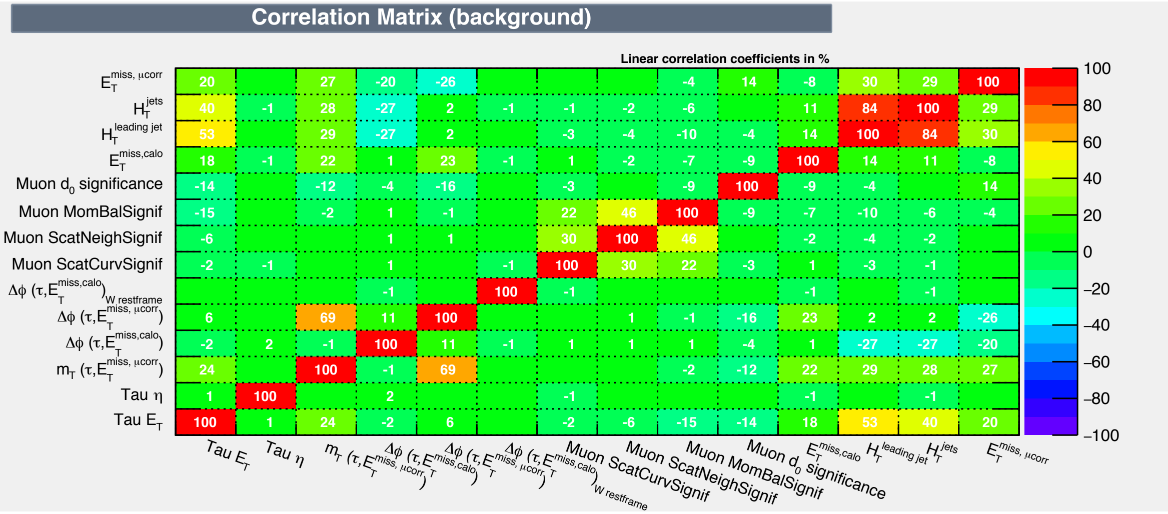


(a) (b)

Fig. 4.11 · Results of the W kinematic BDT training. (a) BDT output score of test and training samples for signal (blue) and background (red). (b) Background rejection versus signal efficiency (ROC curve).



(a)



(b)

Fig. 4.12 · Correlation, in %, between the input variables for signal (a) and background (b) samples.

4.6 MC samples for background

The background estimation in this analysis is performed using data. The simulation samples produced for the background processes would require much higher statistics to achieve the same precision. Nevertheless, the MC background samples are employed as a cross-check and provide useful information about the background characteristics and composition. In this section, several distributions are shown for the different MC backgrounds and compared with the ones obtained from the data.

The samples considered, which have been introduced in Section 4.2, are inclusive $W \to \mu\nu$, $W \to \tau\nu$, $Z \to \mu\mu$ and $Z \to \tau\tau$ decays (including decays with two associated radiative photons), HF ($b\bar{b}, c\bar{c} \to X\mu$)[1] and $t\bar{t}$ decays. A summary of the properties of the MC background samples is given in Table 4.3. At the stage of the analysis represented by the W kinematic BDT training, the largest background contributions are given by HF (65%) and $W \to \mu\nu$ decays (28%). Further rejection of these backgrounds is obtained by applying the selection on the two BDT scores.

As already mentioned in Section 4.2, the MC background samples are weighted to the correct data luminosity. The weight is obtained by the ratio of the integrated luminosity of the data to the luminosity of the MC sample. The latter is calculated by dividing the number of generated events in the MC sample by the cross-section of the specific process (including generation efficiency and higher order corrections if necessary).

To highlight the different background contributions, several distributions are presented for the MC samples. Figure 4.13 shows four variables used as input to the W kinematic BDT training, namely the transverse energy of the LFV tau candidate (Tau E_T), the two E_T^{miss} definitions and $H_T^{\text{leading jet}}$ (see Table 4.2). For a better visualization, the HF sample is scaled down by 50% for all the variables illustrated in this section. These distributions are shown after the di-photon BDT score $>$ 0 cut is applied, where HF and $W \to \mu\nu$ channels are the major backgrounds. Discrepancies are observed in the comparison with data, at least partly due to the limited statistics of the MC background samples.

Di-photon and W kinematic BDT scores for the MC backgrounds are obtained by applying the trained BDTs to these samples. The di-photon BDT score is shown in Figure 4.14a, before any cut is applied on this variable. The W kinematic BDT score[2] (obtained after the di-photon BDT score $>$ 0 selection) is shown in Figure 4.14b.

To conclude this brief overview, mass distributions of the LFV tau candidate are shown in Figure 4.15 in two stages of the selection: after the requirement on muon $|d_0|$, one of the cuts of the $W \to \tau\nu$ selection presented in Section 4.4, and following the di-photon BDT score $>$ 0 cut. The background composition is homogenous in the range that includes the SB and the SR, between 1.1 and 2.9 GeV.

---

[1] HF decays to muons with $p_T < 15$ GeV are missing in the MC sample, while present in data.

[2] While in this section the whole signal and data samples are used, in Section 4.7 the W kinematic BDT score distribution will be shown for the two test samples, signal (restricted to the SR) and SB of data.

Physics process	Cross-section (nb)	k-factor	Generator Efficiency	Generator	$\mathcal{L}_{MC}$ (pb^{-1})
$W \to \mu\nu$+jets (Np0)	8.05	1.19	1	AlpgenJimmy	362
$W \to \mu\nu$+jets (Np1)	1.58	1.19	1	AlpgenJimmy	1328
$W \to \mu\nu$+jets (Np2)	0.48	1.19	1	AlpgenJimmy	6635
$W \to \mu\nu$+jets (Np3)	0.13	1.19	1	AlpgenJimmy	6337
$W \to \mu\nu$+jets (Np4)	0.036	1.19	1	AlpgenJimmy	5955
$W \to \tau\nu$+jets (Np0)	8.03	1.19	1	AlpgenJimmy	352
$W \to \tau\nu$+jets (Np1)	1.58	1.19	1	AlpgenJimmy	1304
$W \to \tau\nu$+jets (Np2)	0.48	1.19	1	AlpgenJimmy	6542
$W \to \tau\nu$+jets (Np3)	0.13	1.19	1	AlpgenJimmy	6333
$W \to \tau\nu$+jets (Np4)	0.035	1.19	1	AlpgenJimmy	5956
$Z \to \mu\mu$+jets (Np0)	0.71	1.19	1	AlpgenJimmy	7813
$Z \to \mu\mu$+jets (Np1)	0.15	1.19	1	AlpgenJimmy	7247
$Z \to \mu\mu$+jets (Np2)	0.049	1.19	1	AlpgenJimmy	6950
$Z \to \mu\mu$+jets (Np3)	0.014	1.19	1	AlpgenJimmy	6529
$Z \to \mu\mu$+jets (Np4)	0.0038	1.19	1	AlpgenJimmy	6662
$Z \to \tau\tau$+jets (Np0)	0.71	1.19	1	AlpgenJimmy	7801
$Z \to \tau\tau$+jets (Np1)	0.16	1.19	1	AlpgenJimmy	7235
$Z \to \tau\tau$+jets (Np2)	0.049	1.19	1	AlpgenJimmy	6974
$Z \to \tau\tau$+jets (Np3)	0.014	1.19	1	AlpgenJimmy	6542
$Z \to \tau\tau$+jets (Np4)	0.0037	1.19	1	AlpgenJimmy	6723
$t\bar{t}$ LeptonFilter	0.21	1	0.54	MCAtNLOJimmy	132791
$t\bar{t}$ allhad	0.21	1	0.46	MCAtNLOJimmy	12614
HF ($b\bar{b} \to X\mu 15$)	199	1	1	PYTHIA8B	25
HF($c\bar{c} \to X\mu 15$)	80	1	1	PYTHIA8B	62

Table 4.3 · Properties of the MC background samples used in the analysis. Cross-section, k-factor (ratio of next-to-leading order (NLO) to leading order (LO) cross-section), type and efficiency of the generator and integrated luminosity of each sample are given. HF samples are inclusive heavy-quarks decays requiring a muon with $p_T > 15$ GeV. (NpX) indicates the number of gluons or quarks produced in addition to the W and Z bosons.

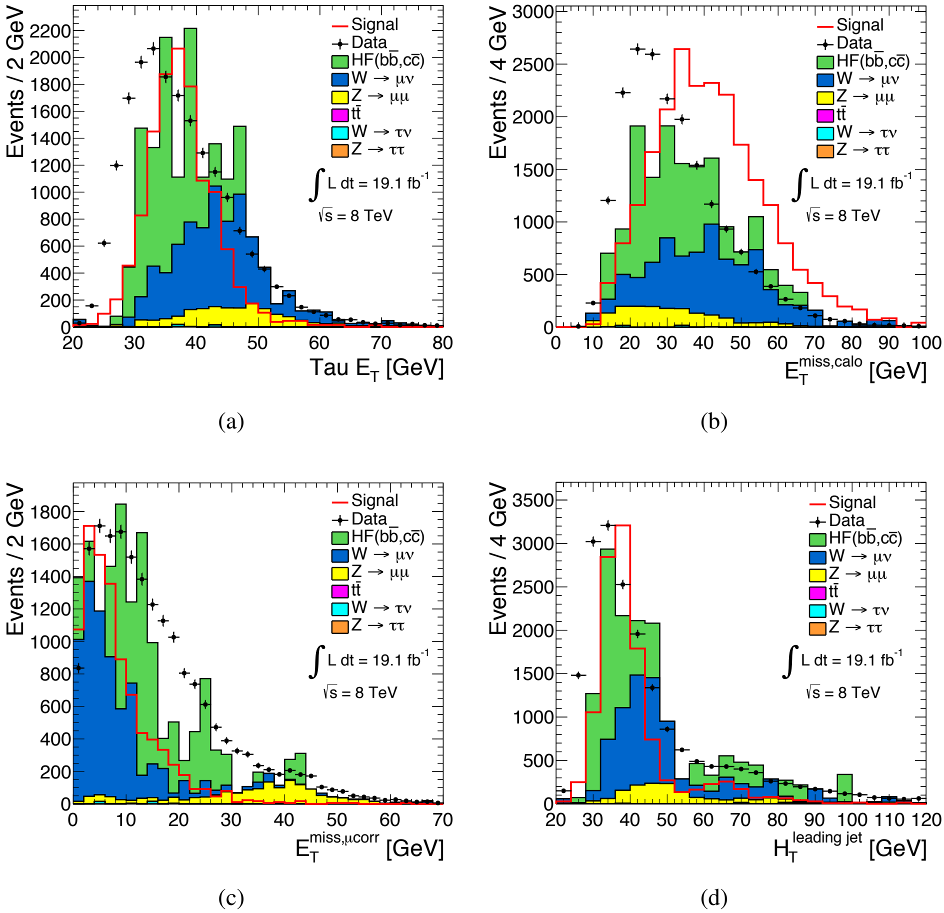
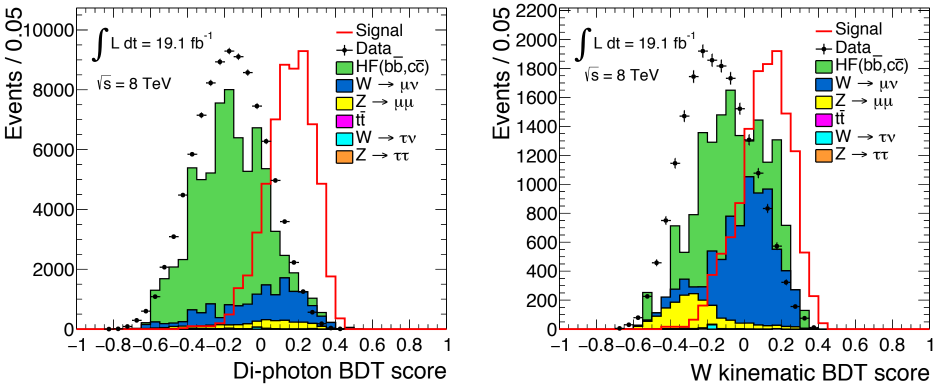
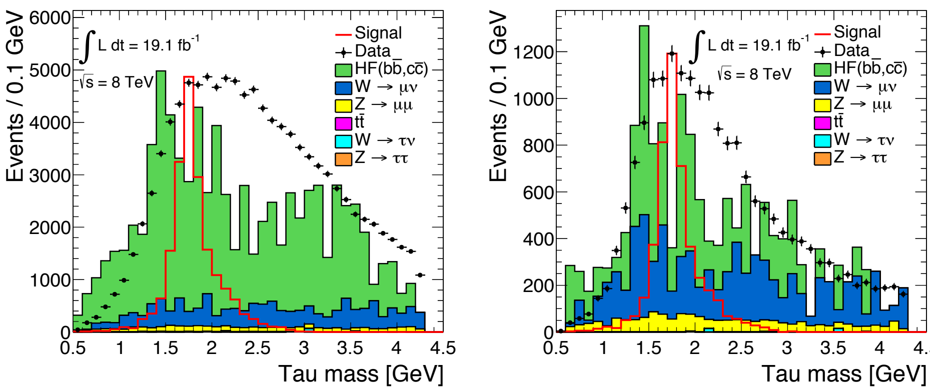


Fig. 4.13 · Distributions of Tau E_T (a), $E_T^{\text{miss,calo}}$ (b), $E_T^{\text{miss},\mu\text{corr}}$ (c) and $H_T^{\text{leading jet}}$ (d). For each variable, the distributions represent data (black dots), MC backgrounds (filled histograms) and signal (red line). For visual matters, the signal distribution is scaled to the data and the HF sample (green) is scaled down by 50%.



