

A SEARCH FOR ASSOCIATED PRODUCTION
OF A SM HIGGS DECAYING INTO TAU
LEPTONS WITH THE ATLAS EXPERIMENT

L. A. Thomsen
Copenhagen University

Thesis submitted for the degree of
Doctor of Philosophy

August 31, 2014

Abstract

Two distinct topics are covered in this thesis: tau hadronic decay identification and the search for a Standard Model (SM) Higgs boson in associated production with a W boson decaying into tau leptons within the ATLAS experiment.

Prior to the work presented in this thesis, tau identification has mainly relied on charged π^\pm , and tracking information and global variables exploiting the energy deposited in the calorimeter. Improvement of the tau identification for hadronically decaying taus is sought using new variables derived from a newly available algorithm capable of π^0 reconstruction. The added information about π^0 's together with a re-optimization of the algorithm results in gain of roughly 30% in rejection of background over the full range of transverse momentum. Furthermore the threshold for tau identification can be lowered from 20 GeV to 15 GeV, which is potentially very valuable for physics analyses involving hadronic tau decays.

The second part of this thesis covers the search for associated production of a SM Higgs boson decaying into tau leptons with the ATLAS experiment. The search procedure is made robust by developing a background estimation method relying on simulation for the irreducible background and a data-driven estimation for the reducible backgrounds. The data are still blinded internally in ATLAS. The result found is an expected upper limit on σ_{VH} of $5 \times \sigma_{SM}$ for 20.3 fb^{-1} which implies that, when combining all WH and ZH final states and including also ZH associated production one should be able to reach sensitivity to SM production of Higgs in association to vector bosons in Run 1 data of LHC. This also provides proof that this channel is important for Higgs property measurements in Run 2 and beyond.

Resume

To adskilte emner er dækket i denne opgave: identificering af hadroniske tau henfald og en søgen efter en Standard Model Higgs boson ved associeret produktion med en W boson, der henfalder til to tau leptoner i ATLAS experimentet.

Før arbejdet, der er præsenteret i opgaven, var tau identifikation hovedsageligt baseret på ladede π^\pm . Forbedringer af tau identifikationen for hadroniske henfald er søgt via nye variable fundet med en ny algoritme, der kan genskabe π^0 partikler. Disse nye variable, og en optimering af algoritmerne, resulterer i en forbedret reduktion af baggrunden på omkring 30 % over hele sprektrret af transvers impuls. Yderligere kan tau partikler nu identificeres ned til 15 GeV istedet for 20 GeV. Disse forbedringer er nyttige for analyser ved ATLAS, der involverer tau henfald.

Den anden del af opgaven omhandler en søgen efter en SM Higgs boson, der henfalder til to tau leptoner ved ATLAS experimentet. Analyse metoden er gjort robust ved at formulere en metode til estimering af baggrunden, der benytter simuleringer for den irreducible baggrund og en data-dreven estimering for de reducible baggrunde. Data er stadig blændet internt i ATLAS. Der er opnået en forventet øvre grænse for σ_{VH} på $5 \times \sigma_{SM}$ ved $20.3 fb^{-1}$. Dette betyder at ved samling af alle WH og ZH kanaler kan der opnåes sensitivitet for SM produktion af Higgs associeret med en vektor boson i run 1 ved LHC. Dette betyder også at denne kanal er vigtig for Higgs målinger ved genopstarten af LHC

Preface

In 2012 it finally happened: the discovery of a Higgs-like particle was made after many years of search. This elusive particle was predicted in 1964 by Francois Englert, Robert Brout, Peter Higgs, Gerald Guralnik, Carl Hagen and Tom Kibble as an important piece in the Standard Model [1] [2] [3] [4]. After the discovery of the Higgs-like particle a Nobel prize was awarded to Peter Higgs and Francois Englert in 2013.

To be a PhD student in this period of time has been a rewarding experience, it has been incredible to be in the field at a time of such a big discovery. I had the chance to be present at the official announcement of the Higgs particle discovery at CERN in July 2012, after many hours of queuing in front of the CERN main auditorium. It has been very interesting to be around for the discussions which have been going on for many years and suddenly changed context very rapidly. "Will the LHC start before the tram is completed? Does the mass window for discovering the Higgs not start to be very small? Will LHC ever find anything? Is there a real sign of a Higgs now? Is it true that CMS now sees a bigger signal? Is there a mass difference between the channels? Does the Higgs-like particle only couple to bosons and not to fermions like taus?... " But surprisingly fast, one question after another was answered, and with the start-up of the LHC in 2015 more questions will be asked and answered. A particle is found which couples to both fermions and bosons, and it seems to have the expected spin [5]. But the main question to ask is still: "Is it really the Standard Model Higgs boson that has been found?"

The Higgs search described in this thesis aims at exploring the potential of the Vector Boson associated Higgs production in current data. This channel is a new channel in the search for Higgs decay in tau leptons in ATLAS. If the investigated channel can enhance the precision of the Higgs boson measurement, and thereby help to determine if the new particle discovered is the Standard Model Higgs Boson, then it should be understood now in Run 1 data and exploited as best possible in Run 2 and later.

Outline of the thesis

This thesis consists of three parts:

1. An introduction and theoretical motivation along with a description of the experimental facility
2. A chapter on tau identification, which covers a study I performed to improve the at that time default ATLAS algorithm on the tau identification. The variables used for identification are presented, and the optimal set of variables are found.

The main part of my contribution to this work is presented in chapters 4 and 5.

- Chapter 3 describes the general principles of the tau hadronic decay reconstruction in ATLAS, and the input variables to be used.

-
- Chapter 4 finds the optimal settings for the Boosted Decision Tree algorithm making a robust algorithm for the identification work.
 - Chapter 5 investigates if the new set of proposed variables give a better performance in the identification of the taus, with respect to the at that time default algorithm.
3. A chapter on the search for associated production of a SM Higgs decaying into tau leptons with the ATLAS experiment

My main contributions to this Higgs search are:

- The trigger and event selection as described in section 6.
By studying the signal and background in simulations, the optimal event selection is found
- The background method as described in section 7.
A lot of effort has been devoted to derive a method to estimate the reducible background, finding the equations which describe the data-driven background estimation. This chapter also introduces a number of important entities: fake-rates, control regions and side-bands.
- The fake-rate of muon and taus as described in chapter 8.
The fake-rates are found in different regions from the signal region or sidebands making it possible to estimate fake lepton contributions correctly into the signal region. Both the muon fake-rate and the tau fake-rate are found in two samples, one applicable to the electro-weak (EW) background contributions and one applicable to non-EW background contributions.
- The content of the applied side-bands is investigated in chapter 9.
By studying this, it is possible to determine which fake-rate is most applicable to the different regions.
- Validation tests of the background estimation method is done in chapter 10.
Tests are done for the muon fake-rate and the tau fake-rate. Further tests are done in a region where the same processes contribute to the background as in the signal region, with roughly the same importance. Lastly a test is done at a pre-selection step, where it is not the same important backgrounds as in the signal region, but it is possible to test the equations used.
- Results and expectation limits are given in chapter 11.
The plots of the distributions for the investigated channel is shown. Unfortunately only the expectation is shown due to the blinded status of the analysis in ATLAS.

This is still work in progress, and currently still blinded internally in the ATLAS collaboration. Therefore some plots are shown without data or with a looser tau identification to get a lower purity.

During my PhD I have worked with the following:

I started my PhD in 2010 working on the ATLAS trigger system. The first six months of my

PhD I was involved in investigating a missing transverse energy plus tau trigger using the so-called missing transverse energy significance variable. During this period I was at CERN and did the work with S. Xella, G. Lewis (New York University), D. Casadei (New York University). This work is not included in this thesis, but was for me an important step to understand some of the big challenges an experiment as ATLAS faces in recording data. The work also contributed to the decision of not to rely on such variables for triggering in the 2012 run as estimated to be less robust than missing transverse energy.

In the summer of 2012 I worked for half a year on identification of tau leptons with mainly, P. H. Sales De Bruin (University of Washington), S. Xella and M. Flechl. (University of Freiburg). The details of this work are presented in this thesis and the study was used in the tau identification of the ATLAS experiment. In the summer of 2013 I was part of the tau validation team for a short duration with K. Bristow (Wits University), T. Bristow (Wits University), and G. Hamity (Wits University).

The last year of my PhD was spend on the Higgs search in the associated production with vector bosons. The working group here consisted mainly of collaborators from Melbourne University, Sydney University and Yale University. I spent three weeks at Yale University in the start of 2014 to facilitate a closer collaboration and settle on the background estimation method. A detailed description of the analysis is included in this thesis.

Acknowledgement

My greatest gratitude goes to my supervisor and friend Associate Professor Stefania Xella. I am very grateful for her support throughout the whole PhD, keeping me going when times were tough, always asking insightful and critical questions, and always willing to spend a great time and offer invaluable advice. Stefania, I really enjoyed working with you, and I know that I will miss (working with) you in the future.

I would like to thank both Almut Pingel and Alejandro (Alex) Alonso Diaz for always investing themselves in all kinds of discussions from tau id to the implementation of RootCore packages. I also want to thank the former PhD students at NBI, Zofia Czyczu and Pavel Jez, for being very helpful in the start-up phase.

The process of bringing my thoughts and results to paper have been facilitated by the meticulous proofreading of Morten Soerensen and Graig Wiglesworth.

Finally I will like to appreciate the entire group at NBI for a pleasant atmosphere during the past years.

Contents

1	Standard Model	1
1.1	Standard Model	1
1.1.1	The Higgs mechanism	3
1.1.2	Status of the Standard model	5
1.2	Search for the Higgs	6
1.2.1	Production mechanisms of the Higgs boson	6
1.2.2	Decay channels of the Higgs boson	8
1.3	Status of the Higgs boson search	8
1.4	Status of the associated production channel	12
1.5	Outlook for the associated production channel	13
2	CERN	15
2.1	LHC	15
2.2	Hadron collider physics	17
2.2.1	Simulations	18
2.3	ATLAS	20
2.3.1	The Inner Detector	21
2.3.2	Calorimeters	23
2.3.3	Muon spectrometer	25
2.3.4	Magnets	27
2.3.5	Trigger	27
2.4	Luminosity	28
2.4.1	Pile-up	28
2.5	Reconstruction of muons	29
2.6	Reconstruction of electrons	30
2.7	Reconstruction of jets	31
2.8	Reconstruction of the transverse missing energy	31
I	Tau identification	33
3	Tau identification inputs	35
3.1	Properties of the τ particle	35
3.1.1	Hadronic τ -decays in the ATLAS detector.	36

3.2	π^0 reconstruction	38
3.2.1	Reconstruction of the topological clusters	38
3.2.1.1	Selection of clusters	39
3.2.2	The number of π^0	40
3.3	Data and preselection cuts	43
3.4	Tau identification	43
3.5	Current variables for tau identification	44
3.6	New variables for tau identification based on π^0 information	48
4	Boosted Decision Tree algorithm	53
4.1	Boosted Decision Tree algorithm	53
4.2	Settings for the Boosted Decision Tree algorithm	55
5	Incorporation of the new tau identification variables and re-optimization	59
5.1	Training with the new tau identification variables	59
5.2	Correlation between the tau identification variables	59
5.2.1	Optimal set of tau identification variables	66
5.3	Performance increase due to the new set of variables	70
II	A search for the SM Higgs produced in association with a vector boson	71
6	A search for the SM Higgs produced in association with a vector boson	73
6.1	Determining the search procedure	73
6.2	Background processes	74
6.3	Data	76
6.4	Monte-Carlo simulations	77
6.5	Defining the basic objects for the search	78
6.5.1	Overlap removal	81
6.6	Trigger selection	81
6.6.1	Trigger matching	82
6.7	Association of the leptons to the Higgs Boson	82
6.8	Optimal event selection	82
6.9	Main background contributions	91
6.9.1	top + anti-top production	91
6.9.2	W production in association with jets	92
7	The background estimation methods	93
7.1	Background types	93
7.2	Irreducible background	93
7.3	Reducible background	93
7.3.1	Definitions	94
7.3.2	Equations for the fake-rate estimation	96

8	Determination of FR_μ and FR_τ	99
8.1	Requirements for a sample enriched in $Z \rightarrow ll + jets$ events	99
8.1.1	FR_μ in the $Z \rightarrow ll + jets$ enriched sample	100
8.1.2	FR_τ in the $Z \rightarrow ll + jets$ enriched sample	102
8.2	Requirements for a sample enriched in multi-jets events	104
8.2.1	FR_μ in the multi-jets enriched sample	105
8.2.2	FR_τ in the multi-jets enriched sample	106
9	The side-band regions of the data-driven background estimation method	109
10	Validation of the data-driven background estimation method	117
10.1	Validation of the data-driven background estimation method in $W \rightarrow l\nu + jets$	117
10.2	Validation of the data-driven estimation method in a SSS control region . . .	120
10.3	Test of the data-driven background estimation method at a pre-selection stage	124
10.4	Validation of FR_μ	126
11	Results	129
11.1	Kinematic distributions in the signal region	129
11.2	Systematic Uncertainties	133
11.2.1	Uncertainty on the data-driven background estimation	133
11.2.2	Experimental uncertainty on simulations	133
11.2.3	Theoretical uncertainties	134
11.3	Expected limits	135
12	Conclusion	139
	Bibliography	141
	Appendices	147
A	Side-bands equation	149
B	Truth parts contamination	151

1

Standard Model

1.1 Standard Model

The Standard Model (SM) is the general theoretical description of the fundamental building blocks of the universe. The SM is best described in a Quantum Field Theory (QFT) framework using three kinds of fields: fermion fields, vector fields and scalar fields [6][7].¹ There are two kinds of particles in the SM;

Fermions - spin 1/2: Matter is made out of the fermions listed in table 1.1.1 which are spin half particles. Fermions can either be quarks or leptons, where the former are in nature confined to groups of quarks, either three quarks (baryons) or a quark and an anti-quark (mesons). An anti-quark have the same properties as the corresponding quark, but with reversed sign of some properties as for example charge. Particles in the SM interact via the forces listed in table 1.1.2. Quarks are the only kind of particles that interact via all three kinds of forces. The other kind of fermions, leptons, do not interact via the strong force and are able to move freely and are not confined. The charged leptons are e , μ and τ , and they can interact with both the weak and the electromagnetic force whereas the uncharged leptons (ν_e , ν_μ and ν_τ) can interact only via the weak force [8].

Bosons - spin 1: Bosons with spin 1 are force carriers and are associated with the interaction forces listed in table 1.1.2. Gluons and W^\pm can also interact with themselves[8].

Boson - spin 0: The Higgs boson. As shown later this particle of the SM allows the other particles, except the neutrino ν , do have mass [8].

¹QFT is the mathematical formalism, based on the principles of quantum mechanics and special relativity, able to describe physics at the subatomic scale

Particles (mass[MeV], Charge, spin)			
Quarks	up(2.4,2/3,1/2)	charm (1270,2/3,1/2)	top(171200,2/3,1/2)
	down(4.8,-1/3,1/2)	strange (104,-1/3,1/2)	bottom (4200,2/3,1/2)
Leptons	e (0.511,-1,1/2)	μ (105.7,-1,1/2)	τ (1777,-1,1/2)
	ν_e (<0.02 · 10 ⁻⁶ ,0,1/2)	ν_μ (< 0.17,0,1/2)	ν_τ (< 15.5,0,1/2)
	First family	Second family	Third family

Table 1.1.1: Fermions

Force	Name (mass[GeV], Charge, spin)	Acts on
Weak	W^\pm (80.4, ± 1 ,1) Z^0 (91.2,0,1)	Quarks and leptons
Electromagnetic	Photon(0,0,1)	Quarks, charged leptons and W^\pm
Strong	Gluon(0,0,1)	Gluons and quarks

Table 1.1.2: Bosons

The SM is described by a Lagrangian density(\mathcal{L}). In classical mechanics a Lagrangian (L) describes the dynamics of a system by the kinetic energy T and the potential V: L=T-V. In QFT the formalism is extended to continuous fields instead of discrete of entities using the Lagrangian density. An important property for the \mathcal{L} of the SM is gauge symmetry.

Gauge symmetry \mathcal{L} admits a local (depending on space-time) gauge symmetry if \mathcal{L} remains unchanged under a local gauge transformation. The important property of this symmetry is that any invariance of \mathcal{L} under such transformation relates to the conservation of a quantity. This is a crucial step in QFT, since it appears as all interactions in nature are associated with fundamental conserved quantities e.g. for the Electromagnetic interaction the conserved quantity is the electric charge [7].

In the SM a notation for the gauge groups is used: U(1) is the notation of the local gauge transform with one force carrier (gauge boson). SU(2) and SU(3) are local gauge transformations with $2^2 - 1 = 3$ and $3^2 - 1 = 8$ force carriers.

The method to obtain the SM \mathcal{L} describing an interaction is similar for all interactions: Consider \mathcal{L} of a free particle, enforce local gauge symmetry and thereby obtain a gauge invariant \mathcal{L} with interaction terms. The \mathcal{L} describing the SM consists of the following terms:

$$\mathcal{L} = \mathcal{L}_{gauge} + \mathcal{L}_{ferm} + \mathcal{L}_H + \mathcal{L}_{Yuk} \tag{1.1}$$

\mathcal{L}_{gauge} describes the kinetic energies and self-interactions of the gauge fields. The free propagation of the gauge fields is allowed due the kinetic energy. \mathcal{L}_{ferm} is the kinetic energies of the fermions and their interactions with the gauge fields. \mathcal{L}_H and \mathcal{L}_{Yuk} are present due to the Higgs coupling. \mathcal{L}_H describes the interaction between the gauge boson and the Higgs boson, and also the Higgs self-coupling. \mathcal{L}_{Yuk} describes the interaction between fermions and the Higgs field. \mathcal{L}_H and \mathcal{L}_{Yuk} will be treated in the next section [1] [2] [3] [4].

A critical point, of the terms \mathcal{L}_{gauge} and \mathcal{L}_{ferm} , is that any fermion mass terms and mass terms for gauge bosons violate the local gauge symmetry. No attempts to introduce such masses in this group have been successful. Therefore this symmetry needs to be broken for the gauge bosons and the fermions to get the observed masses. This is done by the Higgs mechanism [7].

The two important interactions in the SM are:

Quantum chromodynamics (QCD) QCD describes the strong interaction, and is described by the SU(3) group. The strong interaction is what keeps protons and neutrons together in a nucleus. The basic idea is that the quarks interact strongly through the exchange of a massless gauge vector field, a gluon.

Electro-Weak interaction The Quantum Electro Dynamics (QED) is described by the U(1) group, and it describes the interaction of charged particles via photons. The weak interaction is responsible for radioactive decays as for example β -decay. As the U(1) group describes QED interaction and SU(3) describes strong interactions, the SU(2) group describes weak interactions. The electromagnetic and weak forces can be unified and thereby denoted $SU(2)_L \times U(1)$. L denotes left-handed particles [7].

1.1.1 The Higgs mechanism

Particles have mass in nature and to allow the SM to describe this the Higgs mechanism is needed. To understand this Higgs mechanism one can first consider a complex Higgs scalar field, ϕ :

$$\phi = \phi_1(x) + i\phi_2(x) \tag{1.2}$$

\mathcal{L}_H is introduced with the potential $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$:

$$\mathcal{L}_H = (D^\mu\phi)^\dagger D_\mu\phi - V(\phi) \tag{1.3}$$

$$= (D^\mu\phi)^\dagger D_\mu\phi - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{1.4}$$

Where

$$D_\mu\phi \approx \partial_\mu + g^\alpha W^\alpha\phi + g'B\phi \tag{1.5}$$

W^α and B are the four gauge bosons of the SU(2) and U(1) symmetry. g^α and g' are the coupling constants and $D_\mu\phi$ contains the interaction of the Higgs field with the gauge fields.

The potential, $V(\phi)$, is described by a "mexican hat" as shown in figure 1.1.1. There are two cases: $\mu^2 > 0$ and $\mu^2 < 0$. The first mentioned is the simple situation of (two) massive scalar particles with positive masses and an interaction term $\lambda(\phi^\dagger\phi)^2$. In the more interesting case $\mu^2 < 0$ the minimum of $V(\phi)$ is not at $\langle\phi_1\rangle = \langle\phi_2\rangle = 0$, but along a circle of points in ϕ_1 [9].

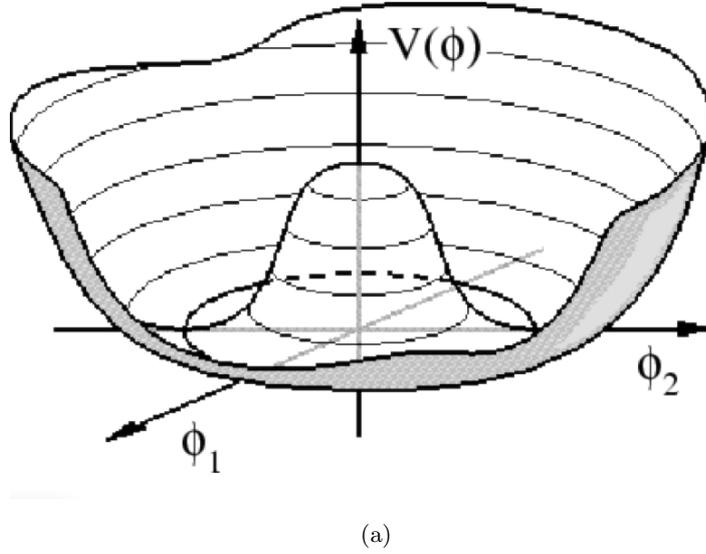


Figure 1.1.1: The so-called mexican hat

The value of the minimum energy is $\frac{\nu^2}{2}$ where $\nu = \sqrt{\frac{-\mu^2}{\lambda}}$ (λ is the coupling constant). The vacuum expectation value $\langle \phi \rangle = \nu$ is 246 GeV, and one has the freedom to choose $\phi_2 = 0$ and $\phi_1 = \nu$.

To understand how this mechanism creates mass the vacuum expectation value of the first term of \mathcal{L} , $(D^\mu \phi)^\dagger D_\mu \phi$, must be considered. The calculations are omitted for the overview given, but can be instead be found here [7]. For the SM one needs to introduce a Higgs doublet of complex scalar fields, rather than a single one. One can again choose $\langle \phi_0 \rangle = \nu$ and the others to $\langle \phi_i \rangle = 0$. Expanding equation 1.4 around the ground state one obtains the mass terms for the Z and W from a mixture of the initial W^α and B gauge fields. A massless term for the photon is also achieved along with interaction terms of the Higgs field with the gauge boson fields. The symmetry is now spontaneously broken meaning equation 1.1 is still gauge invariant, and the vacuum provides the particles with masses by interaction with them. Bottomline one finds that the mass of the W^\pm is $m_W = |g|\frac{\nu}{2}$, the mass of Z^0 is $m_Z = \left[\sqrt{g^2 + g'^2} \frac{\nu}{2} \right]$, and the photon (A) has mass zero [7] [9] [10].

The introduced Higgs boson will have a mass of $m_h = \sqrt{2\lambda}\nu$. The strength of the interactions of the gauge fields (V) to Higgs fields (λ_{HVV}) is given by:

$$\lambda_{HVV} = \frac{2m_V^2}{\nu} \quad (1.6)$$

The last missing point in the SM \mathcal{L} is a term incorporating the interaction between Higgs and the fermions, \mathcal{L}_{Yuk} . This is given by the Yukawa term:

$$\mathcal{L}_{Yuk} = \lambda_{H\psi\psi} (\bar{\psi}_R \phi^\dagger \psi_L + \bar{\psi}_L \phi \psi_R) \quad (1.7)$$

$\lambda_{H\psi\psi}$ is the coupling strength between the fermion fields, ψ , and the Higgs. ψ is split into a left-handed (ψ_L) and a right-handed (ψ_R) component. By estimating \mathcal{L}_{Yuk} in the ground state one obtain the coupling between the Higgs boson and fermions $\lambda_{Hf\bar{f}}$:

$$\lambda_{H\psi\bar{\psi}} = \frac{m_\psi}{\nu} \quad (1.8)$$

It is very important in verifying the Higgs mechanism to verify that both $\lambda_{H\psi\bar{\psi}}$ and λ_{HVV} develop as expected by the SM.

1.1.2 Status of the Standard model

The SM has had many successes over the years, and many predictions have been tested to a remarkable precision. Just one of many measurements are the mass of the bosons W^\pm and Z^0 . For the W^\pm the observed mass is $80.385 \text{ GeV} \pm 0.015 \text{ GeV}$ and the predicted is 80.372 GeV . For the Z^0 the observed mass is $91.1876 \text{ GeV} \pm 0.021 \text{ GeV}$ and the predicted is 91.1875 [11].

It is fair to say that in 2012 a new discovery also contributed significantly to the success of the SM: The discovery of a Higgs-like particle at a mass of about 125 GeV [12] [13]. The status of this search will be described in greater details in section 1.2.1

Even though there are many successful precision measurements of the Standard Model, there are also some problems with the theory. A few of the critical points are:

Dark matter The known matter accounted for in the SM only accounts for 4.9% of the matter in the universe. The dark matter seems to constitute around 26.8% of the universe energy content.

Dark energy The universe expands at a rate which is impossible to explain by the SM. It has been found that for the expansion to be as observed around 68.3% of the universe energy content needs to consist of dark energy

Gravity Gravity is not incorporated in the SM. First of all there are no successful descriptions of anything like graviton bosons, and secondly it is somewhat a puzzle why the force is so much weaker than the other three forces.

Amount of anti-particles The described SM offers no explanation for the preponderance of matter over anti-matter, even though this is an observation in the universe. A speculative explanation is that from the start there was slightly more matter than anti-matter or otherwise a new phenomenon causes such asymmetry and a explanation in the SM is lacking.

Neutrino masses In the SM neutrinos have no mass, but by close studies of neutrino oscillations, it seems they have some mass.

1.2 Search for the Higgs

The most recent success of the SM was in 2012 with the discovery of a Higgs-like particle. In the discovery of the new particle and in the search for the Higgs boson it is very important to ensure all the properties like mass, spin, and couplings of the observed particle are in accordance with the predicted SM Higgs. In section 1.2.1 and 1.2.2 the production modes and branching ratios of the Higgs particle are summarized. Afterwards the status of the Higgs boson experimental search is summarized.

1.2.1 Production mechanisms of the Higgs boson

There are in all four different production modes for the Higgs boson relevant at the LHC:

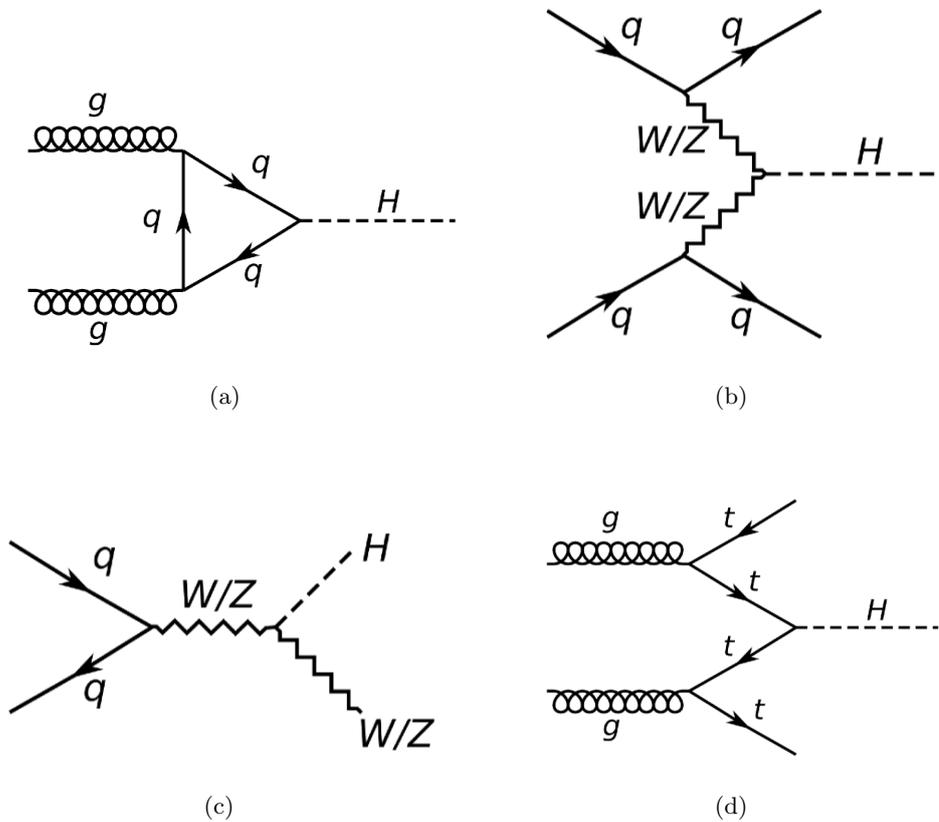


Figure 1.2.1: The production channels of the Higgs at the LHC

gluon-gluon (gg) The main production mode is gluon-gluon fusion, where Higgs bosons are produced via a triangular quark loop as shown in figure 1.2.1(a). The quark loop is necessary since the Higgs boson does not couple to massless particles like gluons. The quarks with the highest mass contribute most to the quark loop, making the top quark the dominant. Gluon-gluon fusion has by far the highest cross-section, in total it is estimated that there is around 0.5 million such events in all 2011+2012 LHC data [14].

The difficulty concerning the gg fusion is that it is difficult to design an efficient trigger, in case the Higgs decay to two τ leptons [15].

Vector Boson Fusion (VBF) The second most dominant Higgs production mode is Vector Boson Fusion, as shown in figure 1.2.1(b). This production mode has a more recognizable signature with two jets with high transverse momentum in the forward direction in addition to the Higgs decay products. This feature makes the signature easier to trigger on and to distinguish from background SM process. The estimate is that there is $40 \cdot 10^3$ Higgs produced from VBF in 2011+2012 LHC data [14].

Associated Production (WH,ZH) The associated production has a small cross-section at LHC compared to the two previous cases, but the signature is very clean. The main principle is that a W or a Z is produced from a quark-quark pair, subsequently radiating a Higgs boson. If the W/Z decay to two leptons they produce a very clean signature in the detector, whereas if they decay to quarks they exhibit a signature that is harder to distinguish. As shown in table 1.2.1 the cross-section at a given m_H is slightly higher for WH than ZH. There would be around $20 \cdot 10^3$ Higgs produced from both WH and ZH combined at the LHC in 2011+2012 data [14].

