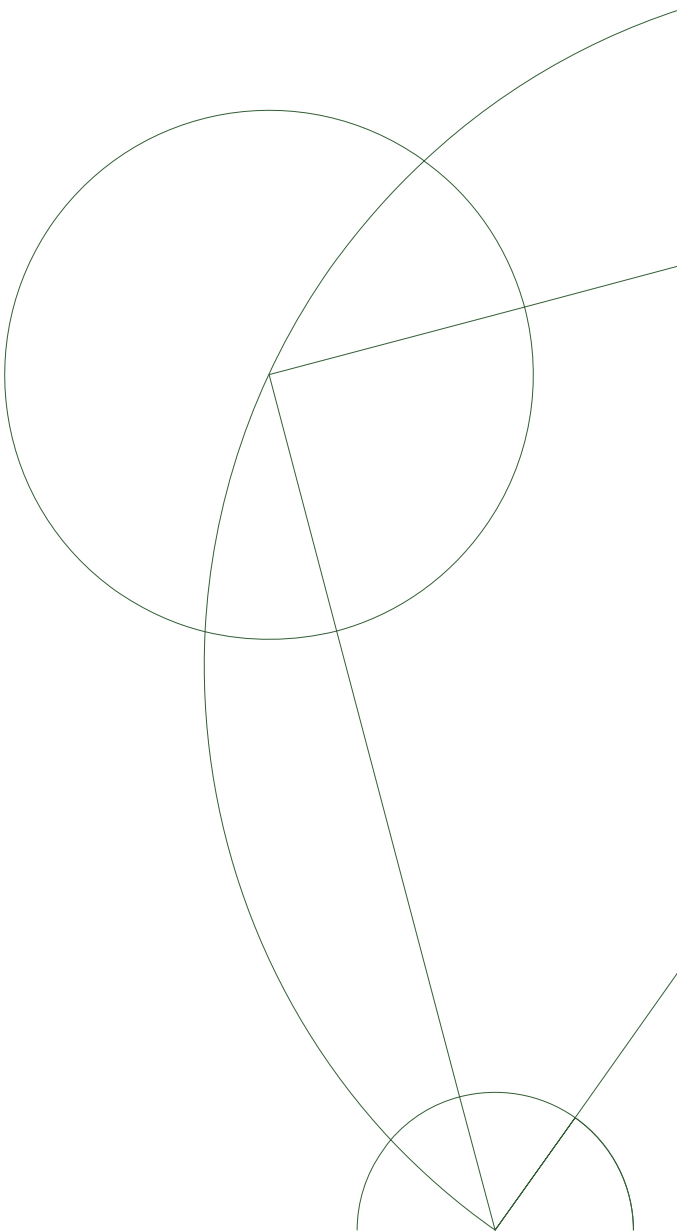
FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Master's Thesis

A Search for Lepton Flavour Violation in $Z \to \tau\mu$ decays at $\sqrt{s} = 13$ TeV with the ATLAS Detector

Sara Svendsen



Supervisor: Assoc. Prof. Stefania Xella
Niels Bohr Institute
University of Copenhagen

January 2, 2017

Abstract

Violation of lepton flavour number conservation is forbidden at tree-level in the Standard Model and highly suppressed at higher orders. While violation of lepton flavour has been observed via neutrino oscillations, it has not yet been measured among charged leptons. The discovery of such processes is therefore a direct sign of new physics beyond the Standard Model.
In this thesis, a strategy for analysing the presence of a lepton flavour violating $Z \to \tau_{\text{had}}\mu$ signature in LHC proton-proton collision data is presented. The analysis is based on $\sqrt{s} = 13$ TeV data collected by the ATLAS Detector corresponding to an integrated luminosity of $\mathcal{L} = 11.8\,\text{fb}^{-1}$. A selection of kinematic cuts are tested in order to locate a phase space region that is enhanced in signal. Two such signal regions are defined, yielding a combined statistical significance of 1.6.
Monte Carlo simulated events are normalised to data by scale factors derived in normalisation regions. Backgrounds that contribute due to misidentification of tau leptons are estimated from data in control regions using a fake factor method.
This results in a background model that is agreement with data at preselection level. Once the proper validation tests have been carried out, the presence of a signal will be evaluated with a likelihood fit in order to detect a signal or set a new upper limit.

Acknowledgements

I would first of all like to thank my supervisor Stefania Xella for keeping me on track and for always asking critical questions. I am very thankful for all the help I have received from Geert-Jan Besjes, from providing the common analysis framework to patiently answering my questions and helping with debugging.
It has been a great experience to work in the $Z \to \tau_{\text{had}}\ell$ analysis team and I've learnt more than I thought possible in the beginning of this project!

I wish to extend my thanks to my fellow students in the office who will all be greatly missed; thanks to Fabian, Milena and Sissel for valuable help on the analysis, for the many coffee breaks and for always being reliable sources of both encouragement and distraction.

Lastly, I would like to thank colleagues, friends and family who have shown their support throughout this year; without the help from Mikkel in particular, I would never have been able to complete this work.

Contents

1 Introduction

Our current knowledge of elementary particles, the most fundamental building blocks of nature, and their mutual interactions are described in the theory known as the *Standard Model*.
The Standard Model aims to describe physical phenomena at the highest accessible energies (smallest scales) providing theoretical predictions that can be tested experimentally. The most notable achievement in recent years was the discovery of the Higgs boson in 2012 [1, 2] but the predictions of the theory continue to excel at passing precision tests.
Despite its success, the Standard Model is considered incomplete. The current theory still leaves questions unanswered and some physical observations are not yet well-described by the theory. These shortcomings do in turn motivate research into new physics sectors *beyond the Standard Model*, such as searches for sterile neutrinos, supersymmetric particles or gravitons, and the nature of dark matter.

One of these unresolved puzzles of the Standard Model is the apparent matter-antimatter asymmetry (baryon asymmetry) of the Universe. The asymmetry is not an initial condition of the early Universe so matter and antimatter would been created in equal amounts. Matter and antimatter eventually annihilated with each other - leaving only radiation. However, a small excess of matter must have survived because everything we observe today is made up almost entirely of matter. Only a relatively small amount of antimatter prevailed.
The mechanism by which this asymmetry occurred is not yet fully understood but the lack of antimatter could be an effect of matter and antimatter behaving differently under charge-parity (CP) transformations (interchange of a particle with its antiparticle and inversion of its spatial coordinates), causing violation of CP symmetry.

The phenomenon of baryon asymmetry is related to flavour physics since flavour mixing in the quark sector introduces the only source of CP violation in the Standard Model. Nevertheless, this source of CP violation alone is not sufficient to generate the observed baryon asymmetry. Additional CP violating processes beyond the Standard Model might therefore present a solution.

Generally, an introduction of new phenomena in the lepton sector could have an important impact on leptogenesis of the early Universe. Leptogenesis describes a possible scenario of baryogenesis that proposes a Standard Model extension including heavy right-handed neutrinos which can decay and spontaneously generate leptons. The extension also allows the seesaw mechanism to provide small masses to the left-handed neutrinos which is required for neutrino oscillation. Calculations with the observed neutrino oscillation parameters are able to reproduce the observed baryon asymmetry [3].

The branching fractions of lepton flavour violating Z decays predicted from the SM are effectively zero. Although neutrino oscillations have been established experimentally, flavour violation among *charged leptons* has not yet been observed, possibly due to a lack of sufficient experimental precision so far. It is therefore not predicted in the SM. The current best direct experimental upper limits (95% CL) on the flavour changing Z decays into charged leptons are:

$$\mathrm{BR}\left(Z \to e^{\pm}\mu^{\mp}\right) < 1.7 \cdot 10^{-6} \text{ [4]} \tag{1.1}$$

$$\mathrm{BR}\left(Z \to \tau^{\pm}e^{\mp}\right) < 9.8 \cdot 10^{-6} \text{ [4]} \tag{1.2}$$

$$\mathrm{BR}\left(Z \to \tau^{\pm}\mu^{\mp}\right) < 1.2 \cdot 10^{-5} \text{ [5]}\,. \tag{1.3}$$

The aim of this thesis is to search for lepton flavour violating decays of the Z boson at a center-of-mass energy of $\sqrt{s} = 13$ TeV, specifically in the $Z \to \tau_{\mathrm{had}}\mu$ decay channel. In terms of the visible decay products of the tau lepton, we consider only the hadronic decay mode $\tau \to \mathrm{had} + \nu$ and not the leptonic $\tau \to \ell\nu\nu\bar{\nu}$ where $\ell = e, \mu$. The two decay modes differ mainly in statistics with hadronic taus being more abundant but also in object identification and their systematic uncertainties.

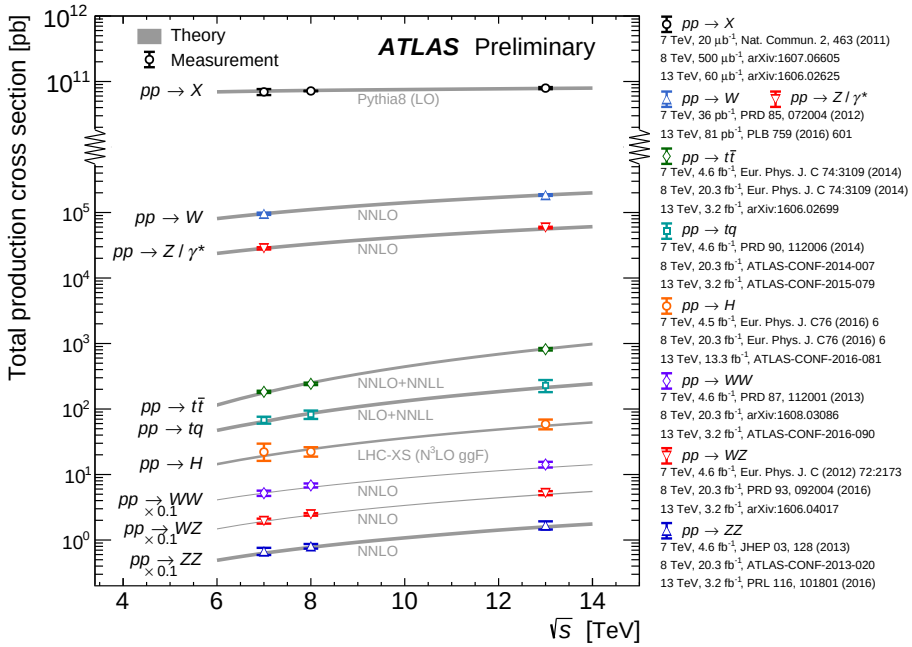


Figure 1.1: Standard Model production cross sections as a function of $\sqrt{s}$. Taken from [6].

The total production cross section in pb from pp collisions is shown in fig. 1.1 as a function of center-of-mass energy in TeV. The production cross section of Z bosons increases with higher energy and with the amount of data collected so far in Run 2, we hope to have enough precision to either see signs of Z LFV or put stronger limits on its occurrence.

Throughout the thesis, the natural unit convention which redefines the speed of light and Planck's constant[1] as $c = \hbar = 1$ will be used. Kinematic variables such as energy, momentum and mass are therefore expressed in terms of electron volt ($1\,\mathrm{eV} = 1.60218 \cdot 10^{-19}$ J). Cross sections are measured in barns: $1\,\mathrm{b} = 10^{-28}\,\mathrm{m}^2 = 100\,\mathrm{fm}^2$.

[1] $c = 2.99792 \cdot 10^{8}$ m/s and $\hbar = 1.05457 \cdot 10^{-34}$ J s, respectively.

2 | Theory

This chapter will describe the particle and interactions of the Standard Model (SM). The general focus will be on electroweak theory and in particular flavour changing processes among weak interactions.
Searches for flavour number non-conserving processes, or flavour changing neutral currents (FCNC), prove a powerful test for new physics due to the suppression of FCNCs at tree level in the SM. Several models beyond the SM propose extensions leading to FCNC cross sections that are orders of magnitudes larger than the SM predictions. The main focus here will be on a model that introduces effective flavour changing Z couplings accessible at energies $M > m_Z$.

2.1 The Standard Model

Elementary particles are the smallest constituents of the Universe, and the Standard Model of particle physics is the generally accepted theory that describes the properties and interactions of the known elementary particles.

The SM was initially established during the 1960s by Glashow, Weinberg and Salam [7, 8] with the successful combination of electromagnetic and weak interactions into a common *electroweak theory*. In the following decades, the SM was expanded to include strong interactions and the Higgs field, providing the mechanism by which other SM particles acquire mass. The theory successfully describes three of the four fundamental forces; the SM does not provide a description of the gravitational force due to its non-renormalisability within the framework of quantum field theory. Overall, the SM provides a successful description of current experimental results in subatomic physics.

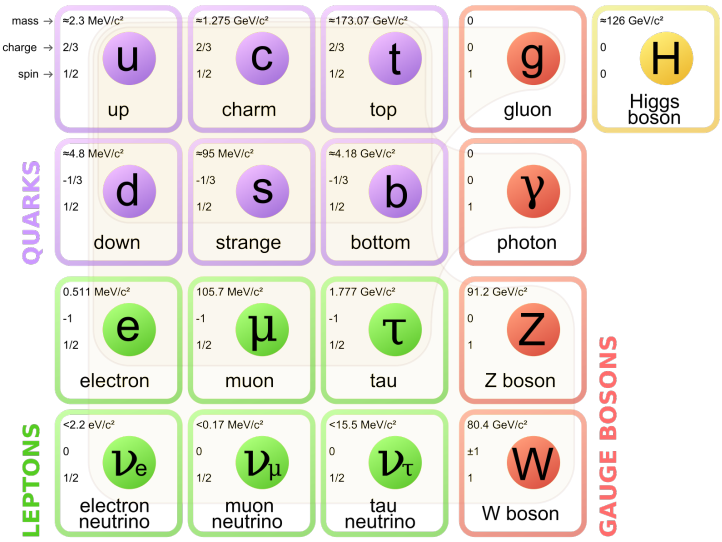


Figure 2.1: Particles of the Standard Model; Leptons and quarks each exist in three generations. The twelve gauge bosons mediate the electroweak and strong forces. The Higgs field causes particles to acquire mass.

2.1.1 Particle content of the SM

The Standard Model consists of fermions with fractional intrinsic spins (leptons and quarks), and bosons carrying integer spins. These are all represented in fig. 2.1. Fermions are usually denoted *matter particles* as they constitute the building blocks of all matter in the Universe whereas the gauge bosons are the mediators of SM interactions. A particle couples only to a gauge boson if it carries the charge of the interaction.
The scalar (spin-0) Higgs boson is distinct since it couples to all other particles, except photons and gluons, causing them to acquire mass through spontaneous symmetry breaking.

Leptons exist in three generations and in either an electrically charged or neutral state. The electron (e) is the lightest of the charged leptons, followed by the heavier muon (μ) and the tau lepton (τ). The nonzero electric charge means that e, μ and τ leptons interact via the electromagnetic force.
For each charged lepton flavour, a corresponding neutrino (ν) exists. In the SM, neutrinos are massless but the observation of neutrino oscillations implies that the three neutrino species are in fact massive [9, 10, 11].
Since all leptons also carry weak charge (isospin charge), they will couple to the weak force.

The quark sector has three generations of up- and down-type quarks. The first generation contains the up and down quarks which are the lightest quarks and the subcomponents of nucleons. The second and third generations include the charm and strange, and the top and bottom quarks, respectively.
Quarks carry colour and anti-colour charge (red, green, blue) in addition to weak and fractional electric charges, causing them to also interact via the strong force governed by quantum chromodynamics (QCD).
The main consequences of QCD are asymptotic freedom and colour confinement. Asymptotic freedom is related to the behaviour of the coupling constant α_S since the strength of the coupling decreases with increasing energies. Confinement means that QCD particles can only exist in bound colourless states which are collectively termed hadrons.
While important to the phenomenological understanding of hadronisation in a collider experiment, i.e the production of jets of hadrons from quarks and gluons from successive quark-antiquark production, QCD will not be further discussed; the treatment of lepton flavour is focused on electroweak interactions which are independent from QCD.

The four types of gauge bosons are collectively known force carriers as they mediate interactions between particles through three of the fundamental forces:

- The photon γ is responsible for **electromagnetic interactions** among electrically charged particles (quarks, charged leptons and $W^{\pm}$).
- The massive $W^{\pm}$ and the Z^0 bosons govern **weak interactions** between particles with weak charge, i.e. all quarks and leptons. Neutrinos *only* interact via the weak force.
- The eight gluons mediate **strong interactions** between particles with colour charge and couple to both quarks and gluons.

Classically, elementary particles are treated as point particles. As the framework of the SM is quantum field theory, however, particles are instead viewed as excitations of a quantum field satisfying the appropriate equations of motion. The dynamics of the fields are formulated in the SM Lagrangian density $\mathcal{L}$.

2.1.2 Electroweak theory

In the electroweak unification, the electromagnetic and weak interactions are invariant under a common SU(2) × U(1) gauge group.
An essential concept in electroweak theory is the Higgs mechanism [12, 13, 14] that is introduced in order to generate the masses for the weak gauge bosons while preserving local gauge invariance. The mechanism is embedded in the local $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry of electroweak theory relating to weak isospin T and hypercharge Y. These quantities yield the electric charge $Q = T_3 + Y/2$ where T_3 is the third component of the weak isospin.
Electroweak gauge symmetry is spontaneously broken as the Higgs field ϕ acquires a non-zero vacuum expectation value $\langle\phi\rangle = \begin{pmatrix}0\\ v\end{pmatrix}$, thus producing the physical electroweak bosons and their masses.

2.1.3 Fermion fields

In terms of a fermion field ψ_i, the electroweak Lagrangian after the symmetry breaking can be written as:

$$\begin{aligned}\mathcal{L} = {} & \bar{\psi}_i \left(i\partial\!\!\!/ - m_i - \frac{m_i H}{v} \right) \psi_i - eQ_i \bar{\psi}_i \gamma^\mu \psi_i A_\mu \\ & - \frac{g_W}{2\sqrt{2}} \bar{\Psi}_{Li} \gamma^\mu \left(T^+ W_\mu^+ + T^- W_\mu^- \right) \Psi_{Li} \\ & - \frac{g_W}{2\cos\theta_W} \bar{\psi}_i \gamma^\mu \left(g_V^i - g_A^i \gamma^5 \right) \psi_i Z_\mu, \end{aligned} \tag{2.1}$$

where H is the Higgs, e is the positron electric charge and Q is the electrical charge in units of e. A_μ, W_μ and Z_μ are the photon field and the physical charged and weak boson fields, respectively. Ψ_i represents the left-handed fermion fields; θ_W the weak angle and g_W the weak coupling constant. $T^\pm = \frac{1}{2}\left(\sigma^1 \pm i\sigma^2\right)$ are the raising and lowering operators of weak isospin and lastly, g_V, g_A are the vector and axial vector couplings. γ^μ are the gamma matrices which satisfy the anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$.

Parity transformation corresponds to spatial inversion through the origin, $\boldsymbol{x} \to -\boldsymbol{x}$; invariance of Dirac fermions under parity transformation is called chiral symmetry. Any spinor can be split into left- and right-handed chiral components. The chiral four-component states are eigenstates of the fifth gamma matrix $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$ with eigenvalues ±1. The projections are defined as:

$$\psi_{L,R} = P_{L,R}\,\psi = \frac{1}{2}\left(1 \mp \gamma^5\right)\psi; \qquad \bar{\psi}_{L,R} = P_{L,R}\,\bar{\psi} = \frac{1}{2}\left(1 \pm \gamma^5\right)\bar{\psi}\,. \tag{2.2}$$

With the exception of right-handed neutrinos which do not exist according to the SM, quarks and leptons are represented by left-handed doublets and right-handed singlets for each generation i:

$$Q_{Li} = \begin{pmatrix} u_{Li} \\ d_{Li} \end{pmatrix}, \quad L_{Li} = \begin{pmatrix} \nu_{Li} \\ l_{Li} \end{pmatrix}, \quad u_{Ri}, \quad d_{Ri}, \quad l_{Ri}\,. \tag{2.3}$$

The Yukawa interactions with the Higgs field generates the fermion mass terms:

$$-\mathcal{L}_Y = \sum_{i,j=1}^{3} \left(\bar{d}_{Li} Y_{ij}^d d_{Rj} + \bar{u}_{Li} Y_{ij}^u u_{Rj} + \bar{l}_{Li} Y_{ij}^l l_{Rj} \right) + \text{h.c.}, \tag{2.4}$$

where Y_{ij} are the Yukawa couplings between the fermions and Higgs field.
Quark masses but also quark mixing arise from CP violation and Yukawa interactions when moving from the flavour base of $\mathcal{L}_Y$ to the mass base Lagrangian, i.e. diagonalising Y_{ij} by means of a unitary transformation [15, 16]. The transformation matrix will be further discussed in section 2.1.5.

2.1.4 Weak interactions

After electroweak symmetry breaking, the weak interaction associated with the local $\mathrm{SU}(2)_L$ gauge symmetry gives rise to the $W^\pm$ bosons and a neutral field W^3_μ. The physical charged W bosons are represented by the linear combination:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp i W^2_\mu\right); \qquad m_W = \frac{1}{2} g_W v\,, \tag{2.5}$$

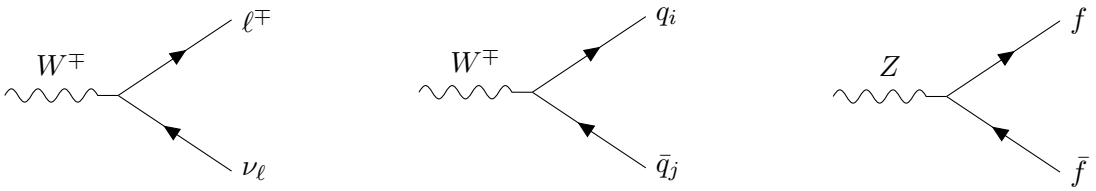
with the mass of the W boson depending on the coupling constant g_W of $\mathrm{SU}(2)_L$ interactions. The W^3_μ gauge field mixes with the B_μ field of weak hypercharge ($\mathrm{U}(1)_Y$), yielding the photon and Z boson fields as linear combinations of the B_μ and W^3_μ:

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu; \qquad m_A = 0; \tag{2.6}$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W^3_\mu; \qquad m_Z = \frac{g_W}{2\cos\theta_W} v\,. \tag{2.7}$$

where the weak angle θ_W can be written in terms of the W and Z boson masses as $\cos\theta_W = m_W/m_Z$. The coupling constant of the Z is defined by $g_Z = g_W/\cos\theta_W$.

The $W^\pm$ bosons couple to fermions differing by one unit of electrical charge in charged-current (CC) weak interactions. The corresponding vertex factors are shown in figs. 2.2a and 2.2b. The weak interaction is parity violating since the $V-A$ structure ($\gamma^\mu - \gamma^\mu\gamma^5$) of the vertex includes the left-handed chiral projector, meaning that only left-handed particles (and right-handed antiparticles) participate in the CC weak interaction. The matrix element V_{ij} in the vertex factor of fig. 2.2b refers to the mixing of quark flavour states as mentioned in the previous section.



$\frac{-g_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)$ **(a)** $\frac{-g_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)V_{ij}$ **(b)** $\frac{-g_Z}{2}\gamma^\mu\left(g^f_V - g^f_A\gamma^5\right)$ **(c)**

Figure 2.2: Standard Model weak interactions of the fermion fields with their corresponding vertex factors.

The physical Z boson mediates weak neutral currents (NC) and strictly couples to fermions of the same generation. The Z couples to both left- and right-handed chiral states but not with

equal strengths[1]. The vector and axial couplings for a type of fermion f are:

$$g_V^f = (g_L^f + g_R^f) = T_3^f - 2Q^f \sin^2\theta_W; \tag{2.8}$$
$$g_A^f = (g_L^f - g_R^f) = T_3^f. \tag{2.9}$$

2.1.5 Flavour mixing in the quark sector

The weak CC vertex factors involving quarks include the matrix element V_{ij}. The weak force couples to the weak quark doublets $\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix}$, of which the down-type quarks d', s', b' are related to the mass eigenstates through the unitary Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.10}$$

With $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$, the CKM matrix can be parameterised in terms of three mixing angles and a single complex phase δ [17]:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}c_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \tag{2.11}$$

The CKM elements are determined using a global fit from all measurements [18] with the most recent (2016) results yielding:

$$\begin{aligned} \sin\theta_{12} &= 0.22497 \pm 0.00069 \\ \sin\theta_{23} &= 0.04229 \pm 0.00057 \\ \sin\theta_{13} &= 0.00368 \pm 0.00010 \\ \delta[^o] &= 65.9 \pm 2.0 \end{aligned} \tag{2.12}$$

The measured values of the mixing angles between quark and mass eigenstates, i.e. the off-diagonal elements of the CKM matrix, are relatively small. Therefore, quark mixing between different generations is suppressed.
Due to the non-zero complex phase δ of the CKM matrix, quark mixing through the weak interaction can lead to CP violation. The quark sector currently provides the only source of CP violation in the SM [16] with observational evidence established from e.g. hadron decays (direct CP violation), mixing of neutral mesons (indirect CP violations) and from interference between meson decays.

2.1.6 Neutral lepton flavour violation

The flavour of a neutrino is determined by the state (lepton flavour) produced in a CC weak interaction as shown in fig. 2.2a.
Contrary to the quark sector, the lepton sector obeys what is called an accidental symmetry of the SM Lagrangian and flavour changing currents are therefore not allowed. Nevertheless, experimental evidence of neutrino oscillations [9, 10, 11] show that transformations between

[1]This asymmetry arises because the associated B_μ field treats left- and right-handed chiral states equally whereas the contribution from W_μ^3 couples only to left-handed states.

neutrino flavour states are possible, given that neutrinos are massive.
Measurements from cosmology provide a non-zero upper limit (95% CL) on the masses of neutrinos: $\sum_{i=1}^{3} m_{\nu_i} \lesssim 0.21\,\text{eV}$ [19] but oscillations to due phase differences require also $m_1 \neq m_2 \neq m_3$ to occur and be observed. Thus, neutrino oscillations depend on the mass-squared splitting Δm_{12}, Δm_{23}, Δm_{13} between the different states.
While measurements of solar and atmospheric neutrinos can determine the mass-squared splittings (see fig. 2.3), the individual neutrino masses cannot be measured directly. The neutrino mass hierarchy is therefore yet to be resolved and the distinction between a normal (NH) and an inverted hierarchy (IH) has to be made.

Neutrino oscillations can be introduced to the SM using a method analogous to quark mixing: A unitary lepton mixing matrix is constructed to relate the weak eigenstates ν_e, ν_μ, ν_τ to a coherent linear superposition of mass eigenstates ν_1, ν_2, ν_3[2]. The mixing of neutrino states is described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix:

$$\begin{pmatrix}\nu_e\\ \nu_\mu\\ \nu_\tau\end{pmatrix} = \begin{pmatrix}U_{e1} & U_{e2} & U_{e3}\\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3}\\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3}\end{pmatrix}\begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix}. \tag{2.13}$$

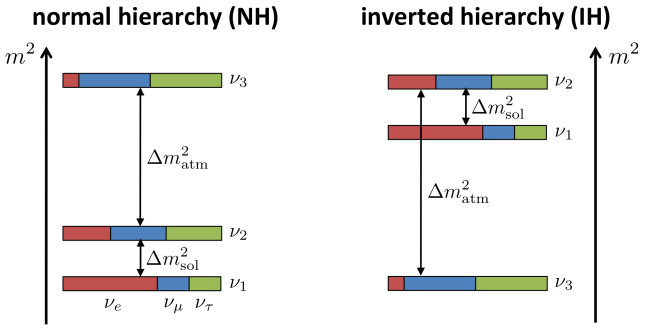


Figure 2.3: Normal and inverted neutrino mass hierarchy scenarios [20].

Any of the three mass eigenstates can be produced in conjunction with a charged lepton but we cannot know which mass eigenstate was produced at the vertex. The neutrino state ℓ will propagate as a superposition of all three mass eigenstates $|\nu_\ell\rangle = \sum_i U_{\ell i}\,|\nu_i\rangle$ until it interacts to produce another flavour of charged lepton.
The PMNS matrix can be parameterised by three mixing angles and a complex phase related to CP violation[3]. Fits to global data measures the following for the parameters (all measured in degrees) [21]:

$$\text{NH}: \quad \theta_{12} = 33.48^{+0.78}_{-0.75}, \quad \theta_{23} = 42.3^{+3.0}_{-1.6}, \quad \theta_{13} = 8.50^{+0.20}_{-0.21}, \quad \delta = 306^{+39}_{-70}; \tag{2.14}$$

$$\text{IH}: \quad \theta_{12} = 33.48^{+0.78}_{-0.75}, \quad \theta_{23} = 49.5^{+1.5}_{-2.2}, \quad \theta_{13} = 8.51^{+0.20}_{-0.21}, \quad \delta = 254^{+63}_{-62}. \tag{2.15}$$

2.1.7 Charged lepton flavour violation

In the context of a model including neutrino mixing as discussed in section 2.1.6, flavour violating decays involving charged leptons can be introduced through loop diagrams, as shown in fig. 2.4. The resulting rates for charged LFV processes are however subject to GIM suppression $\propto (\Delta m_{ij}/m_W)^4$ and due to the smallness of the neutrino masses, the predicted decay rates are extremely small and outside any detector scope [22]:

$$\text{BR}\left(Z \to \tau^\pm \mu^\mp\right) \sim 10^{-54} \tag{2.16}$$

$$\text{BR}\left(Z \to e^\pm \mu^\mp\right) \sim \text{Br}\left(Z \to e^\pm \tau^\mp\right) \lesssim 4 \cdot 10^{-60}\,. \tag{2.17}$$

[2]The mass eigenstates are the stationary states of the free particle Hamiltonian: $\hat{H}\psi = i\frac{\partial \psi}{\partial t} = E\psi$.
[3]Assuming Dirac neutrinos. Two additional phases arise if neutrinos are Majorana particles.

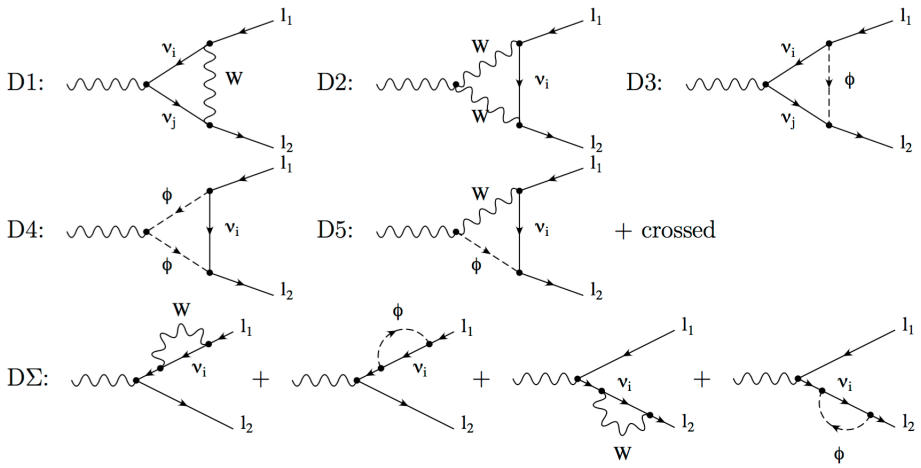


Figure 2.4: Contributing Feynman diagrams for the LFV $Z \to \ell_1 \ell_2$ decays in models including neutrino oscillations [22].