(a) (b)

Fig. 4.14 · Di-photon BDT score (a) and W kinematic BDT score (b) for data (black dots), MC backgrounds (filled histograms) and signal (red line). For visual matters, the signal distribution is scaled to the data and the HF sample (green) is scaled down by 50%.



(a) (b)

Fig. 4.15 · Mass distribution of the LFV tau candidate for data (black dots), MC backgrounds (filled histograms) and signal (red line) obtained after the following selection requirements: (a) muon $|d_0| < 0.5$ mm and (b) di-photon BDT score > 0. For visual matters, the signal distribution is scaled to the data and the HF sample (green) is scaled down by 50%.

A quantitative estimation of the MC backgrounds through the selection preceding the W kinematic BDT training is given in Table 4.4. The number of events of each MC background sample, obtained after the weighting to the data luminosity (19.1 fb^{-1}), is shown together with the number of signal and data events.

Cuts	Signal	Data	$W \to \mu\nu$	$W \to \tau\nu$	$Z \to \mu\mu$	$Z \to \tau\tau$	$t\bar{t}$	HF
Preselection	1909	138762	13232	185	3582	21	304	165334
$m_T(\tau, E_T^{\text{miss,calo}}) > 40$ GeV	1864	112680	12785	116	2841	13	286	137285
Muon $\lvert d_0 \rvert < 0.5$ mm	1862	107104	12277	116	2825	13	256	132766
Muon d_0 significance < 10	1838	95404	12242	116	2822	10	213	115778
Di-photon BDT score > 0	1611	19111	6662	29	1604	5	35	15507

Table 4.4 · Number of events passing the selection cuts for signal and data (whole samples) and MC backgrounds (weighted to the data luminosity). The *Preselection* includes all the cuts described in Section 4.3. The requirement on the di-photon BDT score strongly reduces the main backgrounds (HF and $W \to \mu\nu$ decays).

4.7 SR optimization and background estimation

The background estimation in the SR is performed using a *data-driven* procedure. The evaluation of the expected signal in the SR and the estimation of the number of tau leptons produced via the $W \to \tau\nu$ decay are also presented in this section. Only the test samples defined for the W kinematic BDT training are used, in order to avoid a bias due to the training. The signal (test) sample contains 535 events, once restricted to the SR. The SB data (test) sample consists of 4056 events.

After the training, the selection can be tightened to reduce the remaining background further. The requirements are applied on the scores of the two BDTs previously trained, testing different cuts on both scores. This creates a two-dimensional *grid of points*, which will be used in the optimization procedure to find the combination of cuts that minimizes the expected upper limit on the $\tau \to \mu\gamma\gamma$ branching fraction (see Section 4.9).

With regard to the di-photon BDT score, nine different cuts are tested in a range within 0[1] and 0.2, split in steps of 0.025. Tighter requirements are not applied, since another cut is placed on the W kinematic BDT score. Figure 4.16 shows the di-photon BDT score for signal events in the SR and data events in the SB at this stage of the selection.

At this point, the trained W kinematic BDT is applied to the two samples to evaluate the response. In Figure 4.17, the distribution of the W kinematic BDT score is shown for signal and data samples. Afterwards, a selection is also applied on this BDT score. Seven variations of the cut value are tested, in a range which varies from 0 to 0.3, split in steps of 0.05. In total, the grid is formed by 63 points

[1] This value corresponds to the cut already applied on the di-photon BDT score before the W kinematic BDT training.

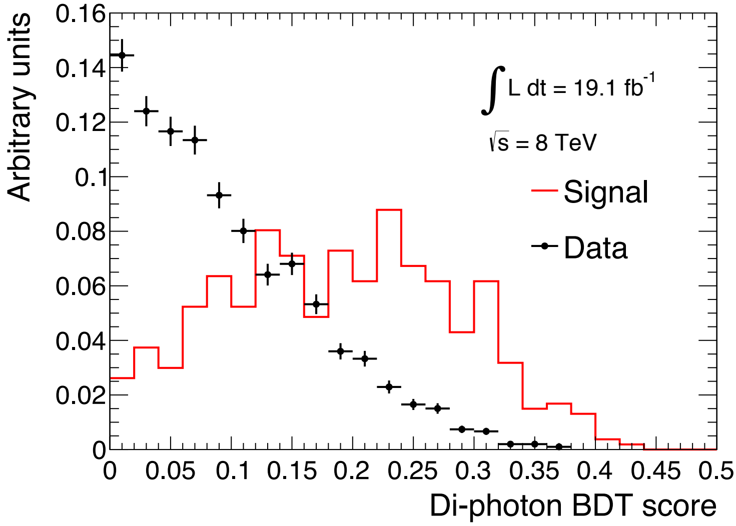


Fig. 4.16 · Di-photon BDT score for SR signal (red) and SB data (black dots) samples.

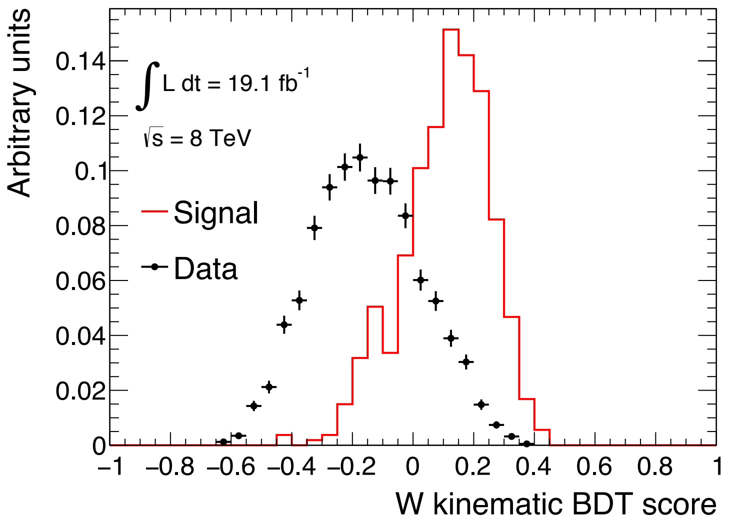


Fig. 4.17 · W kinematic BDT score for SR signal (red) and SB data (black dots) samples.

(given by the combination of the nine cuts on the first BDT score with the seven cuts on the second BDT score), which are used to estimate signal and background contributions.

4.7.1 Background estimation

The background contribution in the SR is estimated starting from the number of remaining data events in the SB test sample at the end of the selection ($N^{\text{test}}_{\text{bkg,SB}}$). The extrapolation to the SR is realized using the *transfer factors*, which are calculated from the data as the ratio of the number of events in a half of the SR sample[1] to the one in the SB test sample for each of the nine variations of the di-photon BDT score cut value.

One of these cuts, namely di-photon BDT score > 0.125, will be used for the evaluation of the systematic uncertainties (see Section 4.8). For this specific cut, the value of the corresponding transfer factor is 0.590 ± 0.027. The method of the transfer factors used for the extrapolation is reliable under two assumptions: the non-correlation between the W kinematic BDT score (on which a cut has to be applied) and the tau candidate mass, which will be shown in Section 4.10, and an analogous background composition in the SR and SB.

As already mentioned, for the last step of the selection, seven variations of the cut value on the W kinematic BDT score are considered. The cut on this variable will be placed at 0.1 to evaluate the systematic uncertainties.[2] The estimation of the number of background events in the SR is obtained by the following expression

$$N^{\text{exp}}_{\text{bkg,SR}} = N^{\text{test}}_{\text{bkg,SB}} \times 2 \times \text{TF}, \tag{4.2}$$

where the factor 2 is used to get an estimation on the whole sample and *TF* indicates the transfer factor. For the point of the grid identified by the cuts aforementioned (di-photon BDT score > 0.125 and W kinematic BDT score > 0.1), which will be labelled as *reference point* from now on, the estimation corresponds to 176 events. The uncertainty associated with this quantity will be discussed in the next section. A graphical representation of the background estimation in the SR for all the points of the grid is shown in Figure 4.18.

4.7.2 Signal efficiency

The amount of the expected signal in the SR is another fundamental quantity to be determined in this search. The signal efficiency, indeed, enters in the calculation of the upper limit on the branching fraction. The MC events of the signal test sample, restricted to the SR, are used for this estimation.

The signal contribution is calculated counting the events in the sample, after the selection is applied on the two BDT scores, as already done for the background. The signal estimation, for each point of the grid, is given by

$$N^{\text{exp}}_{\text{sig,SR}} = N^{\text{test}}_{\text{sig,SR}} \times 2, \tag{4.3}$$

[1] This region of the data sample is then blinded before the selection on the W kinematic BDT score.

[2] The choice of the two BDT score cuts used to estimate the systematic uncertainties is based on the minimization of the expected upper limit. From a first rough estimation, the best result is given by the combination of cuts mentioned in the text (reference point), which is later refined with the optimization procedure.

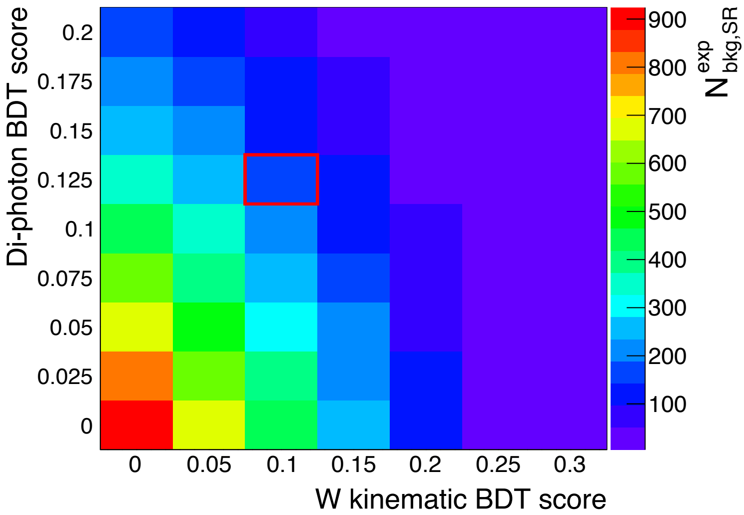


Fig. 4.18 · Expected number of background events in the SR as a function of the nine cuts applied on the di-photon BDT score and the seven cuts applied on the W kinematic BDT score. The reference point is indicated by the red frame.

where the remaining events in the signal test sample are multiplied by two, for an estimation on the whole (test + training) sample. In particular, for the reference point employed in the evaluation of the systematic uncertainties, 452 events are estimated in the SR.[1] Figure 4.19 shows the expected number of signal events in the SR for the complete grid of points.

The signal efficiency is calculated as the ratio of $N^{\text{exp}}_{\text{sig,SR}}$ to the initial number of MC signal events which amounts to 4×10^5, before the generator filter is applied. For the reference point, the signal efficiency corresponds to $\epsilon_{\text{sig}} = 0.11\%$.

4.7.3 Tau leptons production

To calculate the branching fraction of the LFV $\tau \to \mu\gamma\gamma$ channel, the number of tau leptons produced via the $W \to \tau\nu$ decay,[2] denoted as $N_{W\to\tau\nu}$, needs to be estimated. The branching fraction, indeed, is calculated as

$$\text{Br}(\tau \to \mu\gamma\gamma) = \frac{N_{\text{sig}}}{\epsilon_{\text{sig}} \times N_{W\to\tau\nu}}, \tag{4.4}$$

where N_{sig} is the number of signal events in the SR and ϵ_{sig} is the signal efficiency, which also includes the detector acceptance. The estimation of $N_{W\to\tau\nu}$

[1] As a reminder, the initial size of the signal sample corresponds to a $\text{Br}(\tau \to \mu\gamma\gamma)$ of 1.7×10^{-3}.

[2] Additional contributions from other sources of tau leptons are small and therefore not considered in the analysis.