$t\bar{t}$ fusion ($t\bar{t}H$) The last production mode is $t\bar{t}$ fusion, which has a very small cross-section as it requires the production and fusion of two top quarks. Top quarks are very heavy and therefore require a high energy to be produced at the LHC. The total number of Higgs produced by $t\bar{t}$ fusion in all data is only around $3 \cdot 10^3$ [14] [15]

An overview of the cross-section for the different production modes is given in figure 1.2.2 and table 1.2.1.²

Production mode	Order perturbation theory	cross section [pb]
gg	NNLO+NNLO QCD NLO EW	$19.53^{+14.7\%}_{+14.7\%}$
VBF	NLO QCD+EW approx. NNLO QCD	$1.58^{+2.8\%}_{-3.0\%}$
WH	NNLO QCD NLO EW	$0.679^{+3.7\%}_{-4.1\%}$
ZH	NNLO QCD NLO EW	$0.394^{+5.1\%}_{-5.0\%}$
$t\bar{t}H$	NLO QCD	$0.130^{+11.6\%}_{-17.1\%}$

Table 1.2.1: Predicted cross-section for a Higgs mass 125 GeV [17] [18]

²The leading-order (LO) is the lowest order in the theoretical calculation, and is shown as the three level Feynman diagram. The higher order perturbation theory takes $O(\alpha)$ corrections into consideration thereby creating cases like next-to-leading order (NLO) and next-to-next-to-leading order (NNLO)

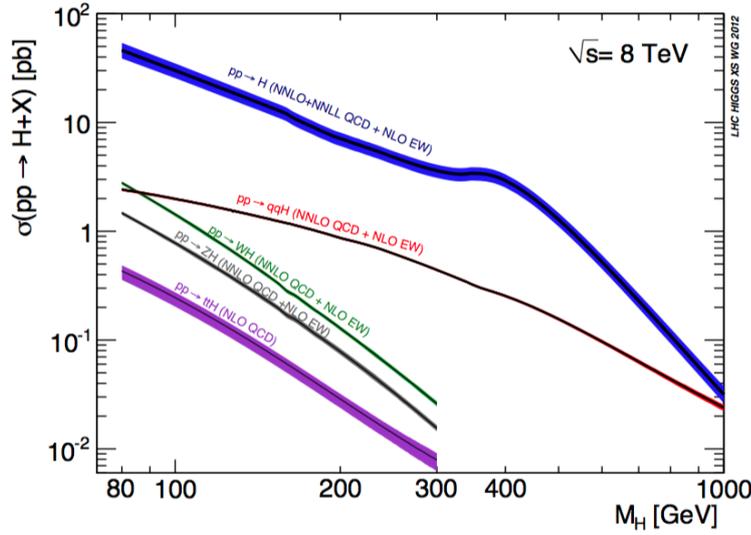


Figure 1.2.2: The production channels of the Higgs boson at the LHC, 8 TeV[16]

1.2.2 Decay channels of the Higgs boson

The possible direct decay products of the Higgs boson are all massive particles, because the Higgs couples to fermions or bosons proportional to their mass signature, as shown in figure 1.2.3.

For bosons (V) the coupling strength, λ_{HVV} , is:

$$\lambda_{HVV} = \frac{2m_V^2}{\nu} \quad (1.6)$$

For fermions (ψ) the coupling to the Higgs boson, $\lambda_{H\psi\bar{\psi}}$, is given by:

$$\lambda_{H\psi\bar{\psi}} = \frac{m_\psi}{\nu} \quad (1.8)$$

The coupling of the Higgs boson to fermions and bosons depends on the masses of these. Heavy fermions have a large branching fraction as shown in figure 1.2.3. Higgs bosons cannot decay to $t\bar{t}$, since the top quarks are too heavy. The main decay mode of the Higgs boson to fermions is $H \rightarrow b\bar{b}$. This is however a very difficult channel to distinguish from the multi-jet background at the LHC. A Higgs boson decaying to two τ leptons is therefore the most important and accessible channel for testing the Yukawa coupling between fermions and the Higgs boson at the LHC.

1.3 Status of the Higgs boson search

Since July 2012 it has been clear that there is a Higgs-like boson particle with a mass around 125 GeV. The main discovery channels was in 2012 $H \rightarrow ZZ$, $H \rightarrow \gamma\gamma$ and $H \rightarrow WW$. These di-boson final state channels have later been updated with all the data from 2012, and

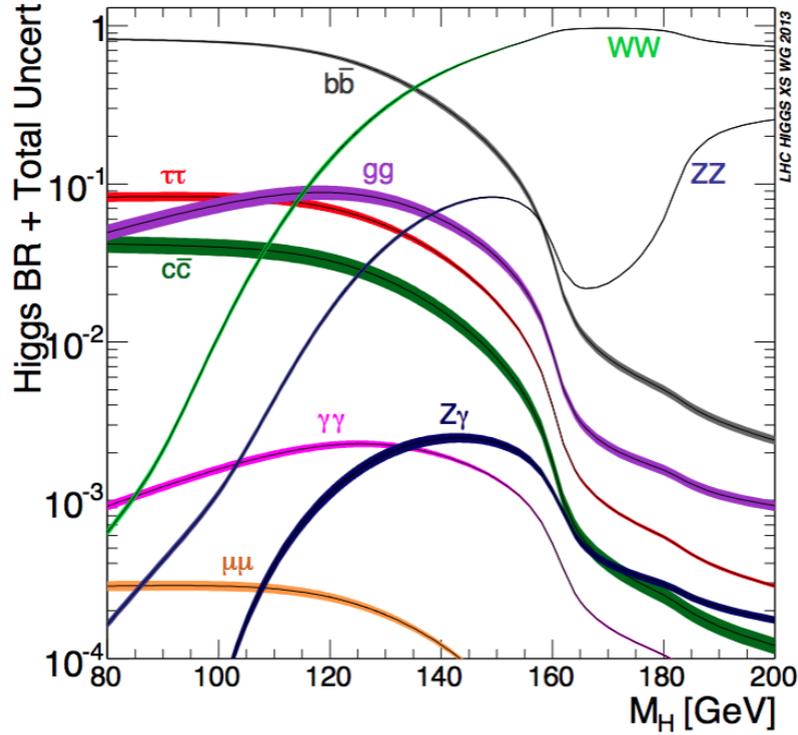


Figure 1.2.3: The SM Higgs boson decay branching ratio as a function of the Higgs boson mass. [16]

have confirmed the discovery of 2012 [19] [20]. Also the spin of the Higgs boson using these decay channels have been investigated and there is now evidence for the spin-0 nature of the Higgs boson, with positive parity being strongly preferred [5].

$H \rightarrow ZZ \rightarrow 4l$:

This search uses the final state where the Higgs boson decays to two Z, which decay to 4 leptons in pairs of two same-flavour, opposite-charge, isolated leptons (leptons are here muons or electrons). The production modes used for investigating this channel are VBF, gg and WH

The reconstructed four-lepton mass spectrum for all these channels is shown in figure 1.3.1. The data is compared to the mass hypothesis of the Higgs at 124.3 GeV, and the significance of the peak is 6.3σ . The mass resolution in these channels are roughly 2 GeV. The mass measured in this decay mode is $124.3^{+0.6}_{-0.5}(stat)^{+0.5}_{-0.3}(syst)$ [19].

$H \rightarrow \gamma\gamma$:

This channel has a very small branching ratio as shown on figure 1.2.3, but a large signal-to-background ratio makes this channel very interesting for Higgs searches. The channel of the $H \rightarrow \gamma\gamma$ is shown in figure 1.3.2. The production modes used for this channel are VBF, gg and WH (VBF and WH are added in 2012 data only). At a $m_H = 126.8$ GeV the significance of the observed peak is 7.4σ . The observed mass is $m_H = 126.5 \pm 0.2(stat) \pm 0.7(syst)$ GeV [21].

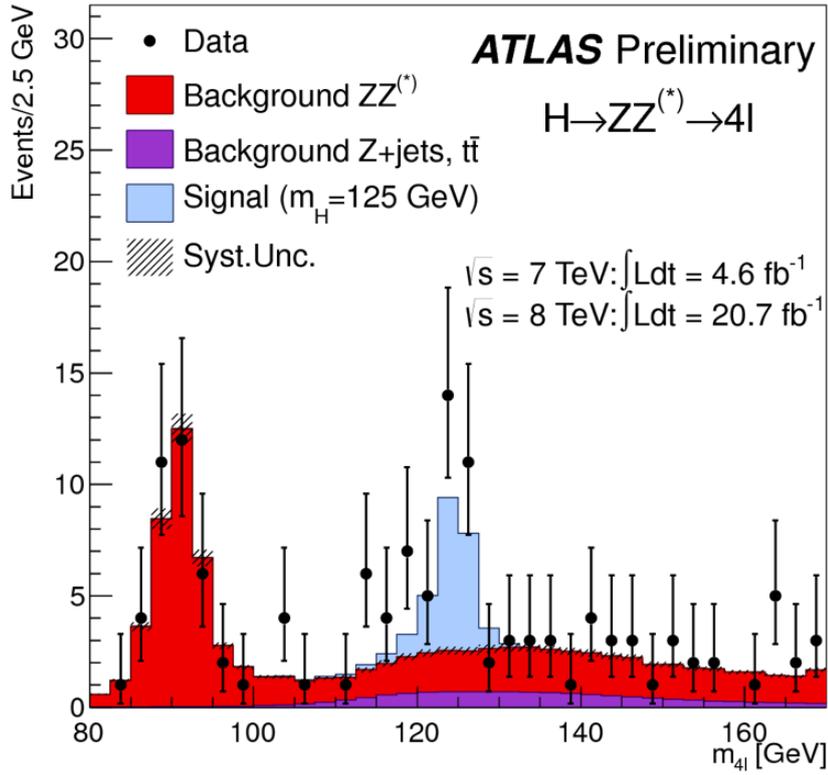


Figure 1.3.1: The reconstructed four-lepton mass spectrum [19]

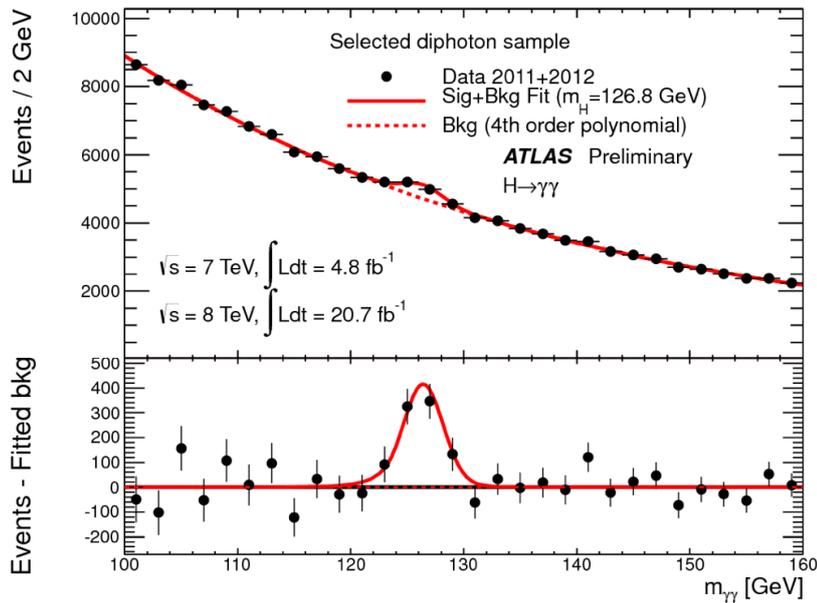


Figure 1.3.2: The reconstructed mass in the di-photon final state [21]

Two of the main channels are those above, but many other channels have been investigated at ATLAS and CMS since the discovery of a Higgs-like boson and the latest combined status is shown in figure 1.3.3 and figure 1.3.4. These figure show the signal strength $\mu = \frac{\sigma}{\sigma_{SM}}$ for

each investigated decay channel. Combining the results, a signal strength of $\frac{\sigma}{\sigma_{SM}} \approx 1$ with $\approx 10\%$ uncertainty is achieved [21] [20].

Since November 2013 there has also been evidence for the fermionic decay of the Higgs-like boson in both ATLAS and CMS [22] [23]. The search was done for the channel in the $H \rightarrow \tau\tau$ and $H \rightarrow b\bar{b}$. In ATLAS the $H \rightarrow \tau\tau$ search evidence comes so far, at more than 3σ level from the $\tau\tau$ final state, focusing on VBF and gg production modes. The $H \rightarrow b\bar{b}$ is done with association production mode and is still inclusive. It is clear that the fermionic decays need most improvement. Furthermore it is clear that the associated production mode is not tested sufficiently.

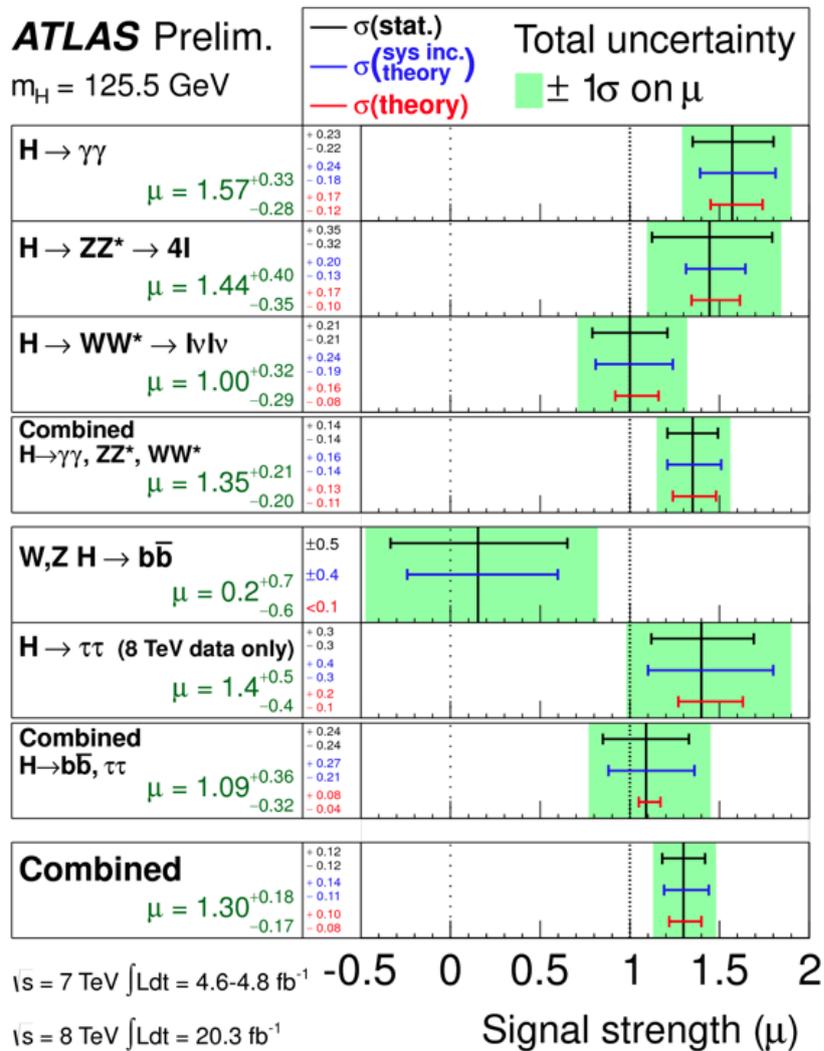


Figure 1.3.3: Higgs result from ATLAS. The signal strength is shown for different decay channels [21]

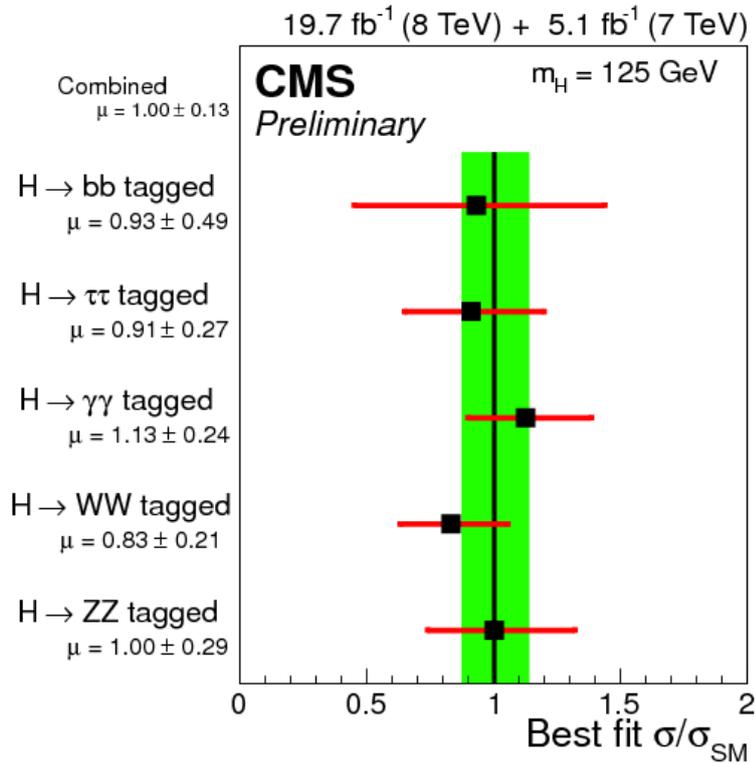


Figure 1.3.4: Higgs result from CMS. The signal strength is shown for different decay channels [20]

1.4 Status of the associated production channel

Before LHC the main search of the associated production was done at the Tevatron [19]. This was a proton-anti proton circular collider, which collided beams at ≈ 1 TeV in center of mass. The expected cross-section times branching ratio for this energy are shown in figure 1.4.1. The figure shows that the associated production is very important for the Higgs searches at the Tevatron. Therefore this channel was used for Higgs boson searches at the Tevatron.

For the searches where $m_H < 130$ GeV, the $H \rightarrow b\bar{b}$ was one of the main modes, and due to the high multi-jet background the associated production was used here. The limit found on the signal strength parameter for this channel at the Tevatron is given in table 1.4.1.

At the LHC the associated production is less important as shown on figure 1.2.2. Therefore other channels have contributed more to the Higgs boson search until now, and less effort have been used to explore the associated production channel. Anyway a search for the associated production was made public already at the CMS experiment for all data collected at $\sqrt{s} = 7$ TeV [25]. The Higgs boson produced from the associated production in the modes $H \rightarrow b\bar{b}$, $H \rightarrow WW$, $H \rightarrow \tau\tau$ and $H \rightarrow \gamma\gamma$ is investigated, and the CMS experiment sets an upper limit of 3.3 times the SM expectation on the cross-section for associated Higgs boson at a Higgs boson mass of 125 GeV.

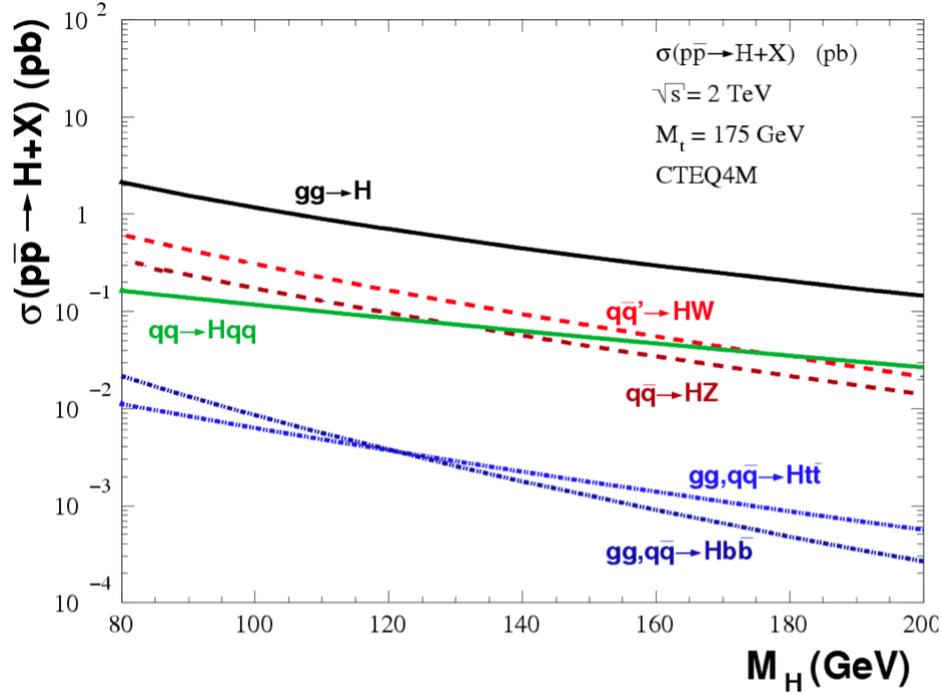


Figure 1.4.1: The production channels of the Higgs boson at the Tevatron [19]

Channel	CDF	D0
	obs(exp.) limit at $m_H=115/125$ GeV	obs(exp.) limit at $m_H=115/125$ GeV
WH ($H \rightarrow b\bar{b}$, $W \rightarrow l\nu$)	3.1 (2.0) /4.9 (2.8)	3.7 (3.2)/ 5.2 (4.7)
ZH ($H \rightarrow b\bar{b}$, $Z \rightarrow \nu\nu$)	2.7 (2.7) /6.7 (3.6)	3.0 (2.7)/ 4.3 (3.9)
ZH ($H \rightarrow b\bar{b}$, $Z \rightarrow ll$)	4.7 (2.7) /7.1 3.9)	4.3 (3.7)/ 7.1 (5.7)
Combined		1.9 (1.6)/3.2 (2.3)

 Table 1.4.1: Expected and observed limits on the signal-strength parameter $\mu = \sigma/\sigma_{SM}$ in SM in the search for $H \rightarrow b\bar{b}$ in D0 and CDF. The entry – indicates that this information is not available [24]

1.5 Outlook for the associated production channel

The associated production mode has a cross-section which is almost two orders of magnitude lower than gluon-gluon fusion, but at the LHC it will become increasingly more useful since it is possible to trigger on leptonic decay products from the W or Z bosons with higher efficiencies, despite the increasing number of collisions per bunch crossing in the new start up of the LHC (from approximately 20 next year to above 100 collisions per bunch crossing towards the end of LHC). Thereby this channel will help a lot in the search of especially $\tau\tau$ and $b\bar{b}$ final states.

Since 2012 the new particle is now found, and a goal for the next run is to do precision measurements of the Higgs-like particle, thereby determining if it is the SM Higgs boson or a more exotic version of the particle. The precision measurements include determination of the signal strength for each production channel and decay mode (also some rare decay modes like $\mu\mu$). The coupling constant between fermion and bosons need also to be tested against the SM expectations. The spin needs to be investigated further in the fermionic decays too. It should also be a goal to test the Higgs self-coupling.

The main goal of this PhD work is to contribute significantly to this effort.



2.1 LHC

The Large Hadron Collider (LHC) is a proton-proton collider located near Geneva in Switzerland at CERN, European Organization for Nuclear Research. The accelerator measures 27 km in circumference making it the biggest particle accelerator in the world. The LHC is housed in a circular tunnel 100 meters underground, originally built for its predecessor the Large Electron-Positron collider (LEP). A goal of the LHC is to reach a higher luminosity and energy than done before in hadron collisions. Luminosity is defined as the number of protons circulating per unit of area and time. A high luminosity is sought, since it enables investigation of physical processes with a small cross-section [26].

Acceleration of the protons starts in the linear accelerator, Linac2, where bunches of 10^{11} protons are accelerated to 50 MeV. From Linac2 the protons are transferred to the PS Booster (PBS) and afterwards the Proton Synchrotron (PS). These two synchrotrons increase the energy to 1.4 GeV and 26 GeV respectively. The final step is to direct the protons to the Super Proton Synchrotron (SPS) which increases the energy to 450 GeV before the protons are diverted into the LHC in both the clockwise and counter-clockwise directions. The LHC then accelerates each proton beam to an energy of up to 7 TeV, thereby achieving a center-of-mass energy of 14 TeV. Figure 2.1.1 shows a sketch of the accelerator complex along with the LHC.

There are several locations, where the beams can collide, and at each location a detector is placed. Six experiments are conducted at these four detectors:

A Toridal LHC ApparatuS (ATLAS): The goal of this experiment is to search for the Higgs particle and investigate other SM physics as well as physics beyond the SM at the energy frontier [27].

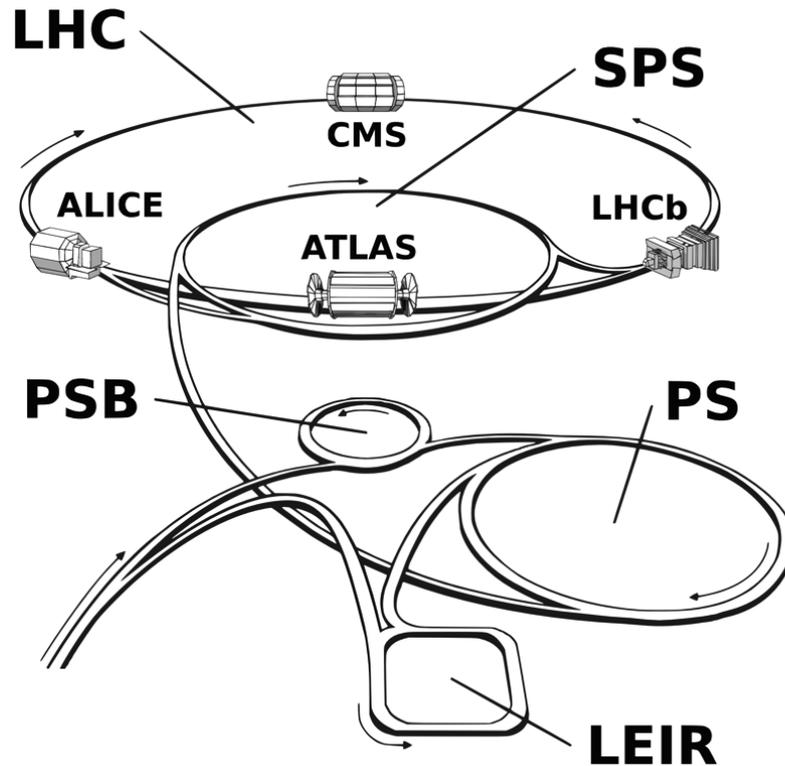


Figure 2.1.1: The LHC and the accelerator complex [26].

A Large Ion Collider Experiment (ALICE) : This detector investigates the quark-gluon plasma supposedly created in the collisions of heavy ions at the LHC. Quark-gluon plasma is a state of matter where quarks and gluons are no longer confined inside the hadrons. Quark-gluon plasma occur at high temperatures and high densities. By studying this state of matter reactions from the start of our universe can be probed [28].

Compact Muon Solenoid (CMS): The goal of this detector is the same as for ATLAS. Both experiments investigate the same physics, and it is thereby possible to verify the results obtained from each of the detectors. The two detectors differ significantly in their detector choices [29].

Large Hadron Collider beauty experiment (LHCb): The goal of this experiment is to conduct a precise measurement of the CP violation and furthermore investigate rare b-decays [30].

Large Hadron Collider forward (LHCf) This experiment is located near ATLAS, and the goal is to study particles generated in the very forward region of collisions [31].

TOTAL Elastic and diffractive cross section Measurement (TOTEM) The purpose of this experiment is to determine the luminosity. It is located near CMS [32].

2.2 Hadron collider physics

The LHC collides protons with protons. The outcome of such collisions is investigated by the detectors, to understand the universe at the smallest scales. Collisions take place at energies so high that the colliding protons are seen as a composite particle instead of point particles. The collision is a complicated process, and the concepts implied in the description of the collisions are:

Hard scattering: The first step in considering an event is the scattering of two partons, where the partons are mostly gluons, u-quarks or d-quarks. Particles from gluon splitting and pair creation may also give rise to partons participating in the hard scattering. The probability of a parton carrying the fraction x of the total proton energy at a momentum transfer Q is given by the Parton Distribution Function, denoted by PDF . Two examples of PDF 's for protons are given in figure 2.2.2. It is clear that the most likely parton to interact are the gluons, but other quarks can also be important at higher x values.

The cross-section for a created resonance like eg. a W-boson can be given by:

$$\sigma_{pp \rightarrow W} = \sum_{ij} \int dx_1 f_i(x_1, Q^2) \int dx_2 f_j(x_2, Q^2) d\sigma(ij \rightarrow W) \quad (2.1)$$

$d\sigma(ij \rightarrow W)$ is the partonic cross-section for creating W over the partons i and j and the SM theory provide the fundamental framework to calculate such cross sections. $f_i(x_1, Q^2)$ and $f_j(x_2, Q^2)$ are the PDF for partons i and j , the sum is over all partons present in the proton [8].

Initial and final state radiation The incoming and outgoing partons radiate photons and gluons before and after the hard scattering. This is the so-called initial state radiation (ISR) and final state radiation (FSR). Ideally these higher order corrections will be incorporated by the theoretical calculation done for the leading scattering, but this is often complicated or even impossible due to divergences [8].

Parton shower: Parton showers are cascades of radiation produced from QCD processes and interactions. The higher order corrections to the tree level process are in simulations sometimes taken into account by parton showering.

Hadronization/fragmentation Isolated colored particles created from partons cannot exist for long and will merge together forming hadrons. The way this is done differ between different simulation implementations as described later [8].

The Underlying event : The rest of the partons from the protons fly around in the detector giving rise to low energy particles, which are referred to as underlying events [8].

Figure 2.2.2 shows a cartoon of the different parts of a collision.

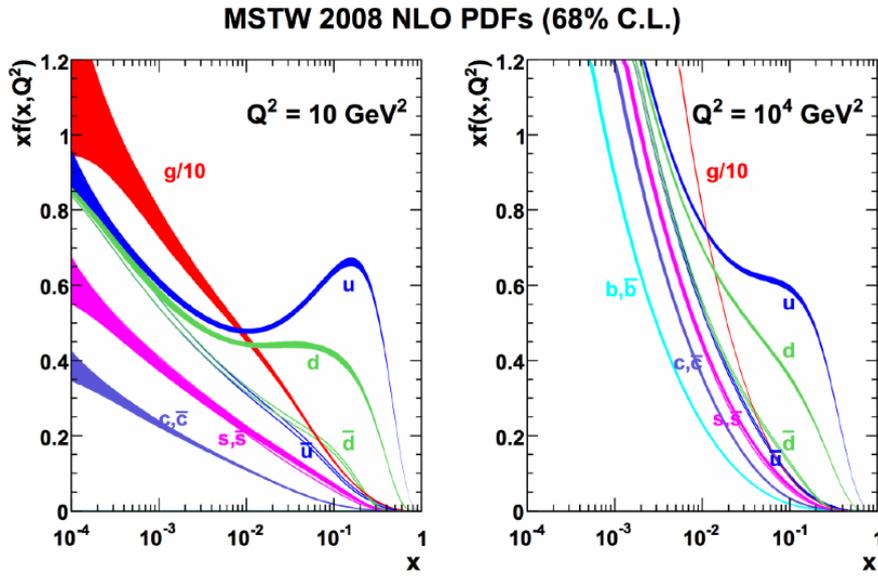


Figure 2.2.1: A PDF at NLO at $Q^2 = 10 \text{ GeV}^2$ and at $Q^2 = 15 \text{ GeV}^2$ [33]. As described in the text x is the fraction of the total proton energy at a momentum transfer Q . $f(x, Q^2)$ is the PDF.

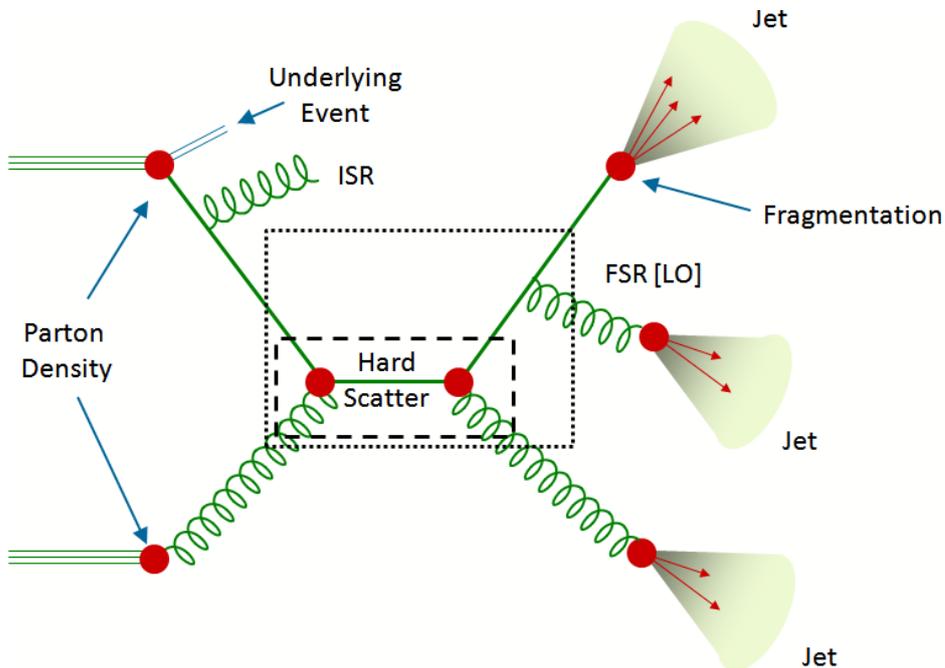


Figure 2.2.2: Cartoon of the different parts of a collision [34]

2.2.1 Simulations

Simulations are an important tool to interpret data from the LHC. Simulations rely on generators which differ in the implementation of the sub-processes explained above:

- **Pythia:** Pythia is a general purpose generator able to simulate parton-parton hard interactions, underlying events, parton shower, hadronization, and particle hadron decay. In Pythia the probability for the incoming parton to arise from another parton is investigated, and this information is used to create ISR and FSR. In Pythia the hadronization is done via the Lund String Model. In this model when two partons move apart there is a strong gluon field between them. The potential between the partons grows linearly with the distance between them. When the energy is sufficient to create a quark/anti-quark pair, this will happen forming color less mesons [35].
- **Sherpa:** Like Pythia, Sherpa is capable of all the steps in the process from hard scattering to hadronization. Sherpa is composed of different phases, where each phase takes care of a small part of the process. By combining each phase the whole event can be considered.