Although cross-generational CC couplings do occur in the SM, FCNCs are forbidden at tree level and are highly suppressed at higher orders by three independent factors [16]:

- Loop suppression: $\propto \alpha_W^2$;
- CKM suppression: the off-diagonal terms of the CKM matrix are small;
- GIM suppression: $\propto (\Delta m_{ij}/m_W)^4$.

As a result, any evidence of transitions between charged leptons prove a powerful sign of new physics phenomena. The SM bosons that could in principle facilitate FCNC and thus induce LFV decays in BSM scenarios are the Higgs and Z bosons.
Various theoretical models beyond the SM, such as the two-Higgs-doublet model (2HDM) [23], supersymmetric models [24] or models including extensions of the lepton and quark sectors can introduce FCNC and predict an increase in the branching fraction several orders of magnitude above the SM expectations.

The SM contains all possible dimension-4 operators, of which none allow for LFV processes. Including a set of all possible higher dimension operators that are suppressed at energy scales $M < m_Z$ decays is therefore a means of introducing effective flavour changing Z couplings. M is here the scale of new physics which could for instance be a new mediating particle.
The lowest order terms that can be introduced are dimension-6 operators which must then have corresponding coupling constants of mass dimension -2, e.g. $\propto g_Z^2/M^2$, in order to give a dimension-4 term to the Lagrangian.
The possible terms that can be added can introduce an effective vertex to the Lagrangian after symmetry breaking giving a $Z\tau\mu$ coupling [25]:

$$\frac{-g_W}{2\cos\theta_W}\begin{pmatrix}\bar{\tau} & \bar{\mu}\end{pmatrix}\gamma^\alpha Z_\alpha \times \begin{pmatrix} g_V - g_A\gamma^5 & m_Z^2[C_{L\mu\tau} + C_{R\mu\tau} - (C_{L\mu\tau} - C_{R\mu\tau})\gamma^5] \\ m_Z^2[C_{L\mu\tau} + C_{R\mu\tau} - (C_{L\mu\tau} - C_{R\mu\tau})\gamma^5] & g_V - g_A\gamma^5 \end{pmatrix}\begin{pmatrix}\tau \\ \mu\end{pmatrix}, \tag{2.18}$$

with $g_V = -\frac{1}{2} + 2\sin\theta_W$ and $g_A = -\frac{1}{2}$, the coefficients $C_{L\mu\tau}$ and $C_{R\mu\tau}$ taken real and where the Z is assumed to be on-shell.
Letting $C_{L\mu\tau} \simeq C_{R\mu\tau} \simeq g_Z^2/(16\pi^2 M^2)$ and $g_V \to 0$, the contribution to the $Z \to \tau\mu$ decays gives an estimate of the branching fraction:

$$BR(Z \to \tau\mu) \simeq 16 m_Z^4 \left(|C_{L\mu\tau}|^2 + |C_{R\mu\tau}|^2\right) \times BR(Z \to \mu\mu) \sim 1.7 \times 10^{-5} \frac{m_Z^4}{M^4} \tag{2.19}$$

The effective Z couplings therefore present a feasible probe of new physics for $M > m_Z$ at energies accessible at the LHC.

3 The Large Hadron Collider and the ATLAS Experiment

The European Organization for Nuclear Research (CERN) is located near Geneva, Switzerland and was founded on 29 September 1954 as a research facility devoted to the study of atomic nuclei and subatomic particle interactions. Since then, many important breakthroughs have been made at CERN, including the discoveries of W and Z bosons (1983), the first creation of antihydrogen (1995), the discovery of CP violation (1999), and the Higgs boson (2012) [1].

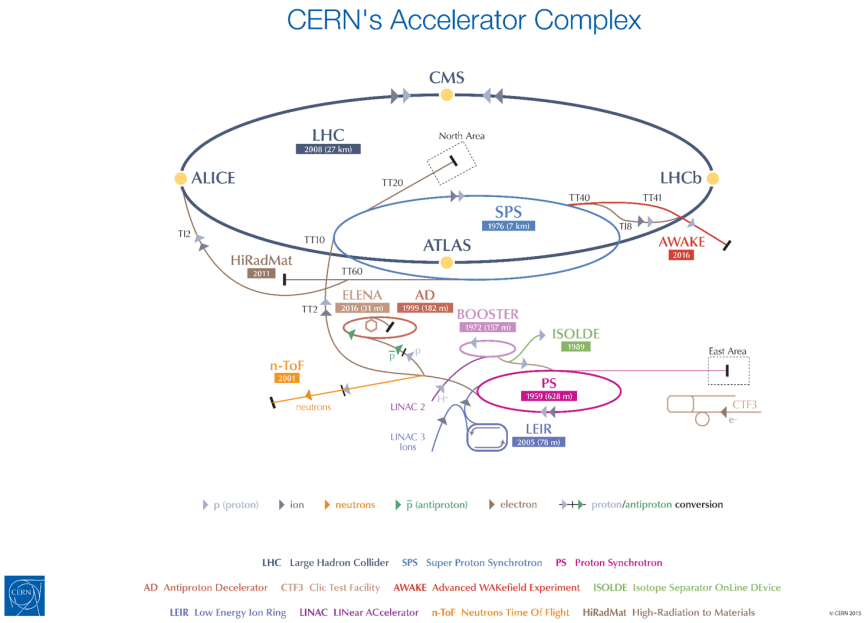


Figure 3.1: Overview of the CERN accelerator complex, including the LHC and its four large detectors [26].

3.1 The Large Hadron Collider

Several accelerators are currently in operation at CERN, the largest collaboration being the Large Hadron Collider (LHC). The accelerator was built into the existing 27 km tunnel beneath the Swiss-French border which was previously occupied by the Large Electron-Positron Collider (LEP). While in use (1989-2000), LEP was the most powerful lepton accelerator (at energies up to 209 GeV) but was shut down in favour of the construction of LHC. The powerful hadron collider has produced collisions of heavy ions and protons since its first research runs in 2010 (pp

collisions at $\sqrt{s} = 7$ TeV) which in turn has inspired a new era of analyses searching for physics beyond the Standard Model.
The proton source for the experiments is hydrogen gas where an electric field is applied to strip electrons off the hydrogen atoms. Protons are then sent through a series of accelerators to increase the beam energy. Linac 2 (fig. 3.1) initially accelerates the protons to 50 GeV after which the beam is gradually pushed to 450 GeV by a series of proton synchrotrons (Booster, PS and SPS).
The protons are split into two counter-circulating beam pipes that are kept at an ultra-high vacuum ($10^{-10} - 10^{-11}$ mbar) and cooled to 1.9 K to avoid protons colliding with residual gas molecules. Superconducting electromagnets inside the LHC ring provide a strong field which guides the proton beams around the accelerator, further accelerating them almost to the speed of light. The superconductors are cooled to 1.9 K using liquid helium [27]. Proton beams in the LHC consist of tight bunches of $\sim 10^{11}$ protons with a 25 ns nominal spacing. Bunches are made to collide at four interaction points along the ring corresponding to the position of the four detectors (ATLAS, CMS, LHCb and ALICE).
Several proton collisions occur at each crossing which is known as pile-up. The pileup corresponds to the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch. With L_{bunch} as the per-bunch instantaneous luminosity, σ_{inel} being the inelastic cross section (80 mb at 13 TeV) and f_r the LHC revolution frequency, the pileup is calculated as:

$$\mu = L_{\text{bunch}}\sigma_{\text{inel}}/f_r. \tag{3.1}$$

The number of events per second N is related to the cross section σ of the studied event and the beam luminosity L by $N = L\sigma$. ATLAS and CMS are designed for high-luminosity experiments and the current design luminosity of LHC is $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ for pp collisions [27].
The two proton beams collide in the center-of-mass frame but since protons are composite particles, the collision is rather an interaction among the constituent quarks and gluons. The center-of-mass energy of the incoming hadrons is converted into production of several particles, depending on the probability of the given process. Of the particles that can be produced from a collision, the only stable ones are electrons, protons and photons (and neutrinos, although largely undetectable).
Any unstable particle will travel a distance $d = \gamma v\tau$, where $\gamma = 1/\sqrt{1 - v^2/c^2}$ is the Lorentz factor, v is the speed of the particle, and τ is the mean lifetime in the particle rest frame. When produced in a high-energy collision, particles with long lifetimes ($> 10^{-10}$ s) such as $\mu^{\pm}$ can travel several metres before decaying. These can be detected directly, whereas most particles with shorter lifetimes decay close to the interaction point and only their decay products can be detected (typically seen as jets from tau leptons or quarks) [28].
The aim of a new physics search is therefore to identify the signature of the phenomenon in question based on careful measurements of the final state particles.

3.2 The ATLAS detector

ATLAS (A Large Toroidal LHC ApparatuS) is one of the four large LHC experiments. The entire ATLAS detector is a $7 \cdot 10^6$ kg cylindrical collection of layered subinstruments which is 46 m in length and 25 m in diameter. Its different components are used to identify particles produced in collisions and measure properties such as their momenta and trajectories. The detector is designed for general-purpose physics experiments and aims to test a wide range of predictions of

the Standard Model. The physics searches beyond the current theory range from unveiling new particles and interactions, to exploring baryon asymmetry, dark matter, the conditions of the early Universe, and the relationship between gravity and the other three fundamental forces.

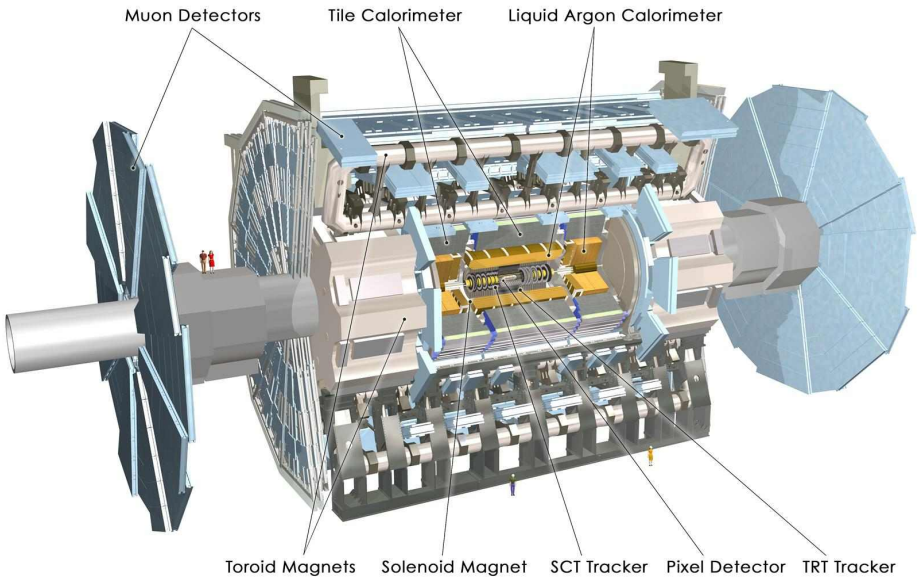


Figure 3.2: A diagram of the ATLAS detector and its subdetector components [29].

3.2.1 Coordinates

A right handed coordinate system with origin at the nominal interaction point (IP) in the center of the detector is used to describe the spatial coordinates of an event. The positive x-axis points to the center of the accelerator ring, y points upwards, and the z-axis lies along the beam line. Particles that travel through the detector are assigned a four-momentum:

$$p = (E, p_x, p_y, p_z)\,, \tag{3.2}$$

where the transverse momentum is $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$.
For the transverse plane, ATLAS uses cylindrical coordinates (r, ϕ), i.e. radial distance from the IP and azimuthal angle. Due to the cylindrical detector design, ATLAS covers the full range of ϕ. The spatial separation of final state particles is defined as:

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}\,. \tag{3.3}$$

In terms of the polar angle, θ, of the particle's three momentum relative to the z-axis, the pseudorapidity can be defined:

$$\eta \equiv -\ln\tan\left(\theta/2\right)\,, \tag{3.4}$$

where large values of $|\eta|$ correspond to forward tracks along the beam line. ATLAS covers the $|\eta| < 4.9$ range. Also commonly used is the rapidity:

$$y \equiv \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right)\,. \tag{3.5}$$

Both $\Delta\eta$ and Δy are useful quantities as they are invariant with respect to boosts along the z-axis.

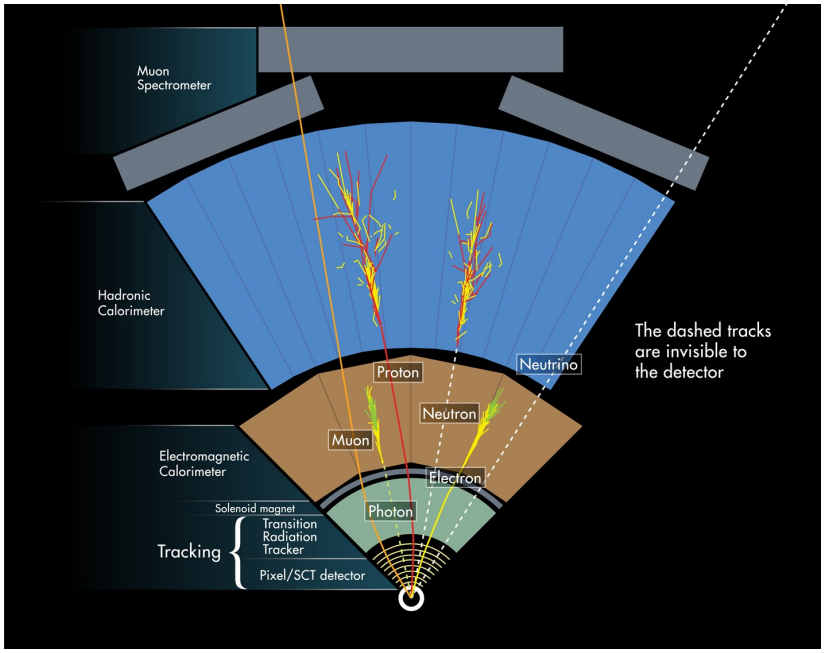


Figure 3.3: Transverse view of the detector components in ATLAS. As different particles travel outward from the interaction point, their typical signatures through the detector are indicated with solid lines, while particles that do not interact with the detector materials are shown with dashed lines.

3.2.2 Detector systems

The ATLAS detector is composed of several subdetector systems that are optimised to identify and track different particles based on the properties of their interaction in matter. The detector components and their pseudorapidity coverage are listed in table 3.1.

Detector		Region	η-range
Tracking	PD, SCT	Barrel + end-caps	$\lvert\eta\rvert < 2.5$
Tracking	TRT	Barrel + end-caps	$\lvert\eta\rvert < 2.0$
Solenoid magnet, 2T	-	-	-
EM calorimeter	LAr	Barrel	$\lvert\eta\rvert < 3.2$
Hadronic calorimeter	TileCal	Barrel + extended	$0.8 < \lvert\eta\rvert < 1.7$
Hadronic calorimeter	LAr	End-caps	$1.5 < \lvert\eta\rvert < 3.2$
Forward calorimeter	LAr	Forward	$3.1 < \lvert\eta\rvert < 4.9$
Toroidal magnets, 4T	-	-	-
Muon spectrometer	MDTs		$\lvert\eta\rvert < 2.7$
Muon spectrometer	CSCs		$2 < \lvert\eta\rvert < 2.7$
Muon spectrometer	RPCs	Barrel	$\eta < 1.05$
Muon spectrometer	TGCs	End-cap	$1.05 < \lvert\eta\rvert < 2.4$

Table 3.1: Pseudorapidity coverages of the ATLAS detector components in order of distance from the beam line [29]. The detector components are further described in the text.

Magnet systems

The two main superconducting magnet systems in the ATLAS Detector are used to bend electrically charged particles around the layers of subdetectors in order to contain particle tracks. Subject to the Lorentz force:

$$\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B}\,, \tag{3.6}$$

particles with non-zero charge q moving at a velocity $\boldsymbol{v}$ in a magnetic field $\boldsymbol{B}$ will be deflected. Hence, the degree of curvature depends on the momentum of the particle, and its direction on the particle charge. Low-momentum particles will be easily deflected while high-momentum particles require strong magnetic fields to deflect significantly. Since the strengths and directions of the magnetic fields are known, the deflection allows particle momenta and charges be measured in certain parts of the detector.
The central solenoid magnet surrounds the inner detector and produces a uniform 2 T magnetic field which is sufficient to bend the trajectories of highly energetic particles. The outer toroid magnets (1 barrel and 2 end-caps) are located between the calorimeters and the muon spectrometer. The toroidal magnetic field holds 4 T and ensures high muon momentum resolution while minimising multiple scattering events.

Inner detector

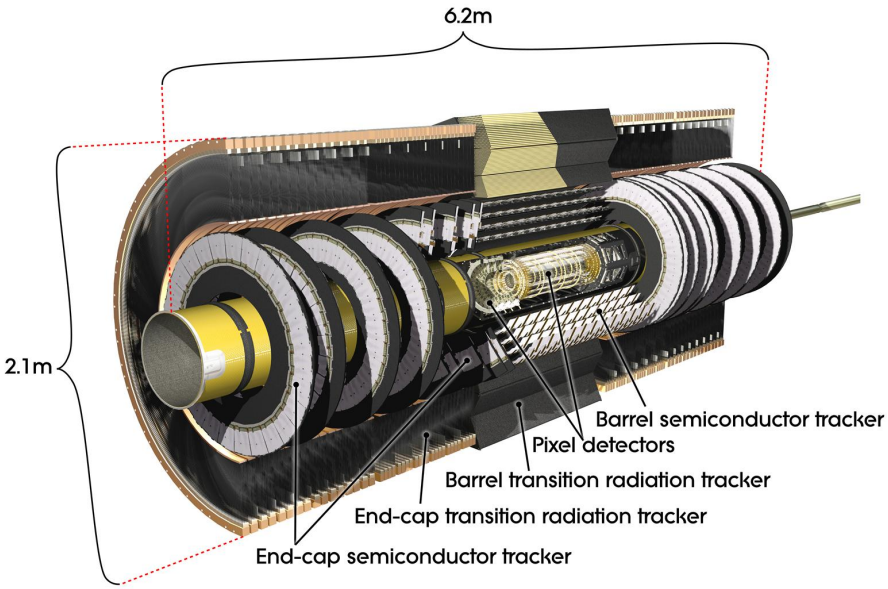


Figure 3.4: The ATLAS inner detector containing the pixel detector, semiconductor tracker and transition radiation tracker.

The inner detector (ID) shown in fig. 3.4 is the compact tracking volume of ATLAS, consisting of three sensor systems in concentric barrel layers: the pixel detector (PD), semiconductor tracker (SCT) and transition radiation tracker (TRT). This design ensures precision tracking at small radial distances and a large number of long-track measurements which combined provide pattern recognition, momentum measurements and electron identification.
Electrically charged particles that either traverse or decay in the tracking detectors leave a series of pixel hits. This provides track reconstruction at great precision which is also extremely useful

in measuring impact parameters and secondary vertex positions near the interaction point (IP). This is used for tagging of jets and hadronically decaying tau leptons.

Precision tracking closest to the IP is provided with the pixel and SCT detector. The pixel detector contains 1744 silicon pixel sensors which are arranged in three layers. In preparation for the 13 TeV center-of-mass energies of Run II, the ATLAS tracker was upgraded during LHC Phase I with a fourth pixel layer between the existing silicon pixel detector and a new, thinner beampipe - the Insertable B-Layer (IBL). Among the many purposes for the IBL are [30]:

- to improve tracking precision by detecting closer to the IP. This improves impact parameter reconstruction and in turn vertex and b-tagging performance.
- to increase tracking robustness in case of pixel failures in the other layers to avoid losing impact parameter resolution.
- to account for high-luminosity effects such as increased pileup which would cause readout ineffeciences and increased fake rates (e.g. if a jet fires a lepton trigger and is wrongly identified as a lepton).

The SCT silicon microstrip tracker is distributed over 4 barrel layers and 18 end-cap discs. More than $6 \cdot 10^6$ readout strips are spaced 80 μm apart on the silicon allowing tracking of charged particles in the plane perpendicular to the beam pipe with an accuracy of 17 μm per layer.

A larger number of hits per track is achieved with the TRT. It consists of $\sim 2 \cdot 10^5$ separate straw tubes which in the barrel are arranged parallel to the beam line and radially in both end-caps. When an electrically charged particle passes through a sequence of straws in the TRT, the gas in each straw is ionised and the resulting free electrons are driven to the central wire which produces a signal.
Electron identification is possible by means of transition radiation detection. A highly relativistic charged particle can emit transition radiation when traversing media of different refractive indices, and the probability depends on $\gamma = E/m$. The TRT straws are imbedded in arrays with polymer fibres (barrel) serving as the transition radiation material. An electron, which is the lightest charged particle and has a larger γ-factor, will be accompanied by more transition radiation photons compared to hadrons. As a result, additional energy is deposited in the straw gas which leads to a higher readout signal [31].

Calorimeters

The ATLAS calorimetry is designed to stop both charged and neutral particles in order to measure their energy deposit through the detector. The calorimeters are sampling calorimeters: They contain cells of an active medium in between layers of high-density material cells. An incident particle interacting with the active material produces additional particles (depending on its energy) that can further interact to create a shower of lower-energy secondary particles. These will ionise the dense material cells which record the degree of ionisation and the location. This enables reconstruction of the shape and energy of the shower. Hence, the energy of the incident particle can effectively be inferred from the total energy deposit.

The ATLAS calorimeter systems are based on the Liquid Argon (LAr) calorimeters operating with lead and liquid argon and the Tile Calorimeter (TileCal) based on scintillating tiles (emitting photons from ionisation). The combination of systems is designed to contain both

electromagnetic and hadronic showers from the outer muon systems.
The inner, electromagnetic calorimeter is dedicated to fully absorb and measure electromagnetically interacting particles, especially electrons and photons, with high granularity. The primary energy loss of electrons is through bremsstrahlung. The calorimeter uses lead and stainless steel with LAr in barrel and end-cap regions, giving a total $|\eta| < 3.2$ coverage.
Strongly interacting particles (primarily protons, neutrons and pions) are measured in the hadronic calorimeters. An accurate hadronic calorimetry is crucial for measurements of jets and missing transverse energy (usually due to neutrinos). The hadronic calorimeter has a TileCal barrel with extended barrel layers, and LAr end-caps.
Lastly, one electromagnetic and two hadronic layers form the forward calorimeters, thus extending the calorimeter range to $|\eta| < 4.9$. These are instrumented with LAr/copper and LAr/tungsten, respectively.

Muon spectrometer

The muon spectrometer (MS) forms the outermost detector parts beyond the toroid magnets, and has four subdetector components: monitored drift tubes (MDT), cathode strip chambers (CSC), resistive plate chambers (RPC) and thin gap chambers (TGC).
The purpose of the MS is to detect charged particles passing through the barrel and end-cap calorimeters. Tracks are deflected in the η-direction by the toroidal field, allowing momentum measurements of charged particles with dedicated precision-tracking chambers: MDTs and CSCs. These are arranged in three cylindrical layers along the barrel and three transverse wheels in the transition (overlap between barrel and end-cap toroidal fields) and end-cap regions. CSCs cover the innermost layer as they give superior time resolution
The system triggers on particles in $|\eta| < 2.4$. A separate system of fast trigger chambers deliver track information on the muon candidate within 15-25 ns. RPCs are used in the barrel and TGCs in the end-caps, and both trigger chamber types provide track coordinates (η, ϕ), beam-crossing identification and accurate measurements of the muon momentum.

Trigger systems

At the design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, the pp interaction rate is ~ 1 GHz. Data recording of events is however limited to ~ 1000 Hz so a highly efficient trigger system is necessary to give sufficient background rejection in real time while retaining a maximum efficiency of events of interest for physics analyses. For the ATLAS upgrade, the increased interaction rate motivated stricter event selection as a definite requirement.
The system has two trigger levels. The level 1 (L1) trigger uses combined hardware information from the calorimeters and RPCs and TGCs of the muon system to make fast decisions to further process events with high-p_{T} muons, electrons, photons, jets, and hadronically decaying taus [29]. This information is categorised into various trigger object definitions, based on e.g. quality and p_{T} thresholds. In particular after the upgrade, higher energy thresholds, object isolation, and topological information were employed in trigger decisions.
The next trigger level is the high level trigger (HLT) which bases decisions on algorithms run on a computing cluster. HLT examines *regions-of-interest* by measured energy and signature type in events accepted by L1. If accepted, the HLT converts the raw data to a reconstructed event and writes it to object data streams.

3.2.3 Data collected in Run 2

The analysis of this thesis is performed on data from pp collisions during Run 2 at the LHC.

With a $\sqrt{s} = 13$ TeV center-of-mass energy, the ATLAS Detector collected data suitable for physics corresponding to an integrated luminosity of 3.2 fb^{-1} in 2015 and 33.3 fb^{-1} in 2016. At the time of this analysis, the total amount of data available corresponded to $\mathcal{L} = 11.8\text{fb}^{-1}$.

Figure 3.5 shows the integrated luminosity delivered by the LHC and measured by ATLAS. The pileup profile, i.e. the mean number of interactions per crossing, for the 2015 and for the 16 April-16 September period of 2016 are shown in fig. 3.6.

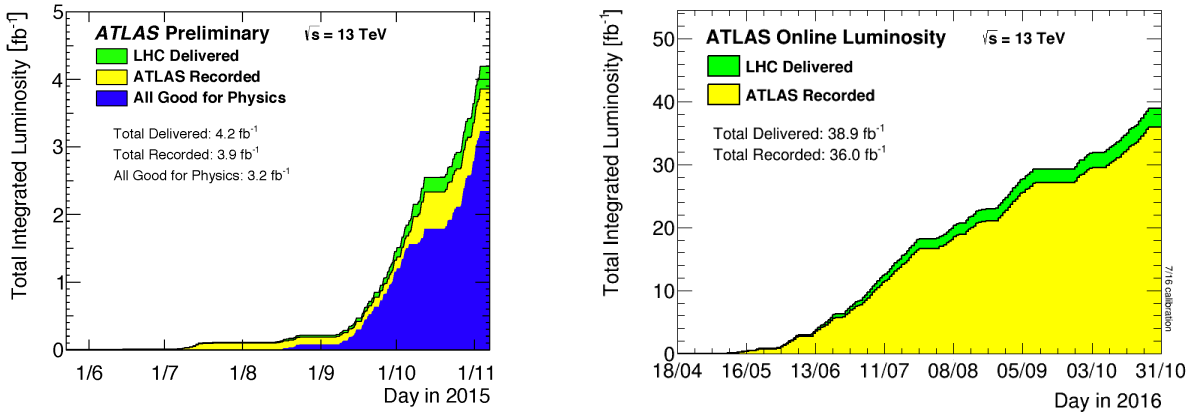


Figure 3.5: Integrated luminosity by day in 2015 and 2016 as delivered by the LHC (green), recorded by ATLAS (yellow) and approved for physics analyses (blue) [32].

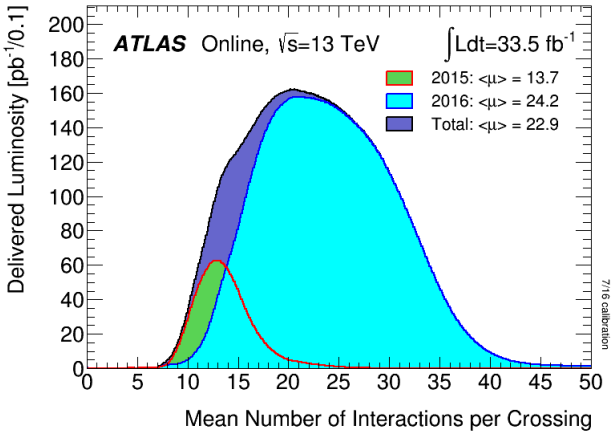


Figure 3.6: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2015 and 2016 (16 April-16 September) datasets [32].

4 | MC simulation and data

The data used in this analysis are events recorded by the ATLAS Experiment during Run 2 at the LHC. The total integrated luminosity at the time of analysis is $\mathcal{L} = 11.8\,\text{fb}^{-1}$ split by the 2015 ($\mathcal{L} = 3.2\,\text{fb}^{-1}$) and 2016 ($\mathcal{L} = 8.6\,\text{fb}^{-1}$) runs, respectively.
Monte Carlo (MC) simulated processes with samples from the MC15c production campaign within the ATLAS simulation framework are used for the signal and SM background modelling. The MC generators for the different samples are briefly described in this chapter.

4.1 Monte Carlo simulated samples

Several different MC generators are used to produce the relevant samples for the analysis. These are all listed in tab. 4.1. The signal $Z \to \tau\mu$ process is modelled with Pythia 8 [33] with the A14NNPDF23LO PDF tuning. All tau decay modes are considered in the generation, and the Z boson width is assumed unchanged by the addition of LFV models. A generator-level lepton filter with $p_\text{T} > 15$ GeV is used. The signal in the following is normalised to the DELPHI upper limit on the branching fraction $\text{BR}(Z \to \tau\mu) = 1.2 \cdot 10^{-5}$.

The Z + jets and W+ jets backgrounds are simulated using Sherpa 2.2 [34]. These samples are sliced in p_T, and $m_{\ell\ell} > 40$ GeV is used in the generation of these samples. As alternative inclusive samples, Powheg interfaced with Pythia 8 is used for Z + jets and W+ jets.

Diboson samples use Sherpa 2.1 for simulation of both hard scattering and parton showering with the A14NNPDF23LO tune used for the PDF [35].

Single-t and $t\bar{t}$ are generated in Powheg [36] interfaced with Pythia 6 [37] for the parton shower.

Detector simulation is performed for all samples via the GEANT4 [38] software to simulate the interaction of the generated particles with the ATLAS detector, and events are subsequently reconstructed. Furthermore, the DAOD_HIGG4D2 derivation originally designed for the $H \to \tau_\ell\tau_\text{had}$ analysis was used.
All MC events are weighted by scale factors for trigger, isolation, ID and reconstruction efficiences.

Process	Generator	Alternative generator
$Z \to \tau\mu$	Pythia 8	-
$Z \to \tau\tau$, $Z \to \mu\mu$, $Z \to ee$ (Z + jets)	Sherpa 2.2	Powheg + Pythia 8
W+jets	Sherpa 2.2	Powheg + Pythia 8
$t\bar{t}$	Powheg + Pythia 6	-
Single-t, Wt	Powheg + Pythia	-
Diboson	Sherpa	-
$H \to \tau\tau$	Powheg + Pythia 8	-

Table 4.1: MC samples used to model signal and SM background processes. The full list of MC samples can be found in Appendix A.

4.2 Data

Events are required to pass the Good Run List (GRL) event-by-event checks for bad detector or data quality. For 2015 and 2016 pp collisions, respectively, luminosity blocks of recorded data are selected using the GRLs:

`data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml`
`data16_13TeV.periodAllYear_DetStatus-v83-pro20-15_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml`

Jets and reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ are cleaned according to the CP group recommendations. Events with a LooseBad jet are rejected.

A candidate event is required to have at least one primary vertex with at least two associated tracks.

The data contains measurements from both the 2015 and 2016 data taking periods of Run 2 where both the pileup weights and integrated luminosities differ. MC events assumes some average amount of pileup $\langle\mu\rangle$ which ultimately affects the reconstruction efficiency. Pileup reweighing is therefore performed in order to correct for possible differences between the profile of instantaneous luminosity used in simulation compared to data.