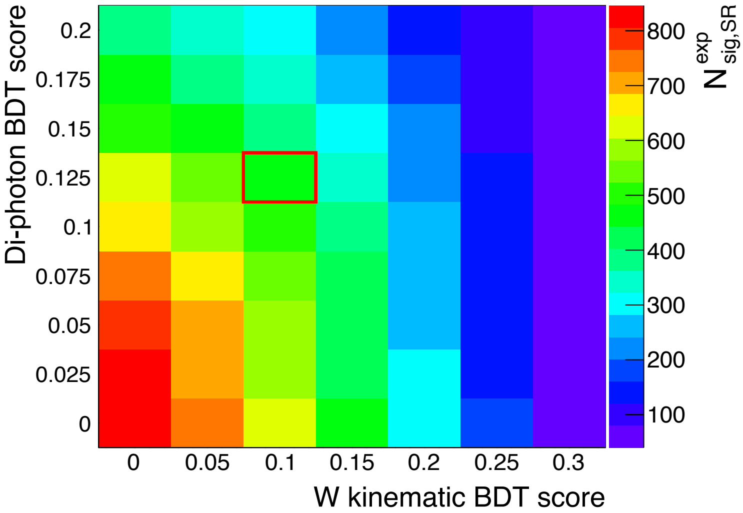


Fig. 4.19 · Expected number of signal events in the SR as a function of the nine cuts applied on the di-photon BDT score and the seven cuts applied on the W kinematic BDT score. The reference point is indicated by the red frame.

is obtained by multiplying the integrated luminosity and the $W \to l\nu$ cross-section[1] at $\sqrt{s} = 8$ TeV. The ATLAS collaboration performed an accurate measurement of the $W \to l\nu$ cross-section at $\sqrt{s} = 7$ TeV [147]:

$$\sigma^{7\,\mathrm{TeV}}(W \to l\nu) = 10.720 \pm 0.003 \pm 0.060 \pm 0.190 \pm 0.130 \text{ nb}, \tag{4.5}$$

where the uncertainties denote the statistical, the experimental systematic, the luminosity and acceptance extrapolation contributions, respectively. The measured cross-section is scaled to 8 TeV using the next-to-next-to-leading order (NNLO) calculations as follows:

$$\sigma^{8\,\mathrm{TeV}}(W \to l\nu) = \sigma^{7\,\mathrm{TeV}}(W \to l\nu)\frac{\sigma^{8\,\mathrm{TeV}}_{\mathrm{NNLO}}}{\sigma^{7\,\mathrm{TeV}}_{\mathrm{NNLO}}} = (12.48 \pm 0.17) \text{ nb}, \tag{4.6}$$

where $\sigma^{8\,\mathrm{TeV}}_{\mathrm{NNLO}} = (12.18 \pm 0.61)$ nb [63] and $\sigma^{7\,\mathrm{TeV}}_{\mathrm{NNLO}} = (10.46 \pm 0.52)$ nb [148]. In the calculation, the small uncertainty on the ratio of NNLO cross-sections is neglected.

Multiplying $\sigma^{8\,\mathrm{TeV}}(W \to l\nu)$ by the integrated luminosity which describes the amount of data for this analysis ($\mathcal{L} = 19.1$ fb^{-1}), the number of taus produced via the decay of the W boson in 2012 corresponds to $N_{W\to\tau\nu} = (2.38 \pm 0.07) \times 10^8$.

[1] The $W \to l\nu$ cross-section is determined by the combination of the measurements performed in the electron and muon channels, assuming lepton universality.

The luminosity uncertainty, taken into account in the calculation, is 2.8% for 2012 data-taking and is obtained following the same procedure described in [149] for 2010 and 2011 data-taking.

4.8 Systematic uncertainties

Several uncertainties are associated with the quantities estimated in the previous section. Here, an overview of the systematic uncertainties considered in the analysis is presented, with an evaluation of their impact.

Uncertainties on background estimation
The systematic uncertainty associated with the extrapolation procedure in the background estimation is derived from the variation of the SB range (defined in Section 4.4.1). Two alternative definitions are studied:

$$1.1 < m(\mu\gamma\gamma) < 1.4 \text{ GeV and } 2.2 < m(\mu\gamma\gamma) < 2.9 \text{ GeV (option A);}$$

$$1.2 < m(\mu\gamma\gamma) < 1.5 \text{ GeV and } 2.1 < m(\mu\gamma\gamma) < 2.7 \text{ GeV (option B).}$$

In the first case, labelled as *option A*, the range is modified to enlarge by factor of two the distance between the SB and the SR on both sides; in the second alternative, labelled as *option B*, left and right sidebands are narrower compared to the nominal definition (300 and 600 MeV versus 400 and 800 MeV, respectively).

Varying the SB range, the number of background events in this region at the end of the selection and the TF (Section 4.7.1) are different. As a consequence, from equation 4.2, a variation in the number of background events in the SR is expected. For the reference point of the grid, already introduced in Section 4.7.1, the following results are obtained:

$$N^{\text{exp}}_{\text{bkg,SR}} = 174 \text{ events (option A)}; \; N^{\text{exp}}_{\text{bkg,SR}} = 184 \text{ events (option B)}. \quad (4.7)$$

The second alternative (option B) presents the largest discrepancy from the nominal value (176 events) and is used to calculate the systematic uncertainty. This difference, normalized to the nominal $N^{\text{exp}}_{\text{bkg,SR}}$, gives an uncertainty of 4.7%.

Another source of uncertainty on the background estimation originates directly from the calculation in equation 4.2. The uncertainty is obtained from the propagation of the errors on TF and number of remaining events in the SB region ($N^{\text{test}}_{\text{bkg,SB}}$). For instance, selecting the reference point of the grid, an uncertainty of 9.5% is obtained

$$N^{\text{exp}}_{\text{bkg,SR}} = (176 \pm 17) \text{ events.} \quad (4.8)$$

The uncertainties associated with the background estimation are used in the HistFitter framework [140], for both the optimization procedure and the final statistical interpretation.

Uncertainties on signal efficiency

The signal efficiency is purely determined from the MC sample. A source of systematic uncertainty is associated with the accuracy of the energy measurements of the objects involved in the analysis. The uncertainty is evaluated shifting each component upward and downward. The shifts are applied separately on the energy deposits in the EM calorimeter, transverse momentum of the jets and $E_T^{\text{miss,calo}}$. Their effects are propagated throughout the selection.

- The first shift is applied on the transverse energy of each of the two photons, γ_H and γ_L, to take into account uncertainties associated with the data-to-MC efficiency ratios and the dedicated correction for photons in proximity of a muon which is used in this analysis, as explained in Section 3.8.1. A conservative 5% shift is applied on the transverse energy of the photons. In addition, the transverse energy in the EM calorimeter measured at the EM scale, which is one of the di-photon BDT input variables (E_T^{EM}), is shifted by 5%.

- The second variation concerns the transverse momentum of the jets, implying an impact on the two H_T variables used as input in the W kinematic BDT training. Several sources contribute to the jet energy scale (JES) uncertainty, as already mentioned in Section 3.5. Since the systematic uncertainties are larger for forward jets than for central jets, a 4% shift is applied on the p_T of jets with $|\eta| \leq 2.5$ and a 7% shift is applied on the p_T of jets with $|\eta| > 2.5$.

- The last shift affects the $E_T^{\text{miss,calo}}$ reconstruction. The systematic uncertainty on this quantity is given by the uncertainties on the individual terms (*soft* and high p_T objects), as described in Section 3.7. A shift of 10% is applied on $E_T^{\text{miss,calo}}$, implying an effect on the W kinematic BDT input variables which involve this quantity.

Muon reconstruction efficiencies in data and MC simulation are in good agreement (see Section 3.2) and the small systematic uncertainties associated with these objects [112] are not considered here. The uncertainty on the second E_T^{miss} definition involved in the analysis, $E_T^{\text{miss},\mu\text{corr}}$, is assumed to be small because it is tracks dominated and therefore is neglected.

The effect of the three shifts is propagated until the end of the selection. A comparison in the number of the remaining events in the signal sample, obtained using the reference point of the grid, is shown in Table 4.5. For this study, all the estimations are obtained directly using the whole signal sample (no split in test and training subsamples).

Each of the three shifts is used to estimate the individual contribution to the total systematic uncertainty, selecting the variation (up or down) which gives the larger discrepancy from the nominal value. The shift applied on the energy deposits in the EM calorimeter results in an uncertainty of 5.4%, the variation concerning the

	Nominal	EM Calo up	EM Calo down	Jet up	Jet down	$E_T^{\text{miss,calo}}$ up	$E_T^{\text{miss,calo}}$ down
$N_{\text{sig,SR}}^{\text{exp}}$	444	434	420	440	443	468	396

Table 4.5 · Number of signal events in the SR at the end of the selection, obtained using the reference point of the grid (di-photon BDT score > 0.125 and W kinematic BDT score > 0.1). The nominal prediction is compared with the results obtained from the up and down shifts applied on the different quantities.

jets produces an uncertainty of 0.9%, while the major contribution is given by the $E_T^{\text{miss,calo}}$ shift with an uncertainty of 10.8%. Finally, these three terms are summed in quadrature to obtain the total energy scale systematic uncertainty for the signal, which corresponds to 12.1%.

Another source of uncertainty on the signal efficiency originates from the trigger efficiency, which has been presented in Section 4.3. The trigger uncertainty is estimated by a weighted average of the uncertainties on the several components of the trigger configuration. A large fraction (54.8%) of the signal events which pass the trigger requirement is selected by at least one of the *TauMass* triggers. The contribution of the EF_tau20_medium1_mu15 trigger (excluding events already selected by the *TauMass* triggers) corresponds to 37.4%. The two *single muon* triggers cover the residual fraction of events (7.8%).

Efficiency measurements for the same set of *TauMass* triggers which is employed in this analysis have been performed in previous studies [150]. The uncertainty, which is evaluated from the disagreement between the efficiencies measured in data and in the MC simulation, is estimated to be 3.5%.

A good agreement is observed between data and MC tau trigger efficiency measurements. The uncertainties on the correction factors, defined to take into account the small difference, are on the order of a few percent [135]. In the LFV $\tau \to \mu\gamma\gamma$ analysis, a 5% uncertainty is associated with the *tau+muon* trigger (EF_tau20_medium1_mu15). The uncertainty contribution related to the *single muon* triggers is neglected, due to the small systematic uncertainties on their efficiency measurements [104] and the little fraction of signal events selected by these triggers in the search.

The trigger systematic uncertainty, which is derived from the weighted average of the uncertainties arising from the set of *TauMass* triggers and the *tau+muon* trigger, corresponds to 4.1%. Finally, the total systematic uncertainty on the signal efficiency is calculated summing in quadrature the uncertainties on the energy scale (12.1%) and the trigger efficiency (4.1%). This leads to a systematic uncertainty of 12.8%.

A MC statistical uncertainty on the number of expected signal events in the SR is also estimated and added separately in the HistFitter configuration. For the reference point of the grid, this uncertainty corresponds to 6.7%.

Other uncertainty sources

The overview of the systematic uncertainties used in the analysis is completed by the uncertainties on the integrated luminosity and the $W \to l\nu$ cross-section at $\sqrt{s} = 8$ TeV (see Section 4.7.3). These two uncertainties contribute to the systematic uncertainty on the estimation of $N_{W\to\tau\nu}$, described in the same section, which corresponds to 3.0%.

In the final upper limit calculation, the systematic uncertainty on $N_{W\to\tau\nu}$ will be summed in quadrature with the uncertainty on the signal efficiency. The resulting value will be used in the HistFitter configuration as the total systematic uncertainty on the signal component.

4.9 Statistical analysis

After the evaluation of the systematic uncertainties, an optimization procedure is used to minimize the expected upper limit on the $\tau \to \mu\gamma\gamma$ branching fraction. All the combinations of cuts on the two BDT scores, represented by the points of the grid introduced in Section 4.7, are investigated to find the one which gives the lowest expected upper limit.

The following step is represented by the final determination of the upper limit and the statistical interpretation of the results (Section 4.10). The expected upper limit on the number of signal events in the SR and on the resulting branching fraction is recalculated relying on the results of the optimization process. Once the SR is unblinded, in absence of a signal excess, the observed upper limit is calculated using directly the number of data events in this region at the end of the selection. The HistFitter framework [140] is used both for optimization and final statistical interpretation. The limit setting is based on the CL$_s$ method [151], an approach commonly used by the LHC experiments.

4.9.1 Statistical procedure

To claim a discovery or set an exclusion limit, a statistical analysis is performed. The starting point is a likelihood function which expresses the probability of the observed data under a specific hypothesis. In its general form, can be expressed as

$$L(\boldsymbol{n}, \boldsymbol{\theta}^0 | \mu, \boldsymbol{b}, \boldsymbol{\theta}) = \prod_i P(n_i | \lambda_i(\mu, \boldsymbol{b}, \boldsymbol{\theta})) \times C_{\text{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta}). \tag{4.9}$$

The likelihood function depends on the number of observed events ($\boldsymbol{n}$) in all regions, indicated by the subscript i, predictions for background processes ($\boldsymbol{b}$), *nuisance parameters* ($\boldsymbol{\theta}$) which parametrize the systematic uncertainties around the central values ($\boldsymbol{\theta}^0$) and *signal strength* (μ). The Poisson distributions (P) indicate the probability to observe n_i events in region i for λ_i expected events. In the LFV $\tau \to \mu\gamma\gamma$ analysis, only the signal region (SR) is included in the likelihood function and therefore, there is only one Poisson distribution. Finally, $C_{\text{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta})$ is typically a product of Gaussian distributions defined to take into account each of the systematic uncertainties.

Two hypotheses are defined to be tested. In case of a discovery, the null hypothesis H_b, defined as background-only hypothesis (signal strength $\mu = 0$), is tested against the alternative hypothesis H_{s+b}, which includes both background and the sought after signal ($\mu \neq 0$). When setting the limits, H_{s+b} plays the role of the null hypothesis, tested against the alternative H_b.

To verify the compatibility of the observed data with the two hypotheses, a test statistic is constructed from a profile likelihood ratio, defined as

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}}(\mu))} & \text{if } \hat{\mu} \geq 0, \\ \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\boldsymbol{\theta}}(0))} & \text{if } \hat{\mu} < 0, \end{cases} \tag{4.10}$$

where single-hatted quantities maximize the likelihood function and double-hatted quantities maximize the likelihood for the specific value of μ. As shown in equation 4.10, $\hat{\mu}$ is set to 0 when it is found to be negative, in order to exclude unphysical results (negative signal rate). Then, the two-sided test statistic is defined as $\tilde{t}_\mu = -2\ln\tilde{\lambda}(\mu)$. Higher values of $\tilde{t}_\mu$ denote a larger incompatibility between the observed data and the tested hypothesis.