In Sherpa the hadronization is done by the Cluster Model instead of the Lund String Model: The idea is that after the parton shower the gluons are split into quark-anti quark pairs, which can be combined into color singlet clusters. This is done by combining the quarks with the one closest to them to form color singlet clusters. The light clusters are directly used as hadrons, whereas heavier clusters decay into lighter clusters [36].

- **Alpgen:** Alpgen is used to calculate the hard process, and is used for process with a large number of final state partons and determines the cross section at leading-order (LO), which is the lowest order of the theoretical calculation. Alpgen is widely used for modeling of W and Z boson and it calculates process with up to six partons in the final state. Alpgen is combined with other generators like Herwig for the hadronization/parton shower part[37].
- **Herwig:** Herwig can, like Sherpa and Pythia, simulate parts of the collision events. For the hadronization it uses the Cluster Model like Sherpa. A main goal for Herwig is to calculate the ISR and FSR precisely [38].
- **McAtNlo:** McAtNlo calculates the hard process with the full next-to-leading-order (NLO) including therefore $O(\alpha_s)$ corrections to the three level. It must be combined with something like Herwig for the parton showering [39].

Other more specialized generators are used for specific parts of the events. MCFM is a generator which calculate processes up to NLO [40]. The generator PHOTOS generates QED radiative corrections [41]. To simulate the decays of polarized τ leptons including spin correlations is used TAUOLA [42].

The final step after generating the events is to simulate the ATLAS detector response. This is done by the GEANT4 framework, which take the particles obtained from the MC simulation and simulate final state corresponding to what is expected from the real collisions, based on detailed information about the detector subparts [43]. With this final step the imprint of the detector is accounted for allowing comparison of simulation with data.

2.3 ATLAS

ATLAS is a general purpose detector, and physically it is the biggest detector at the LHC. The ATLAS detector is 45 m long, more than 25 m high, and weighs more than 7000 tons [44]. The purpose of the ATLAS detector is to identify each particle created in the collisions, and through this reconstruct the physics involved with the primary interaction [27]. Figure 2.3.1 shows the design of the ATLAS detector.

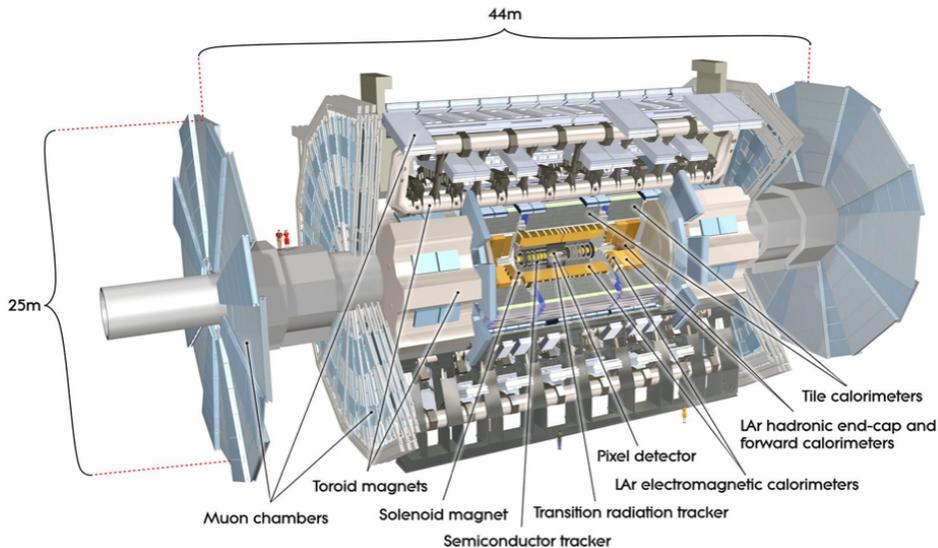


Figure 2.3.1: An overview of the ATLAS detector [44].

To reconstruct the primary interaction it is both necessary to count the number of created particles and determine the energy and momentum of these particles. To achieve this the detector consists of four components placed in concentric shells around the beam pipe. Each component of the detector gives a signal depending on which kind of particle crosses the detector as schematically shown in figure 2.3.2. From this it is also possible to identify different types of particles [44].

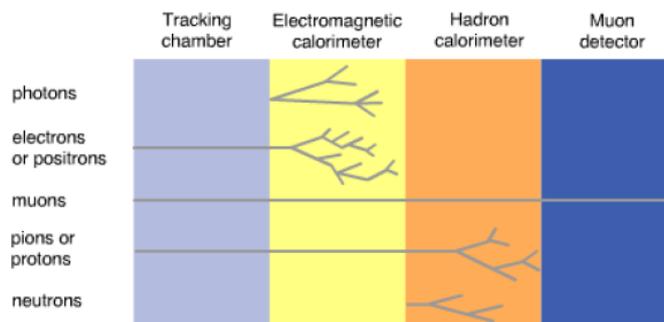


Figure 2.3.2: A schematic cartoon of the components of the ATLAS detector [44].

The coordinate system

The coordinate system used for the ATLAS detector is a right handed coordinate system with the x – axis pointing inwards to the LHC ring center. The z – axis is along the beam line and the y – axis is perpendicular to the plane of the LHC, which is at an angle to vertical because the tunnel is tilted to keep all shafts at a maximum of 150 m while keeping the tunnel in solid rock. The coordinate axis along with the tilt is shown on figure 2.3.3. Furthermore the pseudo-rapidity, η , is introduced, which is a function of the angle of the particles relative to the z – axis:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (2.2)$$

where θ is the angle of the particle direction measured from the positive z – axis. Furthermore the angle ϕ , indicates the angle in the x – y plane wrt positive x-axis.

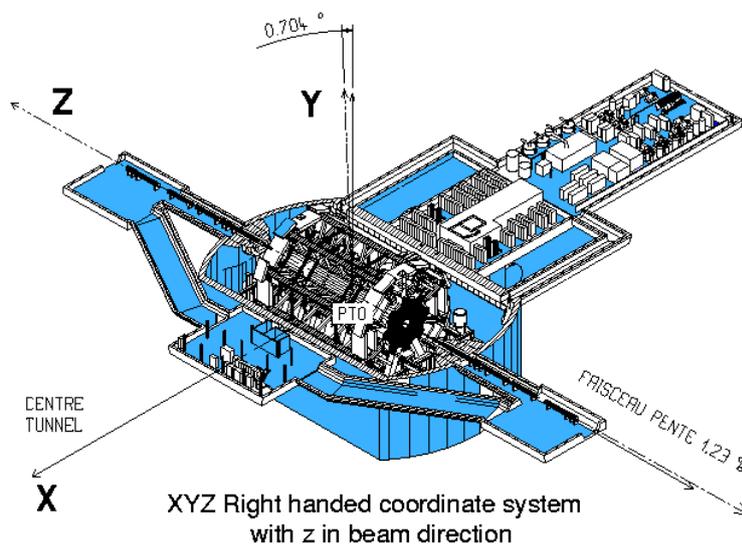


Figure 2.3.3: The coordinate system for the ATLAS detector. Modified from [44]

2.3.1 The Inner Detector

The Inner Detector, ID , tracks charged particles. The ID is placed in a 2T magnetic field to facilitate determination of charge and momentum of the charged particles.

Pixel detector

The Pixel Detector is the innermost part of the ID , and is placed close to the interaction point. The goal of this sub-detector is to make high-precision measurements of the particle properties such as lifetime and impact parameter. (The impact parameter is the distance of closest approach between the track and the primary vertex). The pixel detector is comprised of approximately 80 million pixels covering an area of 1.7 m^2 . The minimal pixel size is

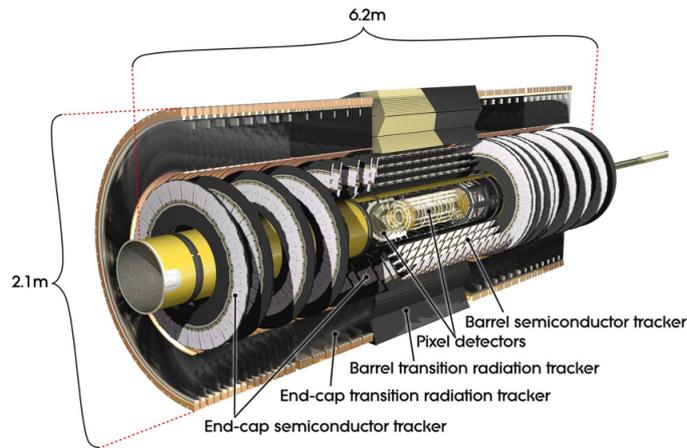


Figure 2.3.4: An overview of the inner detector, consisting of three parts: the pixel detector, the Semiconductor tracker and the Transition Radiation Tracker [44]

$R - \phi \times Z = 50 \times 40 \mu\text{m}^2$ with a resolution of $15 \times 115 \mu\text{m}^2$. R is the radius orthogonal to the beam axis. The pixel detector is built of three overlapping cylindrical barrels at radial positions of 50.5 mm, 88.5 mm and 122.5 mm. The first layer is called the b-layer, since it is very useful for identifying b-quarks by measuring B hadrons lifetime. The end-cap consists of three disks placed perpendicular to the beam axis located in radii between 9 and 15 cm.

Semiconductor tracker

After the pixel detector comes the Semiconductor tracker (*SCT*). This detector is used for tracking, and momentum measurement, impact parameter and primary vertex position determination. The *SCT* detector is designed to give eight precision measurements per track. The *SCT* detector is placed at an intermediate radial distance in the *ID* at $30 \text{ cm} < r < 52 \text{ cm}$. The barrel part of the *SCT* detector consists of 4 cylindrical layers of silicon micro-strip detectors which have a coverage of η up to $|\eta| < 1.4$. The end-caps consist of silicon micro-strips aligned radially, giving the *SCT* coverage up to $|\eta| < 2.5$. The *SCT* have a total of about 6.3 million read-out channels. There is a read-out every $80 \mu\text{m}$ and the positions of charged particles can be recorded to a precision of $17 \mu\text{m}$ in the $r-\phi$ coordinates.

Transition Radiation Tracker

The Transition Radiation Tracker (*TRT*) detector is the outermost component of the *ID*. The *TRT* is used for tracking and momentum resolution, but also for electron/ π separation via Transition Radiation. Transition radiation is produced by highly-relativistic charged particles crossing boundaries with differing dielectric constants. Transition radiation will by design turn-on at Lorentz boost factors between 10^3 and 10^4 . Electrons have a much lower mass than hadrons and consequently a higher Lorentz boost at a given momentum. This makes it possible to separate between electrons and π .

The *TRT* detector consists of a barrel and two end-caps, where both contain a large number

of straw detectors. In the barrel part of the detector the straws are aligned with the z -axis, and in the end-caps the straws are oriented radially. The *TRT* barrel is based on 52544 straw detectors. Each straw has a diameter of 4 mm with a 30 μm tungsten wire at the center. The radial and azimuthal spacing between the straws are 6.8 mm. The gas used in the straw tubes is comprised of 70% Xe, 27 %CO₂ and 3 %O₂. Xe is used for absorption of the transition radiation, CO₂ and O₂ increase the electron drift velocity and facilitate photon-quenching. Photon-quenching is the absorption of photons produced in the avalanche to stabilize gain and prevent damage of the straws. The 73 planes in the TRT barrel are split into three radial sections called layers. Each of these layers are azimuthally divided into 32 modules. The modules in the three layers differ a bit in number of straws, and geometry. The *TRT* measures about 36 hits per track and has a resolution of 125 μm in the r - ϕ plane barrel and 130 μm in endcaps [44].

2.3.2 Calorimeters

The purpose of the calorimeters is to measure the energy of both charged and neutral hadrons.

Electromagnetic calorimeter

The electromagnetic calorimeter (*EM*) is shown in figure 2.3.5 and figure 2.3.6. The *EM* consist of several lead plates which electrons pass through, triggering an electromagnetic shower. The electromagnetic shower is a result of bremsstrahlung from a e^\pm or a photon which turns into a e^+e^- pair. Between the lead plates is liquid argon, which gets ionized by the electromagnetic shower. Since an electric field is applied to the calorimeter the ionized electrons drift towards electrodes which record the energy deposited [45].

To estimate the total energy it is important that the calorimeter provides full coverage out to $|\eta| < 3.2$. This is done by a barrel part covering out to $|\eta| < 1.5$ and an end-cap for the remainder. Unfortunately this results in a small gap at $|\eta| \approx 1.5$ with a slightly worse resolution[45].

There are a four layers in the calorimeter [45].

Presampler The presampler is a single layer of liquid argon before the first sampling. This layer is used to determine the energy loss due to interactions in the material before the first sampling. It has coverage out to $|\eta| < 1.8$ and a very low granularity in $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$

First sampling The goal of the first sampling is to distinguish between a single photon and $\pi^0 \rightarrow \gamma\gamma$. For this purpose a very fine granularity of $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ is needed.

Second sampling Most of the electromagnetic energy will be deposited here, and there is a medium granularity $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the barrel and $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the end-caps. For these layers there is also a granularity in ϕ making it possible to define variables dependent on ϕ . From the shower shape it is possible to distinguish between energy deposits of electromagnetic and hadronic origin.

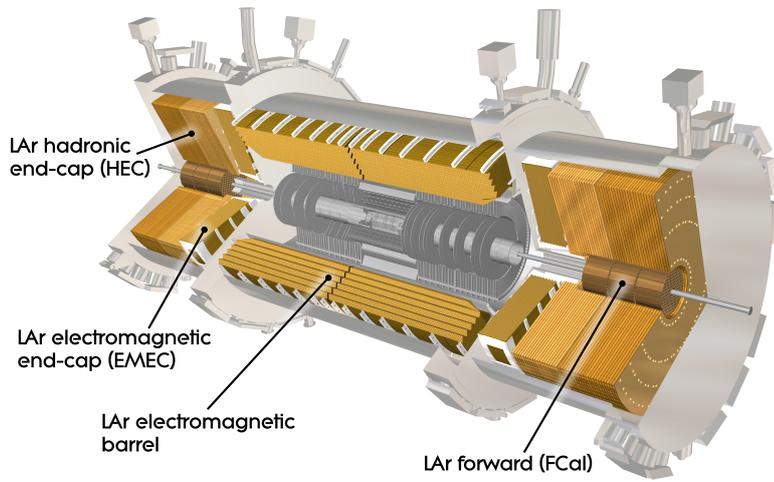


Figure 2.3.5: A schematic view of the ATLAS Liquid Argon calorimeter. [44].

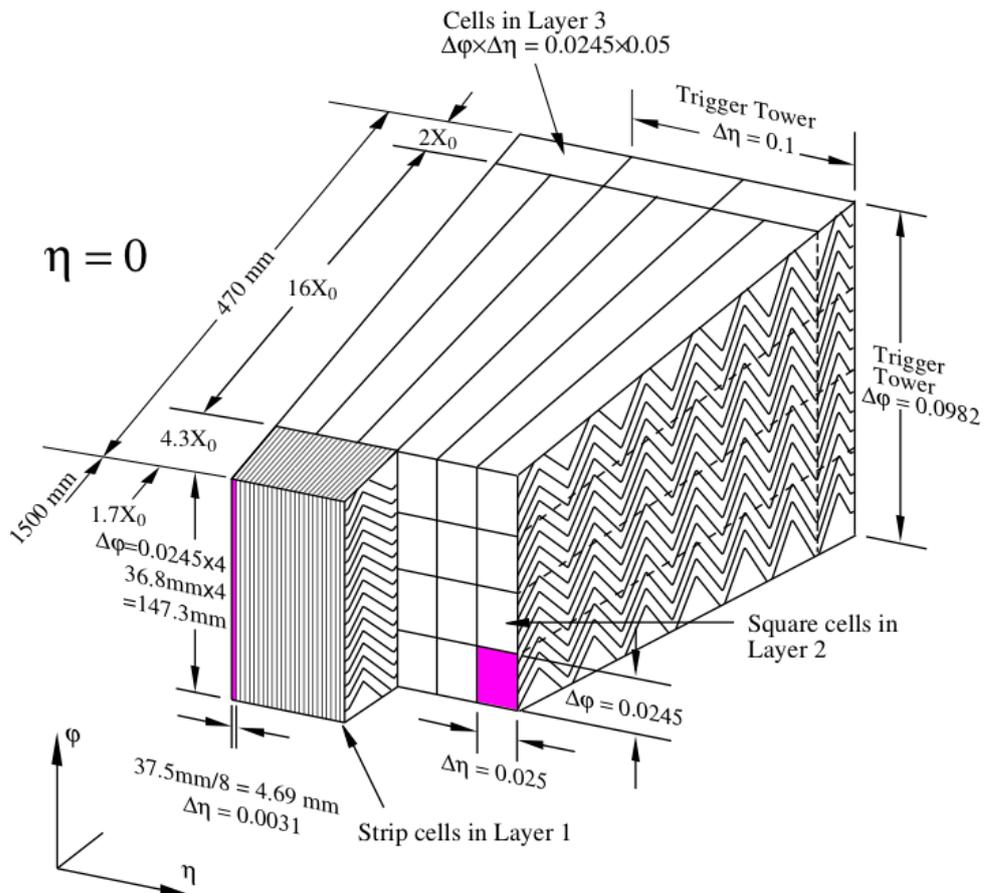


Figure 2.3.6: Illustration of a barrel module of the ATLAS EM calorimeter, illustrating the segmentation of the strip, second and third layers [44].

Third sampling The goal of the third sampling is to measure any remaining electromagnetic energy associated with energy rich particles. The granularity here is $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ and covers $|\eta| < 2.5$

Hadronic calorimeter

The goal of the hadronic calorimeter is to measure the energy of hadrons, and the principle idea is similar to the *EM*. One central difference is that the hadronic calorimeter needs to be denser to absorb all the energy because hadrons are stopped via the strong interactions. Instead of liquid argon, tungsten and copper are used (the latter in the forward direction.) Figure 2.3.7 shows the hadronic calorimeter surrounding the Liquid Argon calorimeter. There is a barrel part of the hadronic calorimeter covering $|\eta| < 1.0$ (Tile calorimeter) and an extended barrel which extends the coverage out to $|\eta| < 3.2$ (HEC). Lastly the coverage is completed by a forward calorimeter covering out to $|\eta| < 4.9$ (FCal). This very large coverage is important to ensure absorption of all the energy of the hadrons.

The granularity of the Had calorimeter is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in $1.5 < |\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ in $2.5 < |\eta| < 3.2$. In the Tile calorimeter the granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for all layers but the last layer where the granularity is $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$. In the FCal, the granularity is given in $x \times y$ (cm) and varies between 3.0×2.6 to 5.4×4.7 .

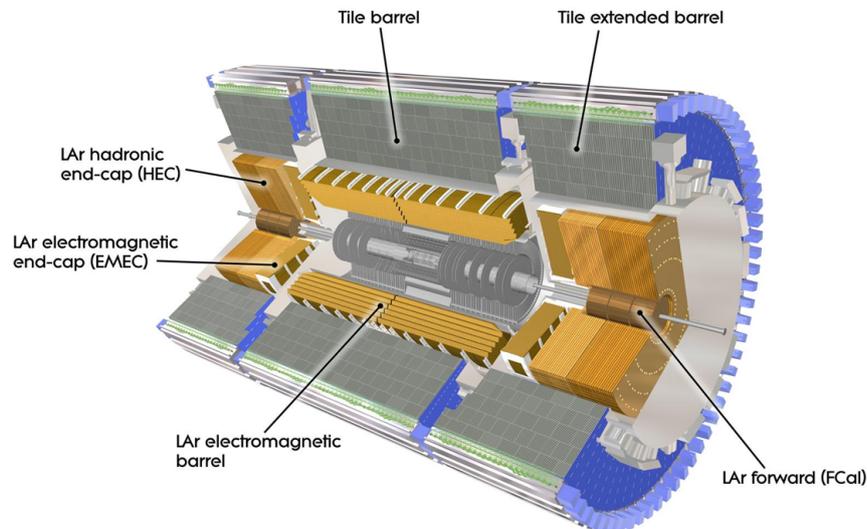


Figure 2.3.7: A schematic view of the ATLAS calorimeter [44].

2.3.3 Muon spectrometer

The outermost sub detector is a muon detector used to identify muons and determine their momentum. The subdetector is shown in figure 2.3.8. Muons and neutrinos are the only particles that reach this outer part of the ATLAS detector. The muon spectrometer detector have two goals: Tracking of muons and triggering on muons. The latter purpose needs to be done fast, whereas for the first mentioned a high precision is desired. For a 10-200 GeV

muon the resolution on the transverse momentum resolution is around 2-4 %, and 10% for a 1 TeV muon. The muon spectrometer gives three points in a coverage out to $|\eta| < 2.7$. The basic working principle of the precision tracking part of the muon spectrometer is the same as for the TRT, namely that it consists of drift tubes with a wire in the middle. The gas in the tube is ionized by the passing of a charged particle. Electrons and charged ions then drift towards the wire in the middle of the straw giving a hit. The tracking part is composed of Monitored Drift Tubes (MDT) out to $|\eta| < 2.7$. There is also a tracking part in the endcap by Cathode Strip Chambers (CSC) via multi-wire proportional wires. These provide coverage out to $1.05 < |\eta| < 2.7$. The change of the tracking part is done due to a very high particle rate and an uneven magnetic field.

Finally two trigger systems are associated with the muon spectrometer, which are the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC). These gas detectors provides fast response to signals with hits detected in the muon spectrometer.

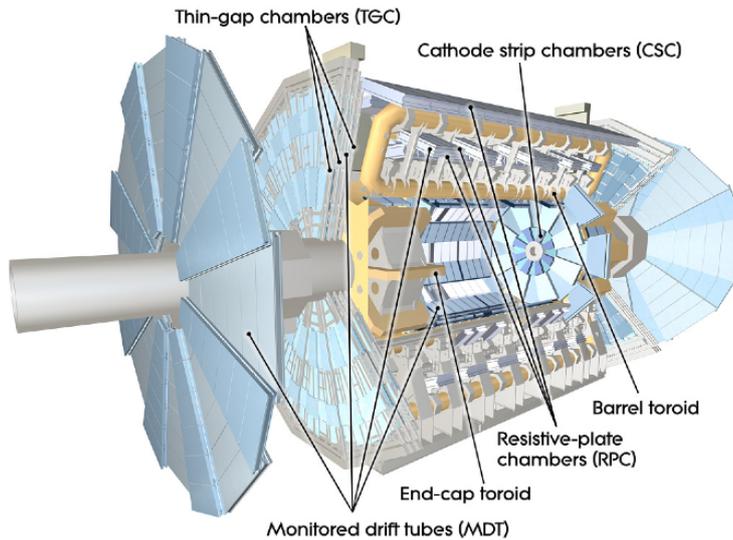


Figure 2.3.8: The muon spectrometer [44].

Resolution

The required resolution of the ATLAS detector is given in table 2.3.1 [46].

Subsystem	Required resolution
Inner detector	$\frac{\sigma_{P_T}}{P_T} = 0.05\%P_T \oplus 1\%$
Electromagnetic Calorimeter	$\frac{\sigma_E}{E} = 10\%/\sqrt{E} \oplus 0.7\%$
Hadronic Calorimeter (barrel, end cap)	$\frac{\sigma_E}{E} = 50\%/\sqrt{E} \oplus 3\%$
Hadronic Calorimeter (forward)	$\frac{\sigma_E}{E} = 100\%/\sqrt{E} \oplus 10\%$
Muon spectrometer	$\frac{\sigma_{P_T}}{P_T} = 10\%atP_T = 1TeV$

Table 2.3.1: The required resolution of the ATLAS detector. The unit is in GeV [46].

2.3.4 Magnets

A complex system of magnets is installed in ATLAS to bend the particles from the collisions for momentum and charge measurement. The system of magnets consists of a Central Solenoid superconducting magnet (CS) and three toroidal superconducting magnets. The CS is used to bend the tracks of particles in the Inner Detector, whereas the latter are used for the muon spectrometer. The CS is 3.4 m in diameter and 5.6 cm thick and produce a field of 2T field.

The toroidal magnets are split into a Barrel Toroid (BT) and two End Cap Toroids (ECT). The BT is 25.3 m in length and 20.1 m in outer diameter, producing a magnetic field is of 0.5-1 T. The structure is shown in figure 2.3.9.

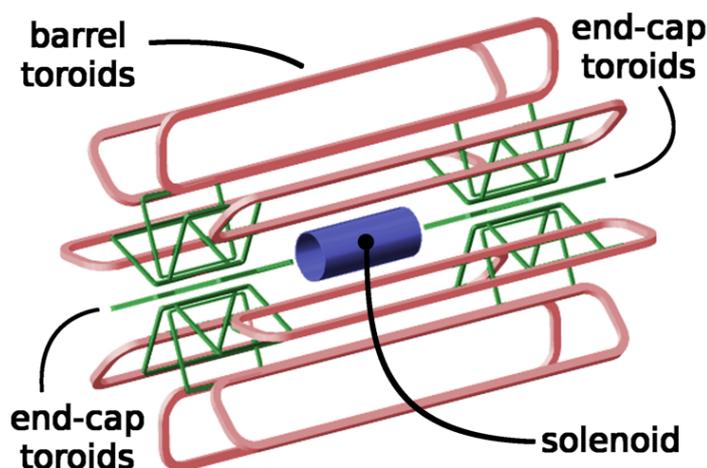


Figure 2.3.9: Magnet system

2.3.5 Trigger

The ATLAS trigger system is one of the biggest challenges, and successes, of the ATLAS experiment. The goal is to select rare events and discard a large number of background events. The trigger in ATLAS is composed of three steps, a hardware based level 1 trigger(L1) and two software based higher lever triggers: Level 2 (L2) and Event Filter(EF)

L1: The signals from the muon and calorimeters are used at this level. From interpreting these and thereby finding high transverse momentum (P_T) objects, a Region Of Interest, (ROI) is made. The time available for L1 to make the decision is about $2.2 \mu\text{s}$ per event. At design bunch crossing rate there is a 40MHz input rate and L1 is capable of reducing this to an output rate of 75 kHz. This huge reduction in so few μs is a very critical point, since wrong decisions can causes several the losses of signal.

L2: At this stage the goal is to find objects like e , γ , μ , τ , jets heavy quarks and transverse missing energy. This is done by analyzing the data in the ROI (even though all data is available at this step it is primarily the ROI information which is used to keep the

reduction time low. For the transverse missing energy the whole event is analysed). The reduction of data is at this step from an input rate of 75 kHz to an output rate of 1 kHz in 40 ms.

EF This level is very close to the reconstruction tools, and a fast analysis is done here determining if the event is of interest. At this step the reduction goes down to a rate of few hundred hertz [44].

2.4 Luminosity

The luminosity delivered to the ATLAS experiment can be measured by

$$\mathcal{L} = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}} \quad (2.3)$$

μ_{vis} is the observed interaction rate per crossing, n_b is the number of protons per bunch, f_r is the proton revolution frequency (11245.5 Hz), $\sigma_{vis} = \epsilon \sigma_{inel}$ where σ_{inel} is the total inelastic cross-section multiplied by the acceptance efficiency ϵ [47].

Luminosity determination is done by two detectors. The first Luminosity measurement uses Cerenkov Integrating Detector, LUCID, which is a segmented Cherenkov detector surrounding the forward beam-pipe 17 m from the interaction point. The second measurement uses the Beam Condition Monitor, BCM, which consists of diamond sensors surrounding the beam pipe at 184 cm from the interaction point. Both of these luminosity detectors determine the luminosity for each bunch crossing[47]. σ_{vis} is calibrated using scans of the beam separation, this is called the Van de Meer scans and is described in [48]

Figure 2.4.1 shows the total integrated luminosities delivered by the LHC in 2012, along with what is recorded by ATLAS. The difference is due to inefficiencies in the data acquisition and detector dead-time. The amount of data which passes the data quality and is good for analysis is also shown. The total integrated luminosity of data which can be used for analysis in 2012 is 20.3 fb^{-1} .

2.4.1 Pile-up

Pile-up is the result of proton-proton collisions happening simultaneously. Pile-up is quantified by the average number of collisions per bunch crossing. Pile-up can be divided in two categories: in time pile-up and out of time pile-up. In time pile-up is caused by concurrent proton-proton collisions within the same bunch crossing as the collision of interest. Out of time pile-up is additional collisions from other bunch crossings than the hard scattering which triggers the event. The pile-up profile seen by ATLAS in 2012 is shown in figure 2.4.2.

The profile is used to weight simulations for comparison with data

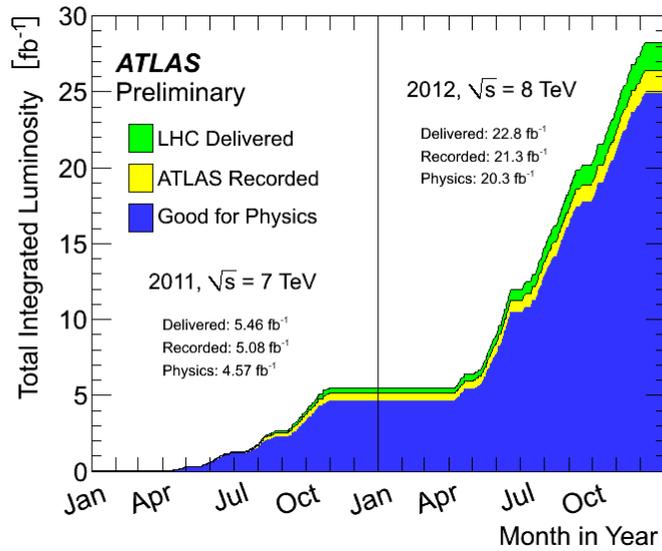


Figure 2.4.1: Cumulative integrated luminosity delivered to ATLAS (green), recorded by ATLAS (yellow), and after requirements of good data quality (blue) in 2012 at $\sqrt{8}$ TeV [26].