5 | Signal and background

5.1 Signal characteristics

In this search for an LFV $Z \to \tau\mu$ signature, the signal of interest is:

$$pp \to Z \to \tau_{\text{had}}\mu \to \text{hadrons} + \nu_\tau + \mu\,. \tag{5.1}$$

The muon originating from the Z decay is expected to be energetic with a p_{T}-distribution peaking around $m_Z/2 \sim 46$ GeV as seen in fig. 5.1 (left). The decay must conserve charge, and since the Z boson is neutral, the charges of the muon and tau candidates must have opposite sign OS(τ, μ). Figure 5.1 demonstrates that by comparing the $p_{\text{T}}(\mu)$ distributions events of with either an opposite-sign (OS) or same-sign (SS) cut. Note that almost no signal events remain with the SS selection and that the background compositions are quite different.
The candidate tau lepton is required to decay hadronically. Unless otherwise specified, the notation τ assumes a hadronic tau lepton in the following.
Due to charge conservation, the $\tau^\pm$ must decay into an odd number of charged hadrons which are usually pions. The charged and neutral decay products are emitted in a narrow cone and leave either 1 or 3 tracks in the detector (an even number of tracks points to a bad object reconstruction):

- $\pi^\pm$ + 1, 2, 3 π^0 (1-prong taus);
- $\pi^\pm\pi^\pm\pi^\mp$ + 0, 1 π^0 (3-prong taus).

From the hadronic tau decay, the single tau neutrino emitted is expected to be almost collinear with the visible decay products. The neutrino passes through the detector and will be reconstructed as missing transverse energy $E_{\text{T}}^{\text{miss}}$.

The signal can be characterised by a range of kinematic properties which will be utilised to distinguish signal from the main backgrounds. These are described in further detail in section 5.3.
In particular, the transverse masses of the visible Z decay products with respect to the $E_{\text{T}}^{\text{miss}}$ are used to define a set of signal regions, i.e. phase space regions in which the $Z \to \tau\mu$ signal is enhanced. This method is based on the 8 TeV analyses for $Z \to \mu\tau_{\text{had}}$ and $H \to \mu\tau_{\text{had}}$ [39].
The correlation between the transverse mass of the $(\mu, E_{\text{T}}^{\text{miss}})$ and $(\tau, E_{\text{T}}^{\text{miss}})$ systems is plotted for signal in fig. 5.2 (left). Signal is focused at low $m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}})$ due to a soft $p_{\text{T}}(\tau)$ spectrum and a small opening angle to $E_{\text{T}}^{\text{miss}}$. Conversely, the average $p_{\text{T}}(\mu)$ and large opening angle to the $E_{\text{T}}^{\text{miss}}$ leads to intermediate values of $m_{\text{T}}(\mu, E_{\text{T}}^{\text{miss}})$.
The right hand side of fig. 5.2 shows the anticorrelation between opening angles in signal. Most events produce the leading (τ, μ) pair back-to-back with the $E_{\text{T}}^{\text{miss}}$ aligned with the tau.

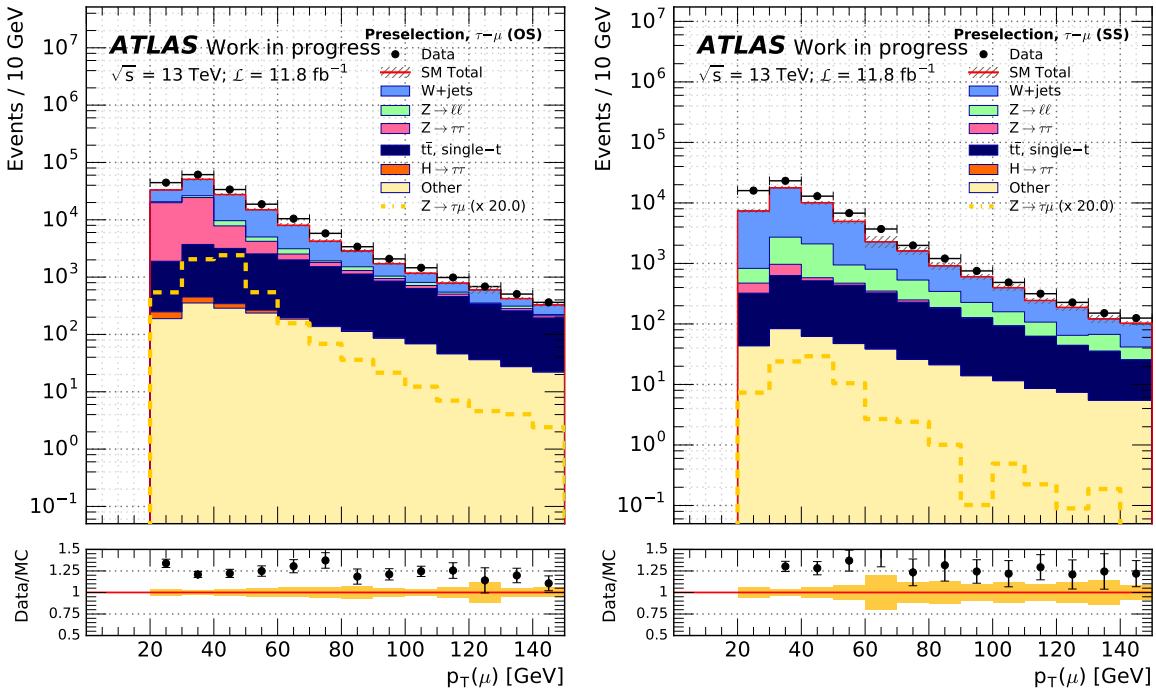


Figure 5.1: Distributions of muon p_{T} events with a reconstructed opposite-sign (τ, μ) pair (left) in comparison to same-sign events (right). Signal events and $Z \to \tau\tau$ background are largely depleted in the SS(τ, μ) selection whereas backgrounds such as $Z \to \mu\mu$ and W+jets that do not contribute with a real tau become dominant. Note that signal is scaled by a factor 20 and that the errors are statistical only.

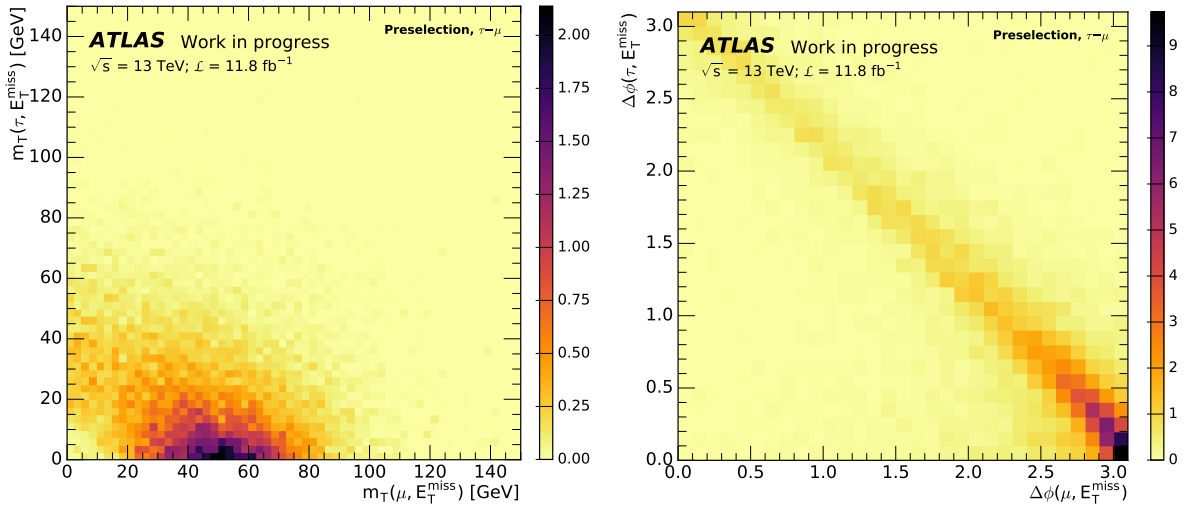


Figure 5.2: Correlations between $(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and $(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ in transverse mass (left) and opening angle (right). Figures show signal $Z \to \tau\mu$.

5.2 Background processes

The signal of interest is expected to be vastly dominated by a number of Standard Model processes that produce the same μ, τ final state. These include electroweak processes ($Z \to \tau\tau$, $Z \to \ell\ell$, single-t, $t\bar{t}$, W + jets, diboson (ZZ, WZ, $WW \to \ell\ell$) and $H \to \tau\tau$), as well as QCD multijets (events with quark- or gluon-initiated jets).
After object reconstruction to identify the final state, these backgrounds can be divided into three categories (due to the rate of misidentification being larger for hadronic taus than for electrons and muon):

- Events with a real muon and τ_{had};
- Events with a real muon and a jet misidentified as a τ_{had};
- Events with a lepton misidentified as a τ_{had}.

The main background contributions arise from $Z \to \tau\tau$, W + jets production where the W decays leptonically, and from $t\bar{t}$ events. These processes will be described in further detail in the following.
Hadronic jets from hard QCD scattering processes contribute to the background composition as well but have to be derived from data. This background and others that result from tau misidentification will be taken into account with a procedure described in section 7.3.

5.2.1 $Z \to \ell\ell$

Figure 5.3 shows the dominant process of Z production with the annihilation of $q\bar{q}$ to create either a photon or a Z boson. The subsequent decay produces a pair of opposite-sign same-flavour leptons.
In events with $Z \to \mu\mu$, a hadronic jet can be misidentified as a tau while one muon can be lost due to acceptance. The $Z \to \mu\mu$ contribution is largely reduced by requiring a good signal tau using a tight BDT requirement and, as seen in fig. 5.1, selecting only reconstructed (τ, μ) pairs of opposite sign.

$Z \to \tau\tau$ events can produce a final state with a real tau and a real muon if one tau decays hadronically and the other leptonically. The non-prompt muon will have a softer p_{T} spectrum but this decay is kinematically similar to the signal and provides one of the largest backgrounds in the analysis. A kinematic variable ($\Delta\alpha$) dedicated to target $Z \to \tau\tau$ background is discussed in section 5.3.

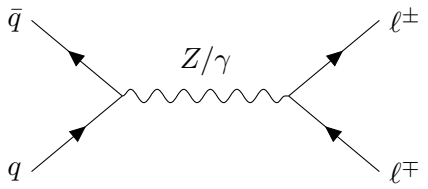


Figure 5.3: Feynman diagram showing the lowest order production of Z bosons and decay into charged leptons.

Since both the muon and the tau candidates are associated with neutrinos, the $E_{\text{T}}^{\text{miss}}$ angle with respect to the visible decay products will be quite different in comparison to signal.
This is seen in the distribution of opening angles to the $E_{\text{T}}^{\text{miss}}$ in fig. 5.6 (top right) which causes the smear of $Z \to \tau\tau$ events across low values of $m_{\text{T}}(\mu, E_{\text{T}}^{\text{miss}})$, $m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}})$.

5.2.2 W + jets

Leptonic W boson decays ($W \to \ell\nu$) can produce a final state with a real charged lepton and missing transverse energy. The high p_{T} muon can originate from either $W \to \mu\nu$ or $W \to \tau\nu$ with a leptonic decay of the tau while a jet produced alongside the W can be misidentified as a hadronically decaying tau. Figure 5.4 shows lowest order Feynman diagram contributing to W + jets production through the single- or dilepton decay modes.
The topology of W + jets events is different from the previously discussed processes since the reconstructed tau and muon do not originate from the same parent. With reference to fig. 5.6, the opening angle between the muon and $E_{\mathrm{T}}^{\mathrm{miss}}$ is always large since the majority of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is associated with jets while the real muon is produced in the W decay. Depending on the number of jets, the tau can be either aligned or anti-aligned with the $E_{\mathrm{T}}^{\mathrm{miss}}$. The distribution of transverse mass is focused around intermediate $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$, $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ values.

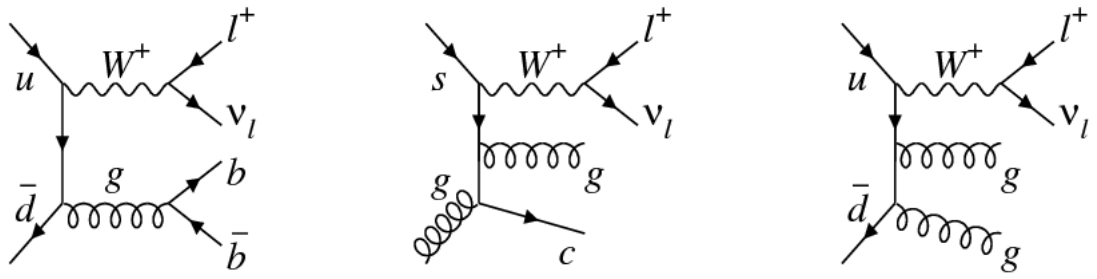


Figure 5.4: Examples of lowest order Feynman diagrams contributing to W + jets. Taken from [40].

5.2.3 $t\bar{t}$

Top quarks are mainly produced in pair production via the strong interaction. $t\bar{t}$ can become a source of background through double leptonic decays of top quarks ($t \to bW \to b\ell\nu$), or when one top decays hadronically ($t \to bW \to bq\bar{q}$) as seen in fig. 5.5. A jet can fake a hadronic tau. $t\bar{t}$ events show no immediate correlation between the $(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and $(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ systems (fig. 5.6). Top backgrounds will therefore not be constrained within this phase space but can be distinguished instead by the multiplicity of b-tagged jets.

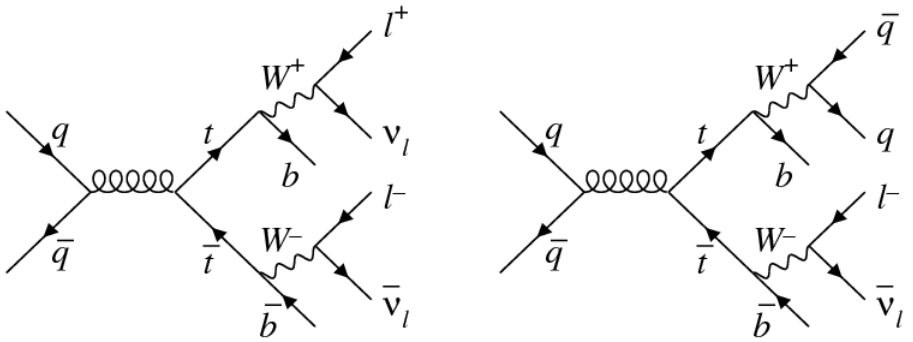


Figure 5.5: Top-antitop production. The final state leptons are produced via leptonic decays of both W bosons (left) and the single-lepton decay mode (right). Taken from [40].

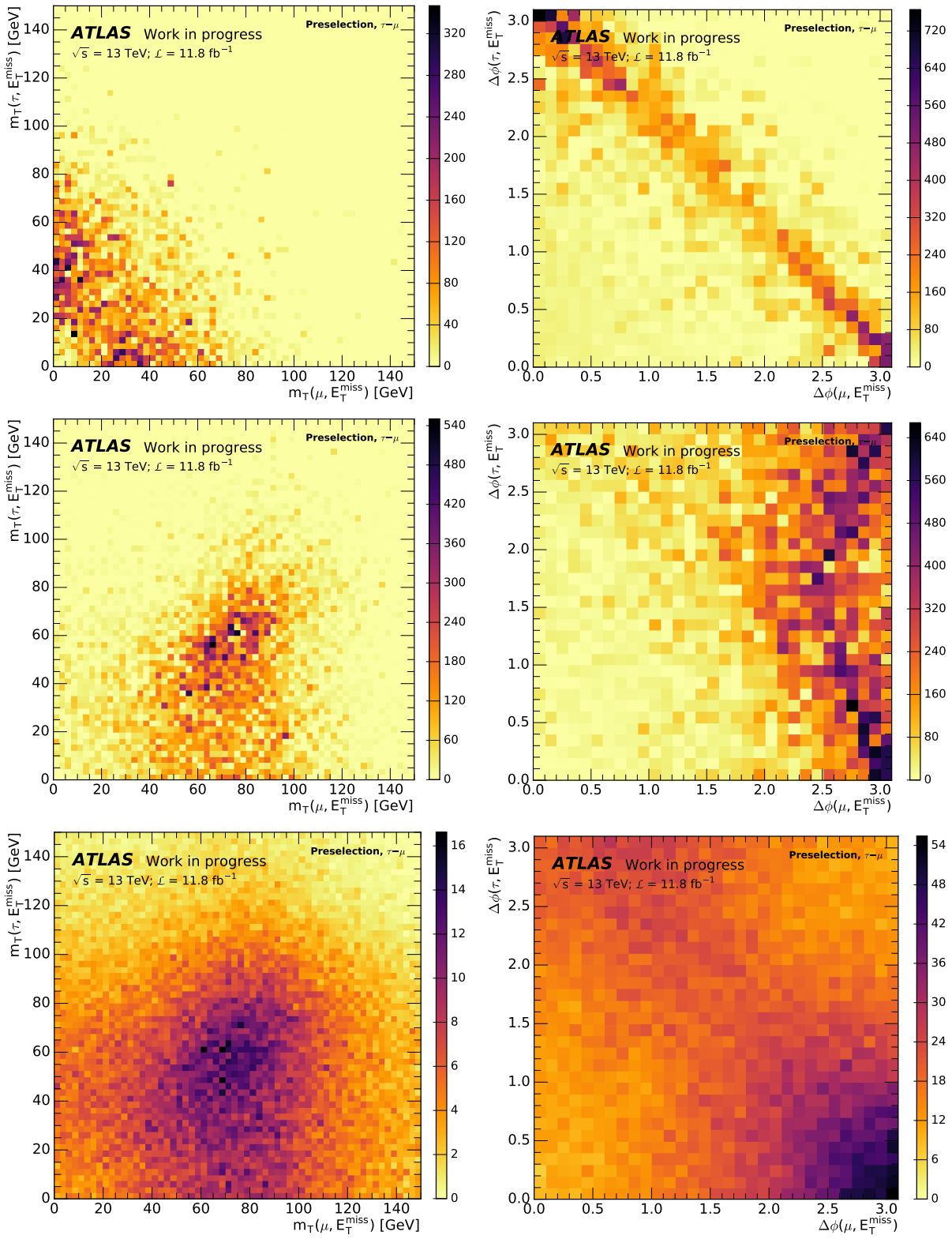


Figure 5.6: Correlations between $(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and $(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ in transverse mass (left) and opening angle in the transverse plane (right). Figures show the three major backgrounds at preselection: $Z \to \tau\tau$ (top); $W \to \mu\nu$ (middle); $t\bar{t}$ (bottom). Note the different scales of the z axes.

5.3 Discriminating kinematic variables

This section describes some of the variables that will be considered to discriminate signal from the various backgrounds in the analysis.

5.3.1 $\Delta\alpha$

The $\Delta\alpha$ discriminator is constructed to exploit differences in kinematics based on the number of neutrinos in an event. The kinematics are described in [25] for the $Z \to \tau_{\text{lep}}\,\mu \to (e\nu\bar{\nu})\mu$ decay, and are specifically targeted towards reducing $Z \to \tau\tau \to (e\nu\bar{\nu})(\mu\nu\bar{\nu})$ background.
Here, the method is adapted to the hadronic tau channel, thus implying a one-neutrino signal and a three-neutrino $Z \to \tau\tau$ background. Since the taus are boosted by at least a factor of $m_Z/2m_\tau$, we can assume collinearity of the tau decay products, and define the tau momentum:

$$p(\tau_{\text{had}}) = p(\tau) + p(\nu) \equiv \alpha\, p(\tau)\,. \tag{5.2}$$

The $Z \to \tau\tau$ process provides an additional tau which can imitate the signal final state by decaying leptonically. This tau is assigned a momentum $p(\tau_{\text{lep}}) = p(\mu) + p(2\nu) \equiv \beta\, p(\mu)$.
Neglecting the lepton final state masses, the following relations can be derived from the Z invariant mass:

$$m_Z^2 - m_\tau^2 = 2\alpha\, p(\tau)\cdot p(\mu) \qquad \text{signal} \tag{5.3}$$

$$m_Z^2 - 2m_\tau^2 = 2\alpha\beta\, p(\tau)\cdot p(\mu) \qquad \text{background.} \tag{5.4}$$

Equations 5.3 and 5.4 define the variable α_1.

With the approximation that $p_{\text{T}}(Z) = 0$, we obtain a second variable that is labelled α_2:

$$\alpha|p_{\text{T}}(\tau)| = |p_{\text{T}}(\mu)| \qquad \text{signal} \tag{5.5}$$

$$\alpha|p_{\text{T}}(\tau)| = \beta|p_{\text{T}}(\mu)| \qquad \text{background.} \tag{5.6}$$

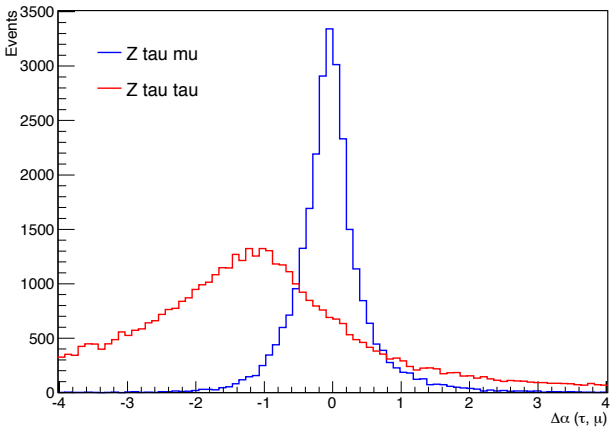


Figure 5.7: $\Delta\alpha(\tau,\mu)$ distributions for $Z \to \tau\mu$ and $Z \to \tau\tau$ after preselection cuts. The $Z \to \tau\mu$ distribution peaks sharply around zero.

$\alpha_1 - \alpha_2$ is used as the discriminating variable between signal and background:

$$\Delta\alpha = \frac{m_Z^2 - m_\tau^2}{2\,p(\mu)\cdot p(\tau)} - \frac{|p_{\mathrm{T}}(\mu)|}{|p_{\mathrm{T}}(\tau)|}\,. \tag{5.7}$$

The signal distribution peaks around $\Delta\alpha = 0$ by construction whereas backgrounds will exhibit some offset due to the β parameter. Figure 5.7 shows the $\Delta\alpha$ distributions of $Z \to \tau\mu$ and $Z \to \tau\tau$ for preselected events. A small amount of irreducible background can occur if the muon from the leptonic tau decay carries almost all the momentum of its parent so that the event has $\beta \to 1$ (i.e. the muon is emitted along the boost of the tau and carries $\sim (m_\tau^2 + m_\mu^2)/2m_\tau$).

5.3.2 Transverse mass

The transverse mass in terms of the lepton and missing transverse energy is defined as:

$$m_{\mathrm{T}}^2 = \left(p_{\mathrm{T}}(\ell) + E_{\mathrm{T}}^{\mathrm{miss}}\right)^2 - \left(\vec{p}_{\mathrm{T}}(\ell) + \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}\right)^2\,, \tag{5.8}$$

which, by assumption of massless decay products, can be rewritten as:

$$m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{2\,E_{\mathrm{T}}^{\mu} E_{\mathrm{T}}^{\mathrm{miss}}\left(1 - \cos\Delta\phi(\mu, E_{\mathrm{T}}^{\mathrm{miss}})\right)}\,. \tag{5.9}$$

5.3.3 Missing mass calculator

The missing mass calculator (MMC) [41] is a technique allowing for a accurate reconstruction of the invariant mass of the (τ, μ) pair. Although developed for events with a $\tau\tau$ final state such as $H \to \tau\tau$, the MMC also has an LFV mode that is currently only tuned for the Run 1 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution.
The main principle of the MMC technique is to attempt to solve a system of equations of multiple unknowns with the number depending on both the final state and the number of neutrinos involved. The MMC is constrained by the tau momentum and mass for leptonic and hadronic tau decay, respectively:

$$m_\tau^2 = [p(\ell) + p(2\nu)]^2, \qquad p^2(2\nu) = E_{2\nu}^2 - m_{2\nu}^2 \tag{5.10}$$

$$m_\tau^2 = [p(\tau_{\mathrm{vis}}) + p(\nu)]^2 \tag{5.11}$$

The system of equations is solved by scanning over all unknowns and choosing the most likely value. Each result is weighted by a probability such that the final result gives $m_{\tau\mu}^{\mathrm{MMC}}$ as the most likely mass. Since this result is correlated with the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, the MMC also considers the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as an input.
The MMC will be used to define a mass window that contains most of the signal in order to optimise cuts on other variables.

5.3.4 d_0 significance

The impact parameter is the distance between the point of closest approach of a track with respect to the primary vertex as illustrated in fig. 5.8. d_0 denotes the impact parameter in the transverse plane, and is computed for ID tracks from combined muons with respect to the primary vertex. The d_0 significance $d_0/\sigma(d_0)$ could prove useful in exploiting the difference in between signatures of prompt muons (from W or Z vertices) and muons with a displaced vertex originating from heavy-flavour jets.

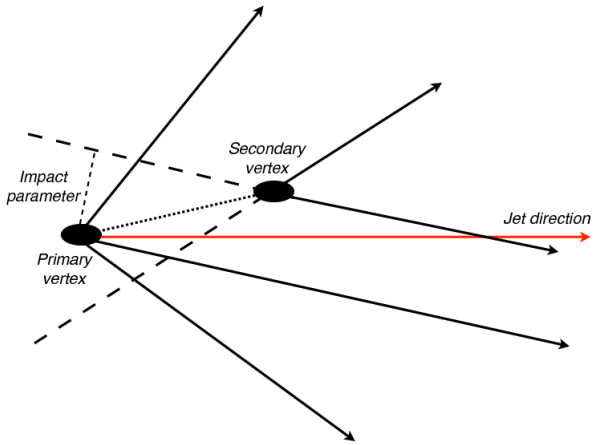


Figure 5.8: Schematic of the definition of impact parameter. Taken from [42].

6 Event selection

This chapter motivates the baseline event selection for the analysis.

6.1 Object definitions

Reconstructed candidate objects include electrons, muons, taus, jets and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and provide the baseline sample of well-identified objects. The object definitions follow the default recommendations for Run 2 analyses [43]. Object calibrations are implemented using SUSYTools [44].

6.1.1 Reconstructed objects

While the signal requires one good hadronic tau and one good muon, objects such as electrons and b-jets are reconstructed and used in order to veto specific backgrounds. The selection criteria for all objects are summarised in tab. 6.1.

Electrons
Electron identification relies on reconstructed tracks in the inner detector that are matched to clusters in the electromagnetic calorimeter. Electron candidates should pass the medium likelihood identification working point, and gradient isolation which is constructed to ensure background rejection efficiencies of $\varepsilon\,(25\,\mathrm{GeV}) = 90\%$ and $\varepsilon\,(60\,\mathrm{GeV}) = 99\%$ for leptons. Electrons with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.47$ excluding the $1.37 < |\eta| < 1.52$ transition region between barrel and end-caps are accepted. The requirements on the transverse and longitudinal impact parameters with respect to the primary vertex $|d_0/\sigma(d_0)| < 5.0$ and $|z_0 \sin\theta| < 0.5$ are implemented to reject cosmic events.

Muons
Muon candidates are identified by reconstructed tracks in the inner detector matched to muon spectrometer tracks (combined muons). Accepted muons are required to pass medium identification related to the number of hits in the inner detector and muon spectrometer [45] and gradient isolation. In addition to having $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$, muons have to satisfy $|d_0/\sigma(d_0)| < 3.0$ and $|z_0 \sin\theta| < 0.5$.

Taus
Hadronic taus produce collimated and isolated jets with energy deposits in both the electromagnetic and hadronic calorimeters. Identification of hadronic taus is therefore based on jets reconstructed with the anti-k_t clustering algorithm [46] using calibrated topological clusters of calorimeter cells with a distance parameter of $R = 0.4$.
The tau jet is expected to have a low track multiplicity due to the dominant hadronic decays into charged mesons. Candidates with either 1 or 3 charged tracks and with unit charge $|q| = 1$

are therefore selected. Furthermore, the baseline τ selection requires candidates with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ excluding the $1.37 < |\eta| < 1.52$ crack region.
Electron overlap removal is performed using the EleOLR of TauSelectionTools. The tool matches leading tau tracks with $p_{\mathrm{T}} > 4$ GeV to electrons within a $\Delta R < 0.4$ cone, and decorates the tau with a likelihood score of matched electrons.
The tau boosted decision tree (BDT) [47] is used to discriminate hadronic taus from quark- or gluon initiated jets. Tau candidates are required to pass a tight BDT-based ID selection.

Jets
Jets are reconstructed using the anti-k_t algorithm with $R = 0.4$. We accept $p_{\mathrm{T}} > 20$ GeV jets within $|\eta| < 2.5$. Jet energy calibration is performed with EM+JES [48]. The jet vertex tagger uses tracking and vertex information to suppress pileup jets, and jets are required to pass the medium JVT working point [49].

***b*-jets**
Identification of b-jets is performed using the MV2c10 tagger [50] in order to reject backgrounds containing b-quarks such as single-t and $t\bar{t}$. A fixed cut working point corresponding to a 77% average efficiency in $t\bar{t}$ events (FixedCutBEff_77) is used.
Selections on the number of b-jets in an event introduces a scale factor to account for misidentification.

Missing transverse energy
Particles that leave no trace in the detector are inferred as missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The $E_{\mathrm{T}}^{\mathrm{miss}}$ in this analysis is reconstructed using calibrated energies of all objects in an event with $E_{\mathrm{T}}^{\mathrm{miss}}$ SoftTerm calculated by the TrackSoftTerm algorithm [51].

6.1.2 Overlap removal

The procedure of overlap removal is performed between pairs of reconstructed objects where objects are kept in the order: muon, electrons, taus, jets and is applied using AssociationUtils [52] and TauAnalysisTools [52]. The condition for overlap removal is based on a certain threshold in $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ to reject the following objects:

- Jets within $\Delta R < 0.2$ of the leading tau track,
- Jets within $\Delta R < 0.4$ of an electron or muon candidate,
- Taus within $\Delta R < 0.2$ of an uncalibrated muon with $p_{\mathrm{T}} > 2$ GeV,
- Taus where the leading track passes loose electron identification.
- Taus within $\Delta R < 0.2$ of an electron or muon candidate,

Object	Requirement
Muons	Medium quality
	$p_{\mathrm{T}} > 25$ GeV
	$\lvert\eta\rvert < 2.5$
	Gradient isolation
	$\lvert d_0/\sigma(d_0)\rvert < 3.0$
	$\lvert z_0 \sin\theta\rvert < 0.5$
Electrons	Medium LLH
	$p_{\mathrm{T}} > 25$ GeV
	$\lvert\eta\rvert < 2.47$, excl. $1.37 < \lvert\eta\rvert < 1.52$
	Gradient isolation
	$\lvert d_0/\sigma(d_0)\rvert < 5.0$
	$\lvert z_0 \sin\theta\rvert < 0.5$
Taus	Tight BDT
	$p_{\mathrm{T}} > 20$ GeV
	$\lvert\eta\rvert < 2.5$, excl. $1.37 < \lvert\eta\rvert < 1.52$
	$\lvert q\rvert = 1$
	$N_{\mathrm{tracks}} = 1, 3$
	EleOLR and MuonOLR
Jets	$p_{\mathrm{T}} > 20$ GeV
	$\lvert\eta\rvert < 2.5$
	Medium JVT WP
b-jet tagging	MV2c10
	FixedCutBEff_77

Table 6.1: Summary of all selections used in the object definitions.

6.2 Preselection

After object definition, we impose preselection cuts in order to ensure that the μ, τ final state is selected.

6.2.1 Trigger

Events are required to have a candidate muon firing at least one of the lowest unprescaled single-muon triggers. Trigger-matching is applied to the leading muon. The muon triggers used are:

`HLT_mu20_iloose_L1MU15`	or	`HLT_mu40`	(2015);
`HLT_mu26_ivarmedium_L1MU15`	or	`HLT_mu50`	(2016).