A test statistic can be used to establish an upper limit on the signal strength, testing the signal-plus-background hypothesis. The definition is given by

$$\tilde{q_\mu} = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \text{if } \hat{\mu} \leq \mu, \\ 0 & \text{if } \hat{\mu} > \mu, \end{cases} \tag{4.11}$$

where the value of $\tilde{q_\mu}$ is set to 0 for $\hat{\mu} > \mu$ (one-sided test statistic), not considering upward fluctuations of the data as representing incompatibility with the hypothesized value of μ. Another test statistic can be defined for the discovery of a new signal, testing this time the background-only hypothesis, as explained in [152].

Setting an upper limit, the level of agreement between the data and the tested hypothesis is quantified by the following p-value:

$$p_\mu = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu)\, d\tilde{q}_\mu, \tag{4.12}$$

where $\tilde{q}_\mu^{\text{obs}}$ is the value of the test statistic observed from the data and $f(\tilde{q}_\mu|\mu)$ is the probability density function of $\tilde{q}_\mu$ assuming a signal strength μ. The distribution $f(\tilde{q}_\mu|\mu)$ can be obtained using approximate formulae as described in [152].

An alternative definition of the p-value is used to prevent any effects of downward fluctuation of the observed data with respect to the background expectation, which could result in a too strong limit, excluding hypotheses to which one has little sensitivity. This modified approach uses the CL_s procedure [151], defining the quantity

$$\text{CL}_s = \frac{p_\mu}{1 - p_b}, \tag{4.13}$$

where the p-value of the signal-plus-background hypothesis (p_μ) is normalized to $1 - p_b$, with p_b indicating the p-value of the background-only hypothesis. In this analysis, the upper limit on the signal strength is set at 90% confidence level. A tested hypothesis is excluded at 90% confidence level if $\mathrm{CL}_s < 0.1$ and the upper limit is determined as the largest value of μ which is not excluded.

4.9.2 Optimization

The optimization process aims to minimize the expected upper limit on the $\tau \to \mu\gamma\gamma$ branching fraction through an iterative procedure which involves all the combinations of cuts on the two BDT scores, identified by the points of the grid. In this phase, the SR of the data sample is blinded. For each combination, different values are obtained for the expected background and signal efficiency, following the strategy described in Section 4.7.

The optimization is performed using the HistFitter framework. The expected upper limit is calculated varying the signal strength in order to find the value of this parameter for which the median CL_s is equal to 0.1 (as this gives the median upper limit at 90% confidence level). The limit is calculated by HistFitter as a *counting experiment*, with the SR constructed as a single-bin region.

The number of events involved in the limit calculation is sufficiently large to allow the approximation of the probability density function of the test statistic with asymptotic formulae. This approach is based on a single artificial data set, called *Asimov* data set, in which all statistical fluctuations are suppressed and physical quantities are replaced by their expectation values. A comprehensive explanation of this method is given in [152].

Several inputs are provided to the HistFitter configuration. The expected number of background events in the SR ($N^{\mathrm{exp}}_{\mathrm{bkg,SR}}$) and the signal prediction in the SR ($N^{\mathrm{exp}}_{\mathrm{sig,SR}}$) vary in each iteration as well as their statistical uncertainties. On the other hand, background and signal systematic uncertainties are estimated using the reference point of the grid, as described in Section 4.8.

HistFitter calculates the expected (median) limit and the $\pm 1\,\sigma$ and $\pm 2\,\sigma$ variations, which take into account the statistical fluctuations. A quantitative estimation of these variations, represented by error bands, will be given in the final upper limit calculation described in the next section. The resulting upper limit on the signal strength is multiplied by the quantity $N^{\mathrm{exp}}_{\mathrm{sig,SR}}$ to obtain the upper limit on the number of signal events in the SR. Finally, the expected limit on the $\tau \to \mu\gamma\gamma$ branching fraction is calculated with the equation 4.4, using the upper limit on the number of signal events in the numerator. Figure 4.20 shows the 90% confidence level expected limits on the branching fraction obtained for each iteration.

At the end of the optimization procedure, the minimal branching fraction is obtained for the following combination of cuts: di-photon BDT score > 0.075 and W kinematic BDT score > 0.1. The corresponding point of the grid is labelled as *optimal point*. In this case, 278 background events are estimated in the SR and the signal efficiency is 0.14% (with $N^{\mathrm{exp}}_{\mathrm{sig,SR}}$ equal to 540 events). The expected

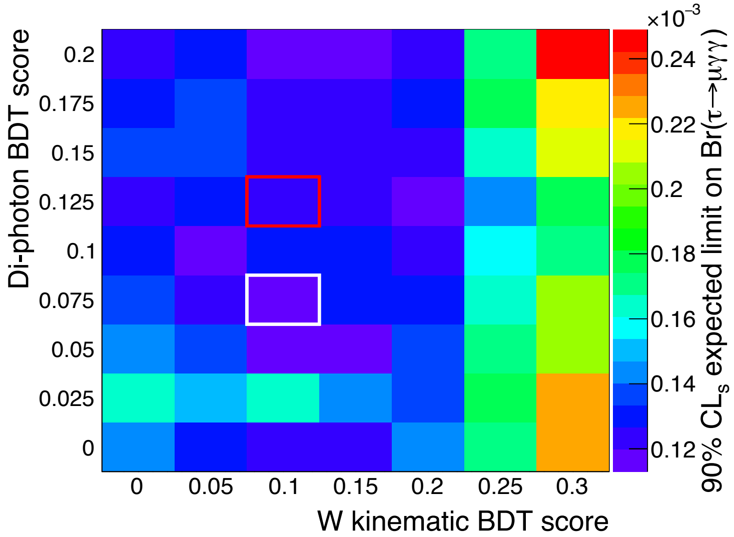


Fig. 4.20 · Expected exclusion limits on the $\tau \to \mu\gamma\gamma$ branching fraction at 90% confidence level. The limit is evaluated for all the combinations of cuts on the two BDT scores included in the grid. The reference point is indicated by the red frame. The lowest expected limit is obtained for the combination of cuts di-photon BDT score > 0.075 and W kinematic BDT score > 0.1, indicated by the white frame and defined as optimal point in the text.

upper limit on the number of signal events in the SR, denoted by $N_{\text{sig}}^{\text{UL}}$, is 36.4 and the corresponding expected upper limit on the branching fraction is set at $\text{Br}(\tau \to \mu\gamma\gamma) = 1.13 \times 10^{-4}$. The final results of the LFV $\tau \to \mu\gamma\gamma$ search will be presented in the next section.

4.10 Results

After the optimal combination of cuts on the BDT scores is selected, the conclusive step of the LFV $\tau \to \mu\gamma\gamma$ analysis is based on the final determination of the expected upper limit on the branching fraction and on the search for the $\tau \to \mu\gamma\gamma$ signal once the SR is unblinded. In absence of a signal excess, the observed upper limit is calculated. As already done in the optimization procedure, the HistFitter framework calculates the upper limit on the signal strength as a one-bin counting experiment, using the asymptotic approximation.

The signal and background systematic uncertainties are recalculated using the optimal point. The resulting contributions are consistent with the ones obtained for the reference point (Section 4.8), with a small reduction of the systematics on the background estimation. Preferring a conservative approach, also the final upper limit determination is performed using the systematic uncertainties calculated for the reference point.

All the systematic and statistical uncertainties considered in the analysis are directly taken into account in the HistFitter configuration. Unlike the optimization, the systematic uncertainty on the number of tau leptons produced during the data-taking (3.0%) is included to be summed in quadrature with the uncertainty on the signal efficiency (12.8%). Therefore, the total systematic uncertainty on the signal component corresponds to 13.1%.

The expected (median) limit is calculated by HistFitter at 90% confidence level, using the CL$_s$ method. Multiplying the expected upper limit on the signal strength by the predicted number of signal events in the SR (540 for the optimal combination of cuts), we obtain

$$N_{\text{sig}}^{\text{UL}} = 36.6^{+18.2}_{-11.4}\,, \tag{4.14}$$

where the $\pm 1\sigma$ variations of the median upper limit are also indicated. Then, the signal efficiency and the $N_{W\to\tau\nu}$ estimation are used to compute the branching fraction. The expected upper limit at 90% CL$_s$ is

$$\text{Br}(\tau \to \mu\gamma\gamma) = (1.14^{+0.57}_{-0.35}) \times 10^{-4}\,. \tag{4.15}$$

Observed limits

At this stage, the region of the data sample which includes the SR can be unblinded. Applying the combination of cuts selected by the optimization procedure (di-photon BDT score > 0.075 and W kinematic BDT score > 0.1), 293 data events are observed in the SR at the end of the selection. This value is compatible with the result of the background estimation (278 expected events), which is obtained using the procedure described in Section 4.7. Therefore, we conclude that we observe no excess with respect to the Standard Model.

The observed number of events in the SR is plugged in the HistFitter configuration to calculate the 90% CL$_s$ upper limit on the signal strength, which is estimated to be $\mu^{\text{UL}} = 0.089$. The corresponding upper limit on the number of signal events in the SR is $N_{\text{sig}}^{\text{UL}} = 48.0$.

Figure 4.21 shows the results obtained for the expected and observed limits, where the CL$_s$ values are given as a function of the signal strength. The upper limit at 90% confidence level is determined by the value of the signal strength for which CL$_s = 0.1$.

The observed upper limit on the $\tau \to \mu\gamma\gamma$ branching fraction is calculated from equation 4.4, using in the numerator the value of $N_{\text{sig}}^{\text{UL}}$ previously obtained. The 90% CL$_s$ upper limit corresponds to

$$\text{Br}(\tau \to \mu\gamma\gamma) = 1.49 \times 10^{-4}\,. \tag{4.16}$$

This result is in good agreement with the expected value aforementioned, $\text{Br}(\tau \to \mu\gamma\gamma) = (1.14^{+0.57}_{-0.35}) \times 10^{-4}$. The observed exclusion limit is within 1σ deviation of the expected upper limit.

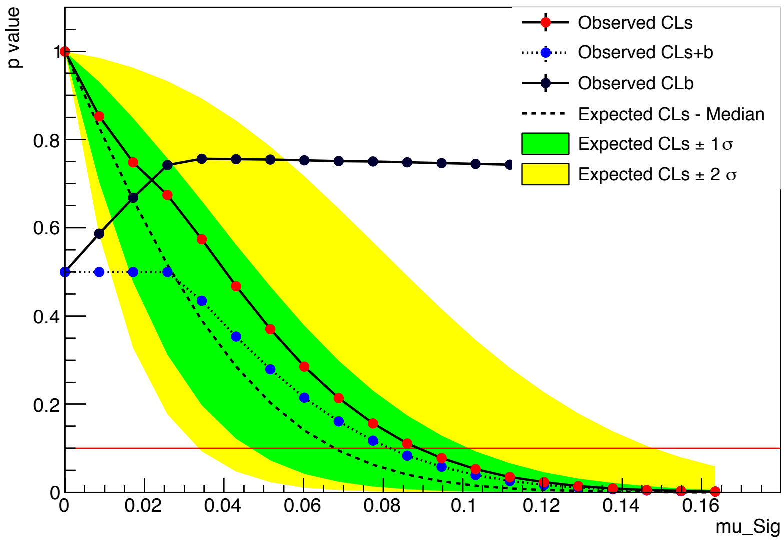


Fig. 4.21 · CL_s p-values as a function of the signal strength. The red markers of the black solid curve represent the observed CL_s. The dashed black curve indicates the expected (median) CL_s with the corresponding 1σ (green) and 2σ (yellow) bands. The observed and expected upper limits are determined at 90% confidence level by the intersection of these curves and the horizontal red line. The CL_{s+b} and CL_b p-values are also shown.

Finally, since the SR of the data sample is now unblinded, the non-correlation between the W kinematic BDT score and the LFV tau candidate mass can be verified. As already mentioned in Section 4.7.1, this condition is necessary for a reliable background estimation. The cut on the W kinematic BDT score has been set at 0.1 in the optimization procedure. Here, three regions of this distribution are considered for the data sample:

- Region 1: $-0.2 \leq W$ kinematic BDT score ≤ -0.1;
- Region 2: $-0.1 < W$ kimematic BDT score ≤ 0;
- Region 3: $0 < W$ kinematic BDT score ≤ 0.2.

Then, the tau mass distributions obtained for these three regions are divided by each other. The resulting distributions, shown in Figure 4.22, present a flat behaviour throughout the range that includes the SB and the SR. The tau candidate mass is observed to be independent of the cut applied on the BDT score. This confirms the validity of the background estimation realized using the transfer factors, calculated before the selection on the W kinematic BDT score (see Section 4.7.1), for the extrapolation to the SR.