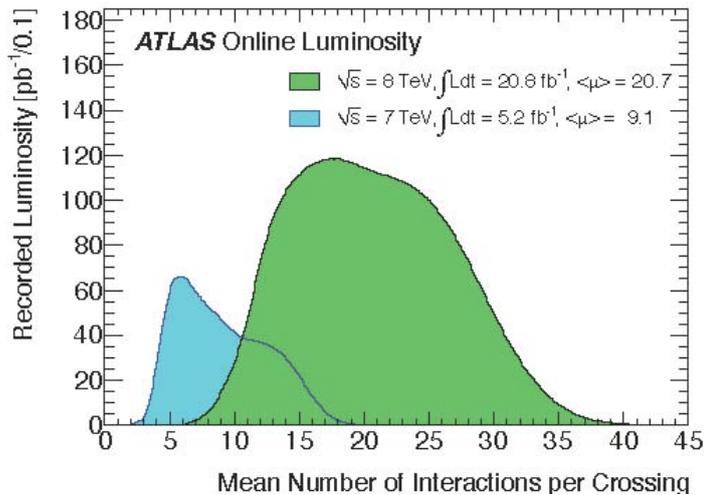


Figure 2.4.2: The average number of collisions per bunch crossing in 2011 and 2012[26]

2.5 Reconstruction of muons

Muons (μ) leave hits in both the inner detector and the muon spectrometer. The reconstruction part of μ 's done in the muon spectrometer, where the signal of a μ is very clean with little background. The inner detector provides a high number of hits per track allowing measurements of the track impact parameters with respect to the primary vertex and improve the momentum resolution. Different kinds of reconstructed μ 's are provided by the muon performance group in ATLAS, the one used in this thesis is the "combined μ ". Combined μ are reconstructed using track segments from both the inner detector and the muon spectrometer, matched by a χ^2 fit, thereby ensuring a good transverse momentum (P_T) resolution.

These combined measurements give a precise reconstruction for a wide range of η and P_T . The efficiency for μ 's to be identified as combined muons in ATLAS is above 97% [49]. The efficiency reconstruction of μ 's differs in data and simulation, and this needs to be corrected. A scale factor for this difference is needed for the performance in simulations to mirror that found in data. The efficiency scale factor is found in a tag and probe analysis on $Z \rightarrow \mu\mu$ events. The "tag and probe" method is a data driven technique used to calculate efficiencies. It selects one μ with some very tight selection, and the other μ has a looser selection, and is then used to calculate the efficiency scale factor. The efficiency scale factor is given by $SF = \frac{\epsilon^{Data}}{\epsilon^{MC}}$, where ϵ^{Data} and ϵ^{MC} are the efficiencies for the muon to be reconstructed in Data and MC respectively [49]. The efficiency scale factors are provided by the ATLAS Muon working group and are close to one.

2.6 Reconstruction of electrons

The reconstruction of electrons (e) uses that the fact that e 's deposit energy in the EM calorimeter and produce a track in the inner detector. Reconstruction is split between the central region $|\eta| < 2.47$ and the forward direction. In the central part the first step in of electron reconstruction is to find the cluster in the EM calorimeter, using the so-called sliding window approach. First it finds a cluster with the size of $\eta \times \phi$ of 3×5 in units of 0.025×0.025 . The cluster is considered if it has an energy above 2.5 GeV. This cluster is selected if it is matched to a track in the inner detector within $\Delta\eta < 0.05$ and on a $\Delta\phi$ accounting for bending due to the magnetic field. If more than one track can be matched to a cluster the track with hits in the SCT is chosen, or otherwise the track with the smallest ΔR . The last step in the reconstruction is the sliding window which re-evaluate the energy of the cluster in a slightly bigger window of $\eta \times \phi$ in 3×7 , corrected for the energy deposited in the material before the EM calorimeter in the detector and energy leakage around the cluster and behind the cluster.

For the forward direction it is not possible to use the inner detector, whereby only cluster information is used to reconstruct e . An e in the forward direction needs to have $E_T > 5$ GeV and very limited hadronic energy [50].

After the reconstruction an identification is done to get the good e^\pm : Three identification points are provided in $|\eta| < 2.47$ in ATLAS: loose, medium and tight. Each working point has slightly different variables enabling the rejection of more background for tighter working points. For loose identification shower-shape variables from the two first layers of the EM calorimeter are used. These variables are motivated in the fact that e 's deposit most of their energy in these two layers. Two further variables are introduced utilizing that e 's deposit almost no energy in the hadronic calorimeter thereby discriminating against the hadronic background. The medium and tight working point make stronger requirements of the tracks and uses more information from the TRT [51]. For the e there are also scale factors provided via a tag and probe analysis on $Z \rightarrow ee$, but to increase the statistics available for determination of these factors $W \rightarrow e + jets$ processes are also used. The details of a similar procedure can be found in [51].

2.7 Reconstruction of jets

The input for reconstruction of jets are topo-clusters found in the calorimeters and calibrated at the LC scale [52]. To reconstruct jets from these clusters, the anti-kt algorithm uses two distances defined for each cluster:

$$d_{ij} \equiv \min(k_{T,i}^{-2}, k_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} = \min(d_i, d_j) \frac{\Delta R_{ij}^2}{R^2} \quad (2.4)$$

$$d_i \equiv k_{T,i}^{-2} \quad (2.5)$$

where $k_{T,i}$ and $k_{T,j}$ are the transverse momentum of cluster i and j , ΔR_{ij} is the distance parameter between the two clusters and R is a parameter that controls the size of the jet (typically set to 0.4 in ATLAS). The next step is combining clusters to jets, and is done by considering the magnitude of these distances:

- $d_i < d_{ij}$: The cluster i is defined as a jet, and removed from the clusters list.
- $d_{ij} < d_i$: The two cluster i and j are merge together to a new cluster

The procedure is run until no more clusters are left. This is called the anti-kt algorithm instead of the kt algorithm where $d_i \equiv k_{T,i}^2$ [53].

2.8 Reconstruction of the transverse missing energy

Transverse missing energy (\cancel{E}_T) is a signature of particles leaving the detector without producing a signal. The SM neutrino is such a particle. Detection of missing transverse energy is a delicate matter in the ATLAS detector since great care is needed to differentiate between missing transverse energy caused by particles traversing the detector unseen and any energy loss due to detector components or imbalance due to energy mismeasurement.

\cancel{E}_T is determined using the fact that in the transverse plane the sum of momenta needs to be zero due to momentum conservation.

\cancel{E}_T is found by summing the contributions from each object found [54].

$$\cancel{E}_T = \cancel{E}_T^\mu + \cancel{E}_T^e + \cancel{E}_T^\gamma + \cancel{E}_T^\tau + \cancel{E}_T^{jets} + \cancel{E}_T^{soft} \quad (2.6)$$

Each term gives the \cancel{E}_T corresponding to an object and is calculated as the negative sum of the calibrated reconstructed objects of that type, projected onto the x and y directions.

A term, \cancel{E}_T^{soft} is added which is comprised of topo-cluster not associated with any particular reconstructed physical object. This term, as well as the jet and tau one, need a correction factor to lower any pile-up effect [54].

Part I

Tau identification

3

Tau identification inputs

3.1 Properties of the τ particle

The τ particle is the heaviest lepton, and can decay hadronically and leptonically. The τ mass is 1776.84 ± 0.17 MeV and the τ lifetime is 0.29 ± 0.05 ps [55]. Due to the short lifetime τ particles never travel further than the beam-pipe, and therefore the ATLAS detector only sees τ decay products. Table 3.1.1 lists the decay channels for the τ particle.

- *The leptonic decay* is split almost equally between muon and electron final states. These final states are easy to detect in the ATLAS detector. The problem with these decays is to distinguish between τ decay products and primary muons/electrons from a hard process like the decay of the Z.
- *The hadronic decay* is split into 1-prong and 3-prong. 1-prong (3-prong) corresponds to 1 (3) charged pion(s) π^\pm , and therefore gives 1 (3) track(s) in the detector. There is also a very small fraction of 5-prong decays. The latter decays are not studied, since the branching ratio is very small and it is hard to distinguish between 5-prong decays and jets. Jets are produced by hadronization of quarks and gluons, and are the major background contribution to tau identification.

The identification work in this thesis will focus on hadronically decaying τ leptons, hereby denoted as $\tau_{had-vis}$, and the $\tau_{had-vis}$ decay products: π^0 and π^\pm . π^\pm are composed of u and d quarks/anti-quarks. They are very long lived and will in be seen as stable particles. π^\pm are used to classify the $\tau_{had-vis}$ as one or three prong.

π^0 is composed of a light quark and the antiquark of the same flavour, and has a very short lifetime. The dominant decay modes for π^0 are $\pi^0 \rightarrow 2\gamma$ and $\pi^0 \rightarrow \gamma + e^+ + e^-$. The first

Decay mode	Branching ratio
Leptonic decay	
$\tau^\pm \rightarrow e\nu_e\nu_\tau$	17.8%
$\tau^\pm \rightarrow \mu\nu_\mu\nu_\tau$	17.4%
Hadronic decay	
One prong (total):	49.5 %
$\tau^\pm \rightarrow \pi^\pm + \nu_\tau$	11.1 %
$\tau^\pm \rightarrow \pi^0\pi^\pm + \nu_\tau$	25.4 %
$\tau^\pm \rightarrow \pi^0\pi^0\pi^\pm + \nu_\tau$	9.2 %
$\tau^\pm \rightarrow \pi^0\pi^0\pi^0\pi^\pm + \nu_\tau$	1.1 %
Three prong (total)	14.6 %:
$\tau^\pm \rightarrow \pi^\pm\pi^\mp\pi^\pm + \nu_\tau$	9.2 %
$\tau^\pm \rightarrow \pi^0\pi^\pm\pi^\mp\pi^\pm + \nu_\tau$	4.3 %
$\tau^\pm \rightarrow \pi^0\pi^0\pi^\pm\pi^\mp\pi^\pm + \nu_\tau$	0.5 %
$\tau^\pm \rightarrow \pi^0\pi^0\pi^0\pi^\pm\pi^\mp\pi^\pm + \nu_\tau$	0.1 %
Five prong (total)	0.1 %

Table 3.1.1: The main decay modes of the τ lepton. [55]

decay mode accounts for more than 99% of the decays and photon detection is used to reconstruct π^0 .

The aim of this identification study is to review the currently available tau identification and in addition to use information from a new π^0 reconstruction which became available recently [56], and to improve the number of good reconstructed $\tau_{had-vis}$ particles. The inner detector has a high precision giving good reconstruction for charged particles. Improvement of $\tau_{had-vis}$ reconstruction can be achieved by utilizing reconstructed π^0 clusters, thereby all the decay products of the τ particles are directly reconstructed.

Table 3.1.1 shows that for 1 prong it is important to split in zero, one and two π^0 cases. For three prong the main decay is with zero or one π^0 . Only rarely will there be more than two π^0 in the decay modes, and therefore these situations will not be considered.

3.1.1 Hadronic τ -decays in the ATLAS detector.

In the ATLAS detector jets and τ_{had} decays give similar signals, and the challenge is therefore to distinguish these signatures. The main differences between the two are shown in figure 3.1.1. The $\tau_{had-vis}$ signature is narrow, has one or three tracks with one track carrying the main part of the momentum and only up to 3 π^0 . A jet signature is wider, has many more tracks and the momentum is spread among more tracks. A jet signature is wider since there will be no reason for a single track to carry the main part of the momentum

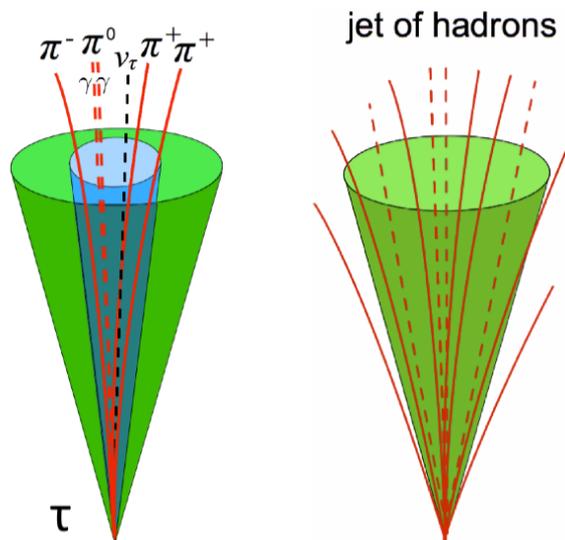


Figure 3.1.1: A simple cartoon showing the differences in signature between $\tau_{had-vis}$ and QCD jets

To identify $\tau_{had-vis}$, first a sample of $\tau_{had-vis}$ candidates are reconstructed and afterwards an identification algorithm is applied to reject fake candidates from QCD jets.

Reconstruction of $\tau_{had-vis}$:

The reconstruction of $\tau_{had-vis}$ is seeded by an anti- k_T jet with a distance parameter of 0.4. This jet is required to have $E_T < 10$ GeV and $|\eta| < 2.5$ [53]. The input to the jet-reconstruction algorithm are topo clusters, which consist of cells from the calorimeter and calibrated as given in [57]. The reconstruction of topo clusters will be described in more detail in section 3.2.1.

An important entity for $\tau_{had-vis}$ is the *tau vertex*, which is the vertex where the $\tau_{had-vis}$ originates from. The tau vertex is chosen as the vertex candidate with the highest ratio of the sum of the P_T of tau tracks matched to the vertex candidate to the sum of the P_T of all tau tracks. This vertex is from now on used to determine the $\tau_{had-vis}$ direction, η^τ and ϕ^τ . This vertex is used instead of the vertex with the maximum sum of P_T of tracks in the event, since it is found to minimize pile-up effects in reconstruction efficiency.

Tracks are added to the $\tau_{had-vis}$ if the tracks are classified as good. The requirements for a good track are: P_T must be above 1 GeV, more than one hit registered in the pixel detector and more than six hits in the pixel and SCT detector combined. Furthermore the transverse distance of closest approach between the track and tau vertex must be less than 1.0 mm. The longitudinal distance of the closest approach is required to be below 1.5mm.

Another important quantity is the *intermediate axis*. This is found by considering the barycenter from the four vectors of the considered clusters (assuming the clusters to

have zero mass) The $\tau_{had-vis}$ four-momentum is then calculated by considering clusters in $\Delta R < 0.2$ around such barycenter. The four-vectors of those clusters are recalculated using the tau vertex coordinate system and the found tau direction provided [58].

3.2 π^0 reconstruction

As the new set of variables used for tau identification is based on the properties of π^0 in the $\tau_{had-vis}$ decay, the definition of reconstructed π^0 becomes important. The first step in identifying π^0 's is cluster reconstruction.

3.2.1 Reconstruction of the topological clusters

Reconstruction of the topological clusters is done in two steps: First the clusters are created, and then they are split in order to resolve individual clusters from the previous step.

The creation of the clusters is illustrated in figure 3.2.1. The whole calorimeter is scanned for cells with an energy significance $\frac{|E|}{\sigma_{noise}}$ above a threshold, and thereby the clusters are created following the steps for the so-called 4-2-0 reconstruction:

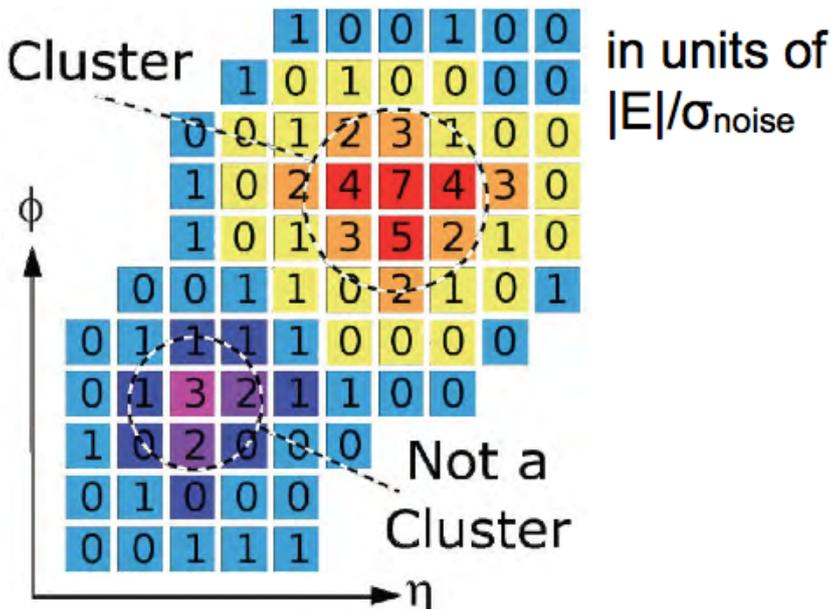


Figure 3.2.1: The reconstruction of the topo clusters[59].

Creating the topological clusters (4-2-0 reconstruction) :

- Identify cells with $\frac{|E|}{\sigma_{noise}} > 4$, and use those as seeds for cluster candidates
- Add neighboring cells to the cluster candidate. Neighboring cells in the η and ϕ space are added as shown in figure 3.2.1 if any of these cells have $\frac{|E|}{\sigma_{noise}} > 2$. Two clusters can merge together at this step.
- Add further neighboring cells to the cluster if $\frac{|E|}{\sigma_{noise}} > 2$

- The adding of cells are stopped when no more cells are found to have $\frac{|E|}{\sigma_{noise}} > 2$.
- The boundary is found at $\frac{|E|}{\sigma_{noise}} > 0$.

Splitting topological clusters :

- The first step in splitting the topological clusters is finding local maxima. A cell is defined as a local maximum when the cell has an energy above a given threshold (typically around 500 MeV) and the energy of the neighboring cells are below a lower threshold.
- The local maximum is identified as a single cluster, and therefore not considered in the next steps.
- A new cluster creation is done on the remaining selected clusters. Since this time the creation of the clusters are on an already selected cluster it is not necessary to use $\frac{|E|}{\sigma_{noise}} > 2$. In this case, if a cell belongs to two clusters the energy will be split between the clusters.

3.2.1.1 Selection of clusters

A score (*pi0score*) is introduced to select the π^0 candidates among the clusters. The *pi0score* is given by:

$$pi0score \approx \frac{E_{PS}}{E_{Had}} \quad (3.1)$$

where E_{PS} is the fraction of energy deposited in the presampler and strip layers of the calorimeters and E_{Had} is the fraction of energy in the hadronic calorimeter. The strip layers are the first and very precise layers in the EM calorimeter. [15] The *pi0score* utilizes the fact that γ 's from π^0 's will mainly deposit energy in the presampler and strip layers, whereas π^\pm will mainly deposit energy in the hadronic calorimeter or the last layers of the electromagnetic calorimeter. The clusters that deposit most energy in the presampler and strip layers get the highest score, and have the highest likelihood of being a true π^0 .

In previous studies a small correction to this score was found to be useful:

$$pi0score \approx \frac{E_{PS}}{E_{Had} + f \cdot E_{cor}} \quad (3.2)$$

where f is an arbitrary factor, that can be chosen at will.¹ E_{cor} is given by equation 3.3.

$$E_{cor} = \sqrt{\left| \frac{E_c}{E_{calo} - E_{track}} - 1 \right|} \quad (3.3)$$

E_c is the energy of the investigated cluster, E_{calo} is the energy sum of all π^0 candidates and E_{track} is the energy of the tracking system. [15] The difference ($E_{calo} - E_{track}$) is a naive

¹ f is set to 1 in the present settings

guess of the π^0 energy. When this naive guess of π^0 energy is far from E_c , this will give a correction to the $pi0score$, and the cluster will be less likely to be identified as a π^0 . As an example, in the situation with $1\pi^\pm+0\pi^0$ the $E_{calo} - E_{track} \neq E_c$ and therefore there is a correction to the $pi0score$, whereas for the situation with $1\pi^\pm+2\pi^0$ the $E_{calo} - E_{track} \approx E_c$, resulting in $E_{cor} \approx 0$.

The clusters with the highest score will be classified as π^0 candidates.

3.2.2 The number of π^0

Determining the number of π^0 is done in two steps via two Boosted Decision Trees, *BDT*. The *BDT* algorithm is described in section 4.1. It is mentioned here without introduction since the π^0 determination algorithm is not optimized in this work. The first *BDT* determines if any π^0 is present, and if so the other *BDT* determines the number of π^0 [15].

In both *BDT*'s the following discrimination variables are used:

- $\frac{E_{track}-E_{Had}}{E_{EM}}$

Where E_{track} is the energy of the tracks from $\tau_{had-vis}$, corresponding to the energy of π^\pm . E_{Had} and E_{EM} is the energy in the hadronic and electromagnetic calorimeters respectively.

This is expected to be a good discriminator since the different decay modes deposit the energy differently in the detector. If there is no π^0 present both the nominator and the denominator is small, and the fraction close to one. Whereas, if both π^0 and π^\pm are present the numerator will be small and the denominator large, and consequently the fraction will be small. The distribution of this fraction is shown for different decay modes in figure 3.2.2(a)

- $\frac{E^{EM}}{P^{track}}$

where E^{EM} is the energy in the electromagnetic calorimeter and P^{track} is the sum of momentum of the $\tau_{had-vis}$ tracks.

The presence of π^0 results in a lower P^{track} and more activity in the electromagnetic calorimeter reducing the ratio as confirmed in figure 3.2.2(b). Therefore this variable is useful for discrimination of π^0 .

- $\frac{E_T^{calo}}{P_T^{lead.track}}$

Where E_T^{calo} is the transverse energy in the calorimeter and $P_T^{lead.track}$ is the transverse momentum of the leading track.

If there is no π^0 the leading π^\pm carry more energy and therefore this fraction is lowered as seen figure 3.2.2(c).

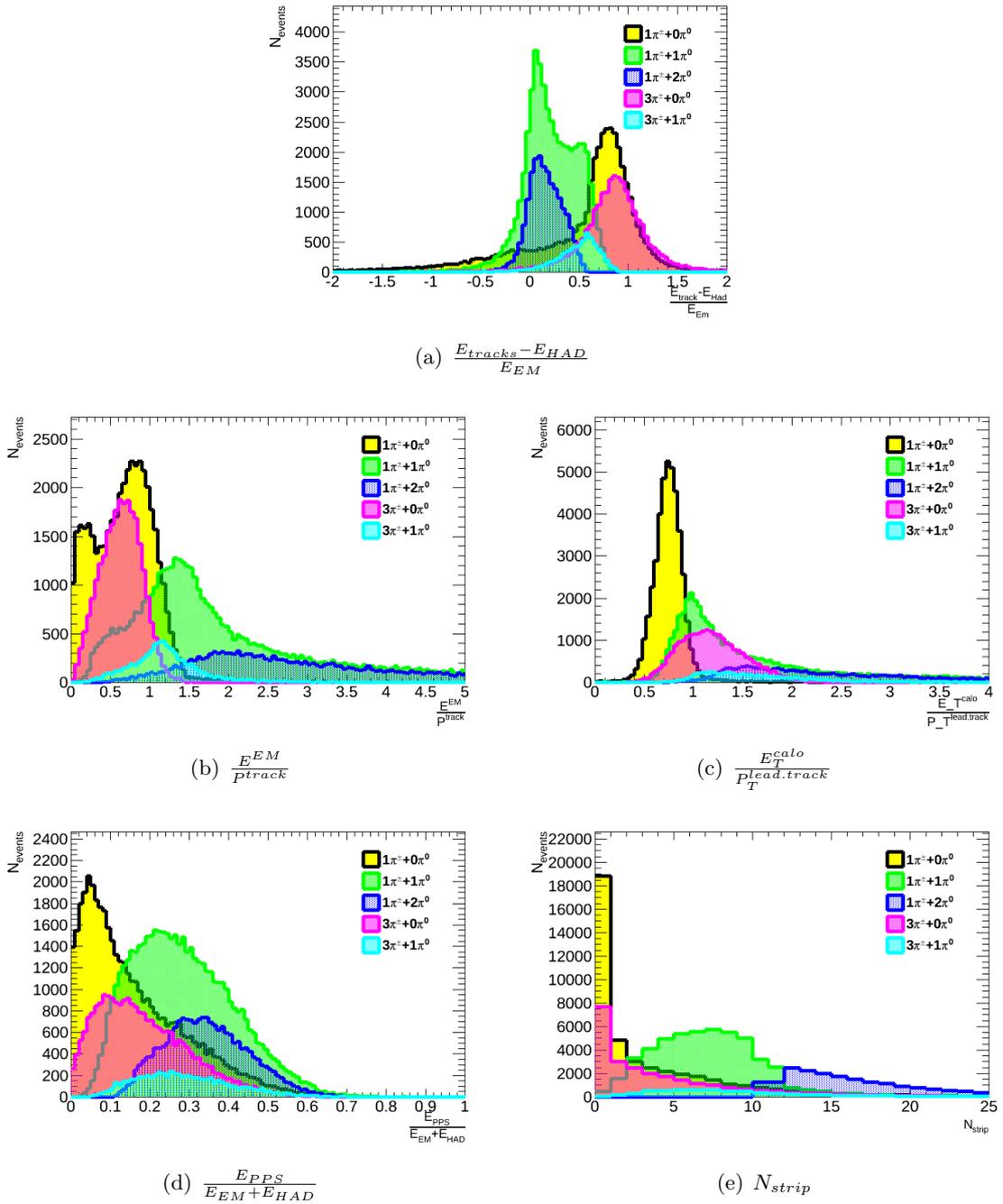


Figure 3.2.2: Distributions of the variables used to distinguish the number of π^0 in the $T_{had-vis}$ decay. The plots are done for $P_T > 10$ GeV, and the cuts mentioned in section 3.3

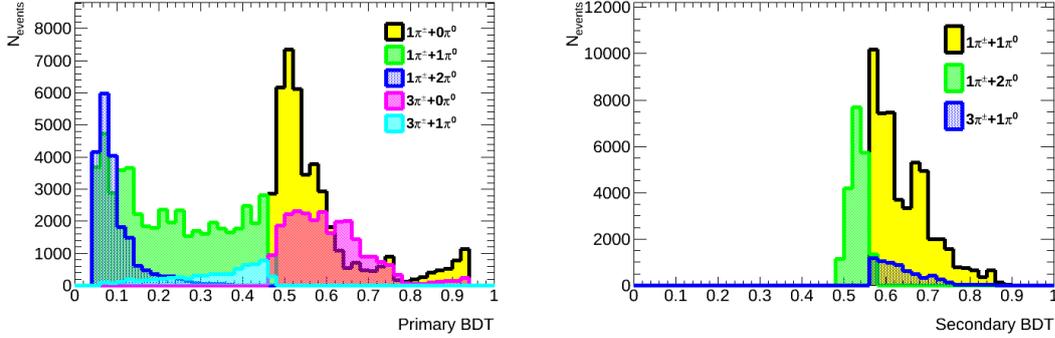
- $\frac{E_{PPS}}{E_{EM} + E_{HAD}}$

The γ from the π^0 decay are more likely to deposit energy in the presampler and strip layer than the π^\pm . Therefore any π^0 present in the decay cause a higher fraction of the energy to be deposited in the pre-sampler and strip layers as seen in figure 3.2.2(d).

- N_{strip}

The number of strips that have more than 200 MeV

This is a good discriminator since only π^0 will deposit this amount of energy in the strip layer. This behaviour is confirmed in 3.2.2(e).



(a) Distinction between zero and above zero π^0 (b) Distinction between one and two π^0 . As explained in the text this figure is made with a subset of the events in (a)

Figure 3.2.3: BDT score to distinguish between zero and above zero π^0 (left) and between one and two π^0 (right)

Figure 3.2.3 shows the BDT score when these variables are applied to simulated events for $\tau_{had-vis}$. A higher BDT score in figure 3.2.3(a) corresponds to a smaller chance of any true π^0 . In figure 3.2.3(b) the distinction between one or two π^0 is shown. The case of $\pi^0 > 2$ is not considered since it is a relatively rare situation and does not garner special interest.

true/reco [%]	$1\pi^\pm + 0\pi^0$	$1\pi^\pm + 1\pi^0$	$1\pi^\pm + 2\pi^0$	$3\pi^\pm + 0\pi^0$	$3\pi^\pm + 1\pi^0$
$1\pi^\pm + 0\pi^0$	83.6	14.7	1.09	0.49	≈ 0
$1\pi^\pm + 1\pi^0$	19.7	56.9	20.4	0.65	0.16
$1\pi^\pm + 2\pi^0$	5.81	46.3	43.3	0.04	0.30
$3\pi^\pm + 0\pi^0$	1.35	1.34	0.25	88.7	7.58
$3\pi^\pm + 1\pi^0$	≈ 0	2.4	1.44	48.0	38.4

Table 3.2.1: Matrix of true vs reconstructed $\tau_{had-vis}$ decay mode

In table 3.2.1 the reconstructed π^0 is compared to the true π^0 . In the table the number of reconstructed π^0 over true π^0 for different numbers of π^\pm and π^0 are given. The training is done on 1 prong and is therefore optimized for the first two entries on the diagonal. From the table it is evident that the reconstruction of the number of π^0 is good.

3.3 Data and preselection cuts

In this study simulated events are used for the $\tau_{had-vis}$ particles and data for the jets. The data was recorded by the ATLAS detector in the summer and fall of 2012. The used data sample was triggered by jets triggers. The following preselection cuts are used on both $\tau_{had-vis}$ and jets :

- $P_T > 15 \text{ GeV}^2$
- $|\eta_\tau| < 2.3$
- Number of $\tau_{had-vis}$ track is 1 or 3

For $\tau_{had-vis}$ these additional requirements are applied:

- Matched to a true tau in $\Delta R < 0.2$
- The true τ particle must have $|\eta_\tau| < 2.5$
- $E_T > 10 \text{ GeV}$

The simulated τ 's are partly from Z(56%) and Z'(44%). The Z' is added so as to train on more events with high transverse momentum. A re-weighting of the P_T^τ is done for simulated Z/Z' to assure accordance with P_T of data. This re-weighting weighs τ 's with low P_T higher than τ 's with high P_T , resulting in Z' being less important

The amount of used data is given in table 3.3.1:

	1 prong	3 prong
Background	$3.42 \cdot 10^6$	$4.67 \cdot 10^6$
Signal	$2.09 \cdot 10^6$	$5.64 \cdot 10^5$

Table 3.3.1: The number of tau candidate used in the study

Half of the events will be used for identification, and the other half will be used for testing.

Compared to previous tau identification the P_T requirement is lowered by 5 GeV. This is done in expectation of improved performance at low P_T by the new variables. It is also expected that the new variables will reduce the background with respect to before. These points can be helpful in studies searching for the Higgs boson or other physics searches with taus.

3.4 Tau identification

The identification is done among the candidates found from the reconstruction. The identification is done by using variables which differ between $\tau_{had-vis}$ and jets. This identification

²There is assumed to be zero mass for the tau, and thereby $E_T = P_T$

can be done via a simple cut based method, but the performance is better using more advanced algorithms like Likelihoods and Boosted Decision Trees (BDT). The identification work in the next chapters will be done using a BDT algorithm.

3.5 Current variables for tau identification

This analysis was carried out in October 2012. The default identification at the time will be referred to as the "current" status. The variables used for $\tau_{had-vis}$ identification are listed in the following. These variables are used since it was found in [58] that this set of variables gives the optimal performance.

Parentheses denote if a given variable is used in 1 prong (1p), three prong (3p) or both (1p,3p)

- Leading track momentum fraction, $corrFTrk$. (1p,3p)

$$FTrk = \frac{P_T^{lead}}{\sum^{\Delta R < 0.2} (E_T^{EM})}$$

P_T^{lead} is the transverse momentum of the track with the highest transverse momentum denoted as leading track. P_T^{lead} is divided by the sum of the transverse energy over all the cells in $\Delta R < 0.2$ around the intermediate axis³. This fraction was previously found to depend on the amount of pile-up in the event, and a linear correction is applied to the fraction.

$$corrFTrk = FTrk + 0.003N_{vtx}$$

N_{vtx} is the number of good vertices in the event, defined as having at least two tracks assigned to it.

This variable is expected to be a good discriminating variable, since the leading track of the $\tau_{had-vis}$ accounts for much of the transverse momentum, whereas the transverse momentum of the jets will be more equally shared by all tracks.

- Central energy fraction, $corrCentFrac$ (1p,3p)

$$CentFrac = \frac{\sum_i^{\Delta R < 0.1} E_{T,i}^{EM}}{\sum_j^{\Delta R < 0.2} E_{T,j}^{EM}}$$

The sum of all the transverse energy deposited in the cells in a cone of $\Delta R < 0.1$ over all the cells in the cone, $\Delta R < 0.2$. ΔR is calculated relative to the intermediate axis.⁴ This fraction is also corrected for pile-up for $P_T^T < 80\text{GeV}$:

$$corrCentFrac = CentFrac + 0.003N_{vtx}$$

$corrCentFrac$ is expected to be a good discriminator between the $\tau_{had-vis}$ particles and the jets, since most of the energy will be in the core for the $\tau_{had-vis}$, whereas it is expected to be more spread out for jets.