L1MU15 requires the object to pass $p_{\mathrm{T}} > 15$ GeV in the L1 trigger system while also passing the HLT threshold. Isolation requirements are reconstructed online by the HLT from the inner detector tracks. Loose isolation requires $\mathrm{pTcone20}/p_{\mathrm{T}}(\mu) < 0.12$ where pTcone20 is the p_{T} sum of ID tracks (excluding the muon track itself) in a $\Delta R = 0.2$ cone. Medium isolation specifies $\mathrm{pTcone30}/p_{\mathrm{T}}(\mu) < 0.06$ within $\Delta R = 0.3$.

Whenever 2015 and 2016 data is used in combination, MC events are randomly treated as results from either year. Only the trigger requirement and pileup weight changes. Figure 6.1 shows muon p_{T} distributions split by 2015 and 2016 data and MC.

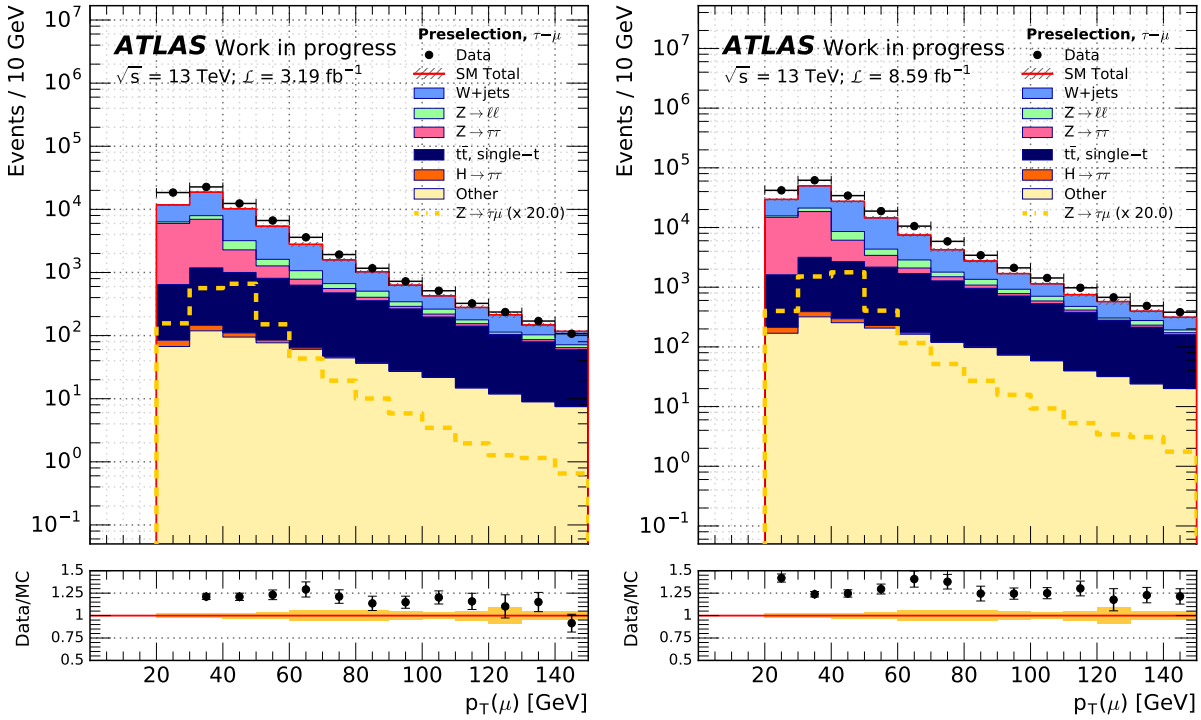


Figure 6.1: Preselection level $p_{\mathrm{T}}(\mu)$ distributions split by 2015 (left) and 2016 (right). The yellow dashed line indicates the $Z \to \tau\mu$ signal scaled by a factor 20. Backgrounds are estimated from MC only (i.e. no fakes from multi-jets are included yet), and no data derived scale factors are applied. Errors are statistical only.

6.2.2 Multiplicity selection

The multiplicity cut vetos events where at least one electron is identified and filters out events with more than one reconstructed muon and tau in the final state.

7 | Analysis

Based on MC simulated signal and background events, a selection strategy for the events of interest is devised. All cuts are optimised to yield the best significance of the signal in the selected region of phase space.
Results are presented with $\mathcal{L} = 11.8\,\text{fb}^{-1}$ of data, and with MC normalised accordingly. In addition, signal MC is normalised to the current upper limit on the branching fraction BR $(Z \to \tau\mu) = 1.2 \cdot 10^{-5}$.

7.1 Signal regions

The purpose of the cut selection and optimisation is to locate a region in phase space that is enhanced in signal. The data of the signal regions (SR) is kept blinded in order to eliminate bias until the analysis strategy and fit model are completed and verified.

After preselection cuts described in the previous chapter, events where a b-jet is identified are vetoed, the reconstructed (τ, μ)-pair is required to have charges of opposite sign. Several kinematic constraints are considered to suppress the background contributions. In particular, we use the topology of the LFV decay in m_T space, as described in sections 5.1 and 5.2, to motivate the selection of two orthogonal signal regions. The signal regions are chosen in such a way because a clear structure is seen at low $m_\text{T}(\tau, E_\text{T}^\text{miss})$ and high $m_\text{T}(\mu, E_\text{T}^\text{miss})$. However, there are also some events present at higher $m_\text{T}(\tau, E_\text{T}^\text{miss})$ and low $m_\text{T}(\mu, E_\text{T}^\text{miss})$. These events are likely to arise from soft taus where the τ, μ opening angle is large and the E_T^miss happens to be aligned with the muon.

7.1.1 Cuts and optimisations

Various kinematic variables are tested for the purpose of separating signal and background. These are shown in fig. 7.1 The cuts are selected and initially optimised for the three main background contributions: $Z \to \tau\tau$, $W \to \mu\nu$ (dominant channel in W+jets events), and $t\bar{t}$. As mentioned above, data is not used in the optimisation.
The variables selected for cut optimisation are $p_\text{T}(\mu)$, $p_\text{T}(\tau)$, $|\Delta\eta|(\tau, \mu)$ and $m_\text{vis}(\tau, \mu)$. These cuts are common to both signal region and are denoted as such. All distributions of kinematics after a cut is applied are shown in appendix B.

The optimal cut values are evaluated by a scan of the selected variables to calculate the relative separation power of the cuts in terms of the statistical significance of the signal with s and b as the signal and background yields, respectively [53]:

$$S = \sqrt{2\left[(s+b)\ln(1+s/b) - s\right]}\,. \tag{7.1}$$

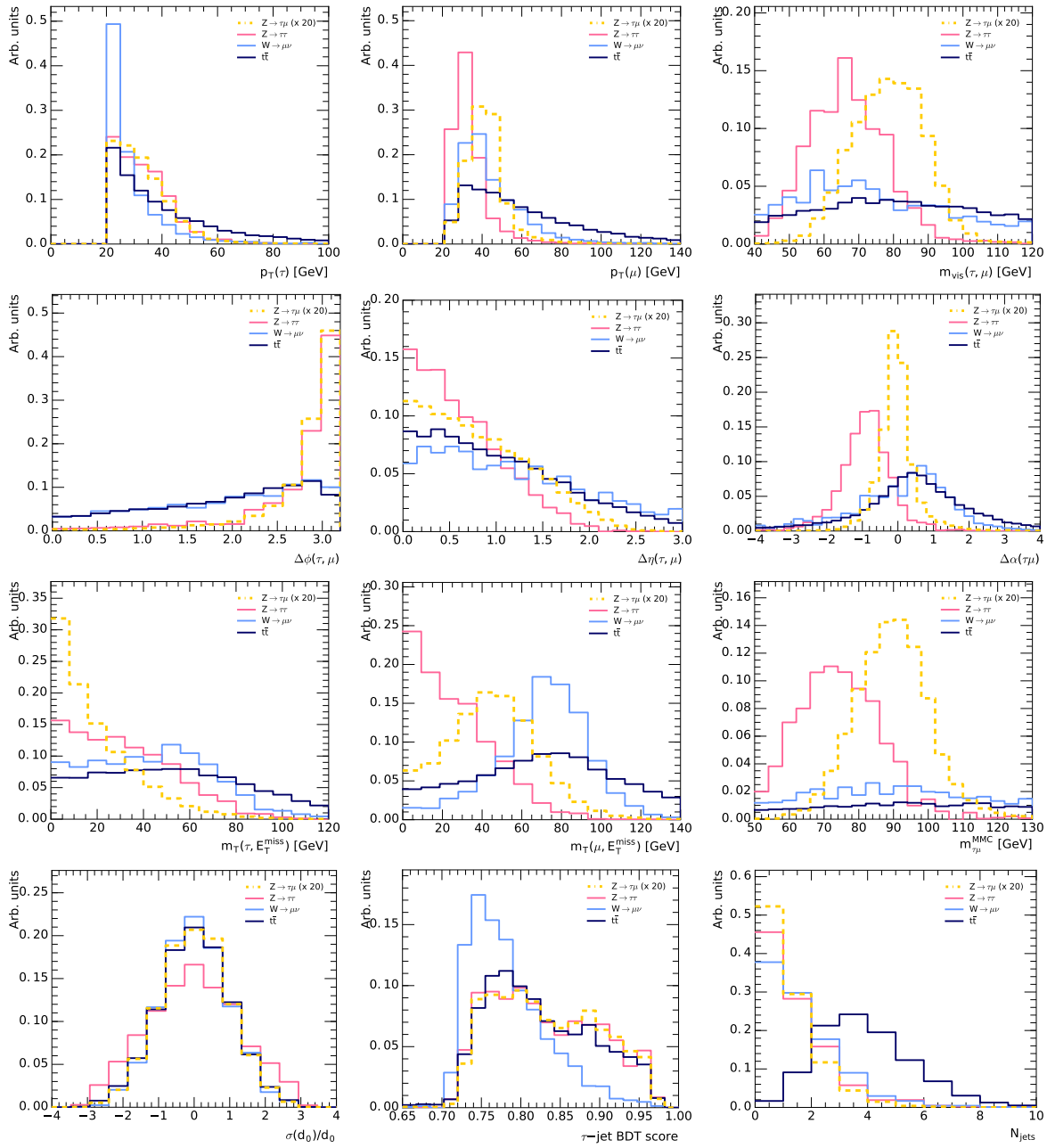


Figure 7.1: Kinematic variables considered for cuts. Figures are normalised to one and show signal and the three main backgrounds after preselection cuts, b-jet veto and OS(τ, μ). Signal is scaled by a factor 20.

7.1.2 Optimisation of cut selections

The cut selection is optimised through a scan of the chosen variables to locate the cut values that yield the largest significance for signal against all MC backgrounds in eq. 7.1. In the one dimensional optimisation plots, the greyed areas indicate the variable range that would be excluded by the cut yielding the maximum significance.

The primary region of interest that will be used is termed signal region 1 (SR1), and an $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ range is optimised to yield the maximum significance. Any signal events left over at low $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ will be included in a signal region 2 (SR2).

In order to optimise the significance in the region of interest, and not in the sidebands, an MMC window which contains approximately 95% of the signal is chosen.

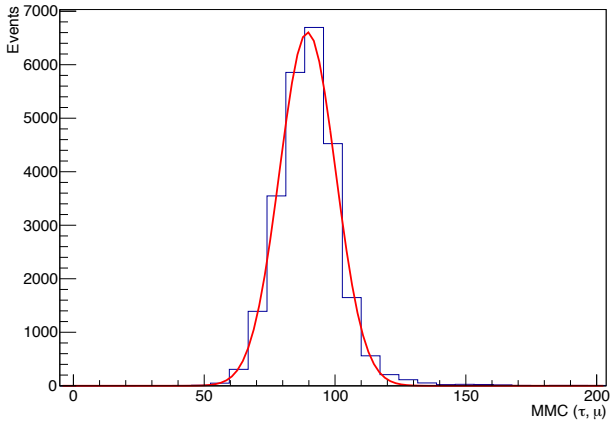


Figure 7.2: $m_{\tau,\mu}^{\mathrm{MMC}}$ distribution of $Z \to \tau\mu$ events from simulation after object selections.

Figure 7.2 shows the $m_{\tau,\mu}^{\mathrm{MMC}}$ distribution of simulated $Z \to \tau\mu$ events immediately after object selection cuts. The distribution is fitted to a Gaussian function with parameters $\mu = 90.17 \pm 0.05$ GeV and $\sigma_\mu = 11.09 \pm 0.04$ GeV. The $m_{\tau,\mu}^{\mathrm{MMC}}$ region of interest chosen is one that is within 2σ of the mean. This is rounded to:

$$70 \text{ GeV} < m_{\tau,\mu}^{\mathrm{MMC}} < 120 \text{ GeV}\,. \tag{7.2}$$

This $m_{\tau,\mu}^{\mathrm{MMC}}$ window contains $\sim 94\%$ of $Z \to \tau\mu$ events after object selection. The cut selection is optimised within this range after applying trigger and multiplicity cuts of preselection, as well as the b-jet veto and OS(τ,μ) requirements.

p_{T} cuts

The first kinematic selections considered are the cuts on $p_{\mathrm{T}}(\tau)$ and $p_{\mathrm{T}}(\mu)$. As seen in fig. 7.1, a lower cut on $p_{\mathrm{T}}(\mu)$ will reduce the amount of $Z \to \tau\tau$ and W + jets events.
Figure 7.3 shows the result of varying the minimum p_{T} cut values simultaneously. The significance for each cut selection is calculated using the sum of all MC backgrounds and the resulting maximum is found at $p_{\mathrm{T}}(\mu) > 30$ GeV and $p_{\mathrm{T}}(\tau) > 20$ GeV.

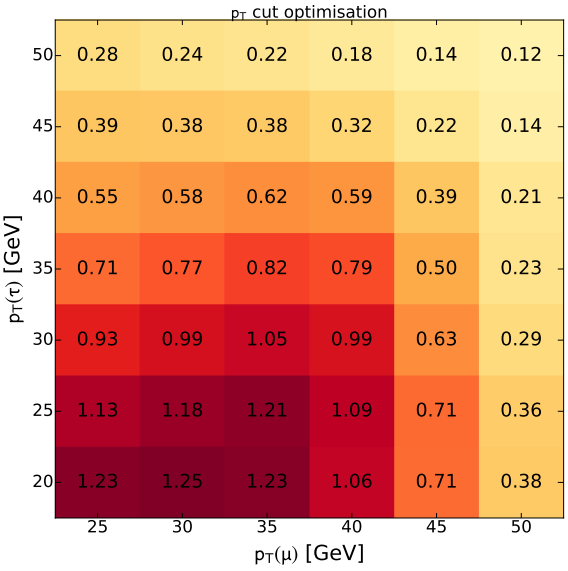


Figure 7.3: Simultaneous significance scans of a lower $p_{\mathrm{T}}(\mu)$ and $p_{\mathrm{T}}(\tau)$ selection. The optimisation is performed within 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV after preselection level, b-jet veto and OS(τ,μ).

$\Delta\eta(\tau,\mu)$ cut

After the p_{T} selection, a $\Delta\eta(\tau,\mu)$ cut was chosen in order to primarily target W + jets events (see the $\Delta\eta(\tau,\mu)$ distribution in fig. 7.1). Figure 7.4 shows the optimisation of an upper bound on $\Delta\eta(\tau,\mu)$. The selected cut is $\Delta\eta(\tau,\mu) < 2.0$.

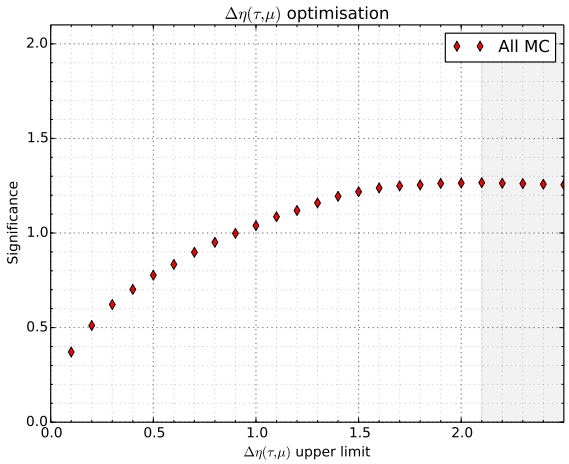


Figure 7.4: Significance scan of upper limit of $\Delta\eta(\tau,\mu)$ for the sum of all MC backgrounds. The optimisation is performed within 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV at preselection level including p_{T} cuts.

m_{vis} cut

Figure 7.5 shows the optimisation of a lower cut on the visible mass after preselection cuts, b-jet veto, OS(τ,μ), $p_{\mathrm{T}}(\mu) > 30$ GeV, $p_{\mathrm{T}}(\tau) > 20$ GeV and $\Delta\eta(\tau,\mu) < 2.0$. A selection of $m_{\mathrm{vis}}(\tau,\mu) > 60$ GeV is used to improve the significance of the signal regions. However, since the cut effectively removes the left sidebands of the mass distributions, the $m_{\mathrm{vis}}(\tau,\mu)$ cut will be removed at a later stage in order to perform a fit to e.g. the $m_{\tau,\mu}^{\mathrm{MMC}}$ distribution.

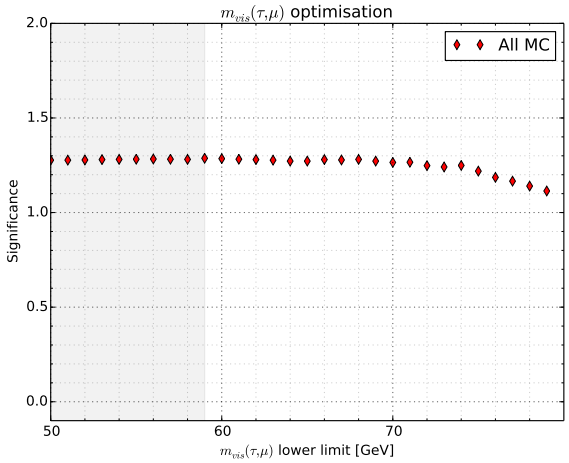


Figure 7.5: Significance scan of lower limit of m_{vis} for the sum of all MC backgrounds. The optimisation is performed within 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV after preselection including OS(τ,μ), b-jet veto, p_{T} and $\Delta\eta(\tau,\mu)$ cuts.

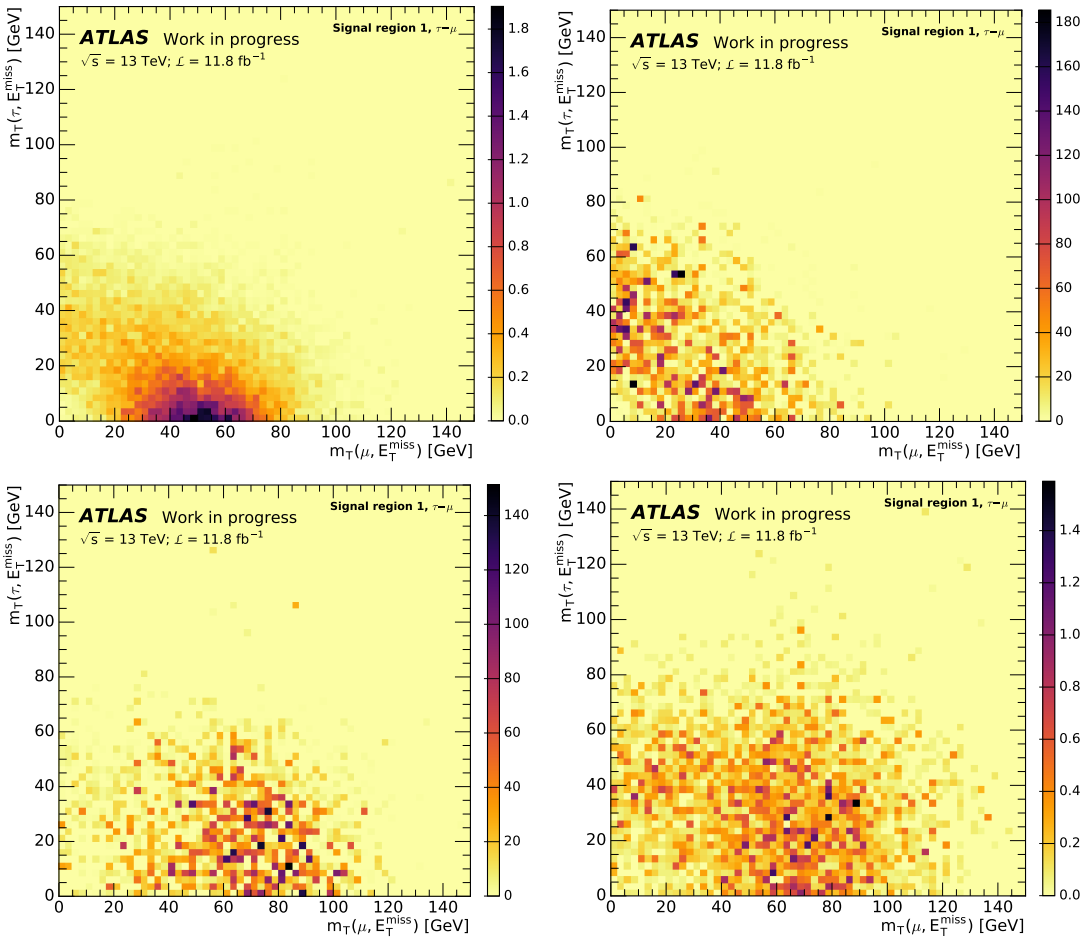


Figure 7.6: Distributions in the m_{T} plane after preselection and common cuts. Events are selected in the 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV window. Figures show MC $Z \to \tau\mu$ (top left), $Z \to \tau\tau$ (top right), W + jets (bottom left) and $t\bar{t}$ (bottom right).

SR1 m_{T} cuts

Figure 7.6 shows the two-dimensional distributions in m_{T} after preselection and all common cuts are applied. Based on the signal distribution located along the $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ axis and at low $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$, a primary signal region SR1 is defined by optimisation of rectangular cuts in the transverse mass plane.

The optimisation of SR1 is performed by varying the $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ lower bound and the $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ upper bounds simultaneously. In the $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ direction, it was required that $m_{\mathrm{T}}(\mathrm{low}) < m_{\mathrm{T}}(\mathrm{upp})$ and that the upper and that the lower cuts had a minimum separation of $\Delta m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) \geq 5$ GeV.

The result of the optimisation is shown in fig. 7.7.

Based on the optimisations in fig. 7.7, a maximum significance of $S = 1.34$ is found for $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 80$ GeV and $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 30$ GeV. At these values, the significance does not depend on the lower bound on $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ given that the bound is less than 25 GeV. The region is therefore defined by 15 GeV $< m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 80$ GeV and $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 30$ GeV.

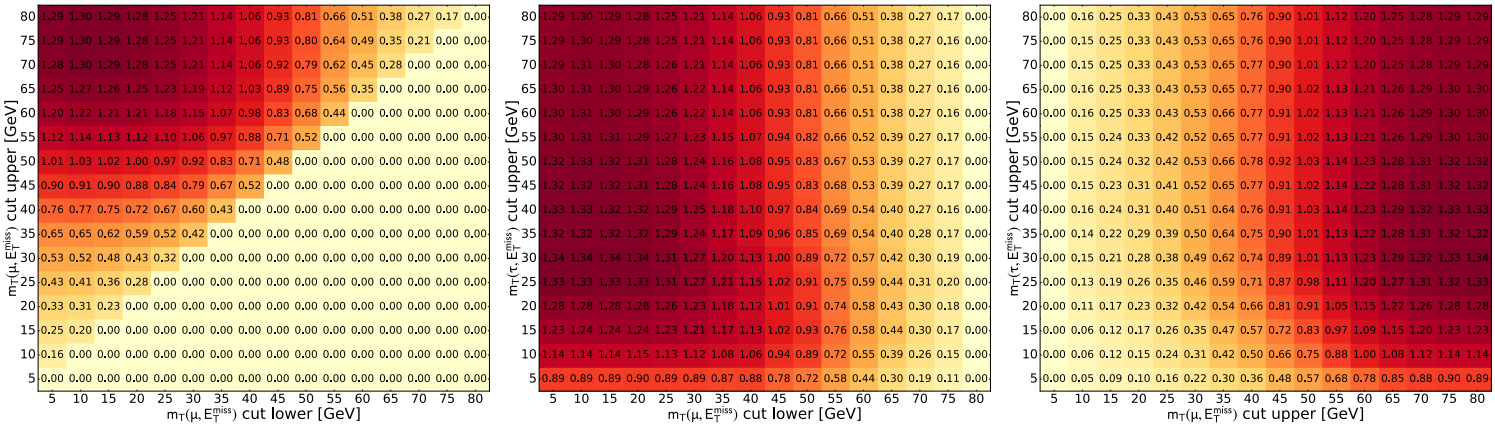


Figure 7.7: Simultaneous optimisation of the three cuts that define the m_{T} window for SR1. The figures show the projection on to the plane. Variation of lower $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and upper $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ with a minimum region width of $\Delta m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) \geq 5$ GeV (left); variation of lower $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and upper $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ (middle); variation of upper $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and upper $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ (right).

SR2 m_{T} cuts

The signal events that are not included in SR1 are selected by defining an SR2 orthogonally to SR1. With $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 15$ GeV, the $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ of SR2 is optimised as seen in fig. 7.8.
The maximum significance is found at approximately $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 60$ GeV. Since the significance quickly converges with increasing upper m_{T} limit and in order to reasonably define a $Z \to \tau\tau$ normalisation region at a later stage, the selected cut is put at $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 35$ GeV. This change excludes a large fraction of $Z \to \tau\tau$ events and only causes a decrease of 0.04 in the significance. The final SR2 selection yields a significance of 0.25.

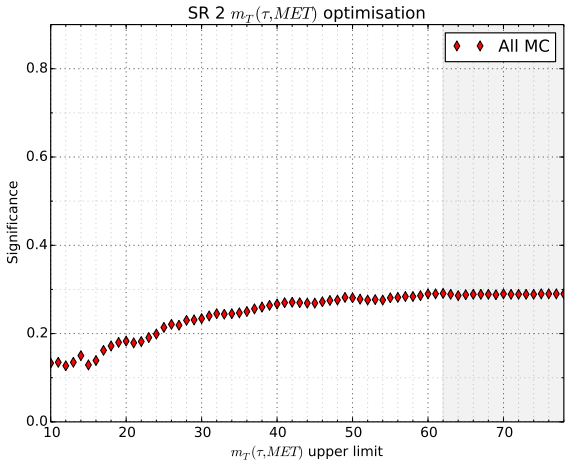


Figure 7.8: Optimisation of $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ upper limit in signal region 2 with $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 15$ GeV.

Table 7.2 gives a summary of the cuts used to select the signal regions and the full cutflow table for the final selection is shown in tab. 7.1. The signal efficiencies for passing the SR1 and SR2 selections are 31% and 2%, respectively.
For both signal regions, distributions of variables of interest $\Delta\alpha(\tau, \mu)$, $m_{\mathrm{vis}}(\tau, \mu)$ and $m_{\tau,\mu}^{\mathrm{MMC}}$ are presented in fig. 7.9.

Selection	$Z \to \tau\mu$	ε_{tot}	S	Other	ε_{tot}	$H \to \tau\tau$	ε_{tot}	$t\bar{t}$, single-t	ε_{tot}	$Z \to \tau\tau$	ε_{tot}	$Z \to \ell\ell$	ε_{tot}	W+jets	ε_{tot}
Initial	499 800.00	1.00	0.00	19 027 800.00	1.00	3 629 600.00	1.00	61 355 800.00	1.00	8 449 750.00	1.00	28 633 100.00	1.00	182 853 380.00	1.00
After obj. def.	50 421.00	1.00	0.00	204 729.00	1.00	348 343.00	1.00	1 309 487.00	1.00	300 415.00	1.00	367 231.00	1.00	861 058.00	1.00
After obj. def. (weighted)	538.84	1.00	0.50	7528.12	1.00	802.23	1.00	70 544.32	1.00	173 027.02	1.00	125 611.23	1.00	763 934.54	1.00
Preselection															
Lowest single-muon trigger	504.61	0.94	0.52	6392.07	0.85	629.74	0.78	60 524.37	0.86	116 079.21	0.67	104 733.23	0.83	658 679.09	0.86
1 tau	323.55	0.60	0.65	2815.09	0.37	412.04	0.51	23 716.89	0.34	71 910.04	0.42	22 109.41	0.18	128 438.86	0.17
1 muon, no electron	297.41	0.55	0.67	2274.22	0.30	272.63	0.34	19 972.18	0.28	46 899.43	0.27	11 922.70	0.09	113 276.61	0.15
Common selection															
b-jet veto	291.15	0.54	0.70	968.54	0.13	263.98	0.33	3425.01	0.05	45 473.28	0.26	11 601.80	0.09	110 119.71	0.14
OS (τ, μ) pair	287.39	0.53	0.80	699.29	0.09	261.79	0.33	2737.09	0.04	44 862.05	0.26	6446.74	0.05	73 503.60	0.10
70 GeV $<$ MMC $<$ 120 GeV	260.33	0.48	1.23	111.83	0.01	168.31	0.21	372.65	0.01	23 663.57	0.14	2863.97	0.02	17 660.48	0.02
$p_{\text{T}}(\mu) > 30$ GeV, $p_{\text{T}}(\tau) > 20$ GeV	237.73	0.44	1.25	99.13	0.01	137.73	0.17	343.04	0.00	16 702.54	0.10	2531.89	0.02	16 078.29	0.02
$\lvert\Delta\eta(\tau, \mu)\rvert < 2.0$	235.32	0.44	1.26	95.93	0.01	135.74	0.17	328.96	0.00	16 692.72	0.10	2400.94	0.02	14 901.21	0.02
$m_{\text{vis}}(\tau, \mu) > 60$ GeV	232.09	0.43	1.27	90.21	0.01	133.28	0.17	312.18	0.00	16 180.06	0.09	2373.48	0.02	14 198.00	0.02
Signal region 1															
15 GeV $< m_{\text{T}}(\mu) <$ 80 GeV	197.76	0.37	1.30	62.22	0.01	91.84	0.11	208.57	0.00	10 848.40	0.06	1912.44	0.02	10 127.26	0.01
$m_{\text{T}}(\tau) < 30$ GeV	164.55	0.31	1.34	33.52	0.00	55.84	0.07	97.73	0.00	7269.88	0.04	1305.47	0.01	6292.72	0.01
Signal region 2															
$m_{\text{T}}(\mu) < 15$ GeV	22.42	0.04	0.29	9.80	0.00	36.92	0.05	25.60	0.00	5209.21	0.03	314.23	0.00	429.57	0.00
$m_{\text{T}}(\tau) < 35$ GeV	13.21	0.02	0.25	3.80	0.00	10.07	0.01	10.32	0.00	2466.75	0.01	171.13	0.00	215.37	0.00

Table 7.1: Cutflow table normalised to 11.8 fb^{-1}. Final selection optimised in 70 GeV $< m_{\tau,\mu}^{\text{MMC}} <$ 110 GeV.

Preselection	
Lowest single-muon trigger	
1 tight tau, 1 muon, no electron	
Common selection	
b-jet veto	
OS (τ, μ) pair	
$p_{\mathrm{T}}(\mu) > 30$ GeV, $p_{\mathrm{T}}(\tau) > 20$ GeV	
$\lvert\Delta\eta(\tau,\mu)\rvert < 2.0$	
$m_{\mathrm{vis}}(\tau,\mu) > 60$ GeV	
Signal region 1	**Signal region 2**
$15\,\mathrm{GeV} < m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 80$ GeV	$m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 15$ GeV
$m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 30$ GeV	$m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 35$ GeV

Table 7.2: Summary of cut selection and signal region definitions.