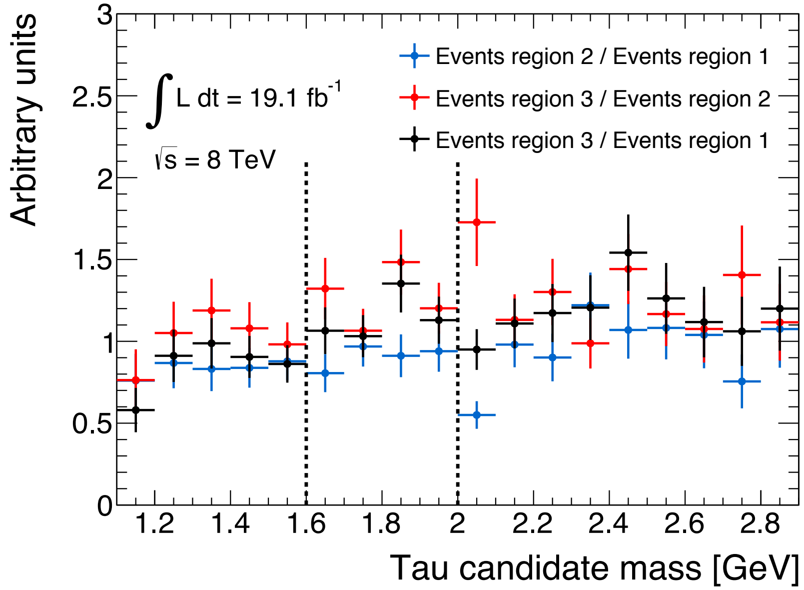


Fig. 4.22 · Ratios between the tau mass distributions of the data sample obtained for the three W kinematic BDT score regions described in the text. The three resulting distributions are given by the ratios of the tau masses obtained for region 2 and region 1 (light blue), region 3 and region 2 (red), region 3 and region 1 (black). The tau mass range includes both the SR, indicated by the two dashed lines (1.6 - 2.0 GeV), and the SB (1.1 - 1.5 GeV and 2.1 - 2.9 GeV).

Conclusions and outlook

Lepton flavour violation in the charged lepton sector is a fascinating topic in high energy physics. In particular, tau lepton decays represent one of the most sensitive probes, according to several scenarios of physics beyond the Standard Model.

In this thesis, a search for the LFV $\tau \to \mu\gamma\gamma$ decay has been presented. The analysis is based on a data sample corresponding to an integrated luminosity of 19.1 fb^{-1} produced in 2012 p-p collisions, collected with the ATLAS detector at $\sqrt{s} = 8$ TeV. The measurement broadens the picture of the LFV searches in tau lepton decays since it represents the first experimental search performed in this channel. The analysis strategy has been illustrated in Section 3.8 and throughout Chapter 4. No excess with respect to the Standard Model expectation is observed. Therefore, exclusion limits are set on the $\tau \to \mu\gamma\gamma$ branching fraction. The limits are calculated by the HistFitter framework using the CL$_s$ method. The observed upper limit is found to be Br$(\tau \to \mu\gamma\gamma) = 1.49 \times 10^{-4}$ at 90% confidence level. The result obtained is in good agreement with the expected limit, which corresponds to Br$(\tau \to \mu\gamma\gamma) = (1.14^{+0.57}_{-0.35}) \times 10^{-4}$.

The channel studied in this search is complementary to other LFV tau decays, such as $\tau \to \mu\gamma$ and $\tau \to 3\mu$. The current limits on these channels have been set by the B factory experiments on the order of 10^{-8} [45, 61]. The ATLAS collaboration set a 90% confidence level upper limit on the $\tau \to 3\mu$ branching fraction corresponding to 3.76×10^{-7} [63].

In 2015, following the first long shutdown, LHC started its Run-2 with an increased centre-of-mass energy of $\sqrt{s} = 13$ TeV. At the end of the second run, ATLAS is expected to collect an integrated luminosity of about 100 fb^{-1}. The large increase of data available would allow to improve the results presented in this thesis. Moreover, a measurement of the $W \to l\nu$ cross-section at $\sqrt{s} = 13$ TeV [153] reported an increment of about 65% with respect to the cross-section estimated for $\sqrt{s} = 8$ TeV used in the analysis (see Section 4.7.3). Due to the substantial background processes affecting the $\tau \to \mu\gamma\gamma$ analysis, mainly heavy-flavour quark decays and $W \to \mu\nu$ decays, the improvement of the upper limit is expected to be proportional to $1/\sqrt{N_{W\to\tau\nu}}$ (where $N_{W\to\tau\nu}$ is the number of taus produced via the $W \to \tau\nu$ decay). Collecting 100 fb^{-1} and considering the increase of cross-section, the upper limit on the $\tau \to \mu\gamma\gamma$ branching fraction should correspond to approximately 5×10^{-5}. In 2023, at the end of the LHC Run-3, the amount of integrated luminosity should reach 300 fb^{-1}, leading to a further improvement.

The progress of the LHC is not the only way to improve the results obtained in this thesis. In the analysis, the $\tau \to \mu\gamma\gamma$ signal sample is characterized by two photons, close to each other, reconstructed in two separated clusters. Another signal sample of the same size has been generated though not employed in the Run-1 search presented here. In this case, the energy deposits of the two photons merge into a single cluster and are reconstructed as only one photon. By construction, the two samples are mutually exclusive. For the sample with merged photons, the trigger efficiency is similar to the one for separate photons, again around 20%, by applying the same trigger selection described in the thesis. The reconstructed LFV tau candidate mass, given by the invariant mass of the tau decay products, is marked by a worse resolution than in the case of two separated clusters. An improvement would then be required in order to define an appropriate signal region for this sample. Radiative W and Z decays represent the main backgrounds for this signal.

A future search for the LFV $\tau \to \mu\gamma\gamma$ channel could be performed using both signal samples following similar strategies. If no significant excess is observed, a combination of results of the two analyses could further improve the current upper limit.

The $\tau \to \mu\gamma\gamma$ analysis presented in the thesis uses tau leptons produced from the decay of W bosons. Additionally, $Z \to \tau\tau$ decays can represent another source of taus to employ in the search. These processes have a cross-section limited to about 10% of the $W \to \tau\nu$ cross-section, but benefit from two taus in the final state.

The quest for lepton flavour violation in tau decays and in the whole charged lepton sector will continue in the next years. A discovery in this field would be another milestone in the history of particle physics.

Bibliography

[1] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. **B122** (1983) 103–116. (Cited on page 3.)

[2] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **122B** (1983) 476–485. (Cited on page 3.)

[3] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around* $95\,GeV/c^2$ *at the CERN SPS Collider*, Phys. Lett. **B126** (1983) 398–410. (Cited on page 3.)

[4] UA2 Collaboration, *Evidence for* $Z0 \rightarrow e^+e^-$ *at the CERN anti-p p Collider*, Phys. Lett. **B129** (1983) 130–140. (Cited on page 3.)

[5] CDF Collaboration, *Observation of top quark production in* $\bar{p}p$ *collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv:hep-ex/9503002 [hep-ex]. (Cited on page 3.)

[6] D0 Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, arXiv:hep-ex/9503003 [hep-ex]. (Cited on page 3.)

[7] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. (Cited on pages 3, 23, and 26.)

[8] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004. (Cited on pages 3 and 23.)

[9] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex]. (Cited on page 3.)

[10] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex]. (Cited on page 3.)

[11] S. L. Glashow, *Partial symmetries of weak interactions*, Nuclear Physics **22** no. 4, (1961) 579 – 588. (Cited on page 3.)

[12] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266. http://link.aps.org/doi/10.1103/PhysRevLett.19.1264. (Cited on page 3.)

[13] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377. (Cited on page 3.)

[14] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509. http://link.aps.org/doi/10.1103/PhysRevLett.13.508. (Cited on pages 3 and 4.)

[15] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323. http://link.aps.org/doi/10.1103/PhysRevLett.13.321. (Cited on page 3.)

[16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587. http://link.aps.org/doi/10.1103/PhysRevLett.13.585. (Cited on page 3.)

[17] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. **B44** (1972) 189–213. (Cited on pages 3 and 5.)

[18] Planck Collaboration, *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.01589 [astro-ph.CO]. (Cited on pages 3 and 11.)

[19] Super-Kamiokande Collaboration, *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567. http://link.aps.org/doi/10.1103/PhysRevLett.81.1562. (Cited on pages 3 and 10.)

[20] SNO Collaboration, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301. http://link.aps.org/doi/10.1103/PhysRevLett.89.011301. (Cited on pages 3 and 10.)

[21] X.-Y. Pham, *Lepton flavor changing in neutrinoless tau decays*, Eur. Phys. J. **C8** (1999) 513–516, arXiv:hep-ph/9810484 [hep-ph]. (Cited on page 3.)

[22] W. Marciano and A. Sanda, *Exotic decays of the muon and heavy leptons in gauge theories*, Physics Letters B **67** no. 3, (1977) 303 – 305. http://www.sciencedirect.com/science/article/pii/037026937790377X. (Cited on pages 3 and 13.)

[23] M. Raidal et al., *Flavour physics of leptons and dipole moments*, Eur. Phys. J. **C57** (2008) 13–182, arXiv:0801.1826 [hep-ph]. (Cited on pages 3 and 13.)

[24] G. Cvetic, C. Dib, C. S. Kim, and J. D. Kim, *On lepton flavor violation in tau decays*, Phys. Rev. **D66** (2002) 034008, arXiv:hep-ph/0202212 [hep-ph]. [Erratum: Phys. Rev.D68,059901(2003)]. (Cited on pages 3, 14, and 15.)

[25] G. J. Besjes, *Pushing Susy's boundaries*. PhD thesis, Radboud University Nijmegen, 2015. (Cited on page 5.)

[26] F. Halzen and A. Martin, *Quarks and leptons: an introductory course in modern particle physics*. John Wiley and sons, 1984. (Cited on pages 6 and 10.)

[27] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory*. Westview Press, 1995. (Cited on page 6.)

[28] SNO Collaboration, *Measurement of the total active B-8 solar neutrino flux at the Sudbury Neutrino Observatory with enhanced neutral current sensitivity*, Phys. Rev. Lett. **92** (2004) 181301, arXiv:nucl-ex/0309004 [nucl-ex]. (Cited on page 10.)

[29] KamLAND Collaboration, *Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion*, Phys. Rev. Lett. **94** (2005) 081801, arXiv:hep-ex/0406035 [hep-ex]. (Cited on page 10.)

[30] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, arXiv:hep-ex/9807003 [hep-ex]. (Cited on page 10.)

[31] Super-Kamiokande Collaboration, *Evidence for an oscillatory signature in atmospheric neutrino oscillation*, Phys. Rev. Lett. **93** (2004) 101801, arXiv:hep-ex/0404034 [hep-ex]. (Cited on page 10.)

[32] K2K Collaboration, *Measurement of Neutrino Oscillation by the K2K Experiment*, Phys. Rev. **D74** (2006) 072003, arXiv:hep-ex/0606032 [hep-ex]. (Cited on page 10.)

[33] T2K Collaboration, *Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam*, Phys. Rev. Lett. **107** (2011) 041801, arXiv:1106.2822 [hep-ex]. (Cited on page 10.)

[34] MINOS Collaboration, *Improved search for muon-neutrino to electron-neutrino oscillations in MINOS*, Phys. Rev. Lett. **107** (2011) 181802, arXiv:1108.0015 [hep-ex]. (Cited on page 10.)

[35] Super-Kamiokande Collaboration, *Evidence for the Appearance of Atmospheric Tau Neutrinos in Super-Kamiokande*, Phys. Rev. Lett. **110** (2013) 181802. http://link.aps.org/doi/10.1103/PhysRevLett.110.181802. (Cited on page 10.)

[36] OPERA Collaboration, *Discovery of τ Neutrino Appearance in the CNGS Neutrino Beam with the OPERA Experiment*, Phys. Rev. Lett. **115** no. 12, (2015) 121802, arXiv:1507.01417 [hep-ex]. (Cited on page 10.)

[37] Troitsk Collaboration, *An upper limit on electron antineutrino mass from Troitsk experiment*, Phys. Rev. **D84** (2011) 112003, arXiv:1108.5034 [hep-ex]. (Cited on page 10.)

[38] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657. (Cited on page 11.)

[39] M. C. Gonzalez-Garcia, M. Maltoni, and T. Schwetz, *Updated fit to three neutrino mixing: status of leptonic CP violation*, JHEP **11** (2014) 052, arXiv:1409.5439 [hep-ph]. (Cited on page 12.)

[40] P. Minkowski, *$\mu \to e\gamma$ at a rate of one out of 1-billion muon decays?*, Physics Letters B **67** no. 4, (1977) 421 – 428. http://www.sciencedirect.com/science/article/pii/037026937790435X. (Cited on pages 13 and 14.)

[41] M. Gell-Mann, P. Ramond, and R. Slansky, *Complex Spinors and Unified Theories*, Conf. Proc. **C790927** (1979) 315–321, arXiv:1306.4669 [hep-th]. (Cited on page 13.)

[42] R. N. Mohapatra and G. Senjanović, *Neutrino masses and mixings in gauge models with spontaneous parity violation*, Phys. Rev. D **23** (1981) 165–180. http://link.aps.org/doi/10.1103/PhysRevD.23.165. (Cited on page 13.)

[43] J. Schechter and J. W. F. Valle, *Neutrino masses in SU(2) $\bigotimes$ U(1) theories*, Phys. Rev. D **22** (1980) 2227–2235. http://link.aps.org/doi/10.1103/PhysRevD.22.2227. (Cited on page 13.)

[44] R. Foot, H. Lew, X. G. He, and G. C. Joshi, *See-saw neutrino masses induced by a triplet of leptons*, Z. Phys. C**44** no. 3, 441–444. http://dx.doi.org/10.1007/BF01415558. (Cited on page 13.)

[45] BaBar Collaboration, *Searches for Lepton Flavor Violation in the Decays $\tau^\pm \to e^\pm\gamma$ and $\tau^\pm \to \mu^\pm\gamma$*, Phys. Rev. Lett. **104** (2010) 021802, arXiv:0908.2381 [hep-ex]. (Cited on pages 15, 18, 22, and 103.)