³The energy is calibrated at the EM energy scale

⁴The energy is calibrated at the EM energy scale.

- The mass of the track system, $massTrkSys$ (3p)

The mass of the track system is the invariant mass of the tracks in both the core and the isolation region. The latter is defined as $0.2 < \Delta R < 0.4$.

The mass of track system is expected to be a good discriminator, since it yield the mass of the charged decay products for $\tau_{had-vis}$, and a more spread out mass signature for the jets.

- The decay length significance of the secondary vertex, $trFlightPathSig$ (3p)

$$trFlightPathSig = \frac{L_T^{flight}}{\delta L_T^{flight}}$$

L_T^{flight} is the reconstructed signed decay length in the transverse plane and δL_T^{flight} is the corresponding uncertainty. The decay length of the jets is expected to be lower than for $\tau_{had-vis}$.

- The sum over tracks weighted by their transverse momentum, $trkAvgDist$ (1p,3p)

$$trkAvgDist = \frac{\sum_i^{\Delta R_i < 0.4} P_{T,i} \cdot \Delta R_i}{\sum_i^{\Delta R_i < 0.4} P_{T,i}}$$

The sum is done over the momentum of all tracks with ΔR defined between the direction of the $\tau_{had-vis}$ and the intermediate axis.

It is expected that ΔR is lower for $\tau_{had-vis}$ than for jets. The reduction will be amplified by the P_T weights.

- The maximal ΔR between the direction of $\tau_{had-vis}$ and the intermediate axis, dR_{max} in the core cone (3p)

For one prong this is almost the same as $trkAvgDist$, and therefore this variable is only relevant for three prong.

- The impact parameter significance of the leading track in the core region, $ipSigLeadTrk$. (1p)

$$ipSigLeadTrk = \frac{d0}{\delta d0}$$

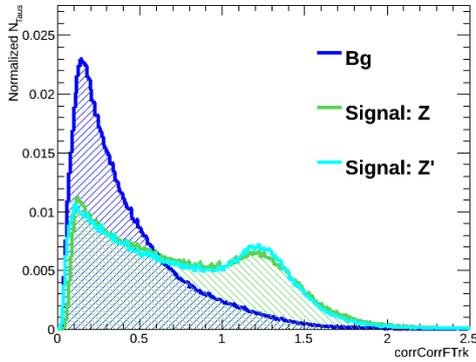
$d0$ is the distance of closest approach between the track and the vertex, and $\delta d0$ is the corresponding uncertainty.

$d0$ is expected to have a wider distribution for $\tau_{had-vis}$ than jets, since the $\tau_{had-vis}$ will travel a finite distance before decaying (due to the $\tau_{had-vis}$ lifetime).

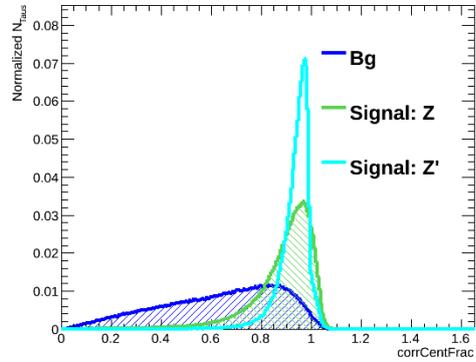
- The number of tracks in isolation region, $N_wideTrk$. (1p)

This quantity is used since the jet signature is broader than the $\tau_{had-vis}$ signature.

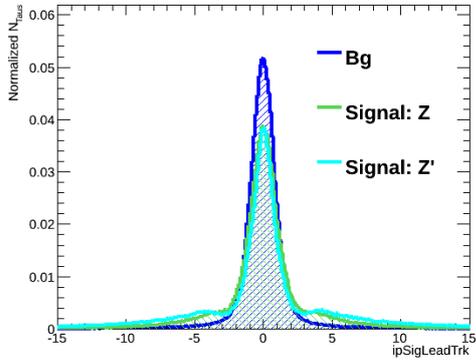
The distributions of the variables for 1 prong and 3 prong are plotted in figure 3.5.1 and figure 3.5.2 respectively.



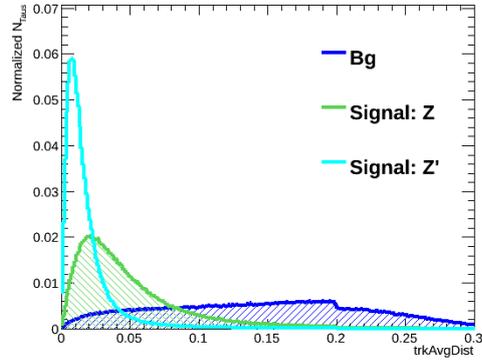
(a) corrFTrk, 1 prong



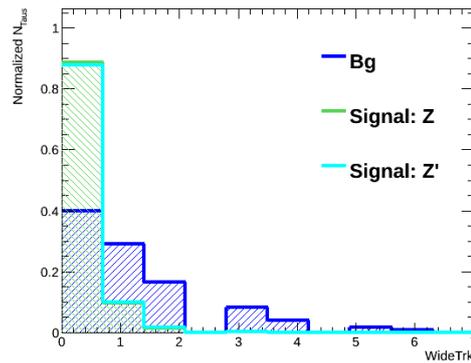
(b) corrCentFrac, 1 prong



(c) ipSigLeadTrk, 1 prong



(d) trkAvgDist, 1 prong



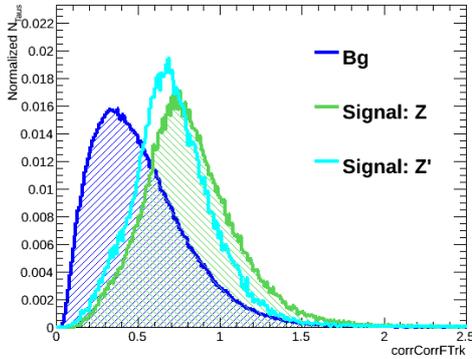
(e) N-wideTrk, 1 prong

Figure 3.5.1: Input variables for tau identification for one prong τ decays

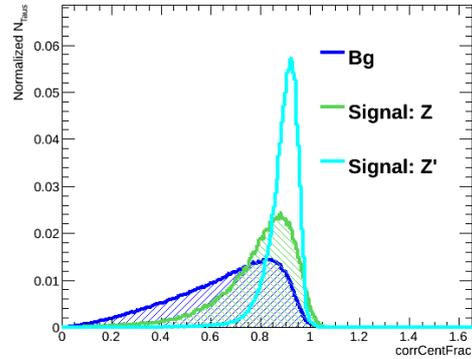
The distributions show that $\tau_{had-vis}$ behave differently depending on whether it is a decay product of a Z or a Z' boson. Due to the difference in mass between the two bosons $\tau_{had-vis}$ particles are boosted differently and this consequently influences the decay products. This is

the reason for using both Z and Z' as signal.

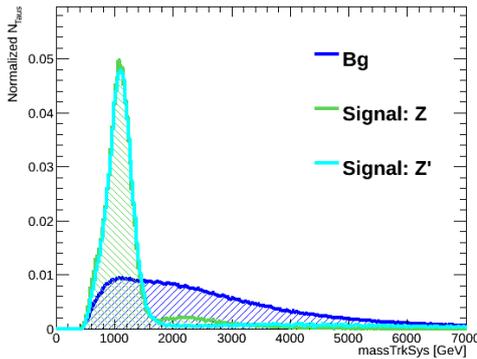
For 1 prong the differences are only seen in trkAvgDist and corrCentFrac and for 3 prong small differences are seen for almost all variables. These differences are due to the relatively higher boost of the Z' $\tau_{had-vis}$ giving it a narrower signature than the Z $\tau_{had-vis}$.



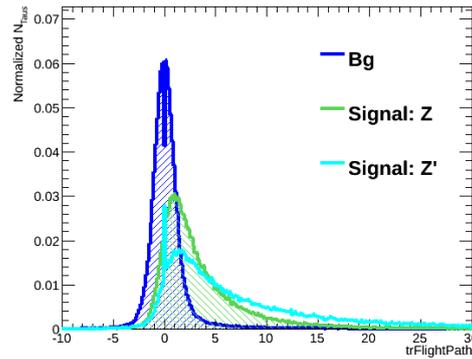
(a) corrFTrk , 3 prong



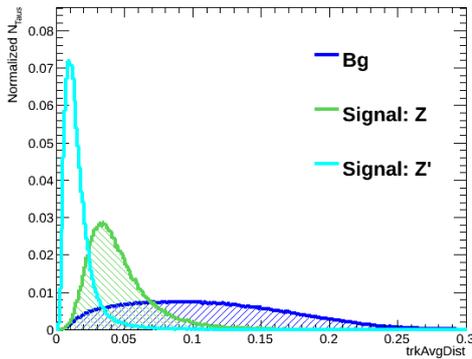
(b) corrCentFrac , 3 prong



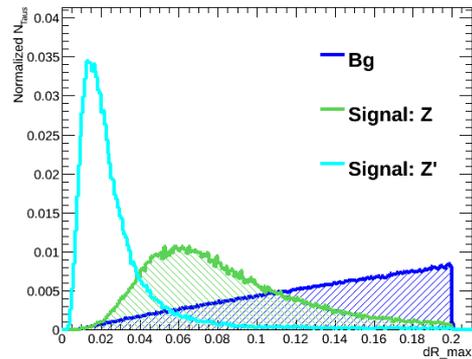
(c) massTrkSys , 3 prong



(d) trFlightPathSig , 3 prong



(e) trkAvgDist , 3 prong



(f) dRmax , 3 prong

Figure 3.5.2: Input variables for tau identification for 3 prong τ decays.

3.6 New variables for tau identification based on π^0 information

The new set of variables tested to improve the identification are:

- The number of reconstructed π^0 , *tau_pi0_n*

The $\tau_{had-vis}$ will decay to zero, one or two π^0 as shown in table 3.1.1. There is no reason for the distribution of π^0 to be the same for jets as for $\tau_{had-vis}$.

Figure 3.6.1(a) and 3.6.1(b) show the distribution of reconstructed π^0 . For 3 prong the discrimination between background and signal is much better than for 1 prong.

- The mass of the track system and π^0 system in the core region, *pi0_vistau_m*

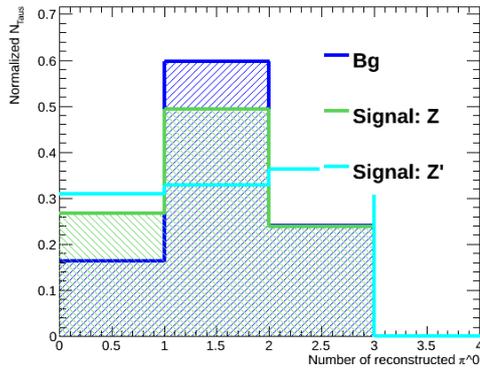
Figure 3.6.1(c) and 3.6.1(d) show the combined mass distributions of the track system and π^0 system in the core region. For 1 prong there are two peaks in the signal mass distribution relatively close to the mass of the $\rho(\pi^\pm\pi^0)$ at 770 MeV and $a_1(\pi^\pm\pi^\pm\pi^\mp, \pi^\pm\pi^0\pi^0)$ at 1260 MeV. For 3 prong the values are on average higher and closer to the mass of a_1 . This behavior is expected from the decay of the $\tau_{had-vis}$ in table 3.1.1.

- The ratio between calorimeter based P_T^c and the combined reconstructed transverse momentum from the found π^0 and tracking system, *PT_ratio*.

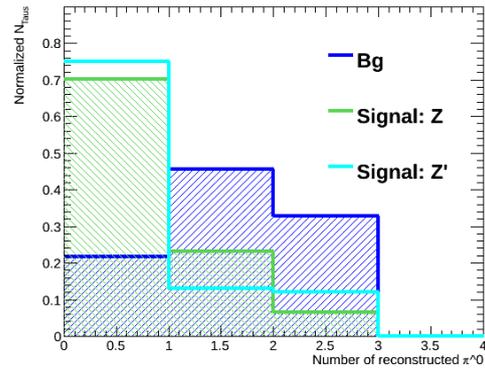
The momentum from the found π^0 is reconstructed from two found π^0 clusters. It was found in previous studies that the best performance was achieved by merging two clusters, even in the case of only one π^0 found.

For $\tau_{had-vis}$ this fraction is expected to be close to one, since it is the momentum of the same object. This is not the case for jets. The ratio is therefore expected to discriminate well between background and signal, which is confirmed on figure 3.6.1(e) and 3.6.1(f) .

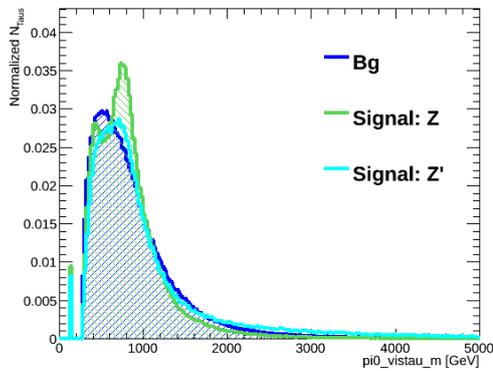
Figure 3.6.2 and figure 3.6.3 show the distributions for one and three prong in the three cases: zero, one or two π^0 . It should be noted that the *pi0_vistau_m* is always found for two clusters, even in the case of only one π^0 . Therefore this will not be different between 1 and 2 π^0 . Differences are present between zero and above zero π^0 .



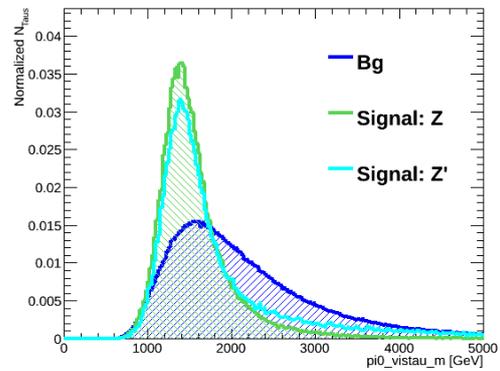
(a) 1 prong



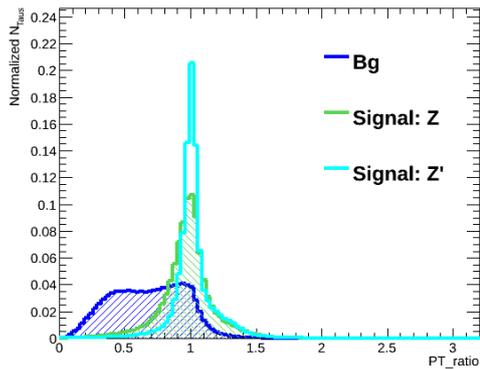
(b) 3 prong



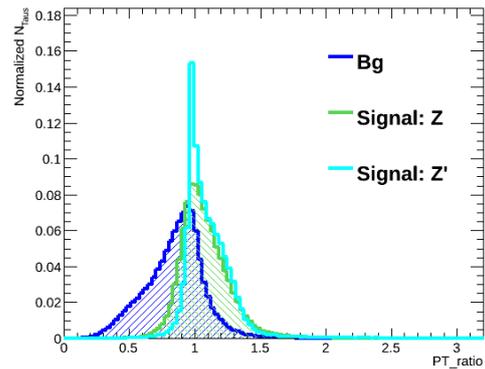
(c) 1 prong, $\pi^0_{\text{vistau}_m}$



(d) 3 prong, $\pi^0_{\text{vistau}_m}$

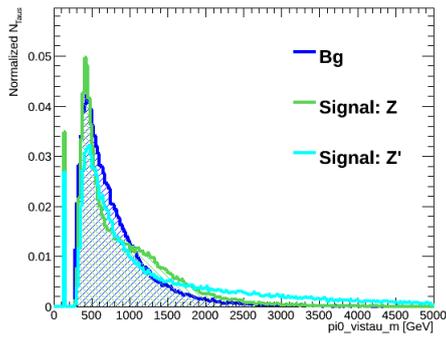


(e) 1 prong, PT_{ratio}

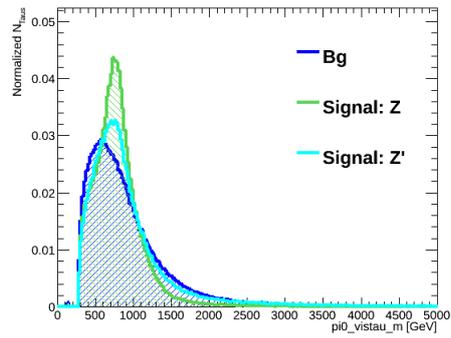


(f) 3 prong, PT_{ratio}

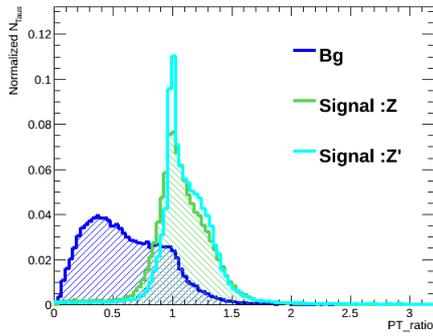
Figure 3.6.1: Figures (a) and (b) show the reconstructed number of π^0 for 1 and 3 prong. Figures (c) and (e) show the mass reconstructed from the π^0 and the tracking system. Figures (d) and (f) show the ratio between the combined π^0 /track momentum and the $\tau_{had-vis}$ calorimeter based momentum.



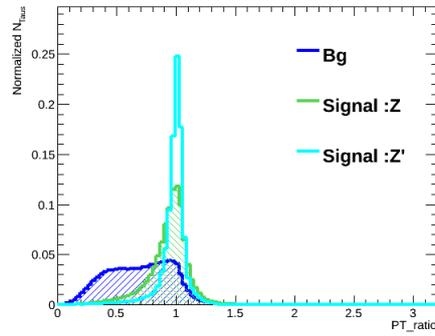
(a) 1 prong, $\pi^0_{\text{vistau_m}}$, 0 π^0



(b) 1 prong, $\pi^0_{\text{vistau_m}}$, 1 or 2 π^0



(c) 1 prong, PT_ratio , 0 π^0



(d) 1 prong, PT_ratio , 1 or 2 π^0

Figure 3.6.2: Distributions for one prong according to the number of π^0 . The first row shows the mass and the last row shows the momentum ratio

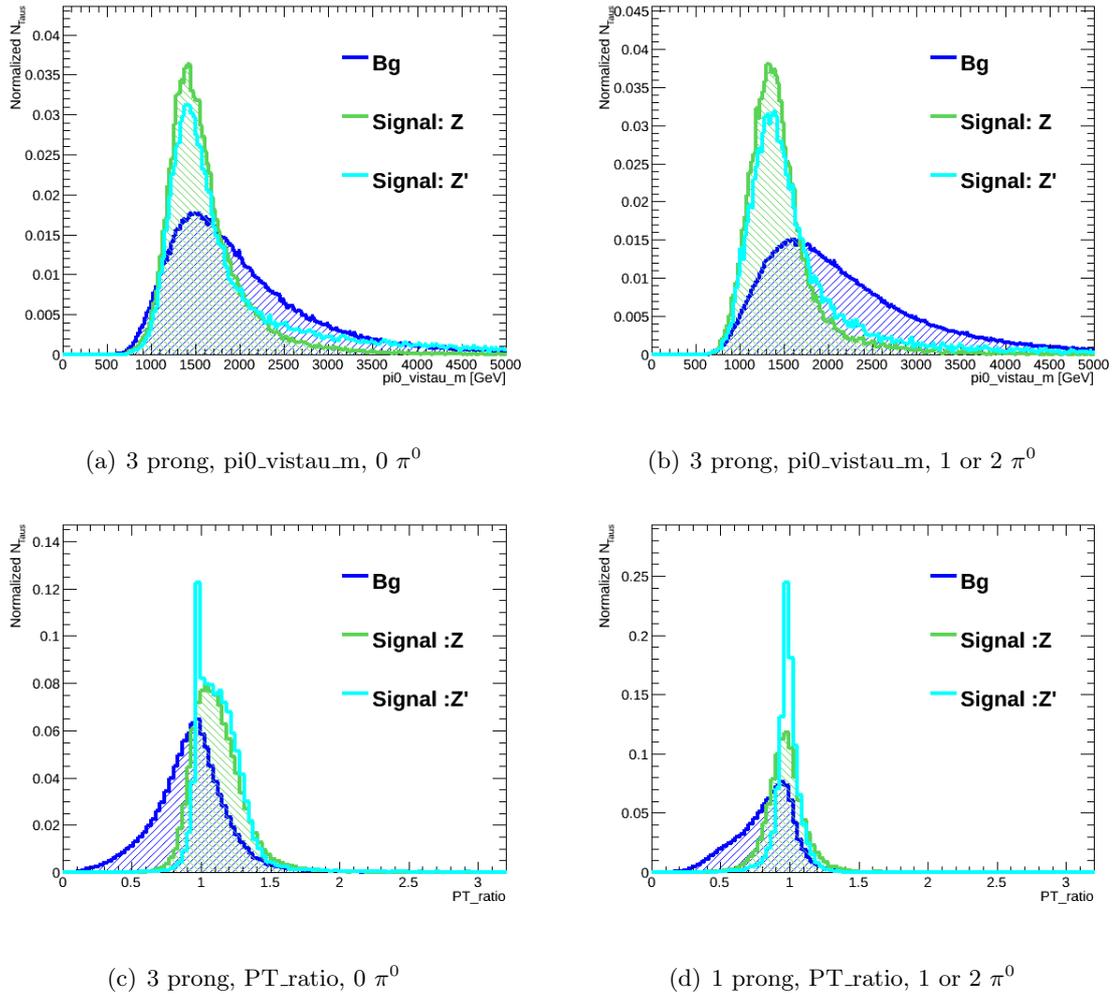


Figure 3.6.3: Distributions for three prong according to the number of π^0 . The first row are the mass and the last row shows the momentum ratio

4

Boosted Decision Tree algorithm

4.1 Boosted Decision Tree algorithm

A Boosted Decision Tree algorithm, BDT, is used to discriminate signal from background. There are several advantages of a BDT algorithm compared to a simple cutting procedure. In the latter it is important that a given variable has a single point with good discrimination in the distributions of the signal and background. In the BDT method this situation is also an advantage, but it will also be able to discriminate in less obvious cases. The split between background and signal is done stepwise, and thereby the finer structure of the variables are used. Other methods like neural networks could also be used, but BDT is chosen since it is expected to give better performance than a cutting base investigation, and is more transparent than neural network since it allows control of the settings.

Figure 4.1.1 illustrates the basics behind BDT training. In the first step all data (background and signal) are collected in a sample called the root node. The variable that best separate background from signal is chosen and splits the sample in root node into two new samples the so-called leaf nodes. Determining the best split can be done in several ways, but in the BDT used in this thesis, GiniIndex is applied. The GiniIndex is a gain defined:

$$gain(node) = p(1 - p) \quad (4.1)$$

p is the purity given by $p = \frac{n_s}{n_s + n_b}$. n_s is the number of signal events and n_b is number of background events. The variable that gives the largest separation gain in the following equation is used, where the largest separation is given by the highest value of G :

$$G = gain(parent-leaf-node) - gain(daughter-leaf-node1) - gain(daughter-leaf-node2)$$

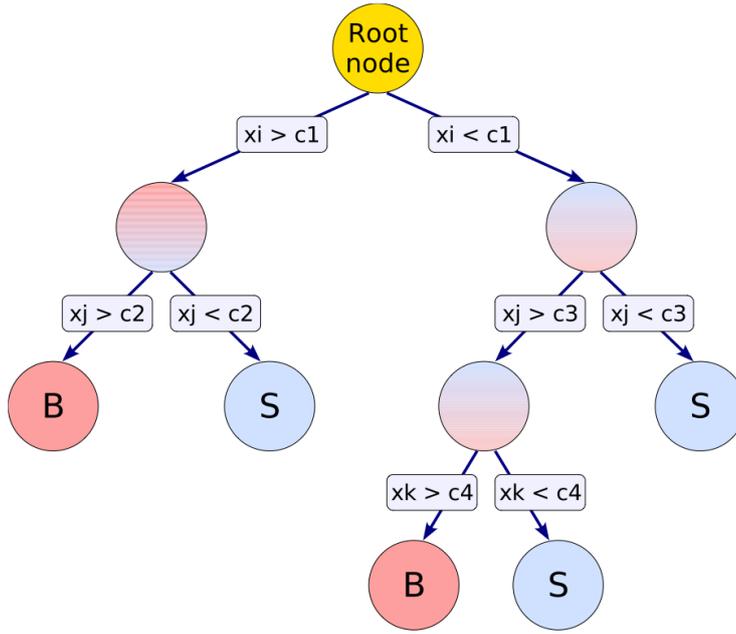


Figure 4.1.1: The diagram illustrates one tree BDT structure [60].

The *parent – leaf – node* is the investigated leaf node and the *daughter – leaf – node1* and *daughter – leaf – node2* are the two new leaf nodes[60] [61]. This splitting procedure is repeated for every leaf node until a given stopping parameter is reached, as for instance a minimum number of events in a node. The same variable can be used repeatedly since a new leaf node contains a new number of events and a new ratio between signal and background. The final leaf is then assigned to be either a signal leaf (*S*) if it contains primarily signal events, or a background leaf (*B*) if it contains primarily background events. The output of this tree is a BDT score for each event which will be either the purity or a binary answer. The latter would be 1 in *S* and -1/0 in *B*. This structure constitutes a tree and is shown in figure 4.1.1 [60].

By boosting the first tree a new tree is obtained. Boosting is done by adding a weight, ω , to signal (background) events that have ended up in a background (signal) leaf. To determine the strength of the boosting a β value is chosen. A β value of zero corresponds to no boosting applied and a high value corresponds to a high degree of boosting for each step. If there is a high degree of boosting the event will get a high weight if classified wrongly. Thereby the wrongly classified event will have a bigger influence on the next tree. If there is no boosting the new tree will be identical to the first tree.

The weight, which is multiplied to a given event k , is calculated by the following for step $i + 1$:

$$\omega^{(i+1)}(k) = \omega^{(i)}(k)\beta^{1-L^i(k)} \quad (4.2)$$

$L^i(k)$ is the loss, which can be interpreted as the fraction classified wrongly.

The collection of trees create a forest. In the end an event will get a BDT score which is the average of the scores it gets in all trees [60].

After the BDT training, BDT testing can be done. An event used for BDT testing needs to be similar to the events used for the BDT training. The training output is applied to the testing events and the performance of the training on new data is determined. It is important to ensure that the training part does not perform better than the testing part, since this would indicate overtraining. In such cases it is always the testing tree that should be trusted.

A final important feature is pruning. When the trees are grown to the maximum size pruning remove all statistically weak nodes. This is done from the bottom and up. The idea is to reduce sensitivity to overtraining and to remove weak nodes [60].

4.2 Settings for the Boosted Decision Tree algorithm

In this section the sensitivity of the BDT to the configuration settings is investigated and the optimal settings of the algorithm is determined. One prong variables are considered, which are the five variables in figure 3.5.1 and the three variables in figure 3.6.1.

Before starting the BDT training a re-weighting of the transverse momentum of the signal relative to jet background is done. This is done to make sure that any gain in performance is from the given variable and not correlated with the momentum of the $\mathcal{T}_{had-vis}$.

The plots in figure 4.2.1 show the performance when varying the algorithm settings. The left column shows the performance and the right column shows if the best performance of the left column have any overtraining. Thereby it is possible to obtain the optimal setting and table 4.2.1 gives the value which was chosen. Table 4.2.1 also states the values currently used. A ranking of 1,2 or 3 is given to each variable depending on the importance of this variable. In the "current" settings there is no pruning, and this has now been added to the new tree.

Figure 4.2.3 shows the comparison between the current and the new BDT settings. The performance of the new settings and the current settings are very comparable. The new tree is, however, much smaller and therefore expected to be more robust and less sensitive to statistical fluctuations

Variable	Importance	Chosen value	Current value	Figure	Comments
Number of bins in the first axis (ncut)	3	140	500	4.2.1(a) and 4.2.1(b)	The range of each variable is set before the training, and therefore this setting is not important for the training performance.
Number of trees	1	500	50	4.2.1(c) and 4.2.1(d)	The performance increases with the number of trees. This is to be expected since the first tree, which has the most misclassified events, has less influence and fewer events are misclassified in every new tree.
Maximum depth of a tree	1	10	-	4.2.1(e) and 4.2.1(f)	The depth of the tree is important to exploit the finer structure of the variables. Figure 4.2.1(f) shows that for a depth of 100 overtraining is present, the same performance is gained by a depth of 12. For a depth of 10 there is a small decrease in performance, but no overtraining.
Maximum number of leaf nodes in a tree	1	200	-	4.2.1(g) and 4.2.1(h)	The number of leaf nodes controls the size of the tree making it possible to keep the tree at a reasonable size to avoid overtraining. The best performance without overtraining is obtained at a maximum number of leaf nodes of 200.
The weight parameter β from equation 4.2 (adb)	3	0.8	0.2	4.2.2(a) and 4.2.2(b)	A small improvement is seen by increasing this variable. Further improvement is not seen at additional increments.
Minimum number of events in a node	2	500		4.2.2(c) and 4.2.2(d)	The minimum number of events is not important due to the very low depth of the tree.