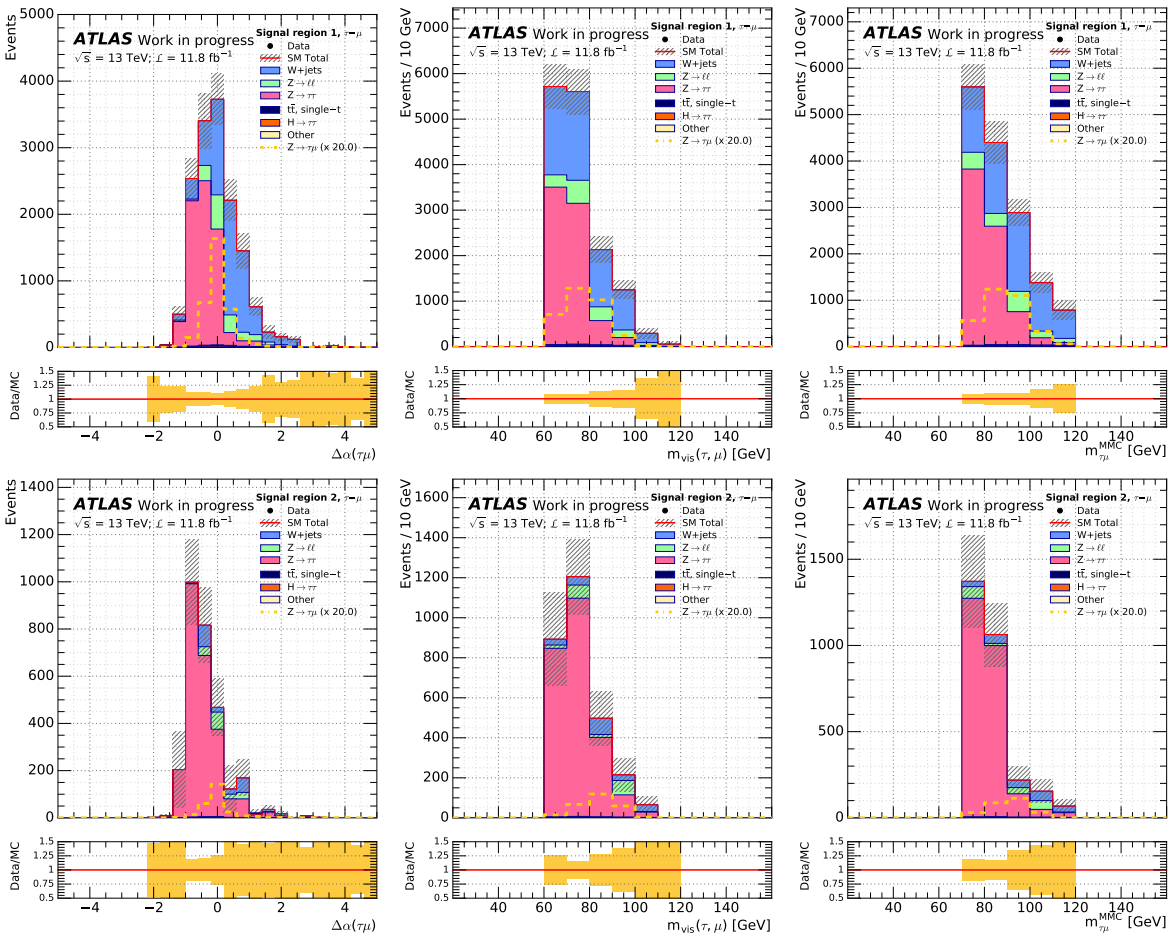


Figure 7.9: Blinded distributions of variables of interest in SR1 (top row) and SR2 (bottom row). Region definitions are summarised in tab. 7.2. Errors are statistical only.

Study of additional cuts in the signal regions

The distributions in fig. 7.9 reveal the main background contaminations in the signal regions. A large fraction of $Z \to \tau\tau$ events are present in both regions. Due to its location in the m_{T} plane, SR1 contains W + jets as well since these events dominate at intermediate to high transverse mass of the $(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and $(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ systems (fig. 7.6).

Therefore, the possibility of including an additional cut for a more targeted background rejection in each signal region was studied. The variables that were considered were $\Delta\phi(\tau,\mu)$, $|\sigma(d_0)/d_0|$ and $|\Delta\alpha(\tau,\mu)|$.
These individual cut optimisations are presented in fig. 7.10 for SR1 in the 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV optimisation range. The best significances are found for $\Delta\phi(\tau,\mu) > 0.0$, $|\sigma(d_0)/d_0| < 2.0|$ or $|\Delta\alpha(\tau,\mu)| < 0.5$. The same cut values were found in SR2.

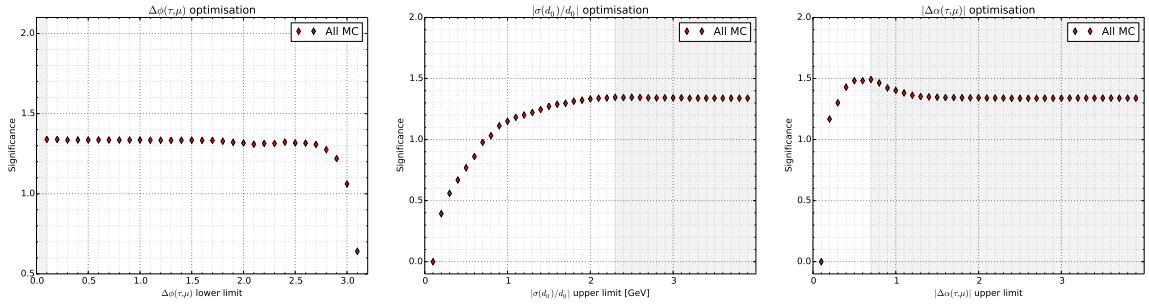


Figure 7.10: Optimisation of variables considered for an additional cut selection in the signal regions. Cut variations of $\Delta\phi(\tau,\mu)$, $|\sigma(d_0)/d_0|$ and $|\Delta\alpha(\tau,\mu)|$ are shown for the SR1 selection (see tab. 7.2) within the 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV optimisation range.

The $\Delta\phi(\tau,\mu)$ and $|\sigma(d_0)/d_0|$ cuts do not discriminate well against background since they leave the significance at the end of each signal region unchanged.
The $|\Delta\alpha(\tau,\mu)| < 0.5$ selection does in fact improve the significance compared to the default selection (see tab. 7.1). After the m_{T} cuts, significances of $S_{\mathrm{SR1}} = 1.51$ and $S_{\mathrm{SR2}} = 0.32$ are found. However, a result of this strict $|\Delta\alpha(\tau,\mu)|$ cut is that it effectively removes the sidebands of the mass distributions, such as$m_{\tau,\mu}^{\mathrm{MMC}}$ shown in fig. 7.11. This additional selection is therefore not used in the further analysis.

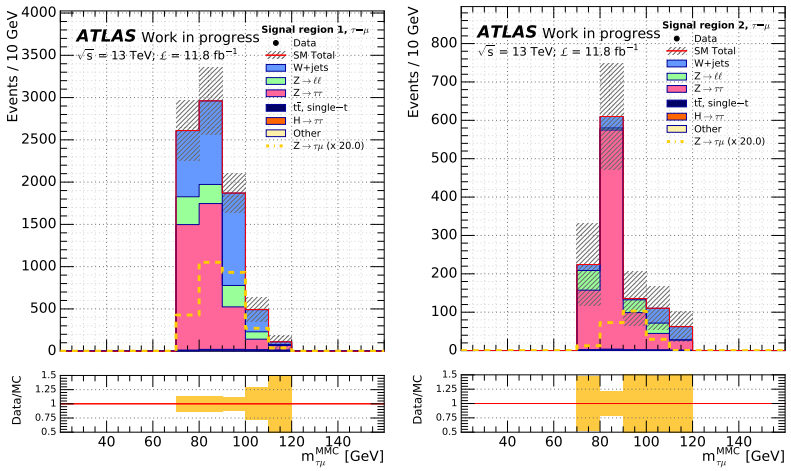


Figure 7.11: $m_{\tau,\mu}^{\mathrm{MMC}}$ distributions of each signal region when additional $|\Delta\alpha(\tau,\mu)|$ selection is introduced.

7.2 Background model

As discussed in chapter 5.2, events containing a jet misidentified as a hadronic tau can produce the final state of interest. The SM background contributions can therefore be split into two categories of relevance:

1. *real* processes in which a true muon and a true hadronic tau are produced ($Z \to \tau\tau$, $H \to \tau\tau$, $t\bar{t}$, single-t, diboson).

2. *fake* processes where the hadronic tau decay is faked by a jet (QCD multi-jets, W + jets, $Z \to \ell\ell$, diboson).

Backgrounds that produce a real final state signature rely on estimation from MC. These contributions are normalised to data using dedicated normalisation regions from which a data-MC scale factor will be extracted.

Since fakes are not well-described in MC, this contribution is also derived from data. Note that *real* processes can also contribute with fakes if a jet is misreconstructed as a tau.
To model background arising from fakes, a fake factor method will be used. The fake factor is essentially the ratio of pass-to-fail events for a given background and is therefore estimated from data in a set of control regions.

A third category containing fake events where a light lepton fakes a hadronic tau ($Z \to \ell\ell$, $H \to \ell\ell$) exists. This type of events is instead handled with the use of MC simulation.

Table 7.3 contains a summary of all regions dedicated to fake estimation and background normalisation and an overview of the m_{T}-based regions is shown in fig. 7.12.

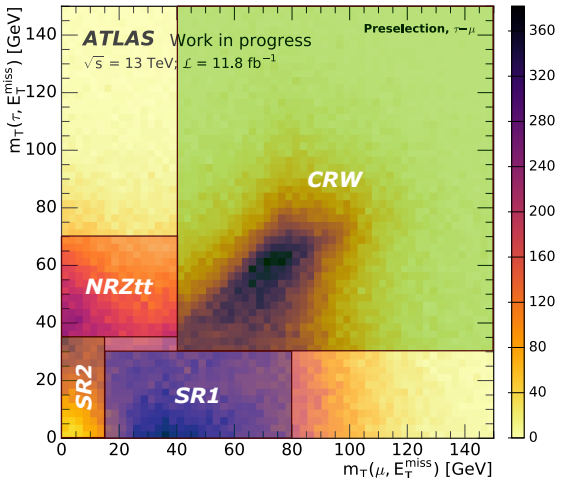


Figure 7.12: Regions that are defined by exclusive rectangular cuts in m_{T} space. The figure shows data at preselection.

Region	Process	Selection	Scale factor
NRZtt	$Z \to \tau\tau$	Preselection + OS(τ, μ)	$k = 0.90 \pm 0.08$
		$p_{\mathrm{T}}(\mu) > 30$ GeV	
		b-jet veto	
		$m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 40$GeV	
		35GeV $< m_{\mathrm{T}}(\tau), E_{\mathrm{T}}^{\mathrm{miss}}) < 80$GeV	
NRT	single-t, $t\bar{t}$	Preselection + OS(τ, μ)	$k = 1.07 \pm 0.02$
		$p_{\mathrm{T}}(\mu) > 30$ GeV	
		≥ 2 b-jets	
CRZll	$Z \to \tau\tau$	Preselection	
		$p_{\mathrm{T}}(\mu) > 30$ GeV	
		b-jet veto	
		No taus, two good muons	
		60 GeV $< m_{\mu\mu} <$ 120 GeV	
CRW	W + jets	Preselection + OS(τ, μ)	
		$p_{\mathrm{T}}(\mu) > 30$ GeV	
		b-jet veto	
		$m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) > 40$GeV	
		$m_{\mathrm{T}}(\tau), E_{\mathrm{T}}^{\mathrm{miss}}) > 30$GeV	
CRT	single-t, $t\bar{t}$	Preselection + OS(τ, μ)	
		$p_{\mathrm{T}}(\mu) > 30$ GeV	
		≥ 2 b-jets	
CRQ	QCD	Preselection + OS(τ, μ)	
		$p_{\mathrm{T}}(\mu) > 30$ GeV	
		b-jet veto	
		Non-isolated muon	

Table 7.3: Selection criteria that define the regions used for background normalisation and fake estimation.

7.2.1 Normalisation regions

The normalisation of the MC expectation to data is performed using a set of normalisation regions. For each of the major backgrounds that contribute with one real lepton and one real hadronic tau, we construct a normalisation region that is enhanced in the process of interest. These backgrounds include $Z \to \tau\tau$ and top (single-t and $t\bar{t}$). Minor backgrounds assume the yield from MC prediction.
The two regions are defined using preselection cuts (trigger and multiplicity). Events with one OS (τ, μ) pair are selected in the definitions.
Both normalisation regions are designed to avoid containing a significant amount of signal events. Their definitions are summarised in tab. 7.4.

NRZtt	NRT
Preselection + OS(τ, μ)	Preselection + OS(τ, μ)
$p_{\mathrm{T}}(\mu) > 30$ GeV	$p_{\mathrm{T}}(\mu) > 30$ GeV
b-jet veto	≥ 2 b-jets
$m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 40$ GeV	-
55 GeV $< m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) < 70$ GeV	-

Table 7.4: Normalisation region definitions.

$Z \to \tau\tau$ normalisation region

$Z \to \tau\tau$ events dominate across low $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ values. However, to avoid an overlap with the signal regions, the $Z \to \tau\tau$ normalisation region (NRZtt) is defined with a rectangular cut that is shifted upwards along the $m_{\mathrm{T}}(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}})$-axis in order to exclude SR2. Events in $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 40$ GeV and $35\,\mathrm{GeV} < m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 80$ GeV are selected. Furthermore, NRZtt requires a b-jet veto and $p_{\mathrm{T}}(\mu) > 30$ GeV.
The NRZtt distributions are shown in fig. 7.14.

Top normalisation region

Single-t and $t\bar{t}$ events are characterised by the presence of at least one b-jet from the top decays. Compared to the other background contributions, a large fraction of $t\bar{t}$ events show $N_{\mathrm{jets}} \geq 2$ and $N_{\text{b-jets}} \geq 2$ (see fig. 7.13). Since both the signal regions and $Z \to \tau\tau$ normalisation region impose a constraint of zero b-jets, the requirement of at least 2 b-jets in the top normalisation region (NRT) excludes these regions. The contamination from other backgrounds is reduced with $p_{\mathrm{T}}(\mu) > 30$ GeV which further removes W + jets and $Z \to \tau\tau$ events.

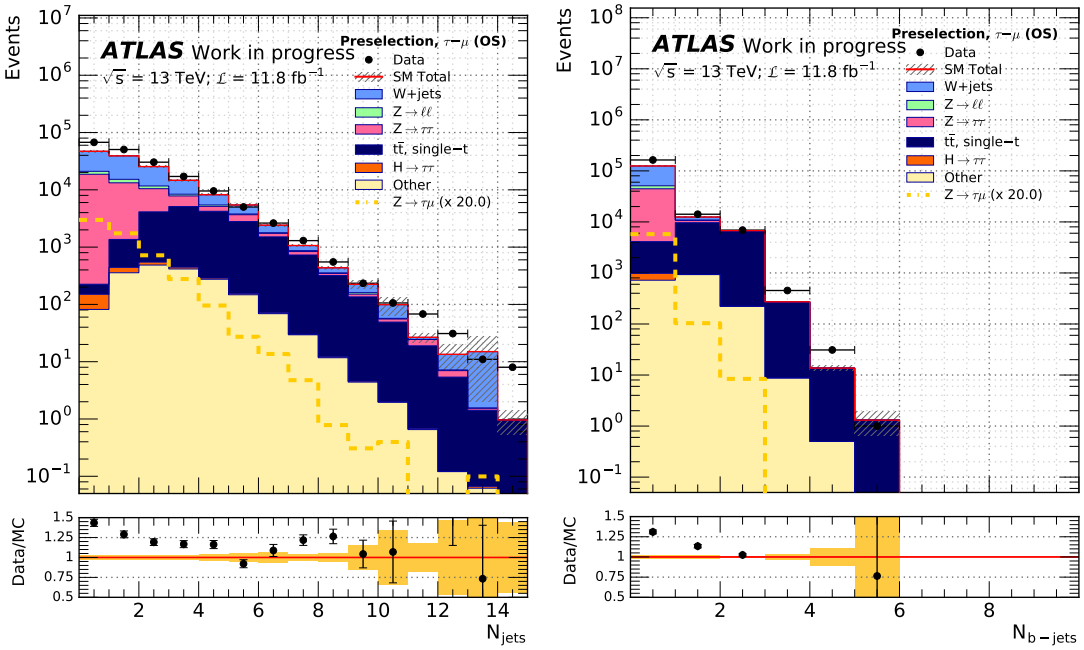


Figure 7.13: Number of jets and b-jets per event after preselection and OS(τ, μ) cuts.

7.2.2 Normalisation region scale factors

Scale factors to normalise MC to data are finally being extracted from a simulatenous likelihood fit of data in the normalisation regions. Only the real (truth-matched) components of the backgrounds are normalised by the derived scale factor. This is done to exclude contributions from fakes which will be estimated and treated separately.

Nonetheless, a crude estimate of the normalisation is carried out by calculating the scale factors using the following method.
From each unblinded normalisation region, a data derived scale factor is extracted to account for the discrepancy between observed data and the background predicted from MC:

$$k_{\mathrm{NR}_i} = \frac{N_{\mathrm{data}} - \sum_{j \neq i} N_{\mathrm{MC}_j, \mathrm{NR}_i}}{N_{\mathrm{NR}_i}} \equiv \frac{N_{\mathrm{res}}}{N_{\mathrm{NR_i}}}, \tag{7.3}$$

where N_{NR_i} is the number of events for the background that the normalisation region is designed for. The sum of other backgrounds remaining in the normalisation region $\sum_{j \neq i} N_{\mathrm{MC}_j, \mathrm{NR}_i}$ is subtracted from the data yield N_{data}.
The statistical errors $\sigma_i = 1/\sqrt{N_{\mathrm{NR}_i}}$ are propagated to estimate the error on the scale factor[1]:

$$\sigma(k_{\mathrm{NR}_i}) = k_{\mathrm{NR}_i} \sqrt{(N_{\mathrm{res}})^{-1} + (N_{\mathrm{NR}_i})^{-1}} \tag{7.4}$$

These normalisation region data/MC scale factors are calculated to make a first check whether there is need to correct MC background, meaning whether k deviates from 1 within its statistical error. The normalisation region scale factors calculated from eq. 7.3 are presented in tab. 7.5.

[1] $\sigma_k = \sqrt{\sum_i \left(\frac{\partial k}{\partial x_i}\right)^2 \sigma_{x_i}^2}$

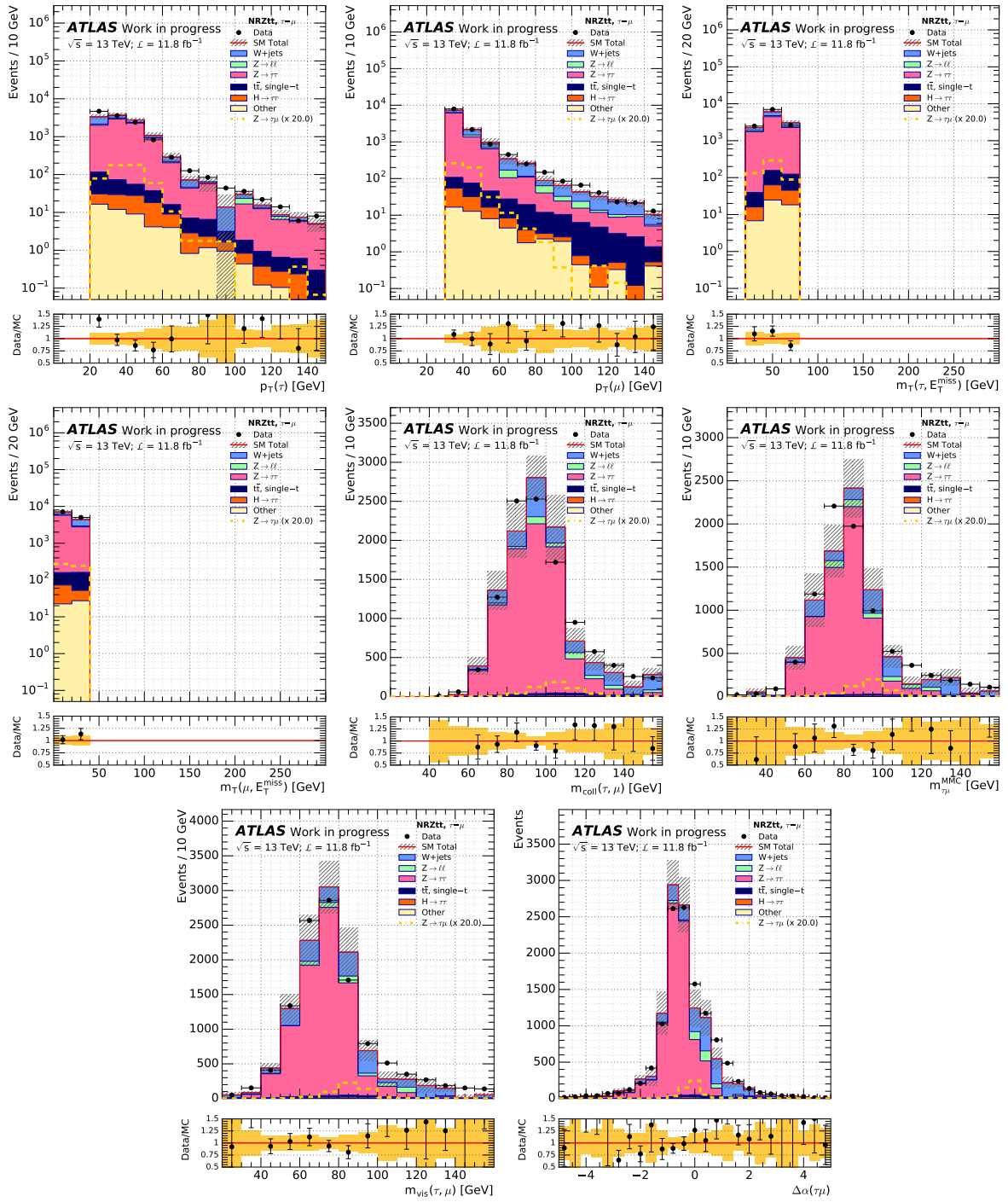


Figure 7.14: Distributions in the $Z \to \tau\tau$ normalisation region. Errors are statistical only.

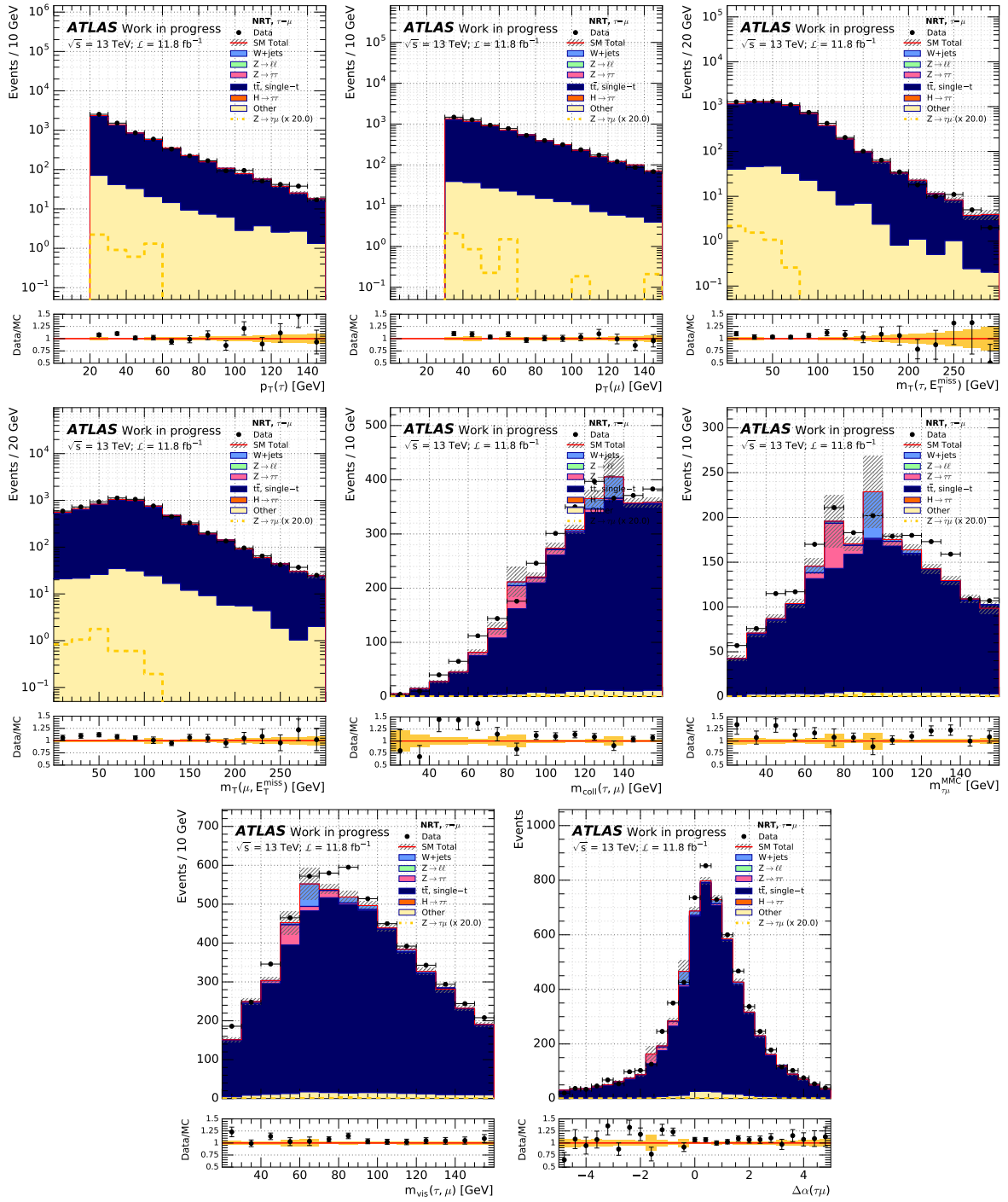


Figure 7.15: Distributions in the top normalisation region. Errors are statistical only.

	NRZtt	NRT
Calculated	$k = 1.08 \pm 0.02$	$k = 1.02 \pm 0.02$
From fit	$k = 0.90 \pm 0.08$	$k = 1.07 \pm 0.02$

Table 7.5: Data-derived scale factors to MC in the normalisation regions based on eq. 7.3 and the region definitions in tab. 7.4.

The HistFitter framework [54] is used to run a background-only fit in the normalisation regions with the normalisation as a free parameter. Both scale factors are rederived in this manner for a more accurate estimate of the data/MC ratio. Figure 7.16 shows $p_{\mathrm{T}}(\tau)$ in both NRs before and after the binned background fit is performed. The resulting scale factors are listed in tab. 7.5.

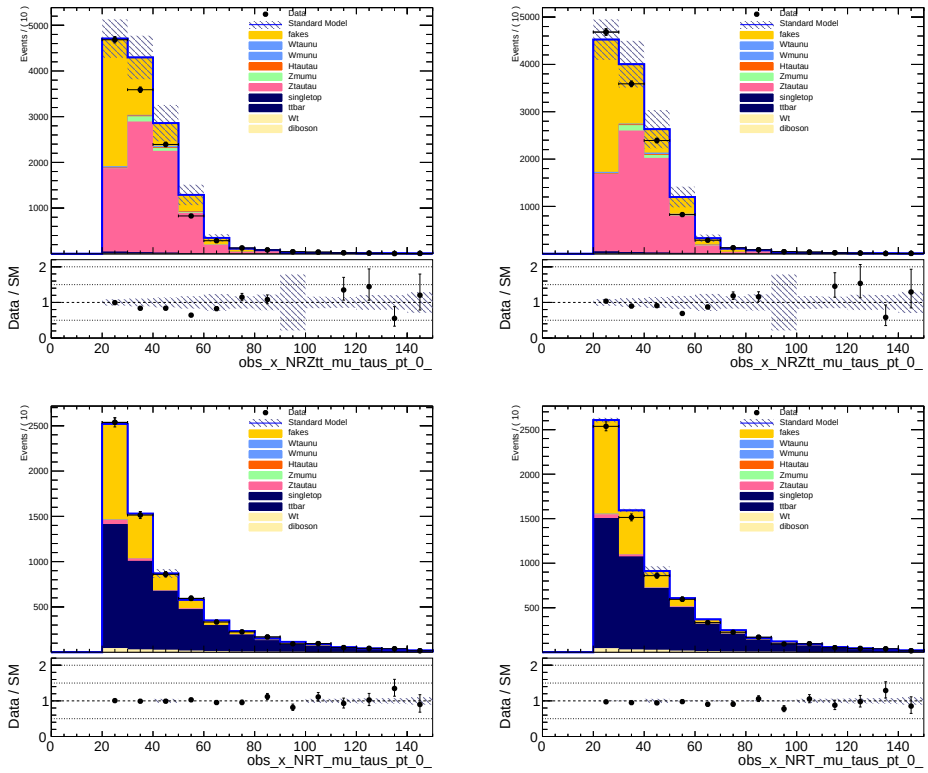


Figure 7.16: $p_{\mathrm{T}}(\tau)$ distribution in NRZtt (top) and NRT (bottom) before and after background fit. Note that background from fakes is included. Errors are statistical only.

Figure 7.17 shows m_{vis} distributions after the scale factors derived from the fit are applied in the NRT and NRZtt regions. In fig. 7.18, a comparison is shown of the distribution at preselection before and after applying the scale factors. The data/MC agreement appears worse after normalisation (right) since the fit model takes background contributions from fakes into account. This background will be estimated in the next section.

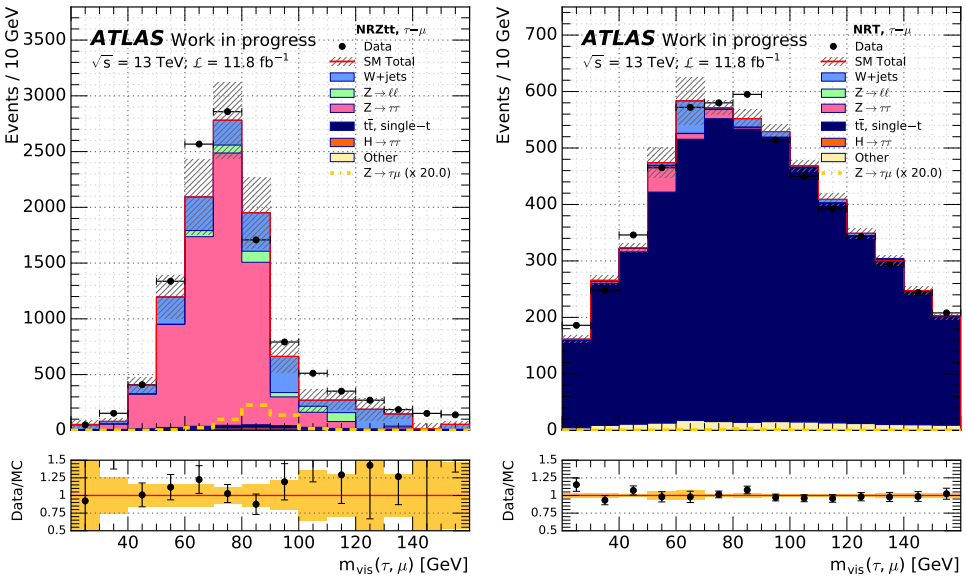


Figure 7.17: m_{vis} distributions showing scale factors applied in NRZtt and NRT.