[46] Belle II Collaboration, *Belle II Technical Design Report*, Tech. Rep. arXiv:1011.0352. KEK REPORT 2010-1, Nov, 2010. http://cds.cern.ch/record/1304162. (Cited on page 15.)

[47] A. Ilakovac and A. Pilaftsis, *Flavor violating charged lepton decays in seesaw-type models*, Nucl. Phys. **B437** (1995) 491, arXiv:hep-ph/9403398 [hep-ph]. (Cited on page 15.)

[48] D. V. Forero, S. Morisi, M. Tortola, and J. W. F. Valle, *Lepton flavor violation and non-unitary lepton mixing in low-scale type-I seesaw*, JHEP **09** (2011) 142, arXiv:1107.6009 [hep-ph]. (Cited on page 15.)

[49] S. P. Martin, *A Supersymmetry primer*, arXiv:hep-ph/9709356 [hep-ph]. [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. (Cited on page 15.)

[50] F. Borzumati and A. Masiero, *Large Muon- and Electron-Number Nonconservation in Supergravity Theories*, Phys. Rev. Lett. **57** (1986) 961–964. https://link.aps.org/doi/10.1103/PhysRevLett.57.961. (Cited on page 15.)

[51] J. Hisano, T. Moroi, K. Tobe, and M. Yamaguchi, *Lepton flavor violation via right-handed neutrino Yukawa couplings in supersymmetric standard model*, Phys. Rev. **D53** (1996) 2442–2459, arXiv:hep-ph/9510309 [hep-ph]. (Cited on page 15.)

[52] A. Gemintern, S. Bar-Shalom, G. Eilam, and F. Krauss, *Lepton flavor violating decays $L \to l\gamma\gamma$ as a new probe of supersymmetry with broken R–parity*, Phys. Rev. **D67** (2003) 115012, arXiv:hep-ph/0302186 [hep-ph]. (Cited on pages 15 and 16.)

[53] R. Harnik, J. Kopp, and J. Zupan, *Flavor Violating Higgs Decays*, JHEP **03** (2013) 026, arXiv:1209.1397 [hep-ph]. (Cited on pages 16 and 20.)

[54] J. I. Aranda, F. Ramirez-Zavaleta, J. J. Toscano, and E. S. Tututi, *Higgs mediated lepton flavor violating tau decays $\tau \to \mu\gamma$ and $\tau \to \mu\gamma\gamma$ in effective theories*, Phys. Rev. **D78** (2008) 017302, arXiv:0804.2652 [hep-ph]. (Cited on page 17.)

[55] HFAG Collaboration, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, arXiv:1612.07233 [hep-ex]. (Cited on page 18.)

[56] Belle Collaboration, *The Belle Detector*, Nucl. Instrum. Meth. **A479** (2002) 117–232. (Cited on page 17.)

[57] BaBar Collaboration, *The BaBar detector*, Nucl. Instrum. Meth. **A479** (2002) 1–116, arXiv:hep-ex/0105044 [hep-ex]. (Cited on page 17.)

[58] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005. (Cited on pages 17 and 23.)

[59] Belle Collaboration, *New Search for tau -> mu gamma and tau -> e gamma Decays at Belle*, Phys. Lett. **B666** (2008) 16–22, arXiv:0705.0650 [hep-ex]. (Cited on page 18.)

[60] I. Angelozzi, M. Bedognetti, S. Dhaliwal, and O. Igonkina, *Search for lepton flavor violating $\tau \to \mu\gamma$*. ATLAS internal note. In preparation. (Cited on pages 18 and 67.)

[61] Belle Collaboration, *Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced $\tau^+\tau^-$ Pairs*, Phys. Lett. **B687** (2010) 139–143, arXiv:1001.3221 [hep-ex]. (Cited on pages 19, 22, and 103.)

[62] LHCb Collaboration, *Search for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$*, JHEP **02** (2015) 121, arXiv:1409.8548 [hep-ex]. (Cited on page 19.)

[63] ATLAS Collaboration, *Probing lepton flavour violation via neutrinoless $\tau \longrightarrow 3\mu$ decays with the ATLAS detector*, Eur. Phys. J. **C76** no. 5, (2016) 232, arXiv:1601.03567 [hep-ex]. (Cited on pages 19, 22, 91, and 103.)

[64] Belle Collaboration, *Search for Lepton-Flavor-Violating tau Decays into a Lepton and a Vector Meson*, Phys. Lett. **B699** (2011) 251–257, arXiv:1101.0755 [hep-ex]. (Cited on pages 19 and 22.)

[65] MEG Collaboration, *The MEG detector for $\mu^+ \to e^+\gamma$ decay search*, Eur. Phys. J. **C73** no. 4, (2013) 2365, arXiv:1303.2348 [physics.ins-det]. (Cited on page 19.)

[66] MEG Collaboration, *Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment*, Eur. Phys. J. C **76** no. 8, (2016) 434. https://doi.org/10.1140/epjc/s10052-016-4271-x. (Cited on pages 19 and 22.)

[67] SINDRUM Collaboration, *Search for the decay $\mu^+ \to e^+e^+e^-$*, Nuclear Physics B **299** no. 1, (1988) 1 – 6. http://www.sciencedirect.com/science/article/pii/0550321388904622. (Cited on pages 19 and 22.)

[68] Mu3e Collaboration, A.-K. Perrevoort, *Status of the Mu3e Experiment at PSI*, EPJ Web Conf. **118** (2016) 01028, arXiv:1605.02906 [physics.ins-det]. (Cited on page 19.)

[69] SINDRUM II Collaboration, *A Search for muon to electron conversion in muonic gold*, Eur. Phys. J. **C47** (2006) 337–346. (Cited on pages 20 and 22.)

[70] OPAL Collaboration, *A Search for lepton flavor violating Z0 decays*, Z. Phys. **C67** (1995) 555–564. (Cited on pages 20 and 22.)

[71] DELPHI Collaboration, *Search for lepton flavor number violating Z0 decays*, Z. Phys. **C73** (1997) 243–251. (Cited on pages 20 and 22.)

[72] ATLAS Collaboration, *Search for the lepton flavor violating decay $Z \to e\mu$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 072010. https://link.aps.org/doi/10.1103/PhysRevD.90.072010. (Cited on pages 20 and 22.)

[73] CMS Collaboration, *Search for Lepton Flavor Violation in Z decays in pp collisions at $\sqrt{s} = 8$ TeV*, Tech. Rep. CMS-PAS-EXO-13-005, CERN, Geneva, 2015. https://cds.cern.ch/record/2019863. (Cited on pages 20 and 22.)

[74] ATLAS Collaboration, *Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector*, Eur. Phys. J. **C77** no. 2, (2017) 70, arXiv:1604.07730 [hep-ex]. (Cited on pages 20 and 22.)

[75] H. Weits, *Looking for lepton flavour violation with the ATLAS detector. A search for $Z \to \tau l$ decays*. PhD thesis, University of Amsterdam, 2016. (Cited on pages 20 and 22.)

[76] G. Blankenburg, J. Ellis, and G. Isidori, *Flavour-Changing Decays of a 125 GeV Higgs-like Particle*, Phys. Lett. **B712** (2012) 386–390, arXiv:1202.5704 [hep-ph]. (Cited on page 20.)

[77] CMS Collaboration, *Search for Lepton-Flavour-Violating Decays of the Higgs Boson*, Phys. Lett. **B749** (2015) 337–362, arXiv:1502.07400 [hep-ex]. (Cited on page 20.)

[78] CMS Collaboration, *Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and $e\tau$ in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-PAS-HIG-17-001, CERN, Geneva, 2017. https://cds.cern.ch/record/2264540. (Cited on pages 20 and 22.)

[79] CMS Collaboration, *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*, Phys. Lett. **B763** (2016) 472–500, arXiv:1607.03561 [hep-ex]. (Cited on pages 20 and 22.)

[80] P. Langacker, *The Physics of Heavy Z' Gauge Bosons*, Rev. Mod. Phys. **81** (2009) 1199–1228, arXiv:0801.1345 [hep-ph]. (Cited on page 21.)

[81] ATLAS Collaboration, *Search for a Heavy Neutral Particle Decaying to $e\mu$, $e\tau$, or $\mu\tau$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **115** no. 3, (2015) 031801, arXiv:1503.04430 [hep-ex]. (Cited on page 21.)

[82] ATLAS Collaboration, *Search for beyond the Standard Model phenomena in $e\mu$ final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-072, CERN, Geneva, Dec, 2015. http://cds.cern.ch/record/2114844. (Cited on page 21.)

[83] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001. (Cited on page 23.)

[84] *LEP design report*. CERN, Geneva, 1984. https://cds.cern.ch/record/102083. (Cited on page 23.)

[85] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002. (Cited on page 24.)

[86] MoEDAL Collaboration, *Technical Design Report of the MoEDAL Experiment*, Tech. Rep. CERN-LHCC-2009-006, CERN, Geneva, Jun, 2009. (Cited on page 24.)

[87] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007. (Cited on page 24.)

[88] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006. (Cited on page 24.)

[89] C. Lefevre, *LHC: the guide (English version). Guide du LHC (version anglaise)*, https://cds.cern.ch/record/1165534, Feb, 2009. (Cited on page 24.)

[90] ATLAS Collaboration, *ATLAS luminosity public results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults. (Cited on pages 26 and 27.)

[91] TOTEM Collaboration, *Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **111** no. TOTEM-2012-005. CERN-PH-EP-2012-354, (2012). http://cds.cern.ch/record/1495764. (Cited on page 26.)

[92] ATLAS Collaboration, *ATLAS Data Summary*, 2016. https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/DataSummary. (Cited on page 27.)

[93] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. https://cds.cern.ch/record/782076. (Cited on page 27.)

[94] J. Pequenao, *Computer generated image of the whole ATLAS detector*, https://cds.cern.ch/record/1095924, Mar, 2008. (Cited on page 28.)

[95] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*. No. ATLAS-TDR-6 in Technical Design Report ATLAS. CERN, Geneva, 1997. (Cited on page 29.)

[96] J. Pequenao, *Computer generated image of the ATLAS inner detector*, http://cds.cern.ch/record/1095926, Mar, 2008. (Cited on page 30.)

[97] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, http://cds.cern.ch/record/1095927, Mar, 2008. (Cited on page 32.)

[98] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*. No. ATLAS-TDR-2 in Technical Design Report ATLAS. CERN, Geneva, 1996. (Cited on page 31.)

[99] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. No. ATLAS-TDR-3 in Technical Design Report ATLAS. CERN, Geneva, 1996. (Cited on page 32.)

[100] J. Pequenao, *Computer generated image of the ATLAS Muons subsystem*, http://cds.cern.ch/record/1095929, Mar, 2008. (Cited on page 33.)

[101] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*. No. ATLAS-TDR-14 in Technical Design Report ATLAS. CERN, Geneva, 1999. (Cited on page 35.)

[102] ATLAS Collaboration, *RunStatsPublicResults*, 2016. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010. (Cited on page 36.)

[103] O. Igonkina and J. Vermeulen, *Trigger for decay of tau to muon and photon*, Tech. Rep. ATL-DAQ-INT-2012-003, CERN, Geneva, Aug, 2012. https://cds.cern.ch/record/1474860. (Cited on page 36.)

[104] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. **C75** (2015) 120, arXiv:1408.3179 [hep-ex]. (Cited on pages 37 and 94.)

[105] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Trans. ASME J. Basic Eng. **D 82** (1960) 35–45. (Cited on page 39.)

[106] D. Wicke, *A New algorithm for solving tracking ambiguities*, (1999) 219–228. LC-TOOL-1999-007-TESLA. (Cited on page 39.)

[107] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, Tech. Rep. ATLAS-CONF-2012-042, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1435196. (Cited on page 40.)

[108] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s}$ *=7 TeV in the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2010-069, CERN, Geneva, Jul, 2010. http://cds.cern.ch/record/1281344. (Cited on page 40.)

[109] W. Waltenberger, R. Frühwirth, and P. Vanlaer, *Adaptive vertex fitting*, Journal of Physics G: Nuclear and Particle Physics **34** no. 12, (2007) N343. http://stacks.iop.org/0954-3899/34/i=12/a=N01. (Cited on page 40.)

[110] K. Grimm, E. Guido, F. Meloni, S. P. Griso, K. Prokofiev, M. Rudolph, and A. Wildauer, *Methods to quantify the performance of the primary vertex reconstruction in the ATLAS experiment under high luminosity conditions*, Journal of Physics: Conference Series **396** no. 2, (2012) 022041.

http://stacks.iop.org/1742-6596/396/i=2/a=022041. (Cited on page 41.)

[111] ATLAS Collaboration, *Vertex reconstruction plots*, http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/IDTRACKING/PublicPlots/ATL-COM-PHYS-2012-474/. (Cited on page 41.)

[112] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. **C74** no. 11, (2014) 3130, arXiv:1407.3935 [hep-ex]. (Cited on pages 41, 42, 43, 44, and 93.)

[113] S. Hassani, L. Chevalier, E. Lançon, J.-F. Laporte, R. Nicolaidou, and A. Ouraou, *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, Nucl. Instrum. Meth. **A 572** (2007) 77–79. (Cited on page 42.)