Table 4.2.1: The settings for the BDT. A relative importance value is defined, where 1 is given to the most important variables in these setting and a 3 is the least important variables. 2 is used for intermediate cases

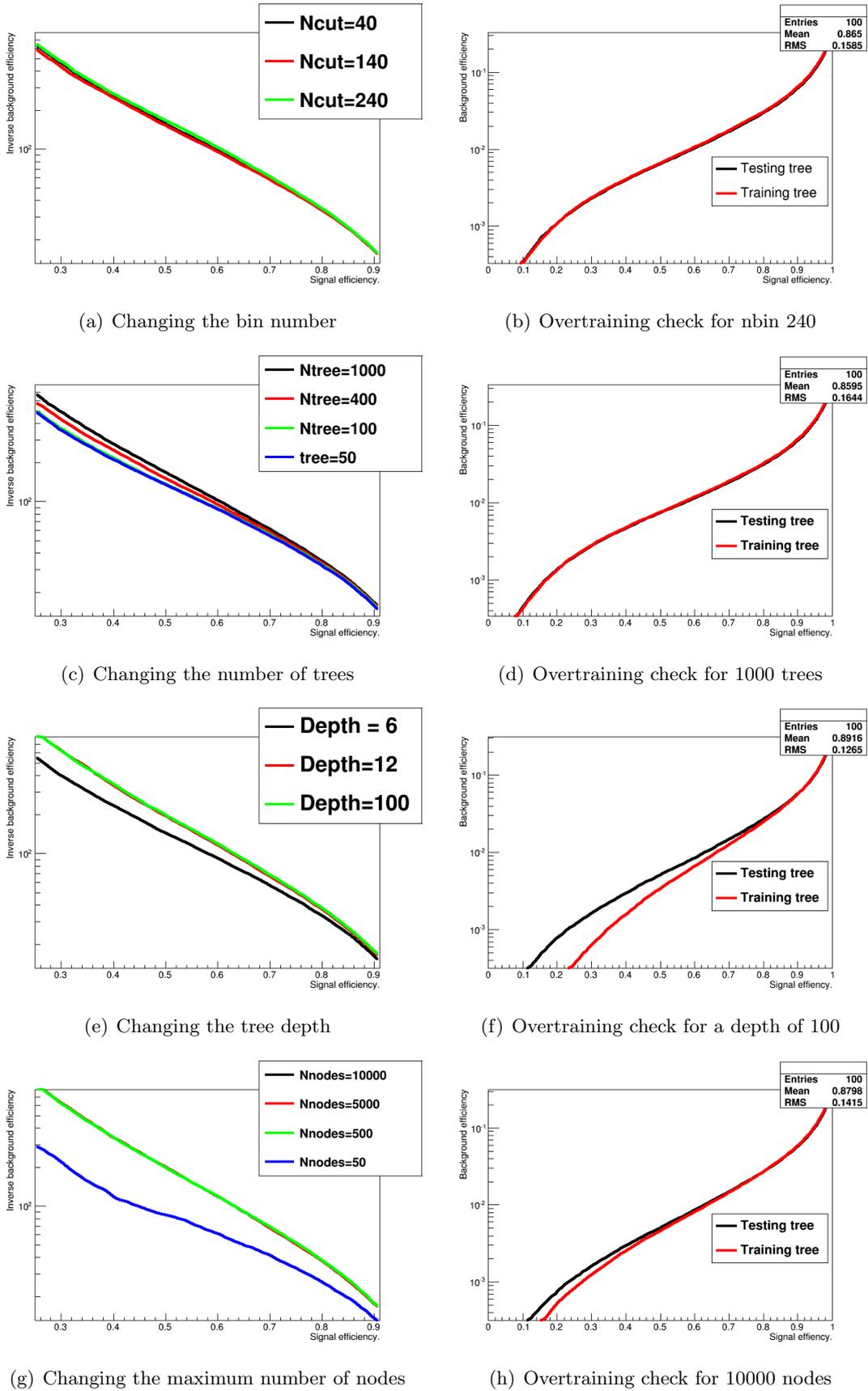
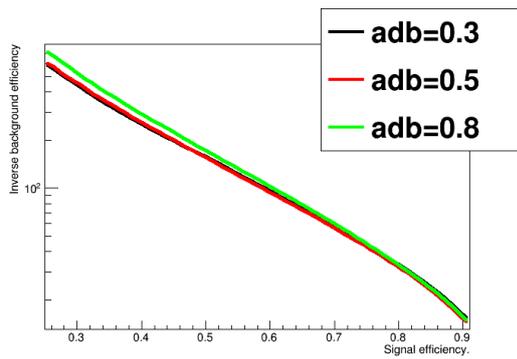
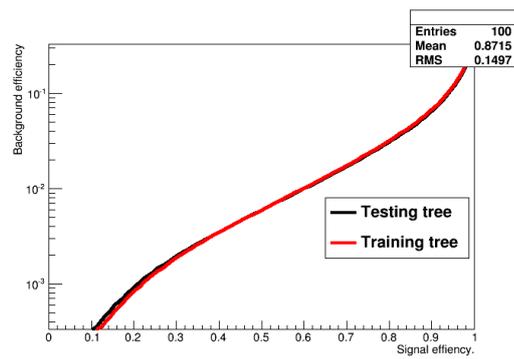


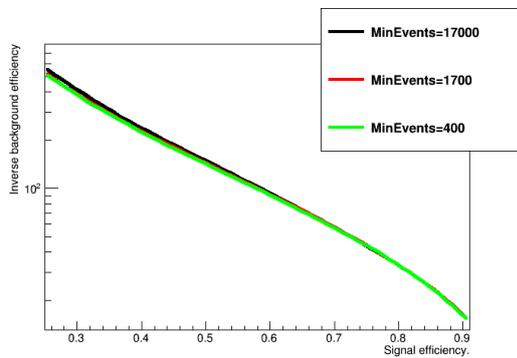
Figure 4.2.1: Investigation of the BDT settings. In (a) the N_{cut} does not alter performance making the plots overlay. In (c) the same performance is gained at 100 or 50 trees. In (e) the same performance is attained by a depth of 12 and 100. The performance is independent of the number of nodes above 500 nodes in (g) Explanations are given in table 4.2.1



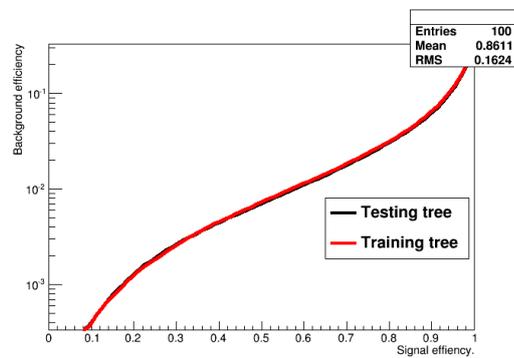
(a) Changing the weight parameter ($\beta = adb$)



(b) Overtraining check for a weight parameter of 0.8



(c) Changing minimum number of events



(d) Overtraining check for 400 event

Figure 4.2.2: Investigation of the rest of the BDT settings. The performance is seen to be independent of the weight parameter in (a) and minimum number of events in (c)

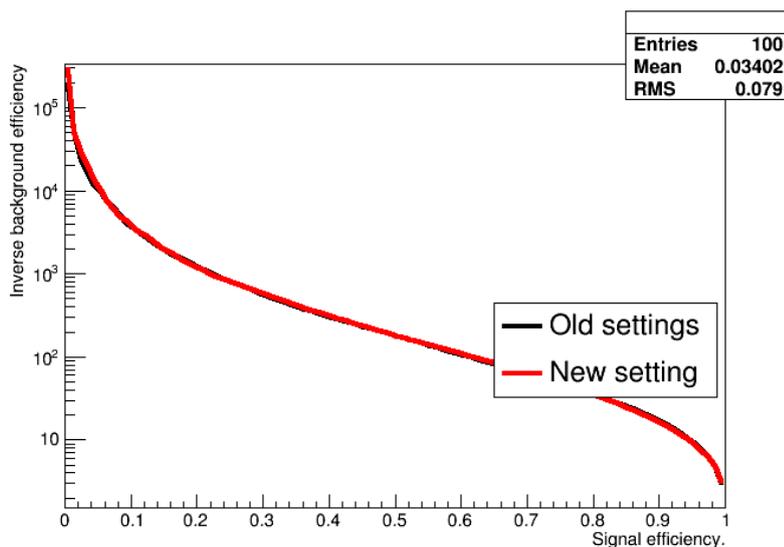


Figure 4.2.3: The performance with the current and the new training

5

Incorporation of the new tau identification variables and re-optimization

5.1 Training with the new tau identification variables

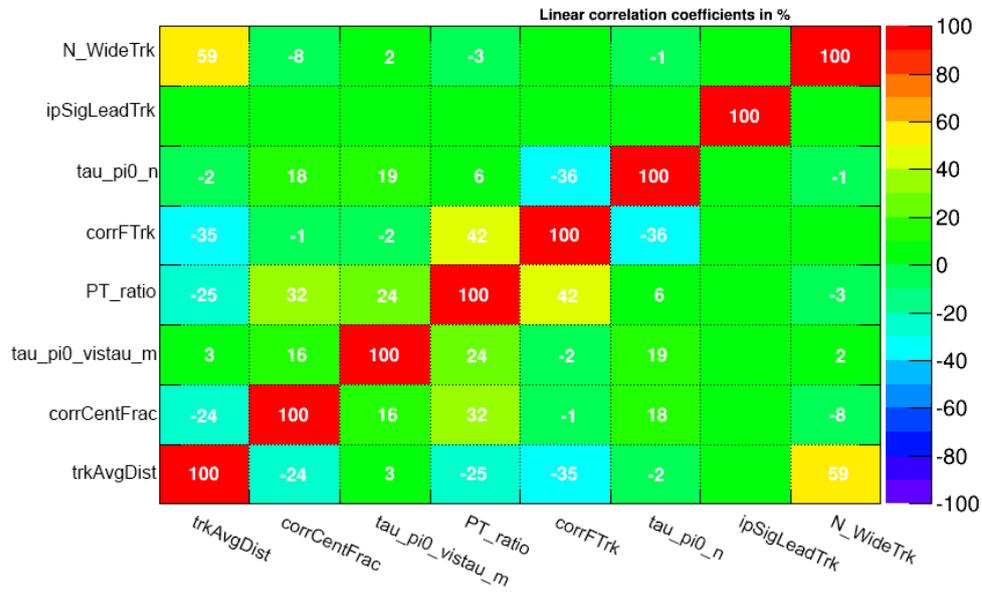
In an earlier study it was found that it is best to perform one training with all the variables together instead of splitting the process into more trainings depending on the number of π^0 . Therefore this procedure will be the one used in the following.

5.2 Correlation between the tau identification variables

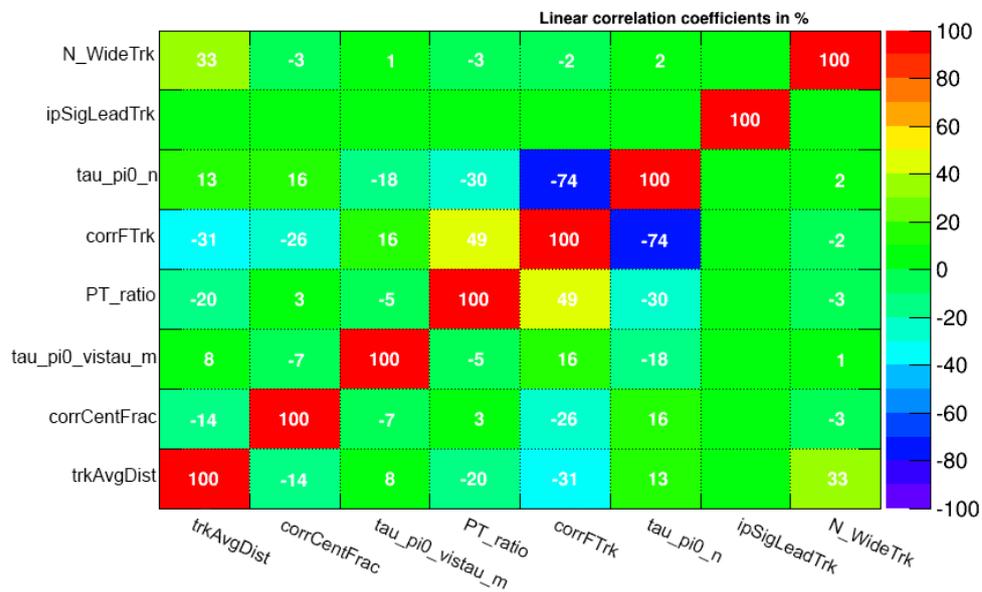
Correlations between variables provides hints for finding superfluous identification variables. If a high correlation between two variables are found the scatter plot is analysed. If the scatter plot shows no clear dependence between the variables, and the same dependence in signal and background, one of the variables is likely to be less important. The training is then attempted without this variable to test if the performance is retained after removing this variable. If two variables have a low correlation there is no need to consider the scatter plot.

Figure 5.2.1 and 5.2.2 show the linear correlations between the variables for one prong and three prong. Scatter plots and profile plots for cases with high correlation are seen in figure 5.2.3, 5.2.4 and 5.2.5. A profile plot shows the mean of the y axis and the spread.

The highest correlations are investigated in more detail in table 5.2.1. In most cases the profile plots look similar and underline a correlation between the variables, but the scatter plots show that differences between the variables do occur. The scatter plots illustrate that for background the distributions are more spread out than for signal as expected.

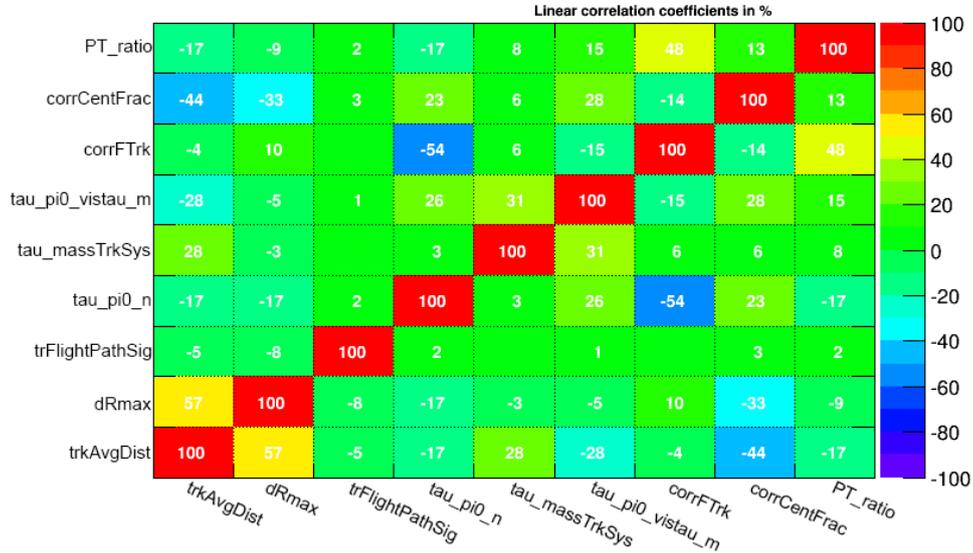


(a) 1 prong, background

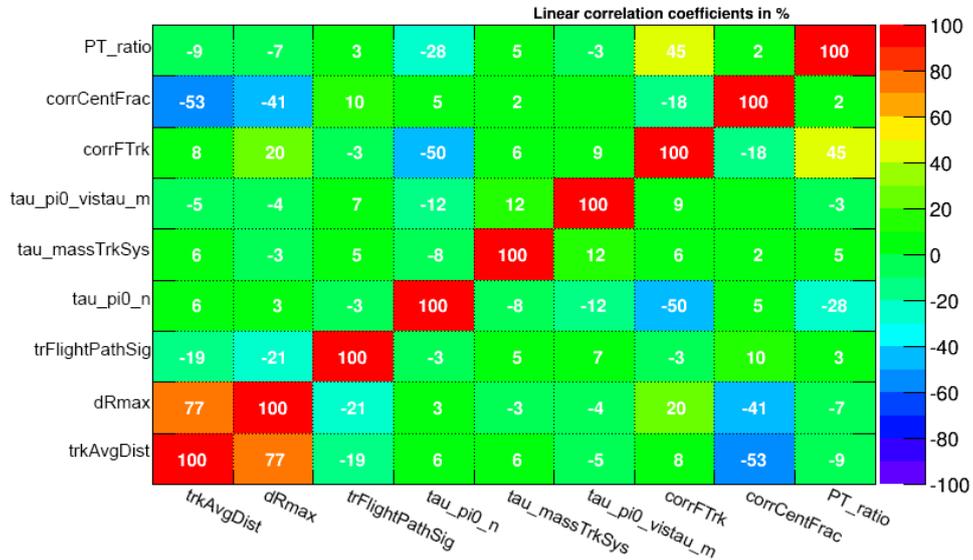


(b) 1 prong, signal

Figure 5.2.1: (a) and (b) show the linear correlation plots for 1 prong background and signal respectively.



(a) 3 prong, background



(b) 3 prong, signal

Figure 5.2.2: (a) and (b) show the linear correlation plots for 3 prong background and signal respectively.

1 prong correlation	Figure	Comments
corrFTrk vs PT_ratio	5.2.3(c) and 5.2.3(a)	
corrFTrk vs trkAvgDist	5.2.3(d) and 5.2.3(b)	
corrFTrk vs tau_pi0_n⁻⁻	-	Due to only 3 bins in the number of π^0 , no scatter plot is made.
trkAvgDist vs N_wideTrk⁻⁻	-	These variables are very likely to be correlated since the ΔR weighted sum of all tracks is calculated out to 0.4, and the number of tracks is found in an isolation annulus from $0.2 < \Delta R < 0.4$. Due to the limited number of bins for N_wideTrk no scatter plots are made.
3 prong correlation		
corrFTrk vs PT_ratio⁻⁻	5.2.4(a) and 5.2.4(b)	
dRmax vs trkAvgDist	5.2.4(c) and 5.2.4(d)	It is very likely for these two quantities to be correlated, since they both use ΔR . dRmax is the maximal distance between the $\tau_{had-vis}$ and the intermediate axis, whereas the latter variable is the sum over all tracks. Therefore the first variable is a part of the latter variable.
corrCentfrac vs trkAvgDist	5.2.5(a) and 5.2.5(b)	
corrCentfrac vs dRmax	5.2.5(c) and 5.2.5(d)	
corrFTrk vs tau_pi0_n	-	Due to only 3 bins in the number of π^0 , no scatter plot is made, but the reason for a correlation is that the higher the momentum fraction of the leading track, the less chance of finding a π^0

Table 5.2.1: List of the variables correlation investigated for variable pairs showing highest correlation. (⁻⁻ indicates that the given variable is later found to be superfluous)

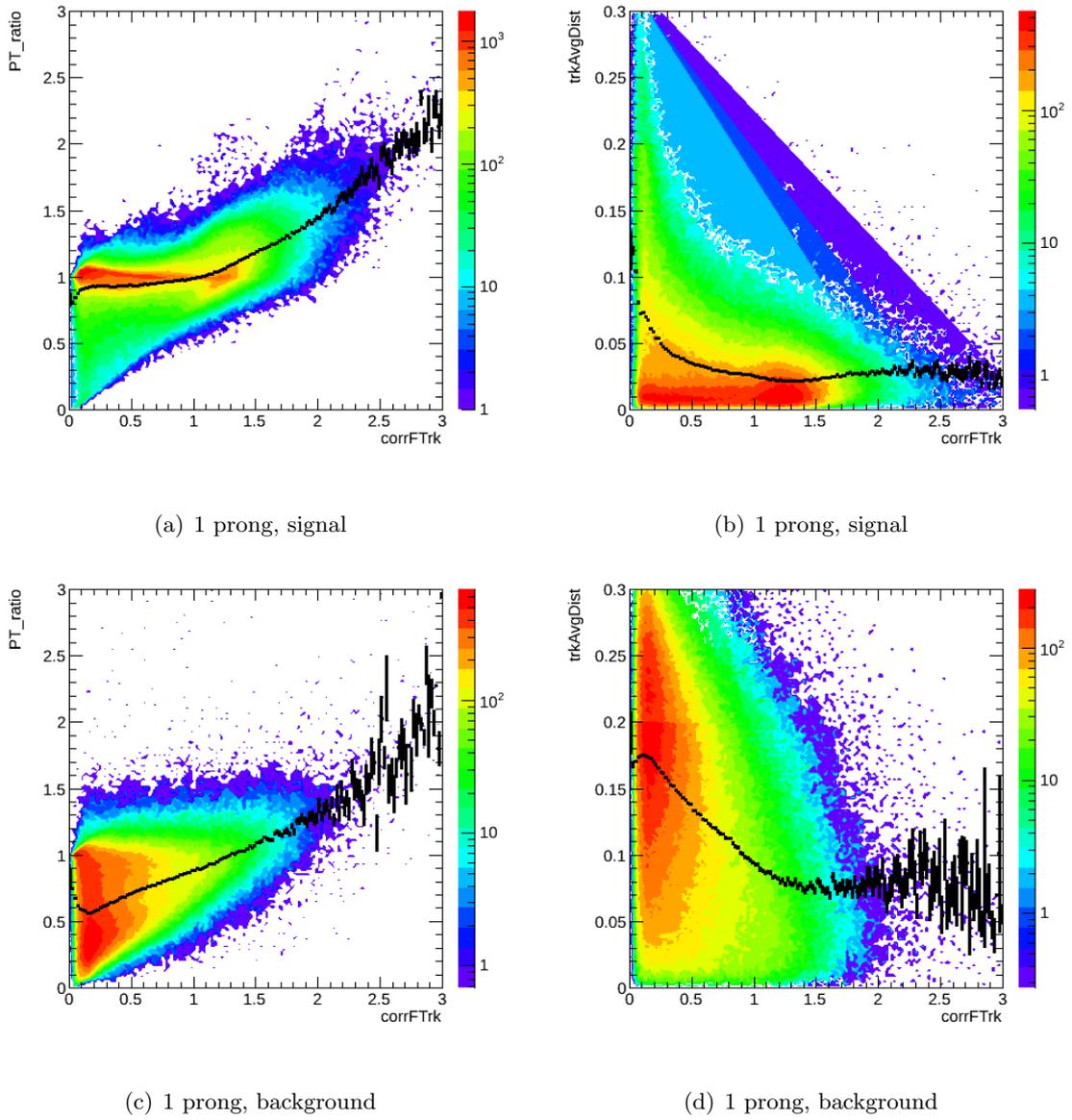


Figure 5.2.3: Scatter plots of the variables showing highest correlations from figure 5.2.1.

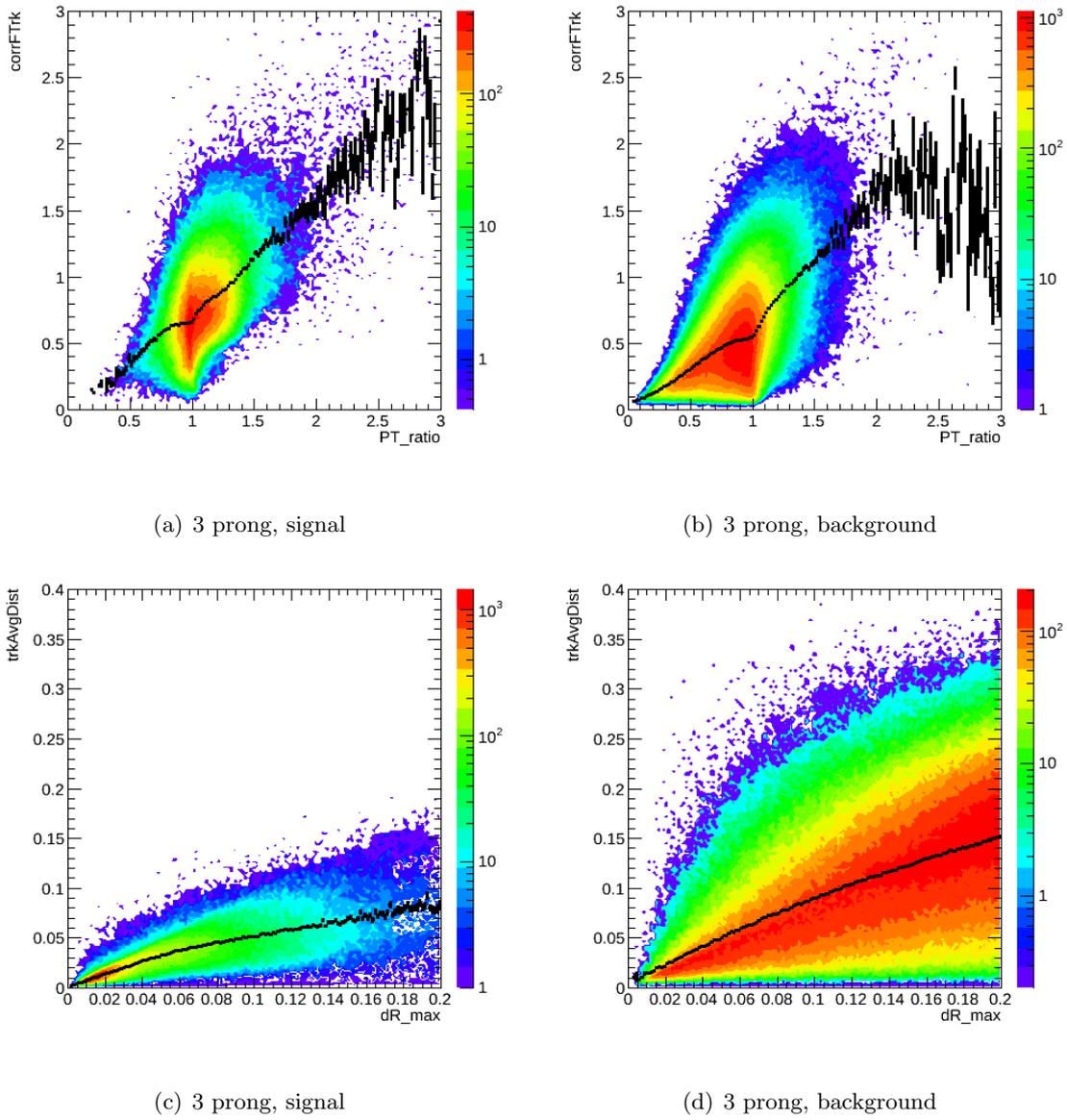


Figure 5.2.4: Scatter plots of the variables showing highest correlations from the figure 5.2.2.

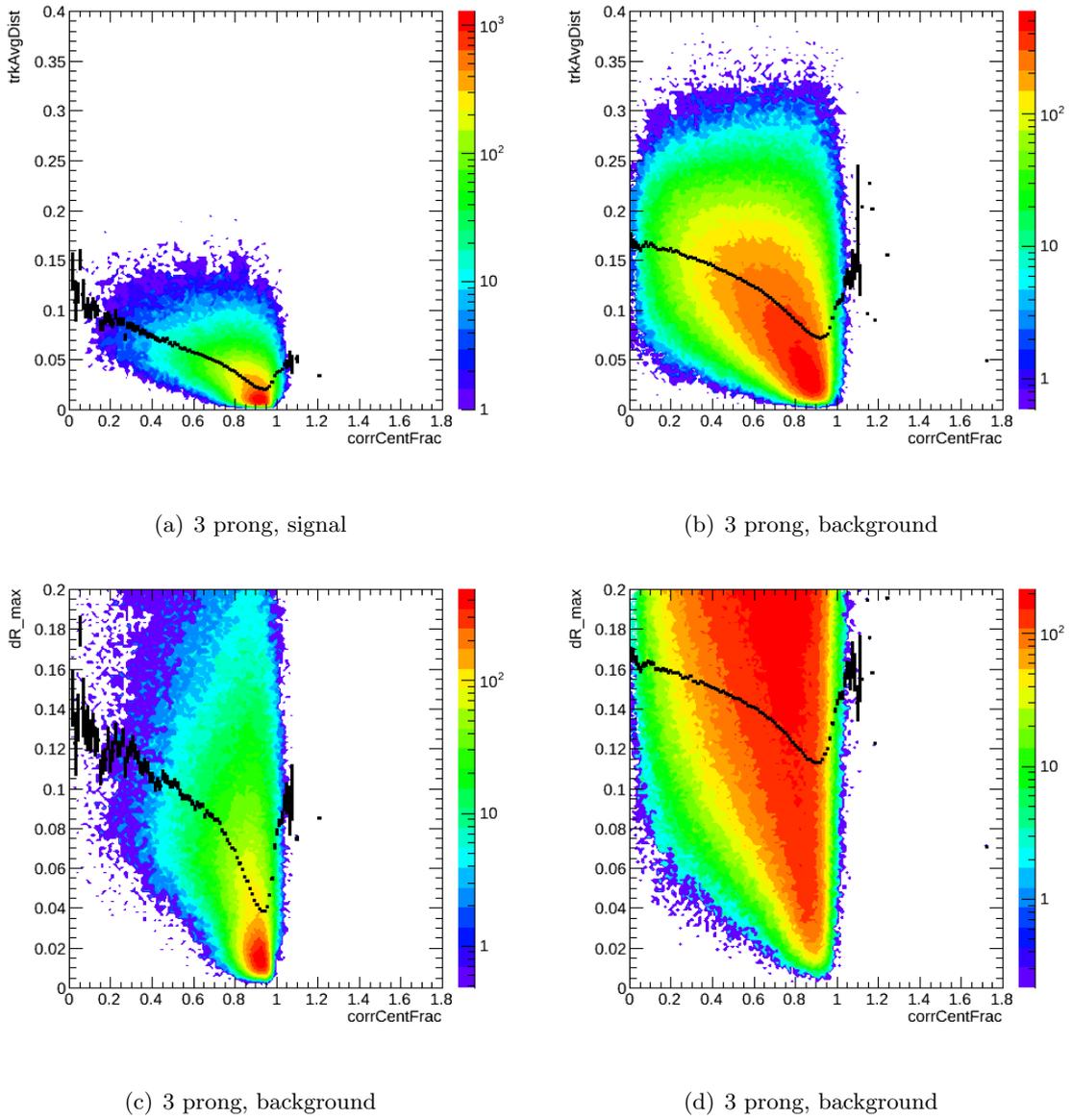


Figure 5.2.5: Scatter plots of the variables showing the highest correlations from figure 5.2.2.

5.2.1 Optimal set of tau identification variables

The goal is to find the optimal set of variables giving the best performance. By using the information from the correlation plots the smallest set of variables yielding optimal performance is sought.

Table 5.2.2 gives the variables used for training 1 prong and 3 prong respectively. The number in the tables correspond to a ranking number which are given by the BDT algorithm. It is the figure of merit and can be taken as a hint of the importance of a variable. In this section this number will be used as reference for the given variable, meaning eg. variable 1 for 1 prong is corrCentFrac .

1 prong	
Rank no.	Variable name
1	corrCentFrac
2	trkAvgDist
3	N_wideTrk
4	PT_ratio
5	corrFTRK
6	ipSigLeadTrk
7	tau_pi0_vistau_m
8	tau_pi0_n
3 prong	
Rank no.	Variable name
1	trkAvgDist
2	corrFTRK
3	PT_ratio
4	tau_massTrkSys
5	dRmax
6	trFlightPathSig
7	tau_pi0_vistau_m
8	tau_pi0_n
9	corrCentFrac

Table 5.2.2: The starting set of variables for one and three prong. The number is the ranking number determined by the BDT algorithm. The number will be used as reference for the variables.

The first two variables are used, since they have been ranked as the most important. The rest of the variable are then added one by one. This progression is shown for 1 prong in figure 5.2.6 and for 3 prong in figure 5.2.7. Figure 5.2.6(b) (5.2.7(b)) is a zoom of the figure 5.2.6(a) (5.2.7(a)).

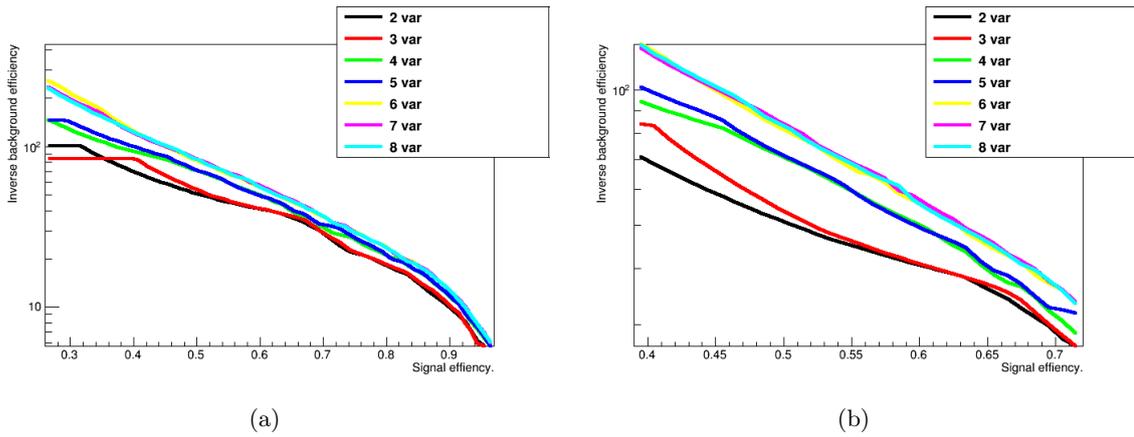


Figure 5.2.6: (a) shows the performance for 1 prong when adding variables one by one. (b) shows a zoom of (a)

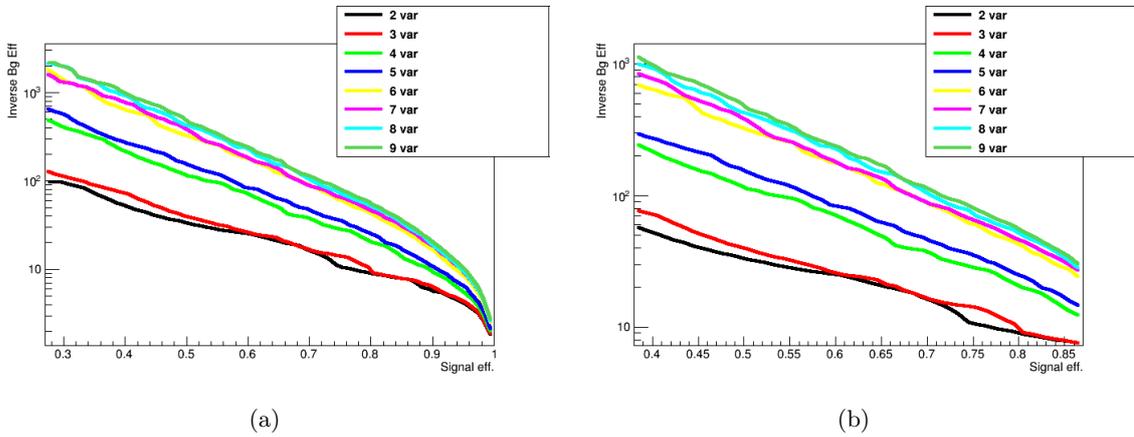


Figure 5.2.7: (a) shows the performance for 1 prong when adding variables one by one. (b) shows a zoom of (a)

- 1 prong

Figure 5.2.6 shows that the three variables `N_wideTrk`, `tau_pi0_n` and `corrFTrk` do not improve the performance. It was found in section 5.2 that `corrFTrk` was highly correlated with the two other variables, and therefore this is expected.