7.3 Estimating background from fakes

The background from jets being misidentified as taus forms a dominant source background in the analysis. To model background arising from tau misidentification (also called fake background), a data-driven fake factor method will be used.

The fraction of fakes in events may depend on whether the jet is quark- or gluon-initiated. Since the quark-gluon fraction is expected to change for low tau BDT score, a dependence on this effect is reduced by excluding events with tau identification looser than the loose WP.

The calculation of fake factors relies on the assumption that the control region-derived fake factors are valid in the signal region or normalisation region they are propagated to and that the amount of QCD contamination in the non-QCD regions is negligible. The reason for the second assumption is that QCD events are not modelled in MC and therefore cannot be subtracted from other regions.

7.3.1 Fake factor method

To determine the amount of background from events with fake taus in a region interest, separate *fail* and *pass* categories of events are constructed.
Failed tau candidates must pass a reverted tau identification selection. The fail category consists of taus with loose or medium but non-tight BDT which ensures a large sample of events where a tau is faked by a jet.
Objects of the pass category are required to pass the tight BDT selection that also defines signal taus.

The expected number of fake events that passes all criteria of a region of interest is estimated from the difference between data that fails tau identification and failed MC events where the tau is real (i.e. *not* truth-matched to a jet). We perform this truth-matching in order to exclude objects that do not come from jet $\to \tau$ but from a true hadronic tau or a lepton and which we

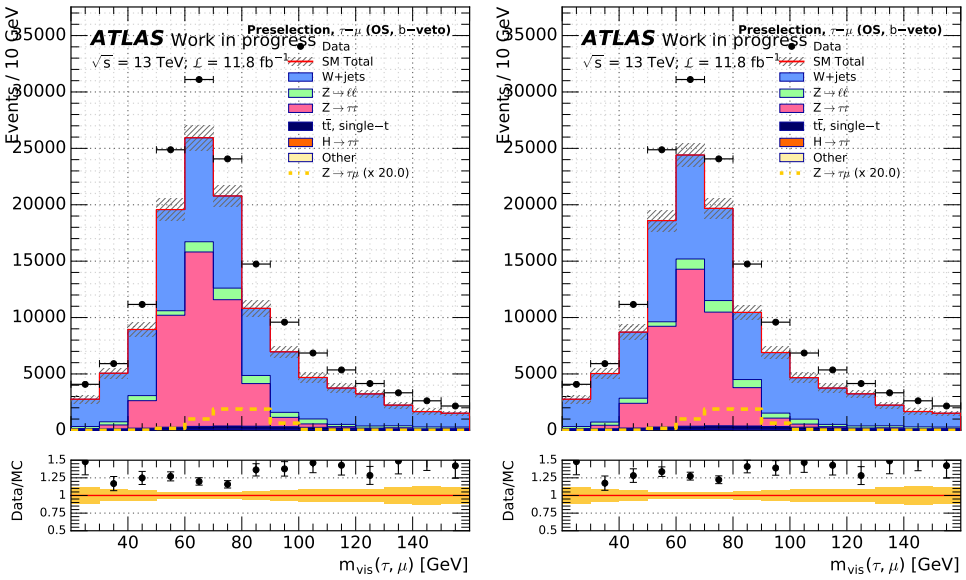


Figure 7.18: m_{vis} distributions after preselection, b-jet veto and OS(τ, μ) without (left) and with (right) scale factors derived from the background fit.

will estimate with MC only. The number of fakes in an SR is thus:

$$N_{\text{fakes}}^{\text{SR}} = F \cdot \left(N_{\text{data}}^{\text{fail, SR}} - N_{\text{MC, not } j\to\tau}^{\text{fail, SR}} \right). \tag{7.5}$$

The difference in equation 7.5 is scaled by the fake factor F which is the number of pass events divided by the number of fail events. The fake factor is defined as the weighted average of individual fake factors F_i derived from various control regions (these will be defined in section 7.3.2):

$$\begin{aligned} F &= \sum_{i=\text{CR}} F_i \\ &= R_{\text{W}} F_{\text{W}} + R_{\text{Zll}} F_{\text{Zll}} + R_{\text{T}} F_{\text{T}} + R_{\text{QCD}} F_{\text{QCD}} \,, \end{aligned} \tag{7.6}$$

where the individual fake factor F_i is the pass-to-fail ratio in a control region CR_i corrected for other truth-matched components of the same process and for other MC processes that pass the CR_i selection:

$$F_i = \frac{N_{\text{data}}^{\text{pass, CR}_i} - N_{\text{MC i, not } j\to\tau}^{\text{pass, CR}_i} - \sum_{j\neq i} N_{\text{MC } j}^{\text{pass, CR}_i}}{N_{\text{data}}^{\text{fail, CR}_i} - N_{\text{MC i, not } j\to\tau}^{\text{fail, CR}_i} - \sum_{j\neq i} N_{\text{MC } j}^{\text{fail, CR}_i}}. \tag{7.7}$$

Determining the fake background requires multiple regions as there are several sources of fakes. For each main process that contributes with fakes, the relative contribution to the total fake factor is given by the fraction of that background in the region in which the amount of fake background is being estimated:

$$R_i = \frac{N_{\text{MC } i,\, j\to\tau}^{\text{fail, SR}}}{N_{\text{data}}^{\text{fail, SR}} - N_{\text{MC, not } j\to\tau}^{\text{fail, SR}}}, \tag{7.8}$$

with the exception of the QCD region which does not have MC, and is instead derived from $R_{\text{QCD}} = 1 - \sum_i R_i$.

In eq. 7.8, a normalisation factor, k_i, can be included to normalise failed events to data: $R'_i = k_i R_i$. The k_i factors for fake events can be derived from each control region by subtracting other processes and MC of the specific process from data:

$$k_i = \frac{N_{\text{data}}^{\text{fail, CR}_i} - N_{\text{MC}\, i,\, \text{not}\, j\to\tau}^{\text{fail, CR}_i} - \sum_{j\neq i} N_{\text{MC}\, j}^{\text{fail, CR}_i}}{N_{\text{MC}\, i,\, j\to\tau}^{\text{fail, CR}_i}} \,. \tag{7.9}$$

This requires rederivation of R_i and R_{QCD}. However, the results for the fake factor estimation in this analysis will assume $k_i = 1$ for all processes in first instance.

7.3.2 Control regions

A set of dedicated control regions are created for W + jets, QCD, $t\bar{t}$ and $Z \to \mu\mu$ to derive fakes from data. The selected control regions (CR) represent the main fake contributions and are kinematically close to the region of interest, except for the inversion of a single cut. The fakes should ultimately be estimated and applied in the signal regions, whereas the results presented in this thesis are produced at preselection level.
All control region selection are listed in tab. 7.3.

Region	Process	Change from SR cut
CRZll	$Z \to \mu\mu$	Two good muons
CRW	W + jets	$m_{\text{T}}(\mu, E_T^{\text{miss}}) > 35$ GeV, $m_{\text{T}}(\tau, E_T^{\text{miss}}) > 40$ GeV
CRT	t, $t\bar{t}$	$N_{b-\text{jets}} \geq 2$
CRQ	QCD	Invert muon isolation (non-isolated)

Table 7.6: The set of control regions used to derive fake factors. The table lists the specific cut that separates the control region from the region in which the fake contribution is estimated.

W + jets
All W + jets events are characterised by a hadronic tau signature faked by jets. The correlation in the m_{T} plane determines the control region definition for W + jets. The largest SR and NRZtt exclusive control region (see fig. 7.12) is obtained by inverting the SR bounds:

$$\begin{aligned} m_{\mathrm{T}}(\mu, E_T^{\mathrm{miss}}) &> 40\,\mathrm{GeV}; \\ m_{\mathrm{T}}(\tau, E_T^{\mathrm{miss}}) &> 30\,\mathrm{GeV}. \end{aligned} \tag{7.10}$$

The distributions in the CRW split by the pass and fail categories can be seen in fig. 7.19.

$Z \to \mu\mu$
Background from $Z \to \mu\mu$ production consists of two components: events where a muon fakes a hadronic tau and events with a jet faking a hadronic tau.
$Z \to \mu\mu$ events are selected in a control region defined by two oppositely charged isolated muons with an invariant mass around the peak of the Z boson ($61\,\mathrm{GeV} < M_{\mu\mu} < 121\,\mathrm{GeV}$). Figure 7.20 shows the resulting pass and fail distributions for the CRZll region.

QCD
QCD background comes from events where one jet fakes a hadronic tau and another fakes a muon.
Signal muons are selected with gradient isolation so a QCD region is obtained by reverting lepton isolation, hence requiring one non-isolated muon. The CRQ distributions are shown in fig. 7.21.

$t\bar{t}$
Except for the requirement of non-tight taus, the definition of the $t\bar{t}$ control region is analogous to the NRT with the requirement of ≥ 2 b-jets. Figure 7.22 shows the CRT distributions.

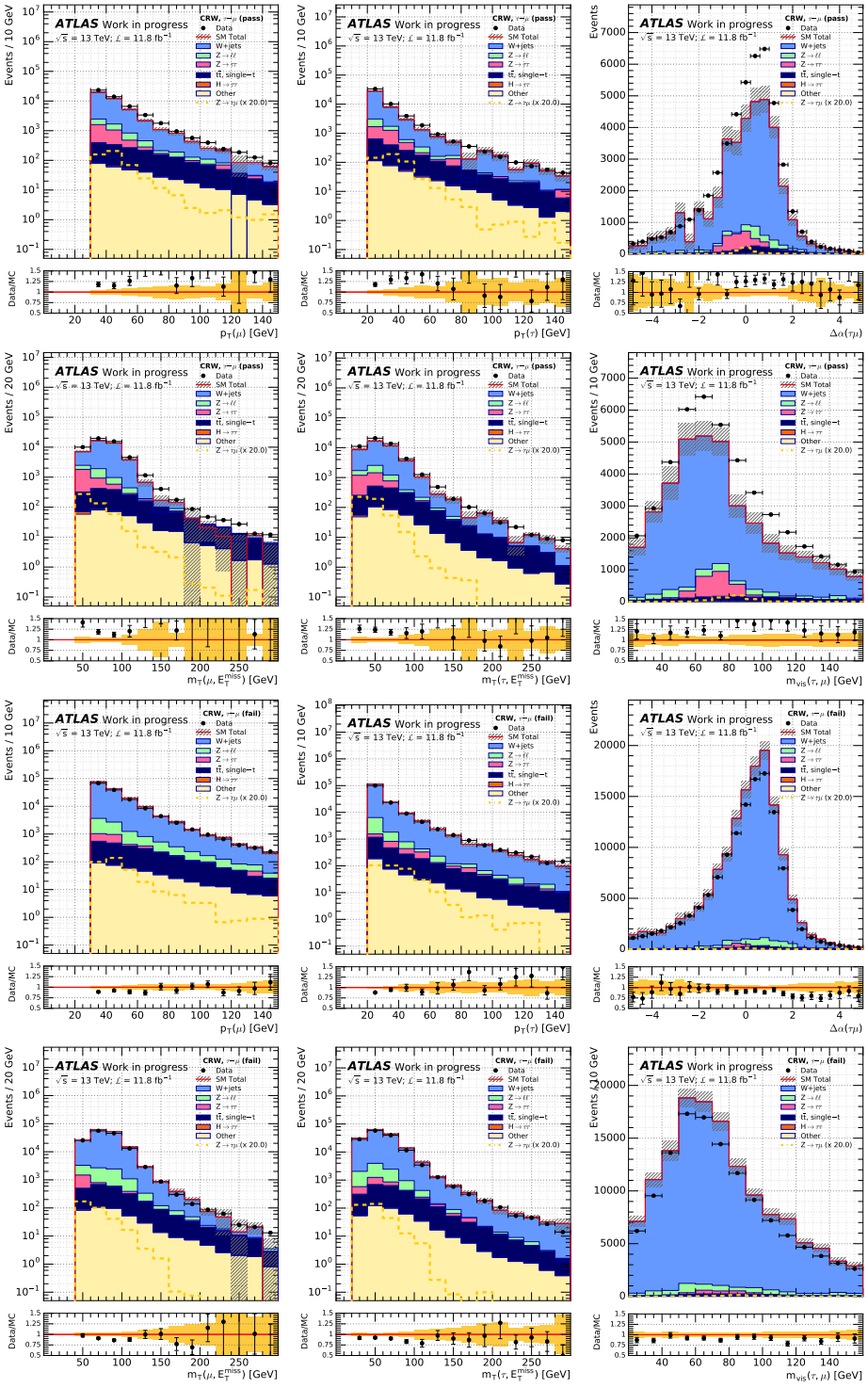


Figure 7.19: Distributions in control region CRW with MC backgrounds and data. The first and second rows contain distributions of the pass category; the third and fourth rows shows fail events. Errors are statistical only.

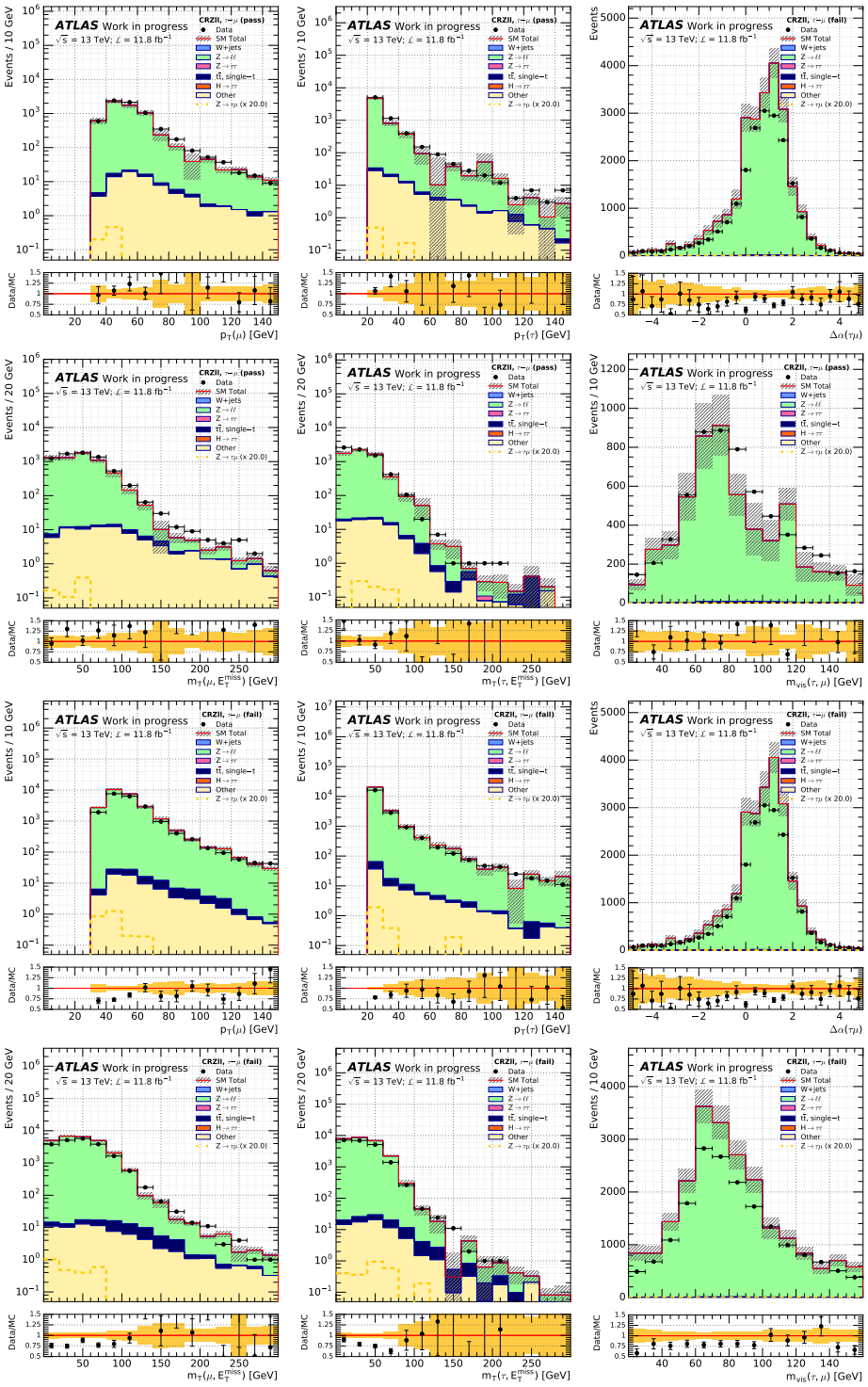


Figure 7.20: Distributions in control region CRZll with MC backgrounds and data. The first and second rows contain distributions of the pass category; the third and fourth rows shows fail events. Errors are statistical only.

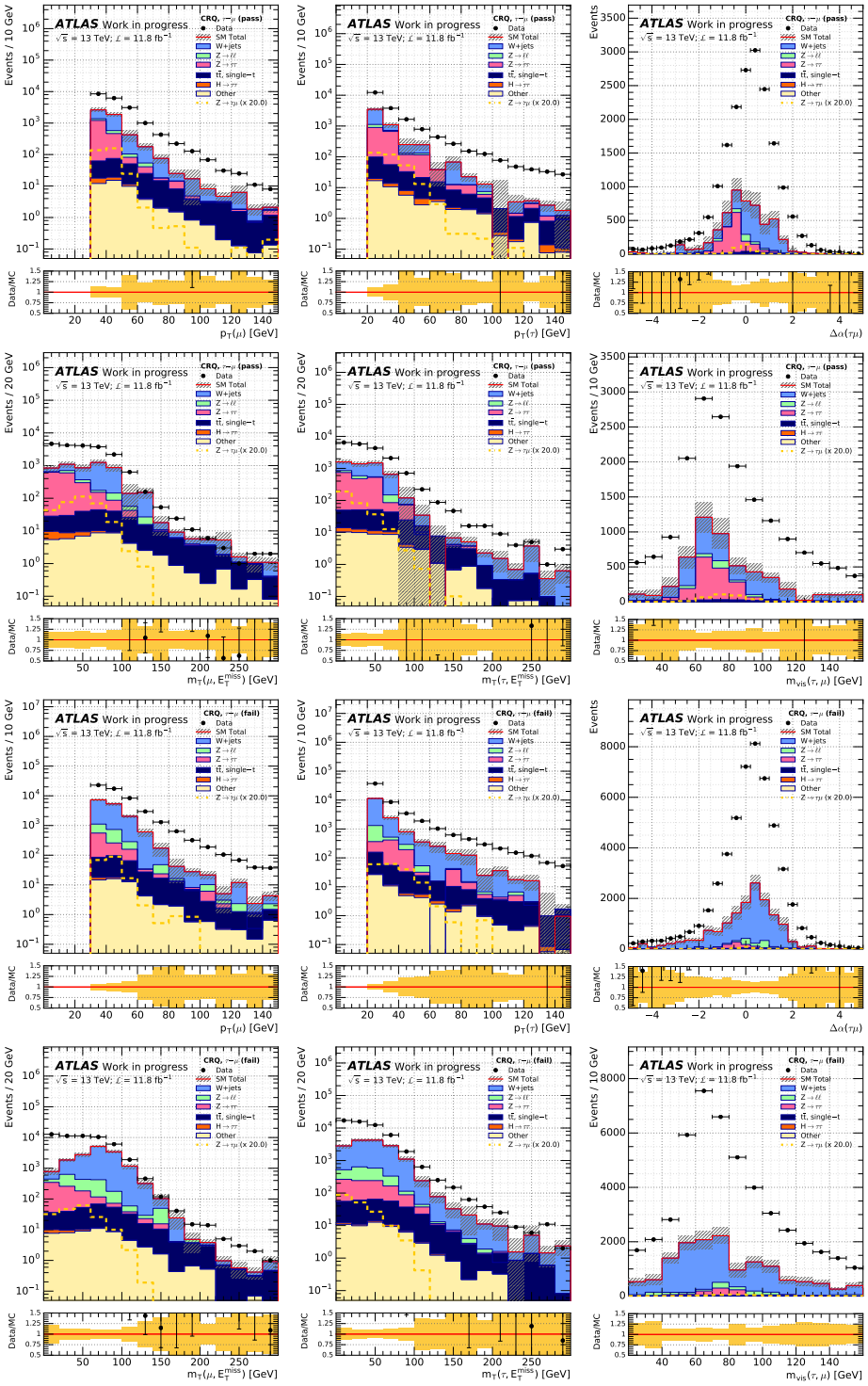


Figure 7.21: Distributions in control region CRQ with MC backgrounds and data. The first and second rows contain distributions of the pass category; the third and fourth rows shows fail events. Errors are statistical only.

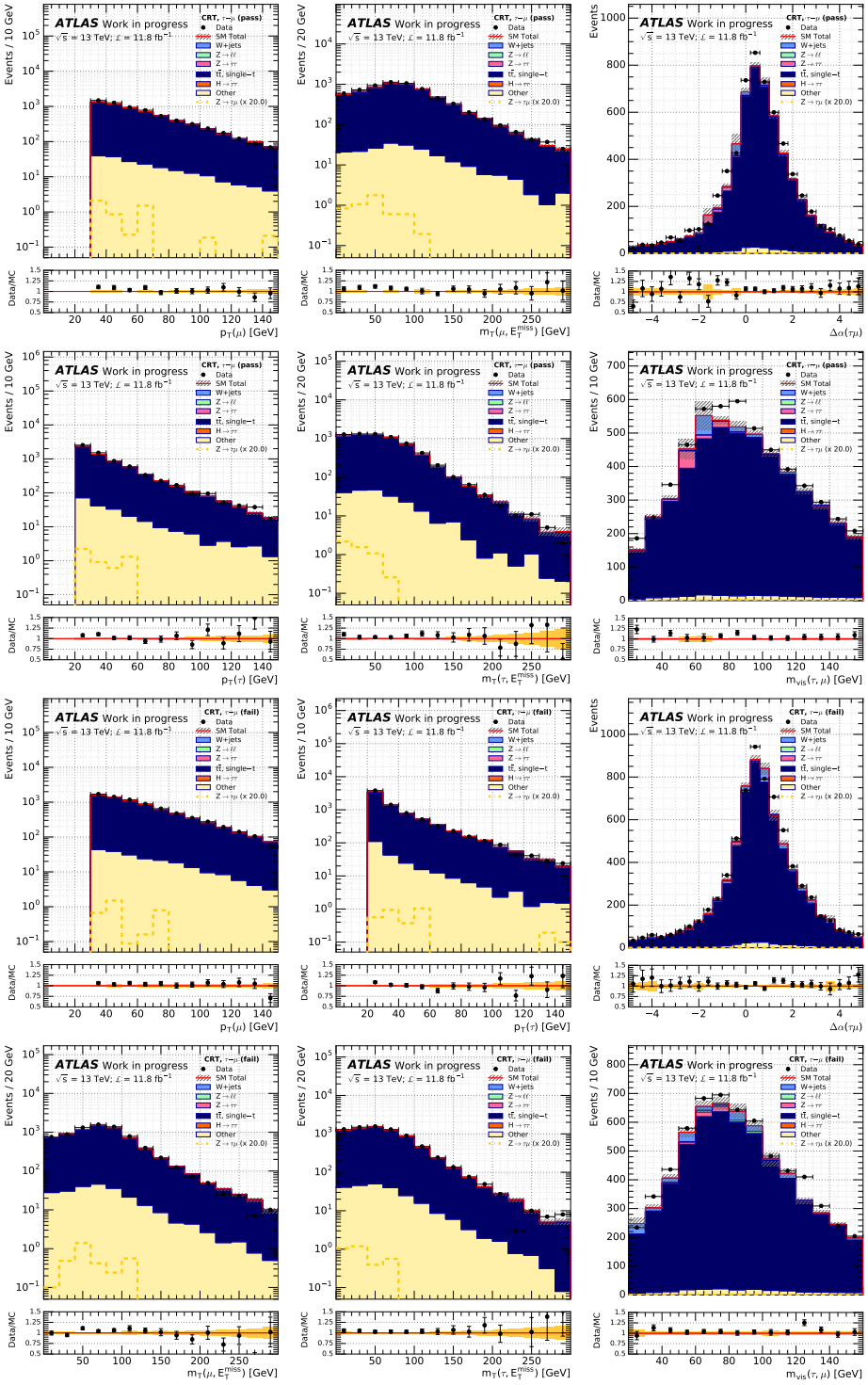


Figure 7.22: Distributions in control region CRT with MC backgrounds and data. The first and second rows contain distributions of the pass category; the third and fourth rows shows fail events. Errors are statistical only.

7.3.3 Results

The regions are split by 1- and 3-prong taus since the rate of misidentification is expected to be larger in events with a 1-prong hadronic tau, compared to events with a 3-prong tau. The fake factors presented are derived in the regions defined control regions according to eq. 7.7.
Figure 7.23 shows the fake factors for events with a 1- or 3-prong tau derived in each control region with $k_i = 1$. The fake factors are binned in p_{T} and are shown with their statistical uncertainties. The large statistical errors are expected to decrease when the full 2016 dataset is used.

The k-factors for normalisation of fakes binned in $p_{\mathrm{T}}(\tau)$ in fig. 7.24 shows that the assumption of letting $k_i = 1$ in the case of 1-prong taus is reasonable. CRW and CRZll k-factors behave similarly in that MC slightly overestimates the amount of fail events at low $p_{\mathrm{T}}(\tau)$ but underestimates at high $p_{\mathrm{T}}(\tau)$. For CRT, the behaviour is reversed.
CRW and CRZll 3-prong failed taus are similar to the 1-prong case, except that the first and last bins yield substantially lower k values than before. With the exception of the high $p_{\mathrm{T}}(\tau)$ bin, the CRT k-factor behaves completely different compared to its 1-prong k-factor. CRT for 3-prong events is instead similar to the CRW and CRZll factors but with a greater offset from 1. This means that MC predicts a lot fewer events in this category compared to what is observed in data.

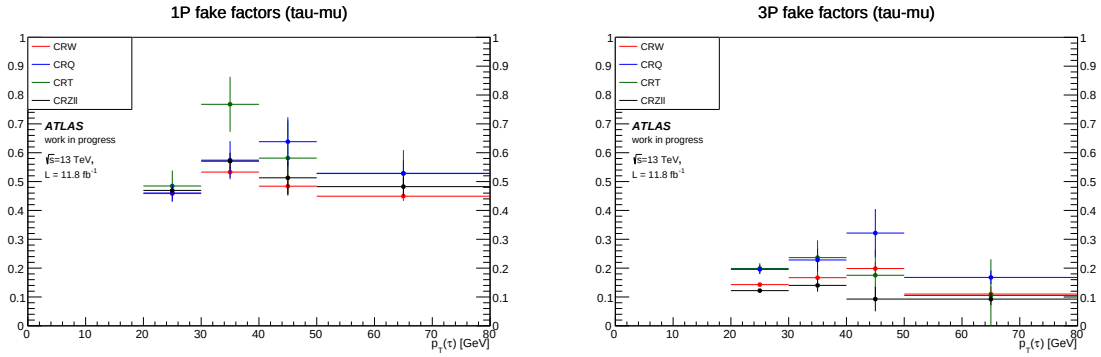


Figure 7.23: $p_{\mathrm{T}}(\tau)$-binned distributions of fake factors for each control region divided by 1- and 3-prong taus. The last bin includes the overflow bins. The errors are statistical only.

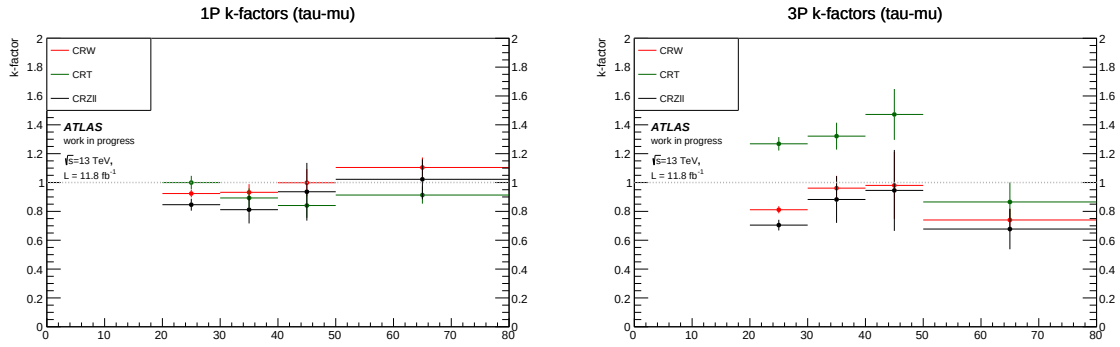


Figure 7.24: $p_{\mathrm{T}}(\tau)$-binned k-factors for fail events in each control region divided by 1- and 3-prong taus. The last bin includes the overflow bins. The errors are statistical only.

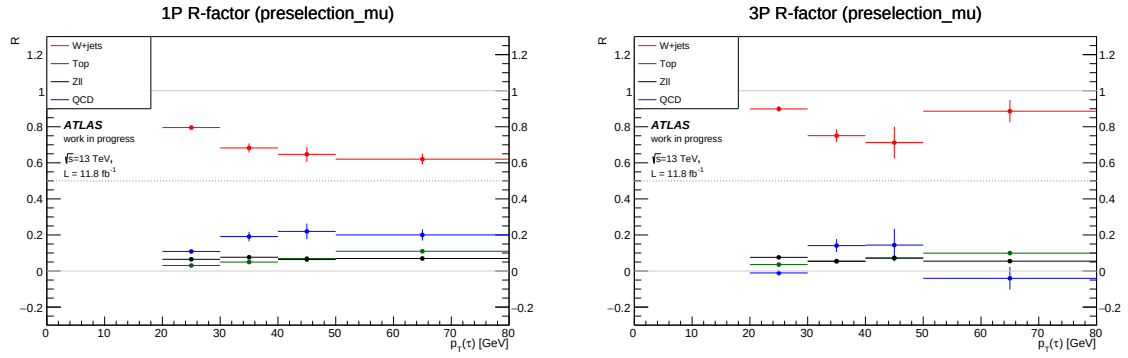


Figure 7.25: $p_{\mathrm{T}}(\tau)$-binned R-factors for each control region process divided by 1- and 3-prong fake taus. The last bin includes the overflow bins. The errors are statistical only.

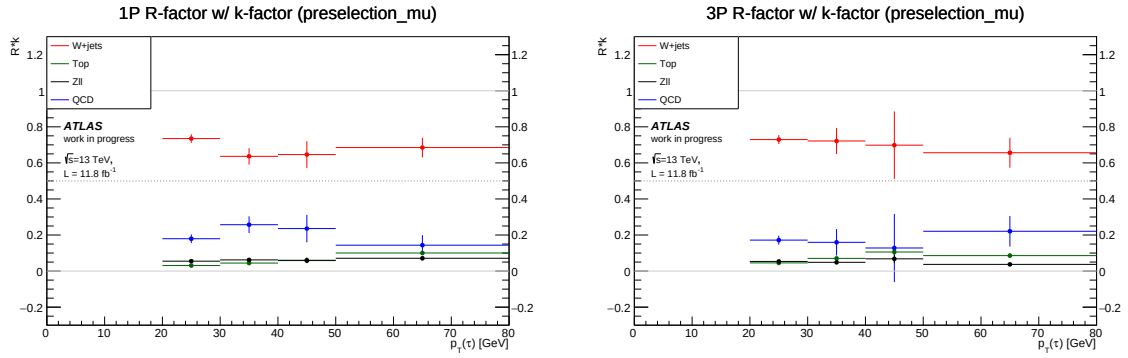


Figure 7.26: $p_{\mathrm{T}}(\tau)$-binned $R \times k$-factors for each control region process divided by 1- and 3-prong fake taus. The last bin includes the overflow bins. The errors are statistical only.