[114] T. Lagouri et al., *A Muon Identification and Combined Reconstruction Procedure for the ATLAS Detector at the LHC at CERN*, IEEE Trans. Nucl. Sci. **51** (2004) 3030–3033. (Cited on page 42.)

[115] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex]. (Cited on page 43.)

[116] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, et al., *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002, CERN, Geneva, Apr, 2008. (Cited on page 43.)

[117] ATLAS Collaboration, *Expected photon performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-007, CERN, Geneva, Apr, 2011. http://cds.cern.ch/record/1345329. (Cited on page 44.)

[118] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, arXiv:1606.01813 [hep-ex]. (Cited on pages 44, 45, 46, and 47.)

[119] ATLAS Collaboration, *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb^{-1} of pp collision data collected in 2011*, Tech. Rep. ATLAS-CONF-2012-123, CERN, Geneva, Aug, 2012. https://cds.cern.ch/record/1473426. (Cited on page 45.)

[120] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. https://cds.cern.ch/record/1449796. (Cited on page 47.)

[121] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. **C74** no. 7, (2014) 2941, arXiv:1404.2240 [hep-ex]. (Cited on page 47.)

[122] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data*, Eur. Phys. J. **C77** no. 3, (2017) 195, arXiv:1612.01456 [hep-ex]. (Cited on pages 48 and 49.)

[123] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, arXiv:1603.02934 [hep-ex]. (Cited on page 48.)

[124] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph]. (Cited on page 48.)

[125] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv:1406.0076 [hep-ex]. (Cited on page 49.)

[126] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C73** no. 3, (2013) 2304, arXiv:1112.6426 [hep-ex]. (Cited on page 50.)

[127] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, arXiv:1510.03823 [hep-ex]. (Cited on page 50.)

[128] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, Tech. Rep. ATLAS-CONF-2015-037, CERN, Geneva, Aug, 2015. http://cds.cern.ch/record/2044941. (Cited on page 50.)

[129] ATLAS Collaboration, *Approved 2012 jet energy scale uncertainty*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013JESUncertainty. (Cited on page 50.)

[130] Particle Data Group Collaboration, K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001. (Cited on page 50.)

[131] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. http://cds.cern.ch/record/1700870. (Cited on page 51.)

[132] ATLAS Collaboration, *Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2013-064, CERN, Geneva, Jul, 2013. http://cds.cern.ch/record/1562839. (Cited on page 51.)

[133] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. **C76** no. 5, (2016) 295, arXiv:1512.05955 [hep-ex]. (Cited on page 51.)

[134] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039 [PHYSICS]. (Cited on pages 51, 56, 58, and 77.)

[135] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, Eur. Phys. J. **C75** no. 7, (2015) 303, arXiv:1412.7086 [hep-ex]. (Cited on pages 51, 52, and 94.)

[136] ATLAS Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012*, Tech. Rep. ATLAS-CONF-2013-044, CERN, Geneva, Apr, 2013. https://cds.cern.ch/record/1544036. (Cited on page 52.)

[137] ATLAS Collaboration, *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector*, arXiv:1609.09324 [hep-ex]. (Cited on page 52.)

[138] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, Tech. Rep. ATLAS-CONF-2013-082, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570993. (Cited on pages 52 and 53.)

[139] I. Antcheva et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **180** (2009) 2499–2512, arXiv:1508.07749 [physics.data-an]. (Cited on pages 56 and 68.)

[140] M. Baak, G. J. Besjes, D. Cote, A. Koutsman, J. Lorenz, et al., *HistFitter software framework for statistical data analysis*, arXiv:1410.1280 [hep-ex]. (Cited on pages 68, 92, and 95.)

[141] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph]. (Cited on page 68.)

[142] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug, 2012. https://cds.cern.ch/record/1474107. (Cited on page 68.)

[143] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph]. (Cited on page 68.)

[144] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293 [hep-ph]. (Cited on page 68.)

[145] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph]. (Cited on page 68.)

[146] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:hep-ph/0204244 [hep-ph]. (Cited on page 68.)

[147] ATLAS Collaboration, *Precision measurement and interpretation of inclusive W^+, W^- and Z/γ^* production cross sections with the ATLAS detector*, arXiv:1612.03016 [hep-ex]. (Cited on page 91.)

[148] ATLAS Collaboration, *Measurement of the $W \to \ell\nu$ and $Z/\gamma^* \to \ell\ell$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **12** (2010) 060, arXiv:1010.2130 [hep-ex]. (Cited on page 91.)

[149] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** no. 8, (2013) 2518, arXiv:1302.4393 [hep-ex]. (Cited on page 92.)

[150] M. Bedognetti, *$Z \to \mu\gamma\gamma$ as a control sample for the LFV $\tau \to \mu\gamma$ analysis at ATLAS*, Master's thesis, University of Amsterdam, 2014. (Cited on page 94.)

[151] A. L. Read, *Presentation of search results: the CL_s technique*, Journal of Physics G: Nuclear and Particle Physics **28** no. 10, (2002) 2693. http://stacks.iop.org/0954-3899/28/i=10/a=313. (Cited on pages 95 and 96.)

[152] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an]. [Erratum: Eur. Phys. J.C73,2501(2013)]. (Cited on pages 96 and 97.)

[153] ATLAS Collaboration, *Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. **B759** (2016) 601–621, arXiv:1603.09222 [hep-ex]. (Cited on page 103.)

Summary

What are the fundamental building blocks of the universe? Which forces govern their behaviour? High energy physics is the branch of science that tries to answer these and other intriguing questions.

The theory known as the *Standard Model* successfully describes the elementary particles and their interactions. The particle content of the Standard Model, shown in Figure S.1, consists of twelve fermions, divided into six quarks and six leptons, four types of gauge bosons (photon, gluon, $W^{\pm}$ and Z bosons) and the scalar Higgs boson. The latter was discovered in 2012, completing this picture. The three interactions described by the Standard Model are the electromagnetic, weak and strong forces.

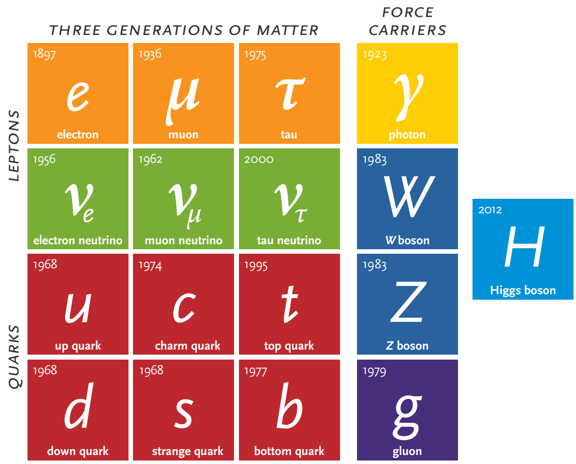


Fig. S.1 · Schematic view of the Standard Model with the three generations of fermions, the gauge bosons and the Higgs boson.

Despite the experimental confirmation of its predictions, the Standard Model does not explain everything. For example, gravity, which is the fourth fundamental force, is not included in the theory. Moreover, only about 5% of the universe is made of the particles of the Standard Model. What is the origin of the other 95%?

In addition, contrary to the original formulation of the Standard Model, the observation of *neutrino oscillations* implies that neutrinos are massive particles and can change from one flavour (electron, muon or tau neutrino) to another. It is natural to ask whether flavour violation can be observed also among the charged leptons (electron, muon and tau). In this sector, lepton flavour violating (LFV) processes are allowed in the Standard Model extension that includes massive neutrinos, but their rates are negligible. On the other hand, several scenarios of physics beyond the Standard Model predict rates within the reach of the experiments. Many searches have been performed during the years, through the investigation of several decay modes, but no evidence of charged LFV has been found so far.

In high energy physics, particle accelerators are employed to perform experimental searches. At the present time, the Large Hadron Collider (LHC) is the most powerful accelerator, designed to collide protons at the centre-of-mass energy of $\sqrt{s} = 14$ TeV. It is located at CERN, between France and Switzerland, in a tunnel of almost 27 km circumference. Four detectors are installed along the LHC ring to record the collisions delivered by the LHC. The ATLAS detector measures Standard Model observables and searches for physics beyond the Standard Model. It consists of several sub-detectors, each with specific properties, which are used to identify and measure the energy and momentum of the different particles produced in the collisions. Furthermore, the ATLAS detector employs a *trigger system* to reduce the amount of the LHC data and select the fraction of interesting events for physics analyses.

In this thesis, a search for the LFV $\tau \to \mu\gamma\gamma$ decay has been presented. It is based on the data collected with the ATLAS detector at $\sqrt{s} = 8$ TeV in 2012, during the first run (Run-1) of the LHC. The data sample corresponds to an integrated luminosity of 19.1 fb^{-1}. The tau leptons used in the analysis are produced by the $W \to \tau\nu$ decay.

Tau leptons provide good opportunities to search for LFV. Several decay channels, such as $\tau \to \mu\gamma$ and $\tau \to 3\mu$, have been studied by different experiments looking for a sign of physics beyond the Standard Model. The $\tau \to \mu\gamma\gamma$ channel represents a new probe of LFV. Theoretical frameworks, such as *supersymmetry* and *Higgs-mediated* LFV, indicate the possibility to investigate this decay, predicting sizeable rates. No experimental searches have been performed in the $\tau \to \mu\gamma\gamma$ channel before the one described here.

In any analysis, it is essential to enhance the *signal* under investigation, reducing the contribution from *background* processes. In the LFV $\tau \to \mu\gamma\gamma$ search, the background is estimated directly from data. Selection criteria are applied to target the decay of the W boson and the $\tau \to \mu\gamma\gamma$ candidate in each event. Several requirements are applied on physical quantities, such as the kinematic properties of photons and muons and the transverse mass of the W boson. Another step in the selection of the events is based on triggers. In particular, dedicated *muon+photon* triggers are used to select the combinations of objects in the final state of the $\tau \to \mu\gamma\gamma$ decay.

For a better discrimination between signal and background sources, key steps of the analysis rely on a multivariate approach. The identification of the pair of photons in the $\tau \to \mu\gamma\gamma$ decay is realized using the multivariate method called *Boosted Decision Tree* (BDT), based on detector characteristics of these objects. At a later stage, the same strategy is adopted to focus on the properties of the $W \to \tau\nu$

decay and target the remaining background, which consists mainly of decays of heavy quarks and $W \to \mu\nu$ decays.

The analysis is designed to be sensitive in the region where a signal excess over the background is expected. The *signal region* is selected relying on the distribution of the LFV tau candidate mass, which is given by the invariant mass of the three tau decay products (one muon and two photons). Figure S.2 shows the mass distribution of the tau candidate for the Monte Carlo *simulation* of the $\tau \to \mu\gamma\gamma$ sample (signal) and the data sample. The signal region is defined as an interval around the nominal tau mass between 1.6 and 2.0 GeV.

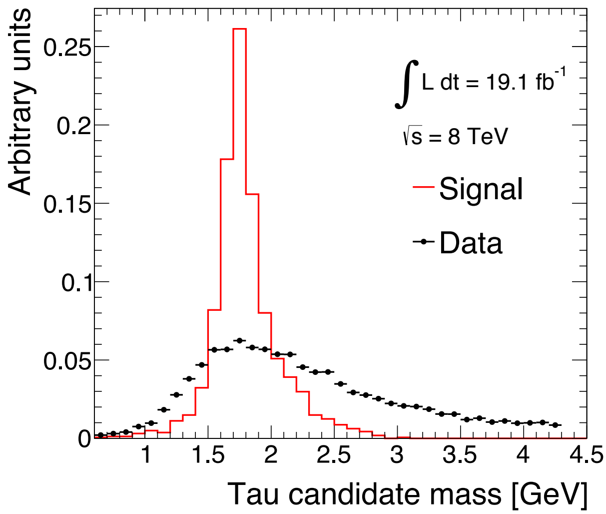


Fig. S.2 · Mass distribution of the LFV tau candidate for signal (red) and data (black dots) samples.

The expected number of background events in the signal region is compared to the number of observed data events in the same region. At the end of the selection, 293 data events are observed in the signal region. This value is compatible with the one obtained from the background estimation (278 expected events). No statistically significant excess with respect to the Standard Model background expectation is observed.

In the absence of a LFV signal, the results of this search are presented in terms of *upper limits* on the $\tau \to \mu\gamma\gamma$ branching fraction. The expected and observed exclusion limits are determined with a confidence level of 90%. The expected (median) upper limit is established to be

$$\mathrm{Br}(\tau \to \mu\gamma\gamma) = (1.14^{+0.57}_{-0.35}) \times 10^{-4}, \tag{S.1}$$

where its variations, expressed in terms of ± 1 standard deviation, are also indicated. The observed upper limit corresponds to

$$\mathrm{Br}(\tau \to \mu\gamma\gamma) = 1.49 \times 10^{-4}. \tag{S.2}$$

Exciting years are in front of us. The LHC is now in its second run (Run-2), which is characterized by an increased centre-of-mass energy of $\sqrt{s} = 13$ TeV. The larger datasets will allow to improve the results presented in this thesis. Looking at the bigger picture, the quest for lepton flavour violation among charged leptons will continue. Is it time for another breakthrough discovery in high energy physics?

Samenvatting

Wat zijn de fundamentele bouwstenen van het universum? Welke krachten bepalen hun gedrag? Hoge energie fysica is de tak van de wetenschap die probeert antwoord te vinden op deze en andere intrigerende vragen.