Figure 5.2.8 shows the performance for 1 prong when removing `N_wideTrk`, `tau_pi0_n` and `corrFTrk` one by one. The performance is independent of `tau_pi0_n` and `N_widetrk`, but drastically lowered when removing the variable `corrFTrk`. Therefore the first two mentioned are superfluous, but the latter needed.

Other combinations were also tried like removing `tau_pi0_vistau_m`, but not without any loss of performance. The optimal set of variables for 1 prong are listed in table 5.2.3.

- 3 prong

Figure 5.2.6 shows that the four variables PT_ratio, dRmax, tau_pi0_vistau_m, and corrCentFrac lead to the lowest increase in performance. Again this was expected, since a high correlation between PT_ratio and corrFTrk + dRmax and trkAvgDist was found. The mass of the track system is important for the performance, and maybe therefore tau_pi0_vistau_m has a smaller influence.

Figure 5.2.9 shows the importance when removing PT_ratio, dRmax, tau_pi0_vistau_m, and corrCentFrac one by one. There is no loss in performance by removing tau_pi0_vistau_m and PT_ratio, but the other variables lower the performance. It is therefore chosen to keep dRmax and corrCentFrac.

Other combination were tried, but no more variables were found to be expendable.

Table 5.2.3 gives the final sets of variables. For 1 prong the new π^0 variable PT_ratio and tau_pi0_vistau_m were found to be give an increase in performance. As expected from figure 3.6.1(a) tau_pi0_n did not have clear discrimination for the 1 prong case.

For 3 prong tau_pi0_n gave a significant improvement, and figure 3.6.1(b) shows this is a good discrimination variable.

1 prong	3 prong
corrCentFrac	trkAvgDist
trkAvgDist	corrFTRK
PT_ratio	tau_massTrkSys
corrFTRK	dRmax
ipSigLeadTrk	trFlightPathSig
tau_pi0_vistau_m	tau_pi0_n
	corrCentFrac

Table 5.2.3: The final set of variables for one prong and three prong in an arbitrary order

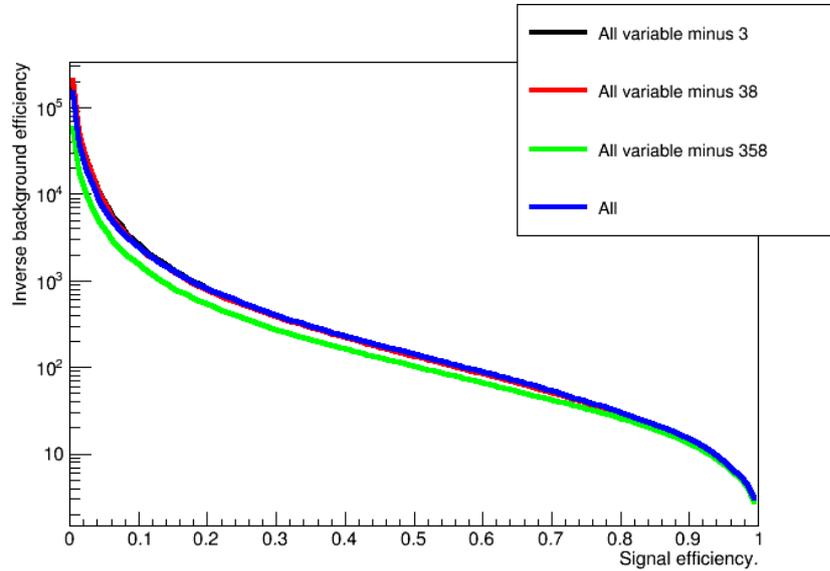


Figure 5.2.8: The performance curve for q prong when removing the variables that give the smallest increase in figure 5.2.6. The Black, blue and red line is exactly on top of each other

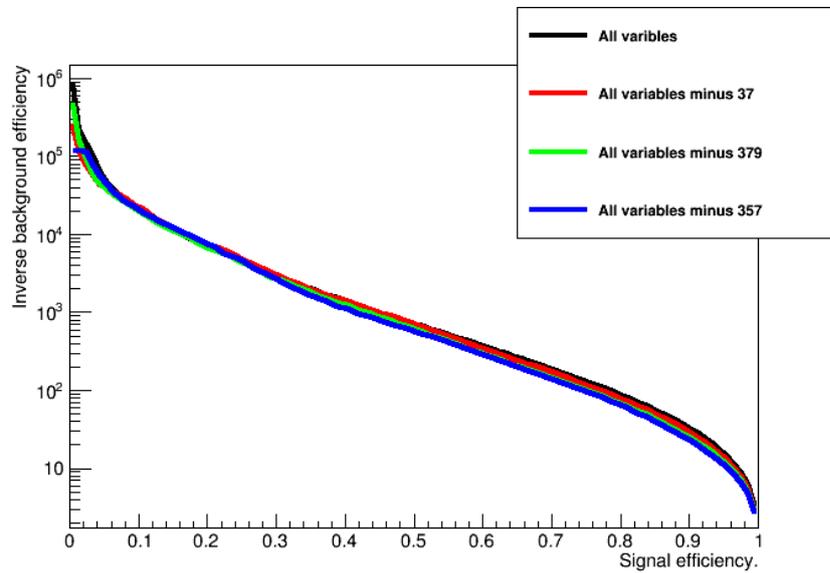


Figure 5.2.9: The performance curve for 3 prong when removing the variables that give the smallest increase in figure 5.2.6

5.3 Performance increase due to the new set of variables

Figure 5.3.1 shows the performance improvement arising from the new variables set BDT re-optimization compared to the "current" method. The number of variables used are for both 1 prong and 3 prong are increased by one, and the final set of variables is given in table 5.2.3.

For both one prong and three prong an improvement is gained as shown in figure 5.3.1. For 1-prong candidates working points are defined for signal efficiencies of 60%, 50% and 30% and for 3 prong candidates similar working points are defined at 65%, 55% and 35%. For both cases in figure 5.3.1 the performance gain is highest at the lowest working point, diminishing as the signal efficiency increases. The performance illustrated in figure 5.3.1 is also obtained using a more restricted BDT making it less prone to statistical fluctuations.

Due to this, and other studies the new π^0 variables were added in ATLAS to re-optimize tau identification for data collected in 2012, thereby improving the jet rejection as addressed in a tau performance paper currently under preparation [62]. The implementation of this new tau ID will also result in a higher signal yield for use in my study in the next sections.

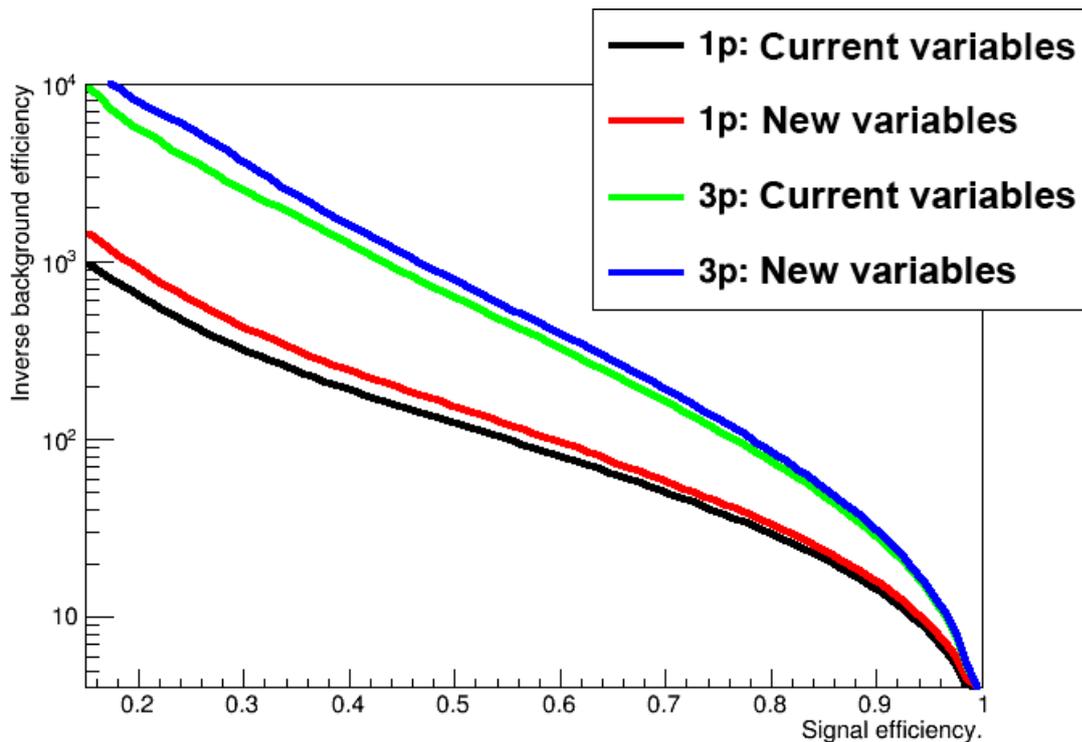


Figure 5.3.1: Performance improvement for tau identification for 1 prong and 3 prong with respect to current set of variables

Part II

A search for the SM Higgs produced in association with a vector boson

6

A search for the SM Higgs produced in association with a vector boson

6.1 Determining the search procedure

As described in chapter 1.2.1 the associated production has a low cross-section compared to other Higgs production channels like gluon gluon fusion (gg) and Vector Boson Fusion (VBF). An investigation of this channel has not previously been done for the ATLAS detector. The channel with a final state including two μ^\pm 's are investigated here, since the efficiency for μ^\pm identification is above 97% in the ATLAS detector [49]. This study provides insight into how useful the WH channel is for determining the Higgs boson properties. The following situation is investigated :

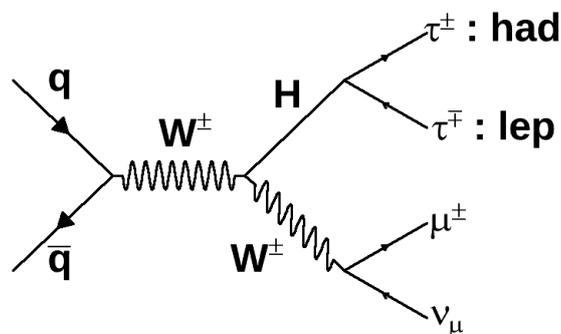


Figure 6.1.1: The Feynman diagram for the investigated channel

- $W \rightarrow \mu\nu, H \rightarrow \tau_\mu\tau_h$

The Feynman diagram of the process is shown in figure 6.1.1. In this channel improvements of $\tau_{had-vis}$ identification as described in chapter 3 can be used to improve the signal yield. This channel is likewise important for future LHC runs where higher pile-up occurrence making trigger efficiency on VBF and gg an increasing concern.

6.2 Background processes

Different processes can contribute as background in the search for the signal in figure 6.1.1. An overview of all expected background contributions is given below:

W+jets :

$W \rightarrow \mu\nu + jets$: This channel contributes to the background if one of the μ^\pm is from the decay of W and two additional jets give rise to the second μ^\pm and the $\tau_{had-vis}$. Figure 6.2.1 shows the tree-level Feynman diagram for this process. A likely scenario is that an incoming gluon splits into two quarks, where one quark contributes to the production of the W boson and the other quark gives origin to a jet. The initial state gluon can emit another gluon which produces another jet. This scenario is likely at the LHC, since it is a proton-proton collider, and as described in section 2.2 the gluon is the most likely parton to scatter.

The branching ratio for the W to decay to μ and a ν_μ is $\approx 10\%$.

$W \rightarrow \tau\nu + jets$: This situation is similar to $W \rightarrow \mu\nu + jets$. Except that the τ now arises from the W and contributes either by decaying leptonically to a μ^\pm (most likely) or decaying hadronically. The ISR/FSR then needs to give rise to two additional leptons.

The branching ratios of the W to decay to a τ and ν_τ is $\approx 10\%$.

$W \rightarrow e\nu + jets$: This process has none of the objects in the wanted final state, and this process is expected to be negligible as background. It is taken into consideration for completeness. The branching ratio for the W to decay to a e^\pm and a ν_e is $\approx 10\%$. It should be noticed that 60% of the time W decays into two quarks, and this is not taken into consideration in simulations, but will be accounted for by the data-driven estimation later.

Z+jets :

$Z \rightarrow \tau\bar{\tau} + jets$: This channel is shown at the tree-level in figure 6.2.2. The Z decay into two τ 's is most likely to contribute if one of the τ 's decay hadronically and one leptonically to a μ^\pm , although there needs to be another μ^\pm , coming from radiation for example.

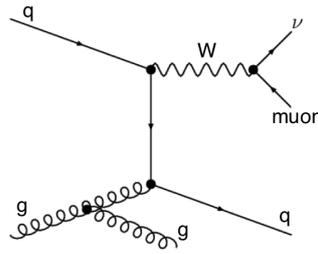


Figure 6.2.1: A tree-level Feynman diagram for $W \rightarrow l\nu + jets$ [63].

$Z \rightarrow \mu\mu + jets$: The process of Z decaying into two μ^\pm 's is expected to be a major part of the background, since it gives the two μ^\pm and then a $\tau_{had-vis}$ can be faked by a jet. If there is a same-sign (SS) requirement for the charges of the two μ^\pm 's the process is expected to be far less likely. This requires a charge misidentification which is very rare

$Z \rightarrow ee + jets$: This process is not expected to contribute, but is taken into consideration for completeness. The situation where the Z decay to two quarks are not taken into consideration in simulations, but will be accounted for by the data-driven estimation later.

Drell Yan (DY): Drell Yan is the process where two quarks annihilate to produce a virtual photon or Z , which then decays to two oppositely charged leptons. The contribution from this process depends on the decay of the Z , and is like $Z \rightarrow ll + jets$

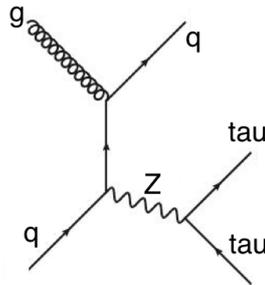


Figure 6.2.2: A tree-level Feynman diagram of the production of Z

Diboson :

WZ: The associated production of the decay of a W and Z gives rise to a signature most alike the one of interest and is shown in figure 6.2.3. The signature of 2 μ^\pm 's and 1 $\tau_{had-vis}$ is obtained if the W decays to a τ and the Z decays to two μ^\pm . The scenario changes if the two μ^\pm are required to have same-sign charge: One of the W can then decay to a μ^\pm and the Z to two τ 's, where one of the τ 's decays hadronically and the other τ decays leptonically to a μ^\pm

- ZZ:** In the case of the associated production of two Z bosons one can decay to two μ^\pm 's and one can decay to two τ 's. The desired signature is obtained if one τ decays hadronically and one is not detected. Again the scenario changes if there is a charge requirement on the two μ^\pm 's: Then the most likely is that at least one Z decays to two τ , where one τ decays to a μ^\pm and the other hadronically, if one object is not detected there is the signature of two SS μ^\pm and one hadronic τ .
- WW:** This process can contribute if one W decays to a $\mu^\pm + \nu$ and the other decay to two quark or to $\tau_{had-vis} + \nu$.

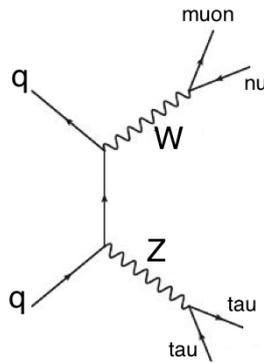


Figure 6.2.3: A tree-level Feynman diagram of the production of WZ

Top :

Single top : A single top is likely to be produced via one of the three channels: s-channel, t-channel or a Wt channel. All three diagrams are shown in figure 6.2.4. The t-channel has a much higher cross-section than the two others and is the dominant production mode. The top will decay to a W and a b quark with probability of about unity, and the decay of W can give rise to a μ^\pm or a τ .

ttbar The production of ttbar is shown in figure 6.2.5. The two tops will decay to two W giving rise to two leptons which can be a μ^\pm or $\tau_{had-vis}$, and for this process to end up as a background the lacking objects will come from one of the b-quarks in figure 6.2.5 .

non-EW background: For this contribution all three leptons are fake, but due to a large cross-section for this process this background can still contribute to the signal region.

6.3 Data

The data used in the analysis was collected in the ATLAS detector in 2012 at a center of mass energy of 8 TeV. The data is subdivided in blocks of length about two minutes. These blocks are called luminosity blocks. All sub-detectors and triggers are required to work optimally for

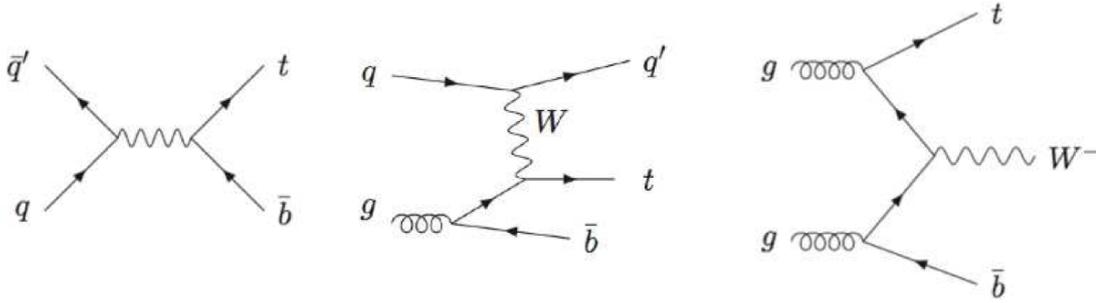


Figure 6.2.4: The single t produced via the s -channel (left), the t -channel (middle) and the Wt channel (right).

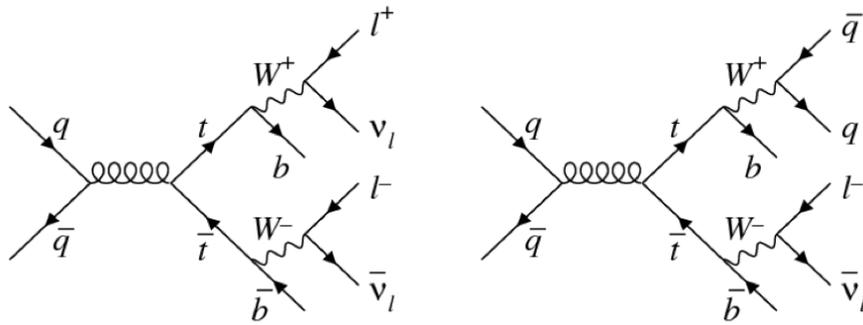


Figure 6.2.5: Feynman diagrams showing the two possible final states for the $t\bar{t}$ decays contributing to this search

a luminosity block to be classified as good. There is a list of the run numbers and luminosity block numbers for good runs where the whole detector is working optimally. This is called the Good Run List, denoted as *GRL*. Only data from the *GRL* are accepted for the analysis, and in this case the selected data has a total integrated luminosity of 20.3 fb^{-1}

6.4 Monte-Carlo simulations

Monte-carlo simulation is used to investigate which processes contribute to the background for the investigated channel. Events are simulated from different generators listed in table 6.4.1. The number of generated events is a crucial parameter in attaining a background estimation by simulation with low uncertainty. Therefore the number of generated events and the cross-section for each process is included in table 6.4.1.¹ In section 2.2.1 the different generators were described, underpinning the choice of generator for a given process. For example McAtNlo is used for top, since the NLO t -channel diagram in figure 6.2.4 is important for the top production.

Many of the background contributions are generated with filters, enhancing some parts of the process over others. This is done to ensure good statistics for the important parts of the

¹The probability of scattering interactions is expressed with the cross-section. The unit is barn (b), corresponding to 10^{-28} m^2 .

Process	Generator, Parton shower	PDF	N_{events}	Cross section [pb]
WH ($m_h = 115 GeV$)	Pythia8 (P)	CTEQ6L1	$3 \cdot 10^4$	0.0316
WH ($m_h = 120 GeV$)	Pythia8 (P)	CTEQ6L1	$3 \cdot 10^4$	0.0256
WH ($m_h = 125 GeV$)	Pythia8 (P)	CTEQ6L1	$5 \cdot 10^5$	0.0201
WH ($m_h = 130 GeV$)	Pythia8 (P)	CTEQ6L1	$3 \cdot 10^4$	0.0151
WH ($m_h = 135 GeV$)	Pythia8 (P)	CTEQ6L1	$3 \cdot 10^4$	0.0110
$W \rightarrow e\nu + jets$	Alpgen	ct10	$12 \cdot 10^6$	12264
$W \rightarrow \tau\nu + jets$	Alpgen	ct10	$12 \cdot 10^6$	12264
$W \rightarrow \mu\nu + jets$	Sherpa	ct10	$36 \cdot 10^6$	10963
$Z \rightarrow ll + jets$	Alpgen+Pythia (T) (P)	CTEQ6L1	$25 \cdot 10^6$	3450
$Z \rightarrow mumu$	Sherpa	ct10	$12 \cdot 10^6$	1106
Drell yan	Alpgen+Jimmy+Herwig (T) (P)	CTEQ6L1	$6 \cdot 10^3$	4347
single top	McAtNlo+Herwig (T) (P)	ct10	$11 \cdot 10^3$	50.9
ttbar	McAtNlo+Herwig (T) (P)	ct10	$15 \cdot 10^6$	130
WZ	Herwig (T) (P)	CTEQ6L1	$2 \cdot 10^6$	22.3
ZZ	Herwig (T) (P)	CTEQ6L1	$24 \cdot 10^4$	6.33
WW	Alpgen+Herwig (T) (P)	ct10	$23 \cdot 10^5$	5.65

Table 6.4.1: The generators used for each background estimation and the number of events generated. A (T) means that also Tauola is used and a (P) that Photos is used

processes. For $W \rightarrow l\nu + jets$ a combination of generators is used. For the $W \rightarrow \mu\nu$ Sherpa samples are used which have higher statistics for the heavy flavor quarks. For $W \rightarrow \tau\nu$ and $W \rightarrow e\nu$ Alpgen samples are applied which filter according to the number of jets. This is done since $W \rightarrow \tau\nu$ and $W \rightarrow e\nu$ are expected to be less important allowing these smaller samples to be used. For $Z \rightarrow ll$ two kinds of generators are used, either the Alpgen samples which filter on the number of jets or the Sherpa samples which filter according to the heavy quarks. If the Sherpa samples are used the notation $Sherpa(b)$ is used for the filtering on the b-quark, $Sherpa(c)$ for the filtering on the the c-quark and $Sherpa(-b/c)$ for the remainder. In the following sections are used the WH sample for $m_H = 125$ GeV for signal.

6.5 Defining the basic objects for the search

The objects: μ^\pm , e^\pm , $\tau_{had-vis}$ and jets are defined by the following

μ^\pm Selection: For a μ^\pm candidate to be selected it must have $P_T^\mu > 8 GeV$ and $|\eta^\mu| < 2.5$.

Furthermore the track from the inner detector is required to pass a set of quality criteria:

b-layer: A μ^\pm candidate must be detected in the b-layer, which is the innermost layer of the pixel detector.

$$N_{b-layer}^{Hits} > 0$$

Furthermore no so-called expected b-layer hits are allowed. An expected hit is where a b-layer hit is expected from the track, but not detected,

Pixel hits: The μ^\pm candidate must be detected in the pixel detector. Expected hits from a dead pixel sensor is also considered, $N_{Pixel}^{DeadHits}$:

$$N_{Pixel}^{Hits} + N_{Pixel}^{DeadHits} > 0$$

SCT hits: The number of hits in the SCT detector must be above four, again expected hits in dead SCT sensors are included, $N_{SCT}^{DeadHits}$:

$$N_{SCT}^{Hits} + N_{SCT}^{DeadHits} > 4$$

Holes in the pixel and SCT detectors A maximum of 3 holes are allowed in the SCT and Pixel detectors combined. A hole is a passed functional sensor which is expected to give a hit from the track, but the hit is not present: ²

$$N_{Pixel}^{Holes} + N_{SCT}^{Holes} < 3$$

TRT hits The TRT criteria depends on the η region:

0.1 < $|\eta|$ < 1.9 :

There must be more than 5 hits including outliers. Outliers must not exceed 90% of the total number of hits, where and outlier is more that 2.5σ from the hit.

$$N_{TRT}^{Hits} + N_{TRT}^{Outliers} > 5$$

$|\eta| < 0.1$ or $|\eta| > 1.9 :$

No restriction on the number of hits, but for more than 5 hits,

$$N_{TRT}^{Hits} + N_{TRT}^{Outliers} > 5$$

only 90 % of the hits can be outliers

The μ^\pm track is found independently in the μ^\pm spectrometer and the inner detector and then combined afterwards. Thereby making it possible to utilize the sensitivity of the inner detector at low P_T^μ and the muon spectrometer at high P_T^μ .

If a track fulfills the above criteria and pass the combined μ^\pm requirements it is identified as a μ^\pm . The identification method has a good performance in a broad range of transverse momentum.

Finally the μ^\pm candidate is required to be isolated:

Isolation in the calorimeter The transverse calorimeter energy in a ring of radius 0.2 not including the μ^\pm must be less than 20% of the P_T^μ

$$\frac{etcone20}{P_T} < 0.2$$

²The difference between an expected hit and a hole, is that for a hole the sensor is functional and it not for an expected hit.

Isolation in the inner detectors The total transverse track-based momentum in a ring of radius 0.4 not including the μ^\pm must be less than 20% of the P_T^μ

$$\frac{ptcone40}{P_T} < 0.2$$

The size of the cones are chosen to be consistent with other ATLAS searches for the Higgs boson in the VBF and gluon-gluon production modes.

e^\pm **Selection:** e^\pm candidates are found by reconstruction of a cluster in the electromagnetic calorimeter. e^\pm candidates are then linked to a track in the inner detector. Selected e^\pm candidates must have $P_T > 8$ GeV

Identification e^\pm is not a part of the final state and a veto is used on e^\pm , since none of these are expected in the signal. A loose identification is used for e^\pm in this thesis, since the idea is that there should be no other objects in the event. In the crack region a medium identification is used, since there is less calorimeter information here demanding a tighter selection to prevent electrons being faked by jets. For the tracking the following criteria are enforced :

Pixel hits: The e^\pm candidate must be detected in the pixel detector:

$$N_{Pixels}^{Hits} > 0$$

SCT hits: The e^\pm candidate must have a good track in the SCT detector:

$$N_{SCT}^{Hits} > 7$$

The cluster and the extrapolated track must not have a $|\Delta\eta| < 0.015$

No problems with the cluster ATLAS data quality group provides information about cluster problems for e^\pm . No cluster problems are allowed for the selected clusters.

Acceptance region $|\eta|$ needs to be less than 1.37 or in the interval $1.52 < |\eta| < 2.47$. In the gap region medium identification is used for e^\pm .

$\tau_{had-vis}$ **Selection:** The $\tau_{had-vis}$ needs to have a $P_T^\tau > 20$ GeV and $\eta^\tau < 2.47$. The latter is required since this is the coverage of the inner detector used for the e^\pm veto.

The identification and reconstruction of $\tau_{had-vis}$ is described in section 5. For $\tau_{had-vis}$ identification a working point corresponding to a signal efficiency of 50% (55%) for 1 (3) prong $\tau_{had-vis}$ is chosen. In section 5 the identification efficiencies are found, these need to be combined with the reconstruction efficiency to get the total efficiency

Furthermore the $\tau_{had-vis}$ must pass a loose e^\pm veto. This is done since $\tau_{had-vis}$ have a signature very similar to e^\pm . The e^\pm veto is implemented with a BDT algorithm, where the signal events are $Z \rightarrow \tau\tau$ and the background events are $Z \rightarrow ee$. A loose veto corresponds to a signal efficiency of 75% [58].

6.5.1 Overlap removal

The object selection contain an overlap removal, which prevents any particle to be in more than one category. μ^\pm are selected first, secondly e^\pm is selected if they do not overlap with the μ^\pm in a 0.2 cone.³ $\tau_{had-vis}$ are required to neither overlap with μ^\pm or e^\pm in a 0.2 cone. Jets are only selected if they do not overlap with any of the other objects in a 0.2 cone.

6.6 Trigger selection

The final state in the investigated channel is composed of 2 μ^\pm and 1 $\tau_{had-vis}$, these are the objects that can fire a trigger. This makes the following triggers possible: a single muon trigger, a di-muon trigger, and a combined tau plus muon trigger.

In this analysis the di-muon triggers (*EF_2mu13*, *EF_mu8_mu18_EFFS*) and single muon triggers (*EF_mu24i_tight*, *EF_mu36_tight*) are used. The *EF_mu36_tight* is used on top of *EF_mu24i_tight* to account for an inefficiency of the latter at high muon momentum due to the isolation requirement.

The motivation for choosing these triggers is given in table 6.6.1. Here the signal yield at truth level is given. The table shows that the inclusion of the di-muon trigger gives 9% more signal yield than the single trigger alone, and since the overall signal yield is expected to be very limited, the di-muon trigger is used on top of the single trigger.

Trigger	Signal Higgs (125 GeV)
EF_mu24i_tight OR EF_mu36_tight + EF_2mu13 OR EF_mu8_mu18_EFFS	100 % (2.56)
EF_mu24i_tight OR EF_mu36_tight	91.0 % (2.36)
EF_2mu13 OR EF_mu8_mu18_EFFS	8.98% (0.23)

Table 6.6.1: The number of events passing the trigger in signal. The basis offline selection is 2 μ^\pm , 0 e^\pm and 1 $\tau_{had-vis}$

It has been investigated if a tau plus muon trigger produces higher signal yield. Only a 2% gain is obtained in the signal yield, and therefore the $\tau_{had-vis}$ plus muon trigger is not used since this trigger comes with an additional uncertainty.

The used trigger are

EF_mu24i_tight is a single muon trigger requiring at least one isolated muon with $P_T^\mu > 24$ GeV where isolation criterion is made with inner detector tracks.

EF_mu36_tight is a single muon trigger requiring at least one muon with $P_T^\mu > 36$ GeV without applying isolation cut.

EF_2mu13 requires two or more muon candidates, each of which passes a single muon trigger with $P_T^\mu > 13$ GeV

³The event is vetoed if any electrons are present

EF_mu18_mu8.EFFS requires at least one muon candidate which passes a single muon trigger with $P_T^\mu > 18$ GeV, and subsequently employs the full scan algorithm to find two or more muon candidates with $P_T^\mu > 18$ and 8 GeV for leading and sub-leading muons

6.6.1 Trigger matching

The offline threshold for P_T^μ needs to be 2 GeV above the trigger P_T threshold. Here the plateau of the trigger efficiency curve is reached, the optimal efficiency is achieved and the uncertainty of the trigger is better known.