The R-factors of fig. 7.25 show the relative contribution of each process to the fake background at preselection level assuming $k_i = 1$.
Generally, the background from fakes is dominated by W + jets events with 63% (1-prong) and 76% (3-prong) contributions.
For both 1- and 3-prong taus, the contributions from $Z \to \mu\mu$ and top compose about 10% and 2% of events, respectively.
26% (1-prong) and 30% (3-prong) of events emerge from QCD fakes. The 3-prong $p_{\mathrm{T}}(\tau)$ distribution of R_{QCD} goes negative in the first and last bins caused by the other two large fake backgrounds (W + jets and $Z \to \mu\mu$) being subtracted from 1. However, both the W + jets and $Z \to \mu\mu$ regions have k-factors that are consistently below 1 in the 3-prong case and the issue could be solved by including the k-factors in the calculation of R_i as shown in fig. 7.26.
R_{QCD} without k-factor does remain greater than zero within the errors and the decision was made to not use k-factors.

8 Results and discussion

Finally, the estimation of the amount of fakes based on eq. 7.5 is propagated to the distributions at preselection and added to the stack of SM backgrounds as shown in fig. 8.1. The distributions show a very good data/MC agreement overall but it should of course be noted that no systematic errors are included at this stage. Based on this background estimation, closure tests to validate the results from the fake factor method will have to be performed as well.

A fit procedure is currently in progress using the signal region selections and the background model described in this thesis. The signal strength μ_{sig} will be used as the parameter of interest and will be determined in the SRs. Since μ_{sig} is defined as the ratio of the fitted signal cross section times the branching fraction times the cross section of the Z, $\mu_{\text{sig}} = 0$ corresponds to the absence of signal while $\mu_{\text{sig}} = 1$ gives the current upper limit $\text{BR} = 1.2 \cdot 10^{-5}$ which the signal is normalised to.

Due to the tuning issue of the $m_{\tau,\mu}^{\text{MMC}}$, the visible mass m_{vis} will most likely be used as a final discriminating variable, using the central region where the signal is expected to evaluate the significance of a possible excess. Figures 8.2 and 8.3 show however the unblinded $m_{\tau,\mu}^{\text{MMC}}$ sidebands of SR1 where the $m_{\text{vis}} > 60$ GeV has been removed. Note that the binning is changed with respect to previously shown distributions. The sidebands in the SRs will be used as additional constraints on the background expectation and the normalisations from the included normalisation regions will be used as free parameters in the fit.
Systematic uncertainties will be considered as nuisance parameters in the fit. Some of these include the experimental uncertainties on trigger efficiency, $E_{\text{T}}^{\text{miss}}$, the tau energy scale and the jet energy scale which can affect the shape and normalisation of signal and background.

The performance of the fit can be tested in regions to validate the extrapolation. If a good agreement with respect to data is found, the central region of the SRs can be unblinded and fitted. If no excess from the SM background-only expectation is found in the signal regions, an upper limit on the branching fraction of the signal can be set using the modified frequentist CL_s method where the confidence level is calculated as $\text{CL}_s \equiv \frac{p_{s+b}}{1-p_b}$ with $1 - p_b$ being the probability of rejecting the background-only hypothesis and p_{s+b} being the p-value corresponding to the signal + background hypothesis.

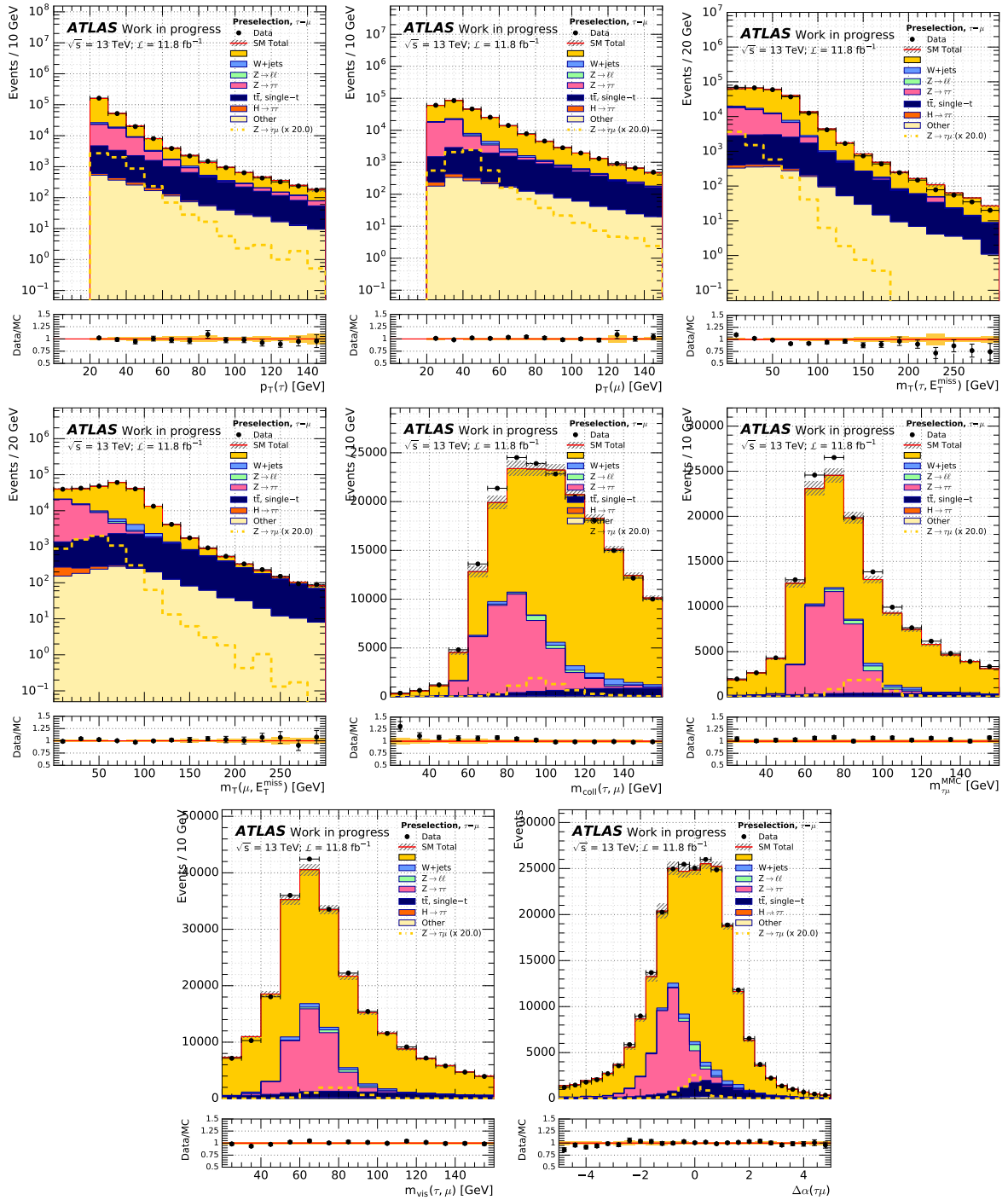


Figure 8.1: Distributions at preselection level with fake background estimated by the fake factor method. Errors are statistical only.

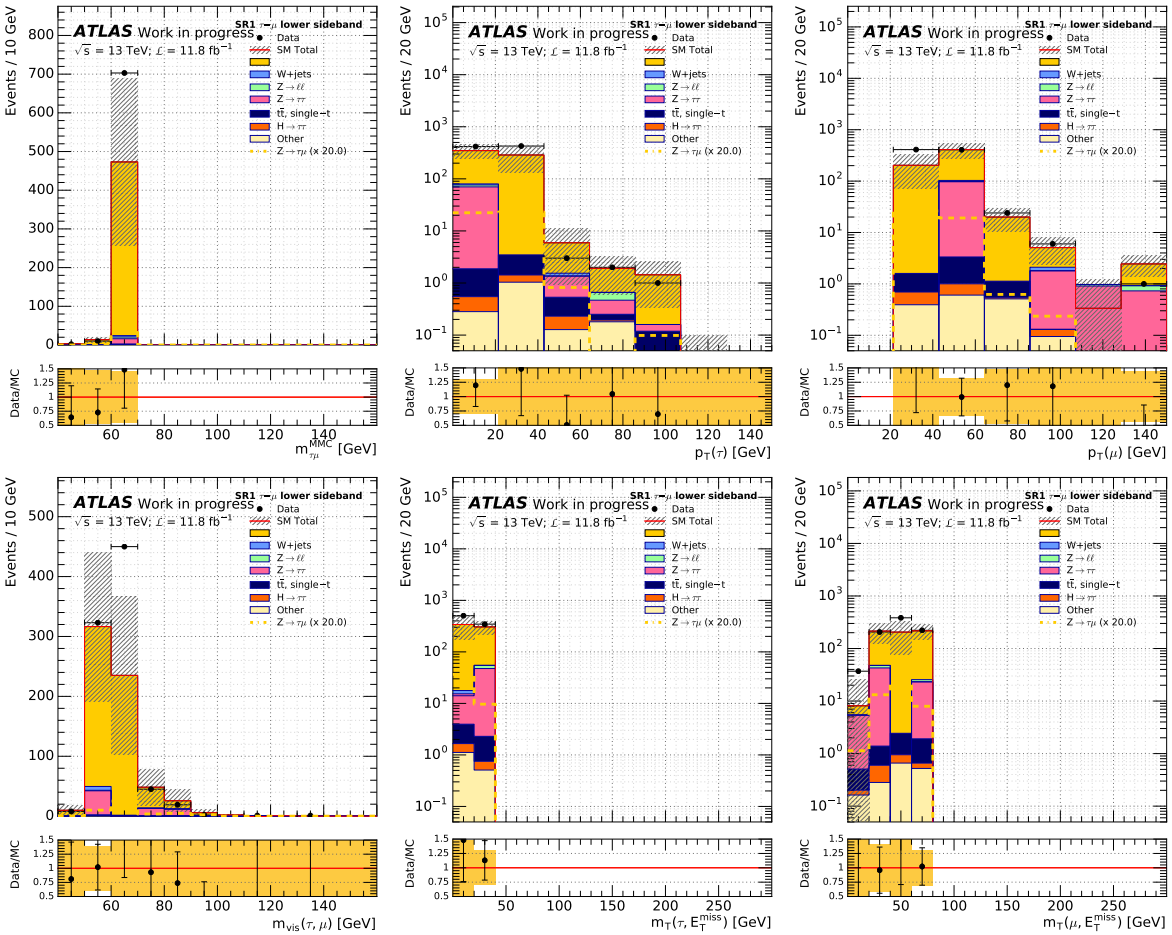


Figure 8.2: Sidebands in $m_{\tau,\mu}^{\mathrm{MMC}}$ for SR1 where the $m_{\mathrm{vis}} > 60$ GeV cut is removed. The figures show the $m_{\tau,\mu}^{\mathrm{MMC}} < 70$ GeV sideband. Errors are statistical only.

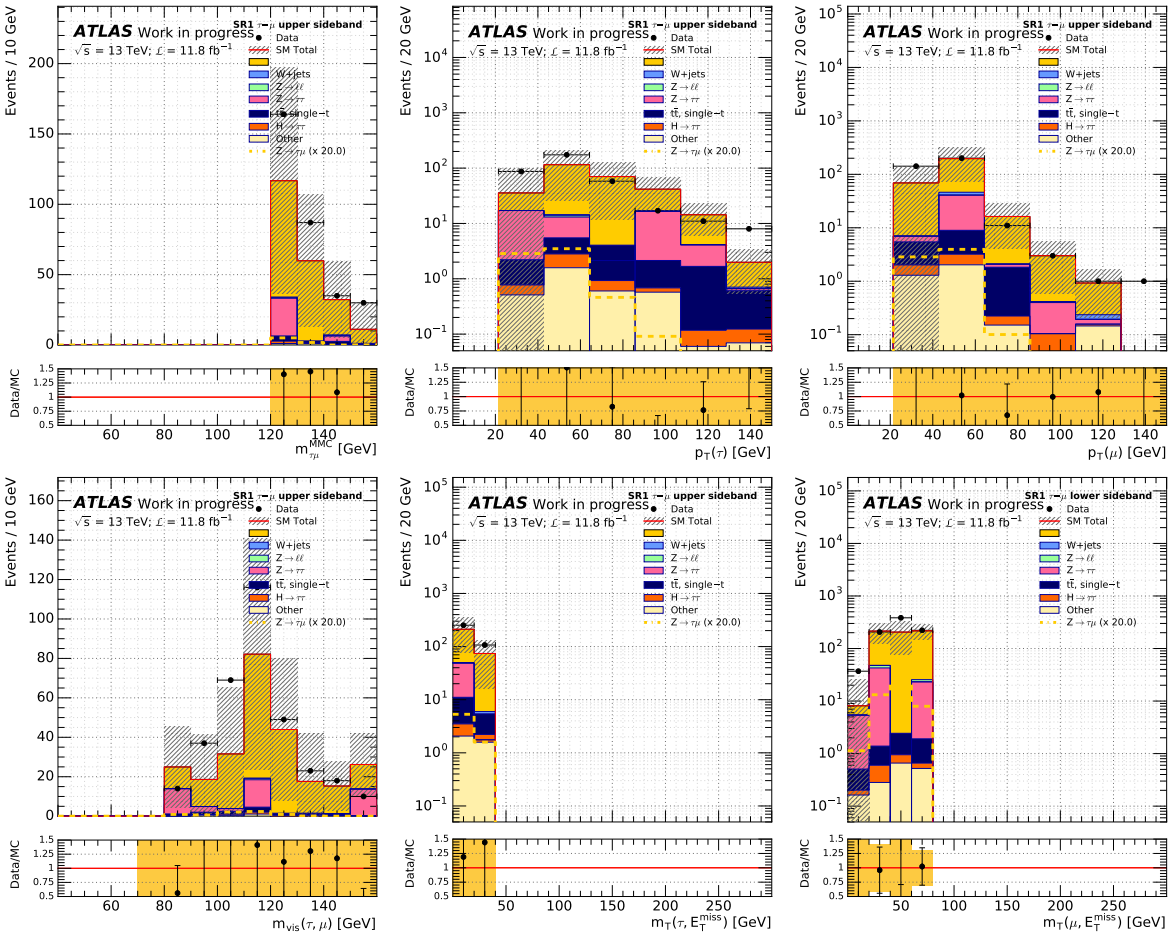


Figure 8.3: Sidebands in $m_{\tau,\mu}^{\mathrm{MMC}}$ for SR1 where the $m_{\mathrm{vis}} > 60$ GeV cut is removed. The figures show the $m_{\tau,\mu}^{\mathrm{MMC}} > 120$ GeV sideband. Errors are statistical only.

9 Conclusion

This thesis presents the event selection and background modelling for a search of a lepton flavour violating $Z \to \tau_{\text{had}}\mu$ signature using proton-proton collision data at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of $\mathcal{L} = 11.8$ fb^{-1} measured by the ATLAS Detector.

A selection of kinematic cuts are optimised in order to reduce the amount of background predicted from Monte Carlo simulation without a major loss in signal yield. Two orthogonal signal regions are defined by exploiting the event topology in $m_{\text{T}}(\mu, E_{\text{T}}^{\text{miss}})$, $m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}})$ space with the primary signal region giving a statistical significance of $S_{\text{SR1}} = 1.34$ and a secondary region giving $S_{\text{SR2}} = 0.25$.

For the main backgrounds that contribute with a real muon and a real tau in the final state ($Z \to \tau\tau$ and t, $t\bar{t}$), the Monte Carlo expectation is scaled by a normalisation factor. This scale factor is derived from a fit to data in normalisation regions that are constructed with a selection to reject other backgrounds and signal. The resulting factors $k_{\text{Ztt}} = 0.90 \pm 0.08$ and $k_{\text{top}} = 1.07 \pm 0.02$ are used to weight the distributions.
$Z \to \mu\mu$, $t/t\bar{t}$, W + jets and QCD processes give rise to fake background due to misidentification of tau leptons. Fake background is estimated from data in dedicated control regions that are designed to be enhanced in jets. A BDT-based tau identification is used to distinguish between events that *pass* or *fail* the tau object definition so that a fake factor can be derived from each region. For the region of interest, the fake estimate is the number of data event times the fake factor.
The resulting background model including normalisation and fakes is in good agreement with data at preselection level. The significance of SR1 combined with the increase in statistics once the full $\mathcal{L} = 36.5$ fb^{-1} dataset is used appears promising for the further analysis.

Bibliography

[1] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716:1–29, September 2012.

[2] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716:30–61, September 2012.

[3] V. Mukhanov. *Physical Foundations of Cosmology.* Cambridge University Press, 2005.

[4] P. Abreu et al. A Search for lepton flavor violation in Z^0 decays. *Phys. Lett.*, B298:247–256, 1993.

[5] DELPHI Collaboration. Search for lepton flavor number violating Z0 decays. *Z. Phys. C*, 73(243), 1997.

[6] ATLAS Experiment Public Results. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults`.

[7] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[8] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[9] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.

[10] Q. R. Ahmad et al. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, Jul 2001.

[11] Q. R. Ahmad et al. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.

[12] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Physical Review Letters*, 13:321–323, August 1964.

[13] R. Brout and F. Englert. Spontaneous Symmetry Breaking in Gauge Theories: a Historical Survey. *ArXiv High Energy Physics - Theory e-prints*, February 1998.

[14] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, 13:508–509, October 1964.

[15] A. Ceccucci et al. Review of the CKM matrix. `http://pdg.lbl.gov/2010/reviews/rpp2010-rev-ckm-matrix.pdf`.

[16] Y. Nir. Lectures on Flavor Physics, Niels Bohr Institute, Nov 7-11 2016. `https://indico.nbi.ku.dk/getFile.py/access?resId=0&materialId=2&confId=897`.

[17] D. Griffiths. *Introduction to Elementary Particles*, volume 2. WILEY-VCH, revised edition edition, 2010.

[18] M. Bona et al. The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD. *JHEP*, 10:081, 2006.

[19] Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. *AAP*, 594:A13, September 2016.

[20] Neutrino Mass Hierarchy. `http://www.staff.uni-mainz.de/wurmm/juno.html`.

[21] M. C. Gonzalez-Garcia, M. Maltoni, and T. Schwetz. Updated fit to three neutrino mixing: status of leptonic CP violation. *Journal of High Energy Physics*, 11:52, November 2014.

[22] J. I. Illana, M. Jack, and T. Riemann. Predictions for $Z \to \mu\tau$ and Related Reactions. *ArXiv High Energy Physics - Phenomenology e-prints*, January 2000.

[23] G. C. Branco et al. Theory and phenomenology of two-Higgs-doublet models. *Physics Reports*, 516:1–102, July 2012.

[24] M. P. Misiak, Stefan and J. Rosiek. *Supersymmetry and FCNC Effects*, pages 795–828. World Scientific Publishing Co, 1998.

[25] S. Davidson, S. Lacroix, and P. Verdier. LHC sensitivity to lepton flavour violating Z boson decays. *Journal of High Energy Physics*, 9:92, September 2012.

[26] F. Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. `https://cds.cern.ch/record/1621583`, Oct 2013. General Photo.

[27] Lyndon R Evans and Philip Bryant. LHC Machine. *J. Instrum.*, 3:S08001. 164 p, 2008.

[28] M. Thomson. *Modern Particle Physics*. Cambridge University Press, 3rd edition, 2013.

[29] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3(S08003), 2008.

[30] M Capeans et al. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, ATLAS Experiment, Sep 2010.

[31] A Vogel. ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates. Technical Report ATL-INDET-PROC-2013-005, CERN, Geneva, Apr 2013.

[32] ATLAS Experiment Public Results. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[33] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[34] T. Gleisberg and other. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.

[35] Multi-Boson Simulation for 13 TeV ATLAS Analyses. Technical Report ATL-PHYS-PUB-2016-002, CERN, Geneva, Jan 2016.

[36] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[37] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[38] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[39] R Clarke, A Pranko, and M Shapiro. Search For Lepton Flavor Violating Decays Of The Higgs Boson in the $H \to \mu\tau \to \mu\tau_{\text{had}}$ Final State With The ATLAS Detector. Technical Report ATL-COM-PHYS-2014-906, CERN, Geneva, Jul 2014.

[40] T. Aaltonen et al. Search for Standard Model Higgs Boson Production in Association with a W Boson Using a Matrix Element Technique at CDF in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev.*, D85:072001, 2012.

[41] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A New Mass Reconstruction Technique for Resonances Decaying to di-tau. *Nucl. Instrum. Meth.*, A654:481–489, 2011.

[42] Andrea Coccaro. Track Reconstruction and b-Jet Identification for the ATLAS Trigger System. *J. Phys. Conf. Ser.*, 368:012034, 2012.

[43] Run 2 Data Analysis.
`https://twiki.cern.ch/twiki/bin/view/Main/Run2DataAnalysis`.

[44] SUSYTools.
`https://svnweb.cern.ch/trac/atlasoff/browser/PhysicsAnalysis/SUSYPhys/SUSYTools/trunk`.

[45] 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. C*, 75:303, 2015.

[46] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008.

[47] TauAnalysisTools.
`https://svnweb.cern.ch/trac/atlasoff/browser/PhysicsAnalysis/TauID/TauAnalysisTools/trunk/doc/README-TauSelectionTool.rst`.

[48] G. Aad et al. Jet energy measurement with the atlas detector in proton-proton collisions at $\sqrt{s} = 7$ tev. *The European Physical Journal C*, 73(3):2304, 2013.

[49] ATLAS Collaboration. Pile-up subtraction and suppression for jets in ATLAS. ATLAS-CONF-2013-083, 2013.

[50] ATLAS Collaboration. Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run. ATL-PHYS-PUB-2016-012, 2016.

[51] ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton–Proton Collisions recorded in 2012 at $\sqrt{s} = 8$ TeV. ATLAS-CONF-2013-082, 2013.

[52] AssociationUtils. `https://svnweb.cern.ch/trac/atlasoff/browser/PhysicsAnalysis/AnalysisCommon/AssociationUtils/trunk/doc/README.rst`.

[53] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *European Physical Journal C*, 71:1554, February 2011.

[54] M. Baak, Geert Jan. Besjes, D. Cote, A. Koutsman, J. Lorenz, et al. HistFitter software framework for statistical data analysis. 2014.

A | MC samples

Table A.1: Signal sample used in the analysis.

DSID	Name
303778	Ztaumu

Table A.2: Single-t and $t\bar{t}$ samples used in the analysis.

DSID	Name
410000	PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad
410011	PowhegPythiaEvtGen_P2012_singletop_tchan_lept_top
410012	PowhegPythiaEvtGen_P2012_singletop_tchan_lept_antitop
410013	PowhegPythiaEvtGen_P2012_Wt_inclusive_top
410014	PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop
410025	PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top
410026	PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop

Table A.3: Summary of diboson samples used in the analysis.

DSID	Name
361063	Sherpa_CT10_llll
361064	Sherpa_CT10_lllvSFMinus
361065	Sherpa_CT10_lllvOFMinus
361066	Sherpa_CT10_lllvSFPlus
361067	Sherpa_CT10_lllvOFPlus
361068	Sherpa_CT10_llvv
361088	Sherpa_CT10_lvvv
361089	Sherpa_CT10_vvvv

Table A.4: $H \to \tau\tau$ samples used in the analysis.

DSID	Name
341123	PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_ggH125_tautaulh
341156	PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_VBFH125_tautaulh

Table A.5: Inclusive Z + jets and W + jets samples used as alternatives to the sliced samples.

DSID	Name
361100	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wplusenu
361101	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wplusmunu
361102	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wplustaunu
361103	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wminusenu
361104	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wminusmunu
361105	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wminustaunu
361106	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee
361107	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zmumu
361108	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Ztautau

Table A.6: Summary of p_{T} sliced $Z \to \tau\tau$ + jets samples used in the analysis.

DSID	Name
363102	Sherpa_NNPDF30NNLO_Ztautau_Pt70_140_CVetoBVeto
363103	Sherpa_NNPDF30NNLO_Ztautau_Pt70_140_CFilterBVeto
363104	Sherpa_NNPDF30NNLO_Ztautau_Pt70_140_BFilter
363105	Sherpa_NNPDF30NNLO_Ztautau_Pt140_280_CVetoBVeto
363106	Sherpa_NNPDF30NNLO_Ztautau_Pt140_280_CFilterBVeto
363107	Sherpa_NNPDF30NNLO_Ztautau_Pt140_280_BFilter
363108	Sherpa_NNPDF30NNLO_Ztautau_Pt280_500_CVetoBVeto
363109	Sherpa_NNPDF30NNLO_Ztautau_Pt280_500_CFilterBVeto
363110	Sherpa_NNPDF30NNLO_Ztautau_Pt280_500_BFilter
363111	Sherpa_NNPDF30NNLO_Ztautau_Pt500_700_CVetoBVeto
363112	Sherpa_NNPDF30NNLO_Ztautau_Pt500_700_CFilterBVeto
363113	Sherpa_NNPDF30NNLO_Ztautau_Pt500_700_BFilter
363114	Sherpa_NNPDF30NNLO_Ztautau_Pt700_1000_CVetoBVeto
363115	Sherpa_NNPDF30NNLO_Ztautau_Pt700_1000_CFilterBVeto
363116	Sherpa_NNPDF30NNLO_Ztautau_Pt700_1000_BFilter
363117	Sherpa_NNPDF30NNLO_Ztautau_Pt1000_2000_CVetoBVeto
363118	Sherpa_NNPDF30NNLO_Ztautau_Pt1000_2000_CFilterBVeto
363119	Sherpa_NNPDF30NNLO_Ztautau_Pt1000_2000_BFilter
363120	Sherpa_NNPDF30NNLO_Ztautau_Pt2000_E_CMS_CVetoBVeto
363121	Sherpa_NNPDF30NNLO_Ztautau_Pt2000_E_CMS_CFilterBVeto
363361	Sherpa_NNPDF30NNLO_Ztautau_Pt0_70_CVetoBVeto
363362	Sherpa_NNPDF30NNLO_Ztautau_Pt0_70_CFilterBVeto
363363	Sherpa_NNPDF30NNLO_Ztautau_Pt0_70_BFilter

Table A.7: Summary of p_{T} sliced $Z \to \mu\mu$ + jets Sherpa samples used in the analysis.

DSID	Name
363364	Sherpa_NNPDF30NNLO_Zmumu_Pt0_70_CVetoBVeto
363365	Sherpa_NNPDF30NNLO_Zmumu_Pt0_70_CFilterBVeto
363366	Sherpa_NNPDF30NNLO_Zmumu_Pt0_70_BFilter
363367	Sherpa_NNPDF30NNLO_Zmumu_Pt70_140_CVetoBVeto
363368	Sherpa_NNPDF30NNLO_Zmumu_Pt70_140_CFilterBVeto
363369	Sherpa_NNPDF30NNLO_Zmumu_Pt70_140_BFilter
363370	Sherpa_NNPDF30NNLO_Zmumu_Pt140_280_CVetoBVeto
363371	Sherpa_NNPDF30NNLO_Zmumu_Pt140_280_CFilterBVeto
363372	Sherpa_NNPDF30NNLO_Zmumu_Pt140_280_BFilter
363373	Sherpa_NNPDF30NNLO_Zmumu_Pt280_500_CVetoBVeto
363374	Sherpa_NNPDF30NNLO_Zmumu_Pt280_500_CFilterBVeto
363375	Sherpa_NNPDF30NNLO_Zmumu_Pt280_500_BFilter
363376	Sherpa_NNPDF30NNLO_Zmumu_Pt500_700_CVetoBVeto
363377	Sherpa_NNPDF30NNLO_Zmumu_Pt500_700_CFilterBVeto
363378	Sherpa_NNPDF30NNLO_Zmumu_Pt500_700_BFilter
363379	Sherpa_NNPDF30NNLO_Zmumu_Pt700_1000_CVetoBVeto
363380	Sherpa_NNPDF30NNLO_Zmumu_Pt700_1000_CFilterBVeto
363381	Sherpa_NNPDF30NNLO_Zmumu_Pt700_1000_BFilter
363382	Sherpa_NNPDF30NNLO_Zmumu_Pt1000_2000_CVetoBVeto
363383	Sherpa_NNPDF30NNLO_Zmumu_Pt1000_2000_CFilterBVeto
363384	Sherpa_NNPDF30NNLO_Zmumu_Pt1000_2000_BFilter
363385	Sherpa_NNPDF30NNLO_Zmumu_Pt2000_E_CMS_CVetoBVeto
363386	Sherpa_NNPDF30NNLO_Zmumu_Pt2000_E_CMS_CFilterBVeto
363387	Sherpa_NNPDF30NNLO_Zmumu_Pt2000_E_CMS_BFilter

Table A.8: Summary of p_{T} sliced $Z \to ee$ + jets Sherpa samples used in the analysis.

DSID	Name
363388	Sherpa_NNPDF30NNLO_Zee_Pt0_70_CVetoBVeto
363389	Sherpa_NNPDF30NNLO_Zee_Pt0_70_CFilterBVeto
363390	Sherpa_NNPDF30NNLO_Zee_Pt0_70_BFilter
363391	Sherpa_NNPDF30NNLO_Zee_Pt70_140_CVetoBVeto
363392	Sherpa_NNPDF30NNLO_Zee_Pt70_140_CFilterBVeto
363393	Sherpa_NNPDF30NNLO_Zee_Pt70_140_BFilter
363394	Sherpa_NNPDF30NNLO_Zee_Pt140_280_CVetoBVeto
363395	Sherpa_NNPDF30NNLO_Zee_Pt140_280_CFilterBVeto
363396	Sherpa_NNPDF30NNLO_Zee_Pt140_280_BFilter
363397	Sherpa_NNPDF30NNLO_Zee_Pt280_500_CVetoBVeto
363398	Sherpa_NNPDF30NNLO_Zee_Pt280_500_CFilterBVeto
363399	Sherpa_NNPDF30NNLO_Zee_Pt280_500_BFilter
363400	Sherpa_NNPDF30NNLO_Zee_Pt500_700_CVetoBVeto
363401	Sherpa_NNPDF30NNLO_Zee_Pt500_700_CFilterBVeto
363402	Sherpa_NNPDF30NNLO_Zee_Pt500_700_BFilter
363403	Sherpa_NNPDF30NNLO_Zee_Pt700_1000_CVetoBVeto
363404	Sherpa_NNPDF30NNLO_Zee_Pt700_1000_CFilterBVeto
363405	Sherpa_NNPDF30NNLO_Zee_Pt700_1000_BFilter
363406	Sherpa_NNPDF30NNLO_Zee_Pt1000_2000_CVetoBVeto
363407	Sherpa_NNPDF30NNLO_Zee_Pt1000_2000_CFilterBVeto
363408	Sherpa_NNPDF30NNLO_Zee_Pt1000_2000_BFilter
363409	Sherpa_NNPDF30NNLO_Zee_Pt2000_E_CMS_CVetoBVeto
363410	Sherpa_NNPDF30NNLO_Zee_Pt2000_E_CMS_CFilterBVeto
363411	Sherpa_NNPDF30NNLO_Zee_Pt2000_E_CMS_BFilter

Table A.9: Summary of p_{T} sliced Sherpa samples for $W \to \tau\nu$ + jets used in the analysis.