De theorie die bekend staat als het *Standaardmodel* beschrijft met succes de elementaire deeltjes en hun interacties. De deeltjes in het Standaardmodel, te zien in Figuur S.1, bestaan uit twaalf fermionen, onderverdeeld in zes quarks en zes leptonen, vier soorten ijkbosonen (foton, gluon, $W^{\pm}$ en Z bosonen) en het scalaire higgsboson. Het laatstgenoemde werd in 2012 ontdekt en voltooide dit plaatje. De drie interacties die door het Standaard Model worden beschreven zijn de elektromagnetische, zwakke en sterke kernkrachten.

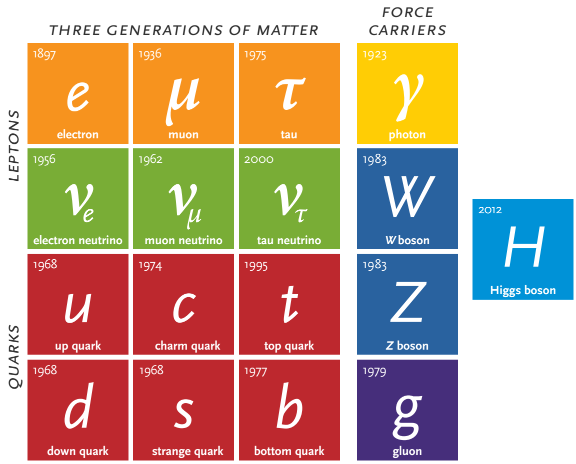


Fig. S.1 · Schematische weergave van het Standaardmodel met de drie generaties van fermionen, de ijkbosonen en het higgsboson.

Ondanks de experimentele bevestiging van zijn voorspellingen, verklaart het Standaardmodel niet alles. Bijvoorbeeld zwaartekracht, de vierde fundamentele kracht, wordt niet beschreven in deze theorie. Overigens bestaat maar ongeveer 5% van het universum uit de deeltjes in het Standaardmodel, wat is de herkomst van de andere 95%?

In tegenstelling tot de originele formulering van het Standaardmodel impliceert de observatie van *neutrino oscillaties* dat neutrino's massieve deeltjes zijn en van smaak (elektron-, muon-, tau-neutrino) kunnen veranderen. Een natuurlijk vervolg is om te vragen of deze smaakschending ook geobserveerd kan worden bij de geladen leptonen (elektron, muon en tauon). In een uitbreiding van het Standaardmodel met massieve neutrino's zijn de processen die de leptonsmaak schenden (LSS) in deze sector toegestaan, alleen komen ze uiterst zelden voor. Aan de andere kant, verscheidene scenario's van fysica voorbij het Standaardmodel voorspellen een frequentie van dergelijke processen die binnen het bereik van de experimenten valt. Door de jaren heen zijn veel zoektochten uitgevoerd in verscheidene vervalskanalen, maar tot nog toe is er voor geladen LSS geen bewijs gevonden.

In de hoge energie fysica worden deeltjesversnellers gebruikt om experimentele metingen te doen. De Large Hadron Collider (LHC) is momenteel de krachtigste versneller, ontworpen om protonen te laten botsen met een zwaartepuntsenergie van $\sqrt{s} = 14$ TeV. Hij bevindt zich op CERN, in Frankrijk en Zwitserland, in een tunnel met een omtrek van bijna 27 km. Vier detectoren zijn geïnstalleerd in de ring van de LHC om de botsingen, geleverd door de LHC, te bestuderen. De ATLAS detector meet Standaardmodel observabelen en zoekt naar fysica voorbij het Standaardmodel. ATLAS bestaat uit verscheidene sub-detectoren, elk met specifieke eigenschappen, die worden gebruikt om de verschillende deeltjes die door de botsingen worden geproduceerd te identificeren en hun energie en impuls te meten. Bovendien maakt de ATLAS detector gebruik van een *triggersysteem* om de hoeveelheid data die de LHC levert te reduceren en alleen de meest interessante gebeurtenissen te selecteren voor verder onderzoek.

In dit proefschrift is de zoektocht naar het LSS $\tau \to \mu\gamma\gamma$ verval gepresenteerd. Het is gebaseerd op de data die de ATLAS detector heeft verzameld bij $\sqrt{s} = 8$ TeV in 2012, tijdens de eerste operationele periode (Run-1) van de LHC. De hoeveelheid data komt overeen met een geïntegreerde luminositeit van 19.1 fb^{-1}. De tau-leptonen die gebruikt worden in deze analyse zijn geproduceerd via het $W \to \tau\nu$ verval.

Tau-leptonen leveren goede mogelijkheden om naar LSS te zoeken. Verscheidene vervalskanalen, zoals $\tau \to \mu\gamma$ en $\tau \to 3\mu$, zijn al onderzocht door verschillende experimenten die op zoek zijn naar signalen van fysica voorbij het Standaardmodel. Het $\tau \to \mu\gamma\gamma$ kanaal is een nieuwe manier om te zoeken naar LSS. Theoretische modellen, zoals *supersymmetrie* en *higgs-mediated* LSS, geven aan dat het mogelijk is om dit verval te onderzoeken, ze voorspellen een aanzienlijke vervalsfractie. Er zijn geen eerdere experimentele zoektochten naar het $\tau \to \mu\gamma\gamma$ kanaal gedaan.

In elke analyse is het belangrijk dat het te onderzoeken *signaal* wordt vergroot, door de bijdrage van *achtergrond* processen te reduceren. In de LSS $\tau \to \mu\gamma\gamma$ meting wordt de achtergrond rechtstreeks uit de data geschat. Selectie criteria worden toegepast om het $W \to \tau\nu$ verval en de $\tau \to \mu\gamma\gamma$ kandidaat in elke gebeurtenis te vinden. Verscheidene eisen worden gesteld aan een aantal grootheden, zoals de kinetische eigenschappen van fotonen en muonen en de transversale massa van het W boson. Een volgende stap in de selectie procedure van de gebeurtenissen is gebaseerd op de triggers. In het bijzonder worden toegewijde *muon+foton* triggers gebruikt om de specifieke combinaties van de objecten in de eindtoestand van het $\tau \to \mu\gamma\gamma$ verval te selecteren.

Sleutelstappen in de analyse, voor een betere scheiding tussen signaal en achtergronden, berusten op een multivariate aanpak. De identificatie van het foton paar in het $\tau \to \mu\gamma\gamma$ verval is gerealiseerd door middel van de multivariate methode genaamd *Boosted Decision Tree* (BDT), gebaseerd op de detector karakteristieken van deze objecten. In een later stadium is dezelfde strategie gebruikt bij het richten op de eigenschappen van het $W \to \tau\nu$ verval, met het doel om de overgebleven achtergronden, die voornamelijk bestaan uit vervallen van zware quarks en $W \to \mu\nu$ vervallen, te reduceren.

De analyse is ontworpen om gevoelig te zijn in het gebied waar verwacht wordt dat het signaal de achtergrond overschrijdt. Het *signaalgebied* is bepaald op basis van de distributie van de LSS tau kandidaat massa, die wordt gegeven door de invariante massa van de drie vervalsproducten (een muon en twee fotonen). Figuur S.2 laat de massadistributie van de tau kandidaat zien voor het Monte Carlo-gesimuleerde $\tau \to \mu\gamma\gamma$ signaal en de gemeten data. Het signaal gebied is gedefinieerd als een regio om de tau massa heen, tussen 1.6 en 2.0 GeV.

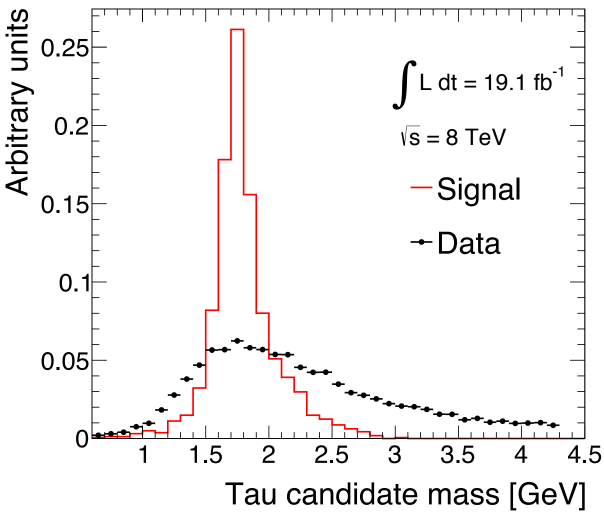


Fig. S.2 · Massa distributie van de LSS tau kandidaat voor het signaal (rood) en gemeten data (zwarte stippen).

Het verwachtte aantal achtergrond gebeurtenissen in het signaalgebied is vergeleken met het totaal aantal gebeurtenissen in de gemeten data in hetzelfde gebied. Aan het einde van de selectieprocedure zijn 293 data gebeurtenissen geobserveerd in het signaalgebied. Deze waarde is compatibel met de verkregen schatting van de achtergrond (278 verwachtte gebeurtenissen). Er is dus geen statistisch significante overschrijding van de Standaardmodel achtergrond verwachting geobserveerd.

In afwezigheid van een LSS signaal worden de resultaten van dit onderzoek gepresenteerd in de vorm van een *bovenlimiet* op de $\tau \to \mu\gamma\gamma$ vervalsfractie. De verwachtte en geobserveerde limieten zijn bepaald op een confidence level niveau van 90%. De verwachte (mediaan) bovenlimiet is

$$\mathrm{Br}(\tau \to \mu\gamma\gamma) = (1.14^{+0.57}_{-0.35}) \times 10^{-4}, \tag{S.1}$$

waar de variaties, uitgedrukt in termen van ± 1 standaardafwijking, ook worden weergegeven. De geobserveerde limiet is

$$\mathrm{Br}(\tau \to \mu\gamma\gamma) = 1.49 \times 10^{-4}. \tag{S.2}$$

Er komen spannende jaren aan. De LHC is nu bezig met de tweede operationele periode (Run-2), die wordt gekarakteriseerd door een verhoogde zwaartepuntsenergie van $\sqrt{s} = 13$ TeV. De grotere datasets maken het mogelijk om de resultaten gepresenteerd in dit proefschrift te verbeteren. Kijkend naar het grotere plaatje zal de zoektocht naar lepton smaak schending bij geladen lepton worden voortgezet. Wordt het tijd voor nog een baanbrekende ontdekking in de hoge energie fysica?

Acknowledgements

Finally I am arrived at the end of the PhD journey. It was tough, sometimes very tough, but I also enjoyed a lot this important period of my life. First of all, I would like to thank my supervisors Olya Igonkina and Paul de Jong.

Olya, thanks for all your support during my PhD. You gave me your help since day one and even if sometimes we had some discussions I would have never finished this thesis without you. I appreciate your patience, commitment and enthusiasm.

Paul, thank you for your precious advices, especially in the last two years. You always indicated me the right way to keep going with your knowledge and suggestions.

I would like to thank the members of my committee for spending their time to read my thesis: Gerhard, Eric, Wouter, Auke Pieter, Ronald, Gerco and Sami. A special thanks goes to Sami, who helped me a lot to get started and during the year at CERN with her kindness being always available to answer my questions.

During the PhD I have met a lot of people in (and outside) the ATLAS Nikhef group, between The Netherlands and CERN who made these years happier. Thank you all! In particular, I would like to mention Joern, Hartger, Ingrid, Lydia, Vince, Koen, Broos, Marc, Hannah, Birgit, Geert-Jan, Wouter, Hegoi, Daniel, Peter V., Nika, Pier-Olivier, Marcus, Noam, Rosemarie, Rolf, Tim, Antonia, Tristan and Stefan.

A special thanks to the italian-greek community who shared with me many moments, inside and outside Nikhef: Antonio, Andrea, Afroditi, Stergios, Panos, Lucrezia, Corinne, Chrysa, Pierfrancesco, Gabriele, Matteo, Emilia, Dimitra, Vassilis, Nikos, Alice, Sabrina, Veronica, Fabrizia, Luca, Francesco, Enrico, Elena, Priscilla, Michele D., Michele M. and Jacopo.

I would like to spend few more words for some of them.

Thanks to my friend, ex housemate and paranymph Antonio. It is always nice talking to you about physics, but even more discuss about sports, politics or movies, comparing our different opinions in front of a beer.

Thanks to my friend and (de facto) paranymph Andrea. We had a lot of fun moments to remember during these years. The endless ping pong matches in front of big crowds, the time spent together in summer schools and conferences and much more.

Thanks to my friend and paranymph Stergios. You are the "number one". In case of a doubt, you are the guy to follow.

Thanks to you, Afro. We started this journey together and I was lucky to meet such a nice person to share moments throughout the years. Now I finish this journey and you will finish it soon as well! Thanks to my friends Pierfrancesco and Lucrezia. It was nice to meet you and I am happy you were part of this adventure.

I would like to mention my italian friends I have met in the past before coming to Amsterdam. Maybe, in some cases, our roads will not cross again, but I will never forget about you. I would also like to say thanks to Alba, who was my housemate during the crazy period of the thesis writing. I wish you all the best!

I owe special thanks to my family because they always supported me. Especially I say thanks to my parents, Loretta and Mauro, to be always there when I need it. I hope you can be proud of this achievement and of my past 30 years.

Finally, I want to say thanks to two persons with the same name (starting with G). For different reasons, you are special for me and you always will be.

Ivan
Amsterdam, 2017