DSID	Name
363331	Sherpa_NNPDF30NNLO_Wtaunu_Pt0_70_CVetoBVeto
363332	Sherpa_NNPDF30NNLO_Wtaunu_Pt0_70_CFilterBVeto
363333	Sherpa_NNPDF30NNLO_Wtaunu_Pt0_70_BFilter
363334	Sherpa_NNPDF30NNLO_Wtaunu_Pt70_140_CVetoBVeto
363335	Sherpa_NNPDF30NNLO_Wtaunu_Pt70_140_CFilterBVeto
363336	Sherpa_NNPDF30NNLO_Wtaunu_Pt70_140_BFilter
363337	Sherpa_NNPDF30NNLO_Wtaunu_Pt140_280_CVetoBVeto
363338	Sherpa_NNPDF30NNLO_Wtaunu_Pt140_280_CFilterBVeto
363339	Sherpa_NNPDF30NNLO_Wtaunu_Pt140_280_BFilter
363340	Sherpa_NNPDF30NNLO_Wtaunu_Pt280_500_CVetoBVeto
363341	Sherpa_NNPDF30NNLO_Wtaunu_Pt280_500_CFilterBVeto
363342	Sherpa_NNPDF30NNLO_Wtaunu_Pt280_500_BFilter
363343	Sherpa_NNPDF30NNLO_Wtaunu_Pt500_700_CVetoBVeto
363344	Sherpa_NNPDF30NNLO_Wtaunu_Pt500_700_CFilterBVeto
363345	Sherpa_NNPDF30NNLO_Wtaunu_Pt500_700_BFilter
363346	Sherpa_NNPDF30NNLO_Wtaunu_Pt700_1000_CVetoBVeto
363347	Sherpa_NNPDF30NNLO_Wtaunu_Pt700_1000_CFilterBVeto
363348	Sherpa_NNPDF30NNLO_Wtaunu_Pt700_1000_BFilter
363349	Sherpa_NNPDF30NNLO_Wtaunu_Pt1000_2000_CVetoBVeto
363350	Sherpa_NNPDF30NNLO_Wtaunu_Pt1000_2000_CFilterBVeto
363351	Sherpa_NNPDF30NNLO_Wtaunu_Pt1000_2000_BFilter
363352	Sherpa_NNPDF30NNLO_Wtaunu_Pt2000_E_CMS_CVetoBVeto
363353	Sherpa_NNPDF30NNLO_Wtaunu_Pt2000_E_CMS_CFilterBVeto
363354	Sherpa_NNPDF30NNLO_Wtaunu_Pt2000_E_CMS_BFilter

Table A.10: Summary of p_{T} sliced Sherpa samples for $W \to \mu\nu$ + jets used in the analysis.

DSID	Name
363436	Sherpa_NNPDF30NNLO_Wmunu_Pt0_70_CVetoBVeto
363437	Sherpa_NNPDF30NNLO_Wmunu_Pt0_70_CFilterBVeto
363438	Sherpa_NNPDF30NNLO_Wmunu_Pt0_70_BFilter
363439	Sherpa_NNPDF30NNLO_Wmunu_Pt70_140_CVetoBVeto
363440	Sherpa_NNPDF30NNLO_Wmunu_Pt70_140_CFilterBVeto
363441	Sherpa_NNPDF30NNLO_Wmunu_Pt70_140_BFilter
363442	Sherpa_NNPDF30NNLO_Wmunu_Pt140_280_CVetoBVeto
363443	Sherpa_NNPDF30NNLO_Wmunu_Pt140_280_CFilterBVeto
363444	Sherpa_NNPDF30NNLO_Wmunu_Pt140_280_BFilter
363445	Sherpa_NNPDF30NNLO_Wmunu_Pt280_500_CVetoBVeto
363446	Sherpa_NNPDF30NNLO_Wmunu_Pt280_500_CFilterBVeto
363447	Sherpa_NNPDF30NNLO_Wmunu_Pt280_500_BFilter
363448	Sherpa_NNPDF30NNLO_Wmunu_Pt500_700_CVetoBVeto
363449	Sherpa_NNPDF30NNLO_Wmunu_Pt500_700_CFilterBVeto
363450	Sherpa_NNPDF30NNLO_Wmunu_Pt500_700_BFilter
363451	Sherpa_NNPDF30NNLO_Wmunu_Pt700_1000_CVetoBVeto
363452	Sherpa_NNPDF30NNLO_Wmunu_Pt700_1000_CFilterBVeto
363453	Sherpa_NNPDF30NNLO_Wmunu_Pt700_1000_BFilter
363454	Sherpa_NNPDF30NNLO_Wmunu_Pt1000_2000_CVetoBVeto
363455	Sherpa_NNPDF30NNLO_Wmunu_Pt1000_2000_CFilterBVeto
363456	Sherpa_NNPDF30NNLO_Wmunu_Pt1000_2000_BFilter
363457	Sherpa_NNPDF30NNLO_Wmunu_Pt2000_E_CMS_CVetoBVeto
363458	Sherpa_NNPDF30NNLO_Wmunu_Pt2000_E_CMS_CFilterBVeto
363459	Sherpa_NNPDF30NNLO_Wmunu_Pt2000_E_CMS_BFilter

Table A.11: Summary of p_{T} sliced Sherpa samples for $W \to e\nu$ + jets used in the analysis.

DSID	Name
363460	Sherpa_NNPDF30NNLO_Wenu_Pt0_70_CVetoBVeto
363461	Sherpa_NNPDF30NNLO_Wenu_Pt0_70_CFilterBVeto
363462	Sherpa_NNPDF30NNLO_Wenu_Pt0_70_BFilter
363463	Sherpa_NNPDF30NNLO_Wenu_Pt70_140_CVetoBVeto
363464	Sherpa_NNPDF30NNLO_Wenu_Pt70_140_CFilterBVeto
363465	Sherpa_NNPDF30NNLO_Wenu_Pt70_140_BFilter
363466	Sherpa_NNPDF30NNLO_Wenu_Pt140_280_CVetoBVeto
363467	Sherpa_NNPDF30NNLO_Wenu_Pt140_280_CFilterBVeto
363468	Sherpa_NNPDF30NNLO_Wenu_Pt140_280_BFilter
363469	Sherpa_NNPDF30NNLO_Wenu_Pt280_500_CVetoBVeto
363470	Sherpa_NNPDF30NNLO_Wenu_Pt280_500_CFilterBVeto
363471	Sherpa_NNPDF30NNLO_Wenu_Pt280_500_BFilter
363472	Sherpa_NNPDF30NNLO_Wenu_Pt500_700_CVetoBVeto
363473	Sherpa_NNPDF30NNLO_Wenu_Pt500_700_CFilterBVeto
363474	Sherpa_NNPDF30NNLO_Wenu_Pt500_700_BFilter
363475	Sherpa_NNPDF30NNLO_Wenu_Pt700_1000_CVetoBVeto
363476	Sherpa_NNPDF30NNLO_Wenu_Pt700_1000_CFilterBVeto
363477	Sherpa_NNPDF30NNLO_Wenu_Pt700_1000_BFilter
363478	Sherpa_NNPDF30NNLO_Wenu_Pt1000_2000_CVetoBVeto
363479	Sherpa_NNPDF30NNLO_Wenu_Pt1000_2000_CFilterBVeto
363480	Sherpa_NNPDF30NNLO_Wenu_Pt1000_2000_BFilter
363481	Sherpa_NNPDF30NNLO_Wenu_Pt2000_E_CMS_CVetoBVeto
363482	Sherpa_NNPDF30NNLO_Wenu_Pt2000_E_CMS_CFilterBVeto
363483	Sherpa_NNPDF30NNLO_Wenu_Pt2000_E_CMS_BFilter

B | Figures after each cut selection

Figure B.1: Distributions at preselection including b-jet veto and OS(τ, μ) cut.

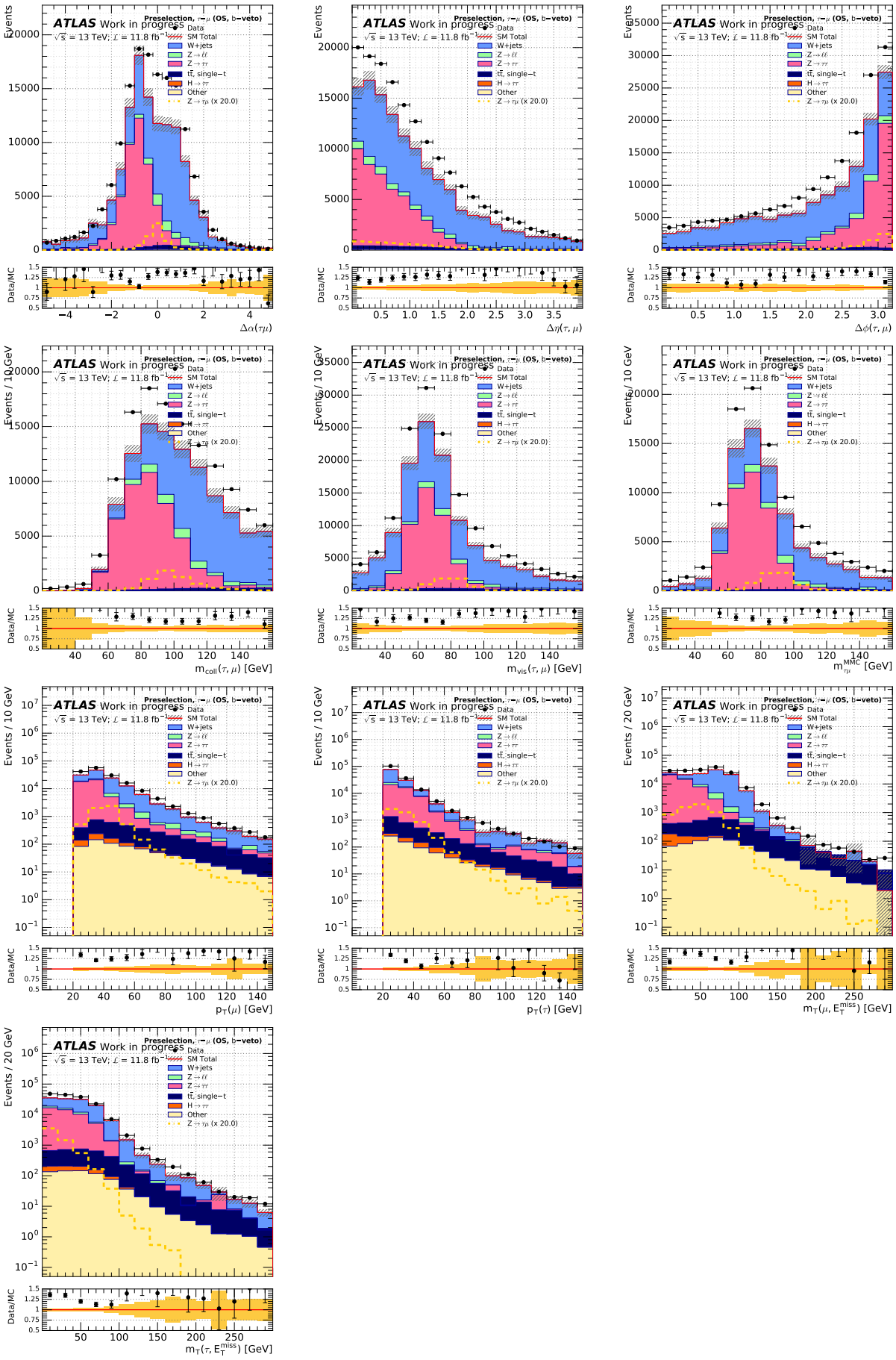


Figure B.2: Distributions in the 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV window.

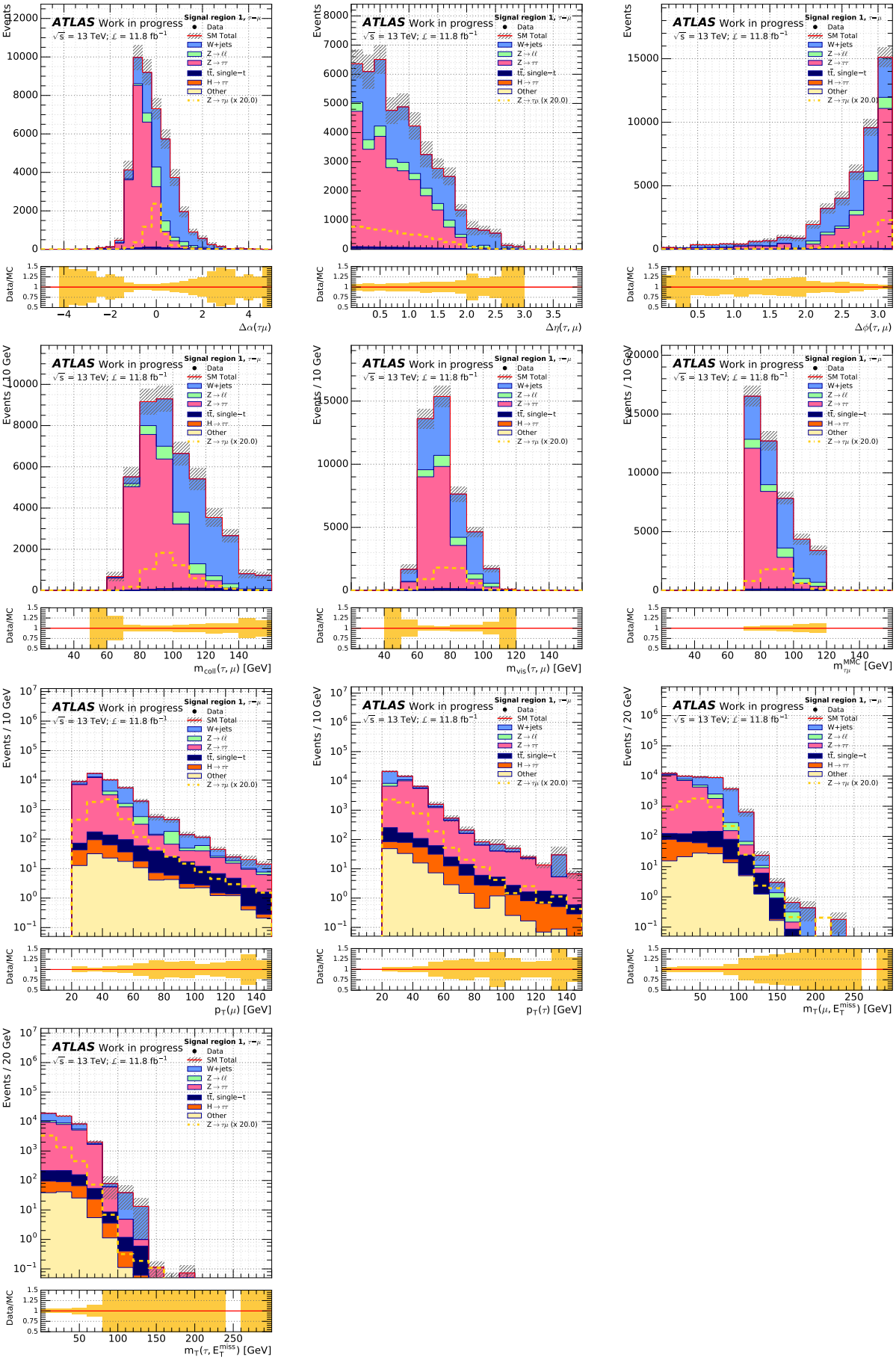


Figure B.3: Distributions in 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV after the $p_{\mathrm{T}}(\mu) > 30$ GeV, $p_{\mathrm{T}}(\tau) > 20$ GeV cuts.

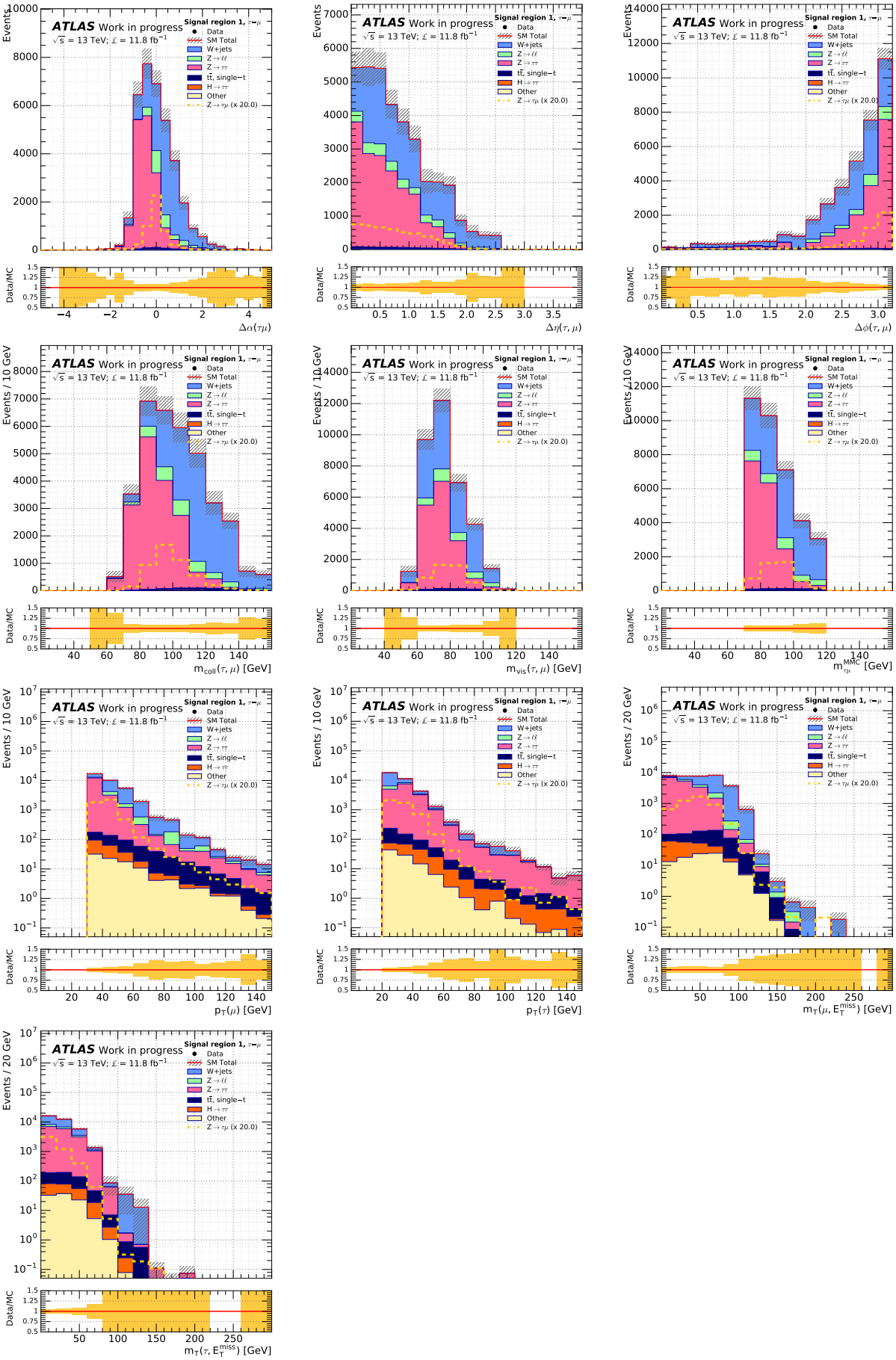


Figure B.4: Distributions in 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV after $p_{\mathrm{T}}(\mu) > 30$ GeV, $p_{\mathrm{T}}(\tau) > 20$ GeV and $|\Delta\eta| < 2.0$ cuts.

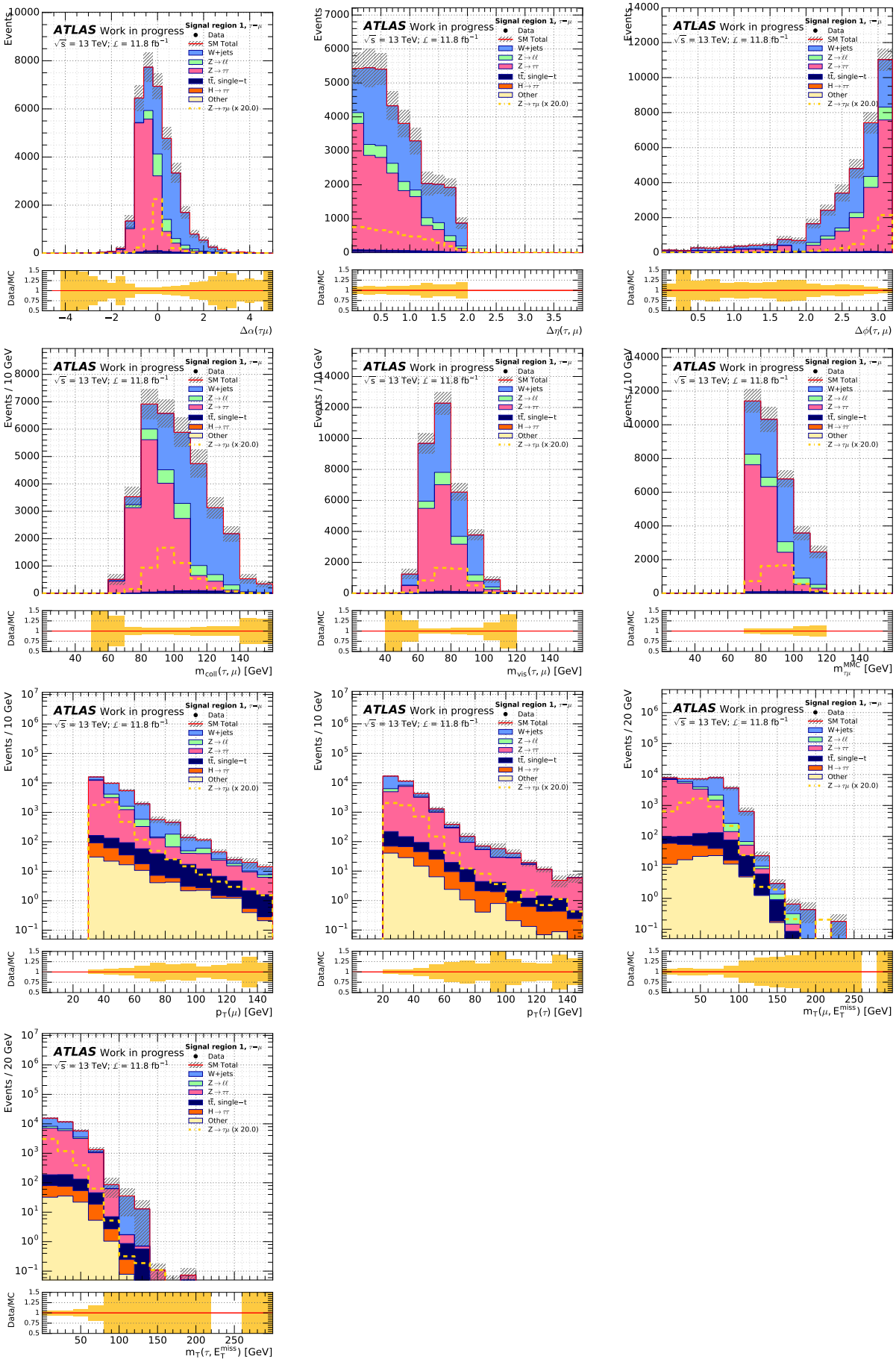


Figure B.5: Distributions in 70 GeV $< m_{\tau,\mu}^{\mathrm{MMC}} <$ 120 GeV after $p_{\mathrm{T}}(\mu) > 30$ GeV, $p_{\mathrm{T}}(\tau) > 20$ GeV, $|\Delta\eta| < 2.0$ and $m_{\mathrm{vis}} > 60$ GeV cuts.

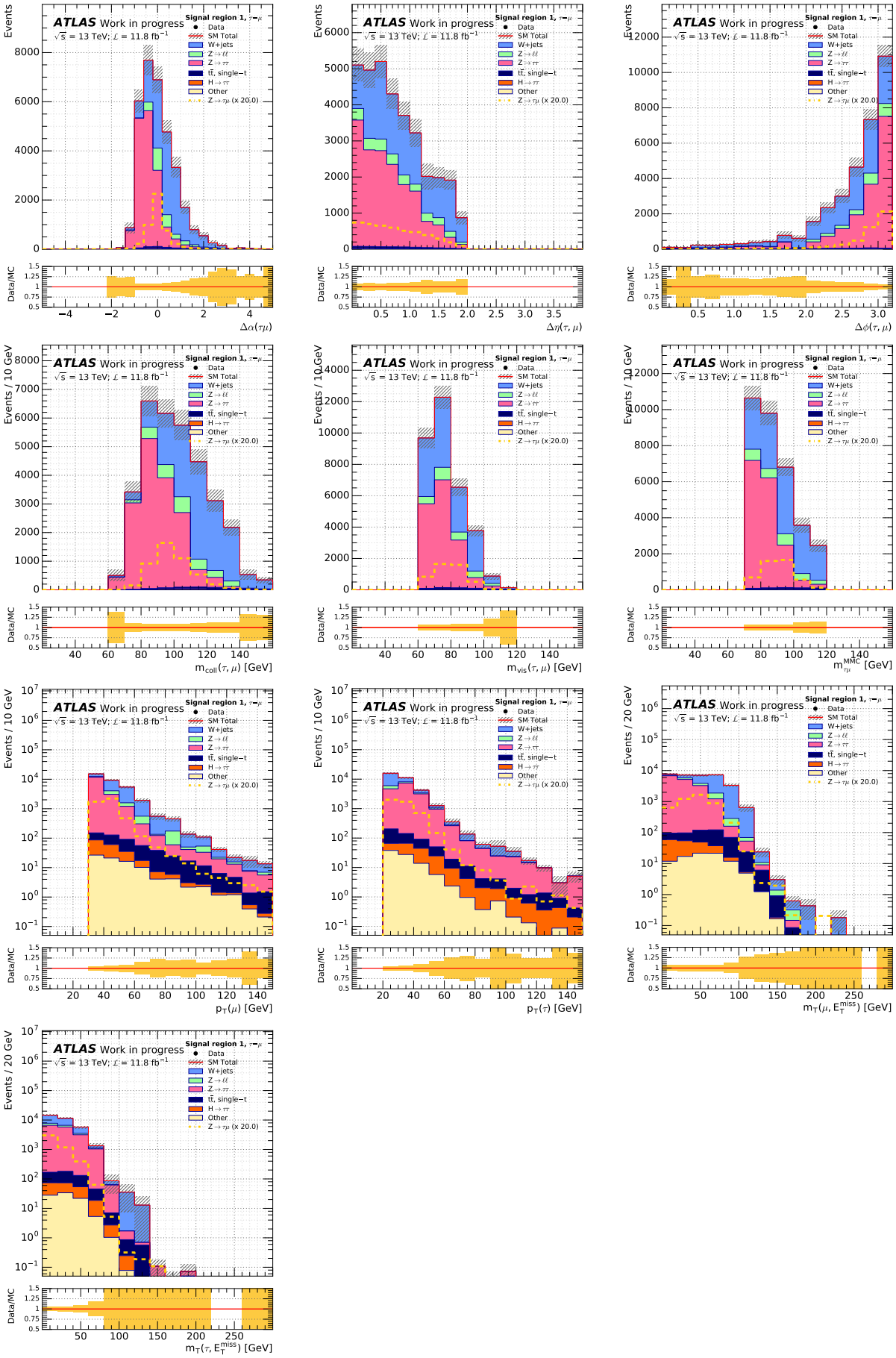


Figure B.6: Distributions in signal region 1 after all common cuts.

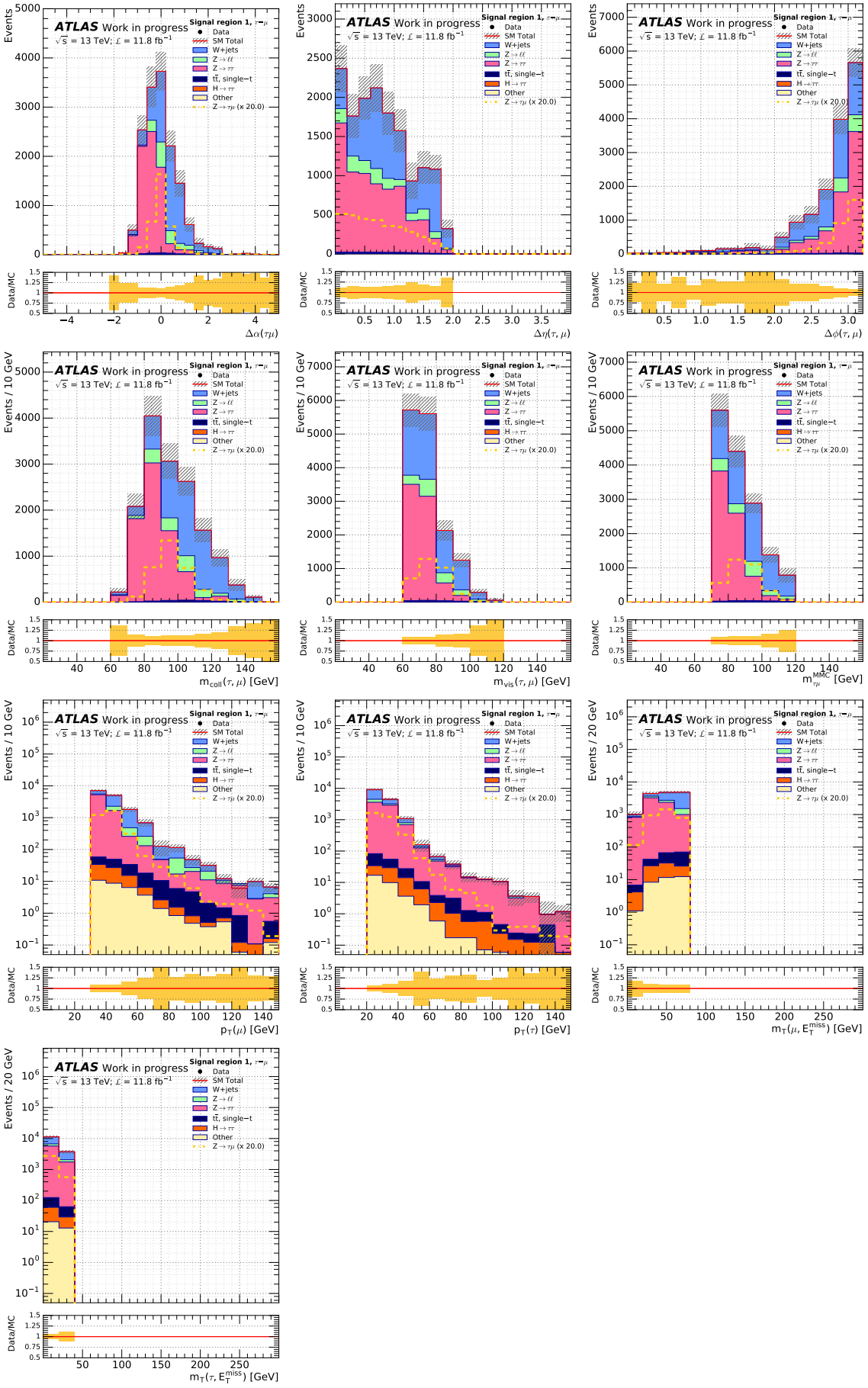


Figure B.7: Distributions in signal region 2 after all common cuts.