Once it is determined which trigger is fired, and a match can be made to a given offline μ^\pm . This μ^\pm needs to have a P_T^μ 2 GeV above the trigger P_T

6.7 Association of the leptons to the Higgs Boson

It is important to decide the association of the leptons to the Higgs Boson decay products. The μ^\pm from W is expected to have a higher P_T^μ since it comes directly from W with the only decay partner being a neutrino. It may therefore be assumed that the leading μ^\pm is from W, and the sub-leading μ^\pm is from the Higgs Boson.

This assumption has been tested in simulation and found to be correct in 75% of the time.

6.8 Optimal event selection

Notice: Since the analysis is ongoing, and still blinded internally in ATLAS, the stacked histograms in this section are done with a cut on a tau BDT score above 0.3 instead of the medium tau identification requirement. The signal over background plots are done entirely with simulation, and thereby the medium requirement is applied

A series of cuts are applied to achieve the highest ratio of signal events to background events. The plots in figure 6.8.1 show the kinematics of the objects with no charge requirements applied to the two μ^\pm and the $\tau_{had-vis}$. The plots show a good agreement between data and simulation.

Based on the plots in figure 6.8.1 the following can be considered :

Charge requirement: The plots in figure 6.8.1 show that $Z \rightarrow \mu\mu$ is the dominant background contribution. A charge requirement is applied to eliminate these background events by requiring the two μ^\pm to be of same sign, SS . This will also eliminate half of the signal events, but the signal over background is simply too low for two opposite sign μ^\pm events. The two SS μ^\pm will have opposite charge to the $\tau_{had-vis}$. Therefore this charge requirement is applied in the signal region, denoted as OSS . Table 6.8.2 show that by restricting the charges reduce the contribution from $Z \rightarrow \mu\mu + jets$ by more than 99%. Figure 6.8.2 shows the MC simulation and data comparison at this stage. There is an overshoot of data due to non-EW background or mis modelled EW backgrounds which becomes more visible now.

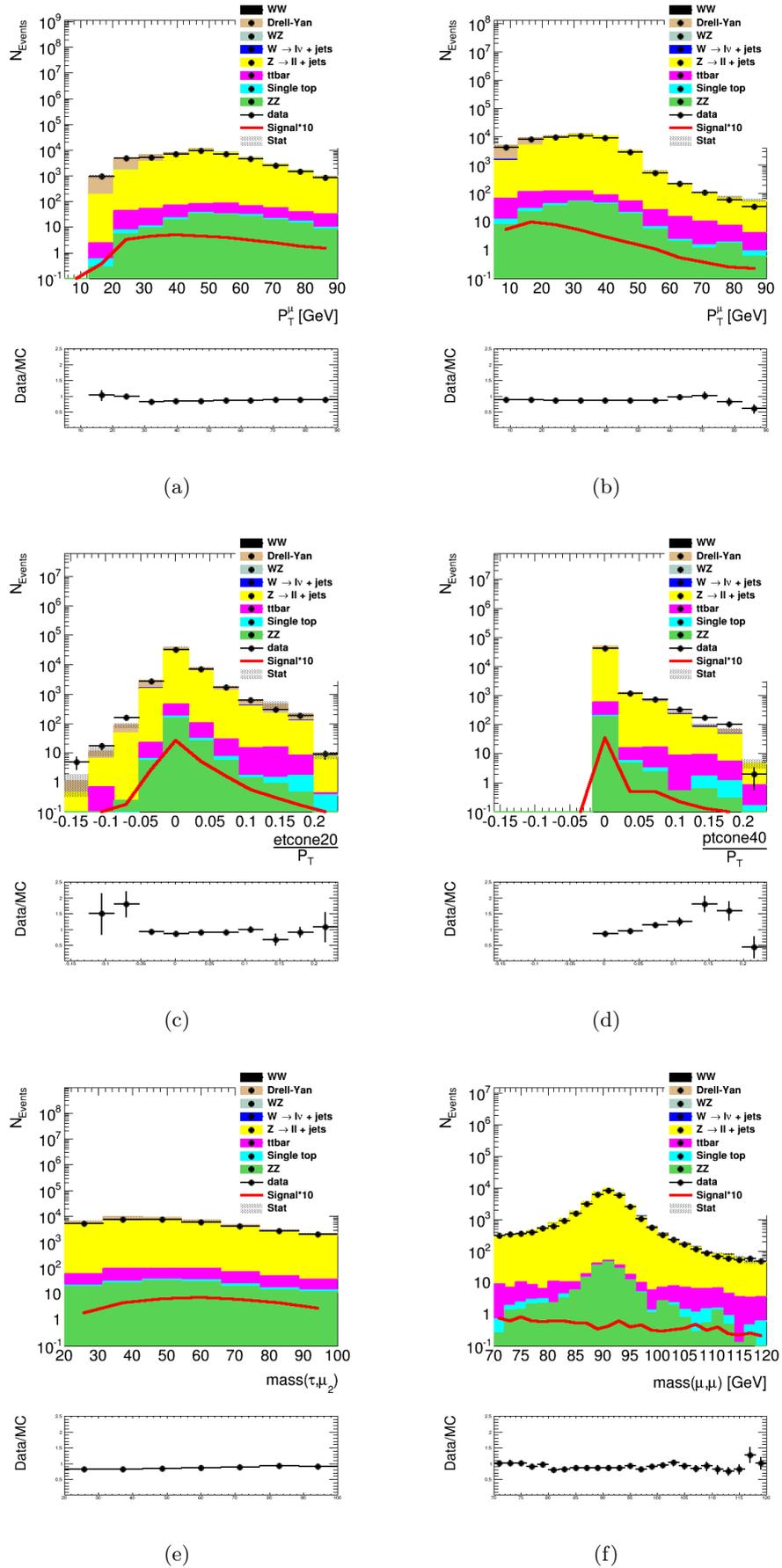


Figure 6.8.1: P_T^μ for the leading μ^\pm (a) and subleading μ^\pm (b), the isolation for the subleading μ^\pm (c),(d), the visible mass of the subleading μ^\pm and the $\tau_{had-vis}$ (e) and the mass of the two μ^\pm 's (f)

Stronger μ^\pm selection: Figure 6.8.3 shows the $\frac{S}{\sqrt{B}}$ dependence on the cut on the isolation energy for the sub-leading and leading μ^\pm , which are shown in figure 6.8.2(a) and figure 6.8.2(b). The x-axis on the plot in figure 6.8.3 gives the maximum isolation value cut, and the y-axis $\frac{S}{\sqrt{B}}$ where S is the number of signal events and B is the number of background events. In these plots the μ^\pm are required to have the same charge. For background the simulated events from table 6.4.1 are used.

Calorimeter isolation in figure 6.8.3(a) and figure 6.8.3(c) The highest $\frac{S}{\sqrt{B}}$ is obtained at $etcone20/pt \approx 0.05$ for the leading μ^\pm and $etcone20/pt \approx 0.08$ for sub-leading μ^\pm .

Tracking isolation in figure 6.8.3(b) and figure 6.8.3(d) The isolation in the inner detectors is calculated in a ring of 0.4. Below 0.08 the maximum value is obtained, but the variation is smaller for the tracking isolation.

A cut of 0.08 is chosen for both tracking and calorimeter isolation and both the leading and the sub-leading μ^\pm . Through this choice more than 97% of the signal is maintained and a higher signal over background ratio is obtained.

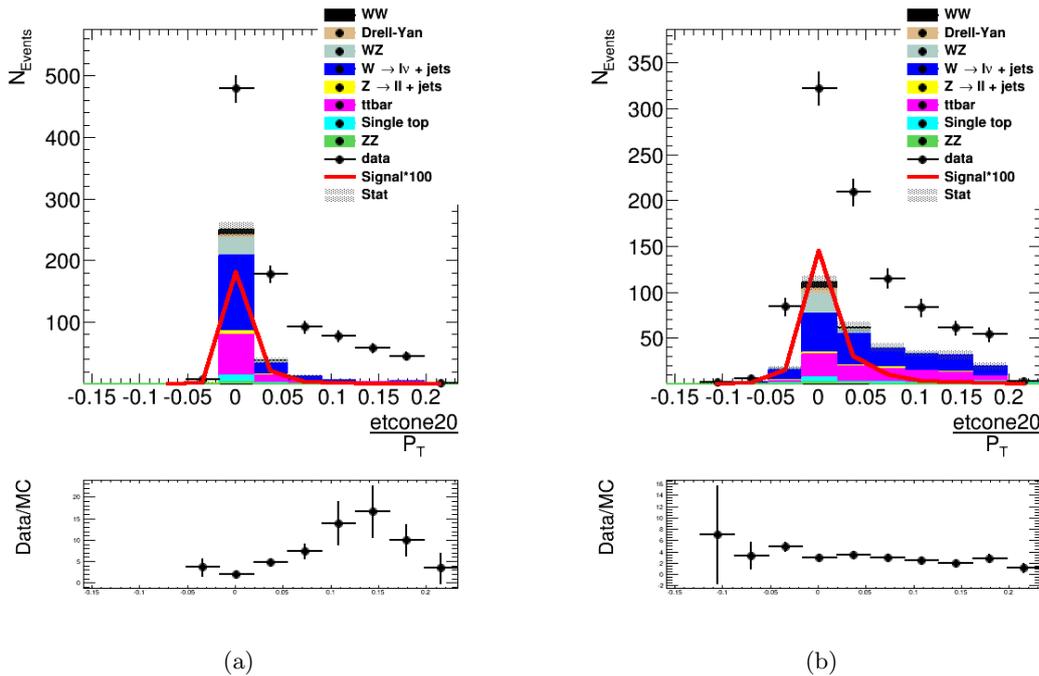


Figure 6.8.2: The isolation in the calorimeter for the leading μ^\pm (a) and sub-leading μ^\pm (b).

Transverse impact parameter The z_0 and $\frac{d_0}{\sigma(d_0)}$ are shown in figure 6.8.4(c) and 6.8.4(d). The $\frac{d_0}{\sigma(d_0)}$ is found not to be a good cut, since it removes too much signal. The longitudinal impact parameter is required to fulfill $|z_0| < 0.4$ mm, to eliminate some of the non-EW background events. The values given in table 6.8.2 show that this cut eliminates the remaining contribution from $Z \rightarrow \mu\mu + jets$ to the background, while also significantly reducing the $W \rightarrow l\nu + jets$.

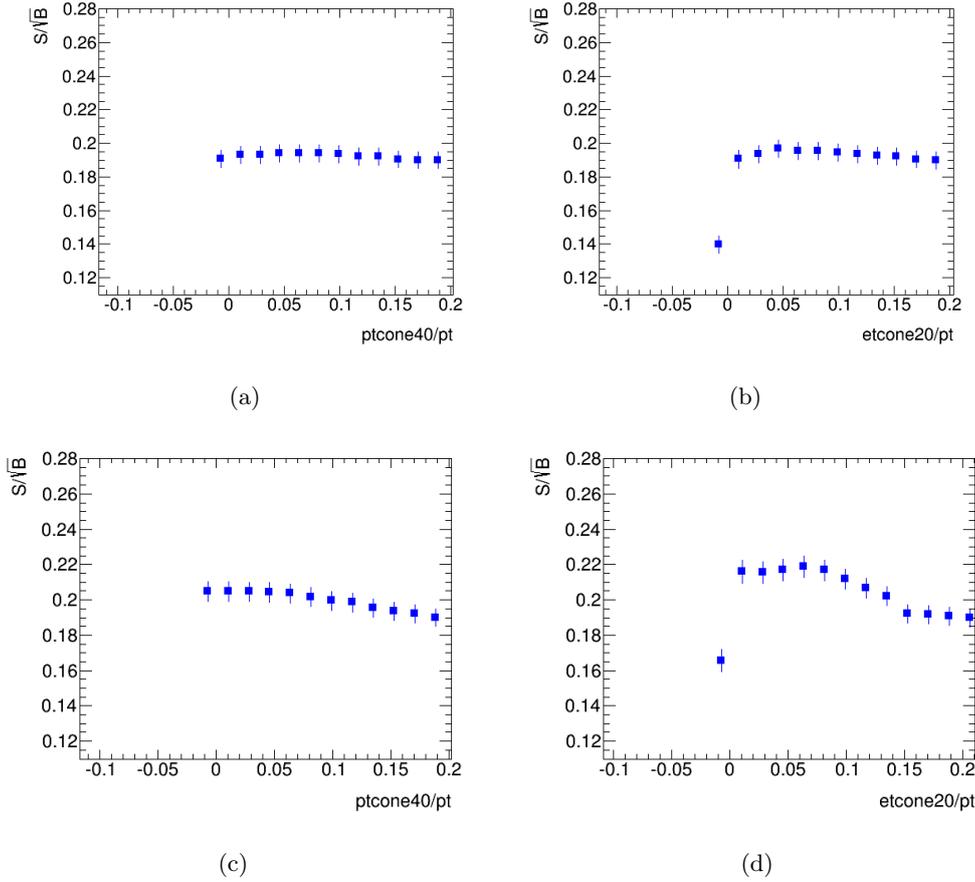


Figure 6.8.3: The $\frac{S}{\sqrt{B}}$ as a function of a given cut on the isolation variables for the leading μ^\pm ((a), (b)), and for the sub-leading μ^\pm ((c), (d)). The value on the x axis is the upper cut applied.

The isolation and impact parameter requirements are implemented and the comparison between MC simulation and data at this stage is shown in figure 6.8.4

Transverse momentum of the μ^\pm : The plots in figure 6.8.5(a) and 6.8.5(b) show $\frac{S}{\sqrt{B}}$ as a function of the lower limit on P_T^μ . The stack histogram are shown in figure 6.8.4(a) and 6.8.4(b). A small gain is achieved by increasing the cut value for the transverse momentum on the sub-leading μ^\pm , but to maintain signal yield this value is kept at 8 GeV.

The plots in figure 6.8.4 show that a stronger requirement of the P_T^μ remove some non-EW background events. These events are not in the simulated background, and therefore the variation in $\frac{S}{\sqrt{B}}$ is expected to be higher than shown here.

Missing energy: One of the decay products of the W^\pm boson and one of the decay products of the $\tau_{had-vis}$ causes energy to be lost without detection. Therefore missing energy could be useful for identifying the signal. The ratio $\frac{S}{\sqrt{B}}$ is investigated as a function of the lower cut on \cancel{E}_T in figure 6.8.6(a) and the distribution is in figure 6.8.7(a). A cut on \cancel{E}_T is found not to produce any gain in $\frac{S}{\sqrt{B}}$ and is therefore not used.

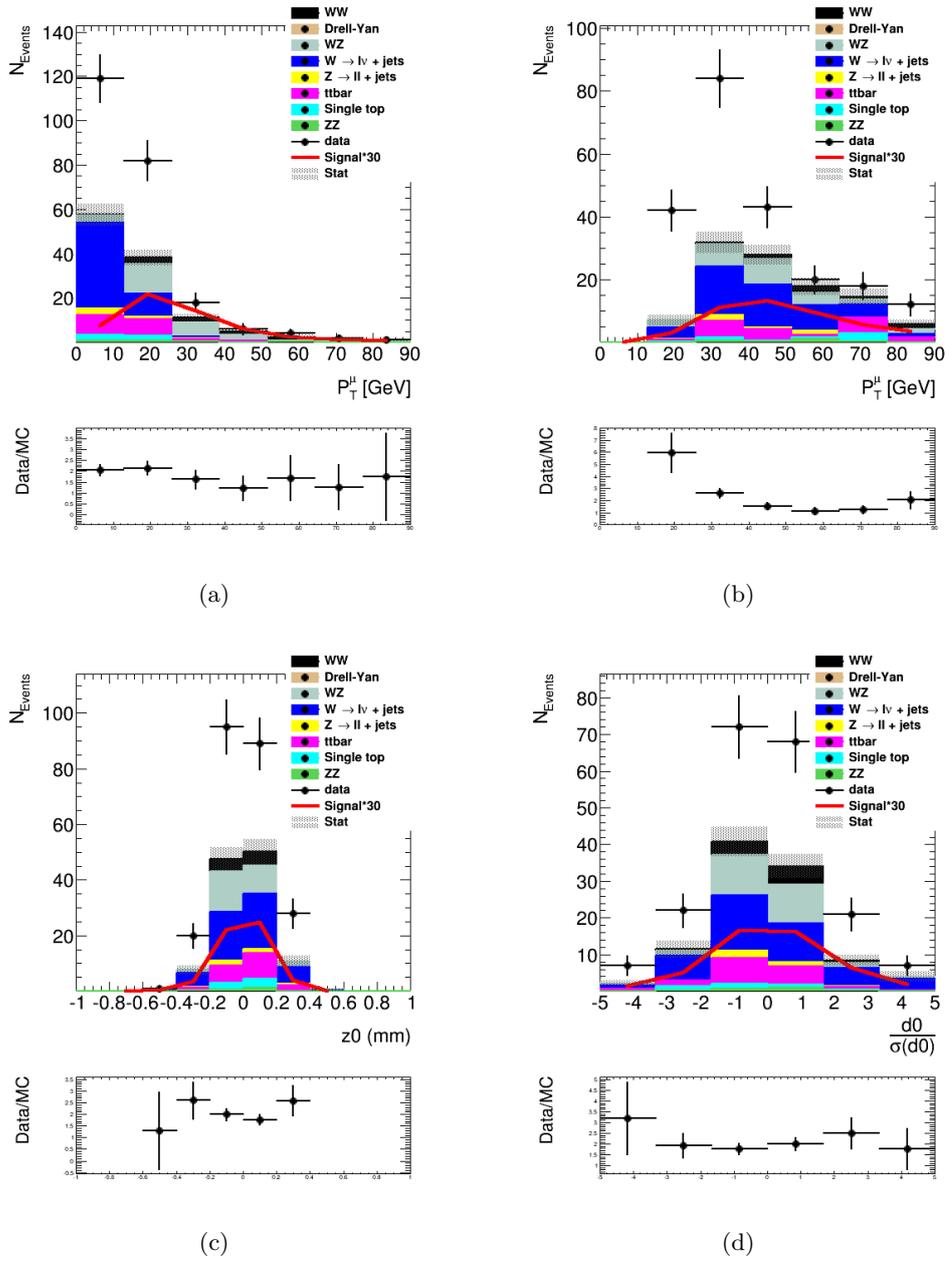


Figure 6.8.4: P_T^μ for the leading μ^\pm (a) and sub-leading μ^\pm (b). z_0 for the sub-leading μ^\pm (c) and $\frac{d_0}{\sigma(d_0)}$ for the sub-leading μ^\pm (d)

Sum of the transverse energy The sum of the transverse energy of the two μ^\pm and the $\tau_{had-vis}$ is investigated in figure 6.8.5(b). The stack histograms are shown in figure 6.8.7(b). A cut at 90 GeV is chosen conserving over 95 % of the signal. Table 6.8.2 shows that this will mainly remove $W + jets$ and DY events. Also non-EW background events are expected to be rejected by cutting on the P_T sum, and since this is not a simulated background the effect of this cut is higher than the $\frac{S}{\sqrt{B}}$ graph indicate.

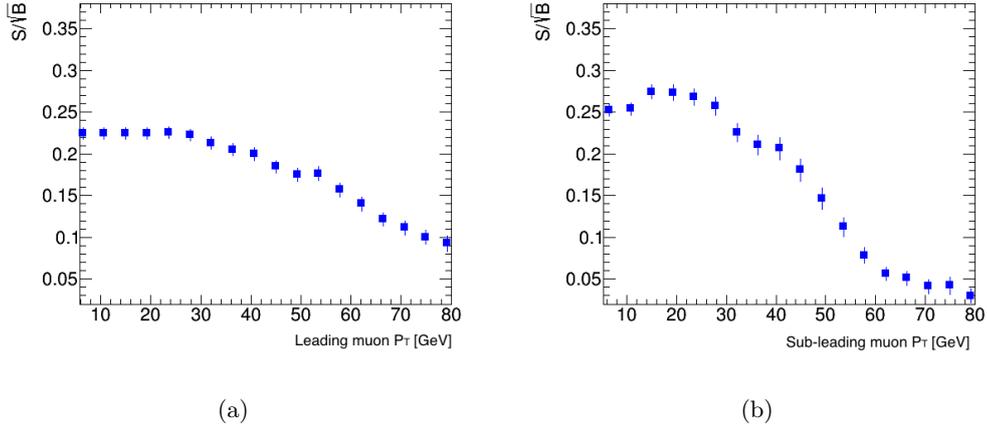


Figure 6.8.5: $\frac{S}{\sqrt{B}}$ for a given cut of the transverse momentum of the leading μ^\pm (a) and sub-leading μ^\pm (b). The value on the first axis is a lower cut.

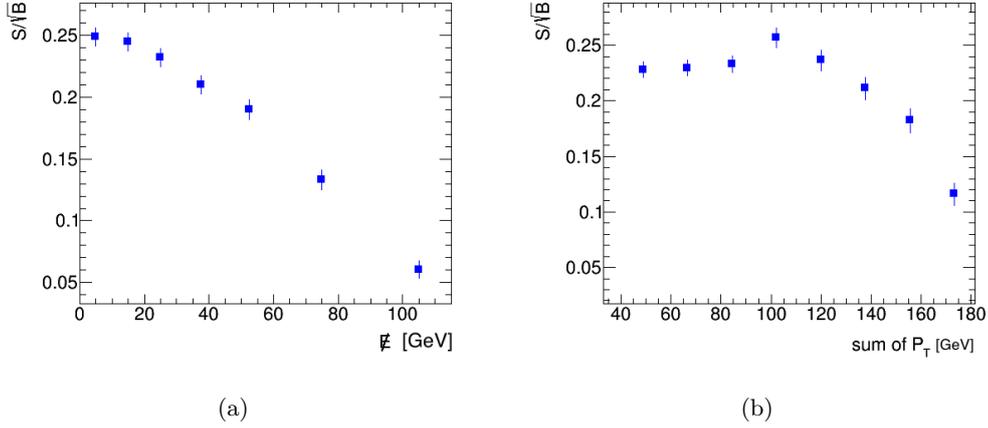


Figure 6.8.6: (a) shows $\frac{S}{\sqrt{B}}$ as a function of a lower cut on the missing energy. 6.8.6(b) shows $\frac{S}{\sqrt{B}}$ as a function of a lower cut on the sum of the momentum of the two μ^\pm and the $\tau_{had-vis}$

Stronger P_T^r : The P_T^r in *OSS* is shown in figure 6.8.7(c). It is clear from this figure that signal has a relatively high P_T^r and therefore it is not useful to lower P_T . To preserve the already limited signal yield it is decided to keep the $P_T^r > 20$ GeV requirement.

b-jet veto The b-jets are found by considering the characteristics of b-decays such as large impact parameter and displacement of the secondary vertex. In this analysis the MV1 tagging is used and a lower cut of 0.789 is applied, which results in an efficiency of 70%[64]. This is applied for all jets which have $P_T^{jet} > 20$ GeV and $\eta^{jet} < 4.5$. Finally an event is rejected if a b-jet is found in the event.

The plots in figure 6.8.7(d) and figure 6.8.7(e) show the distributions before and after the b-jet veto. It can be seen here and in table 6.8.2 that the b-jet veto reduces the effect of $t\bar{t}$ significantly.

A data driven background estimation is chosen for the background contributions which are simulated with too low statistics, such as Z+jets, W+jets, ttbar, and for any non-EW background. It is expected that this gives a more reliable estimation of the background, and a smaller uncertainty compared to simulation. This will be described in chapter 7.

It table 6.8.1 is a summary of the cuts used to get the best S/\sqrt{B} .

Charge requirement	The two μ^\pm must be SS and OS to the $\tau_{had-vis}$
Impact parameter	$ z_0 < 0.4$ mm
Isolation requirement for the two μ^\pm	etcone20/pt<0.08 and ptcone40/pt<0.08
Sum of the transverse momenta	> 90 GeV
b-jet veto	A veto is enforced on events with b-jets

Table 6.8.1: A summary of the cuts used to get the best S/\sqrt{B}

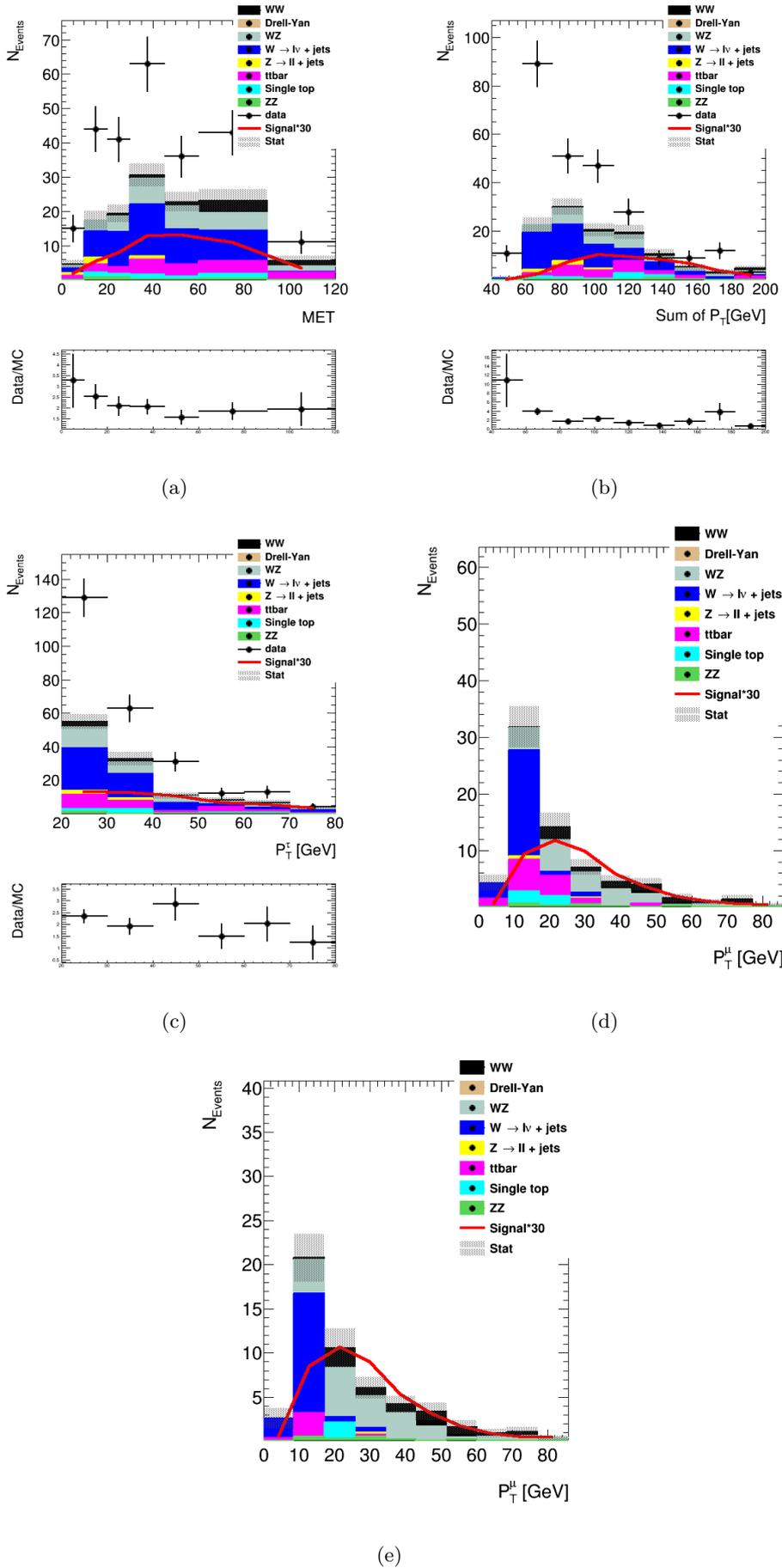


Figure 6.8.7: The \cancel{E}_T in (a), the sum of P_T for the three objects (b), P_T^l (c), the sub-leading μ^\pm after the P_T sum over 90 GeV (d) and OSS after the b-jet veto (e)

	Trigger	Veto e^\pm	$2 \mu^\pm$	$1 \tau_{had-vis}$	OSS	$z0^\mu < 0.4$	$\sum P_T^i > 90 \text{ GeV}$	b-jet veto
Data	20782540 ± 4559	20577932 ± 4536	10202269 ± 3194	39348 ± 198				
WW	18506 ± 81	18362 ± 80	6935 ± 49	51.7 ± 4.1	1.07 ± 0.51	1.07 ± 0.51	1.07 ± 0.51	0.75 ± 0.40
ZZ	6725.57 ± 35.67	6648.65 ± 35.47	5011.28 ± 30.87	170.52 ± 5.71	1.88 ± 0.56	1.88 ± 0.56	1.73 ± 0.54	1.54 ± 0.52
WZ	12456 ± 50	12326 ± 50	7509 ± 39	340 ± 8.3	18.45 ± 1.96	17.12 ± 1.88	17.12 ± 1.88	11.71 ± 1.56
$W \rightarrow \mu\nu + jets$	187909 ± 780	186041 ± 778	11521 ± 265	123 ± 24	25.1 ± 6.3	4.3 ± 1.3	3.56 ± 1.22	1.14 ± 0.55
$W \rightarrow \tau\nu + jets$	22414 ± 216	22287 ± 216	469 ± 42.12	6.70 ± 3.59	0.57 ± 0.57	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$W \rightarrow e\nu + jets$	2449 ± 59	2433 ± 59	0.20 ± 0.20	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
DYZ	1032176 ± 8799	1030321 ± 8795	779092 ± 7854	7993 ± 375	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$Z \rightarrow \mu\mu + jets$	9994414 ± 5490	9911854 ± 5468	8417030 ± 5087	37830 ± 219	5.62 ± 4.98	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$Z \rightarrow \tau\tau + jets$	131926 ± 61	131504 ± 610	59339 ± 422	320 ± 21	2.87 ± 1.21	0.91 ± 0.68	0.30 ± 0.30	0.00 ± 0.00
$Z \rightarrow ee + jets$	264 ± 13	260 ± 13	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
ttbar	168259 ± 237	166381 ± 236	26871 ± 94	306 ± 10	6.73 ± 1.7	4.84 ± 1.33	2.84 ± 0.95	1.40 ± 0.68
Top	29782 ± 146	29444 ± 145	3237 ± 36	22.6 ± 2.8	1.39 ± 0.64	0.91 ± 0.52	0.00 ± 0.00	0.00 ± 0.00
WH	24.25 ± 0.17	24.06 ± 0.17	8.20 ± 0.10	3.20 ± 0.06	1.59 ± 0.04	1.49 ± 0.04	1.34 ± 0.04	1.14 ± 0.04

Table 6.8.2: The number of expected events by simulations. Data observed is also shown upto a stage where the expected amount of signal is still low enough to consider the analysis blinded.

6.9 Main background contributions

As shown in table 6.8.2 $t\bar{t}$ and $W \rightarrow \mu\nu + jets$ are two of the main contributors to the background. Since neither of these contain three good leptons it is important to investigate how these processes enter into the background in order to provide the correct estimation. The investigation is based on MC simulations.

6.9.1 top + anti-top production

There are two possible final states for $t\bar{t}$ decays to contribute to the signal region as shown in figure 6.2.5. In $t\bar{t}$ decay one of the μ^\pm is from W decay as in the signal events. The second μ^\pm can arise from different processes, which are either a W, a hadron in q/\bar{q} jets or from b-quark decays.

From the simulated $t\bar{t}$ events it is evaluated which of the two processes in figure 6.2.5 contributes most when the two μ^\pm have the same sign of charge, SS , or opposite signs of charge, OS . In table 6.2.5 the relative fractions in the final state are given.⁴ When the μ^\pm 's have the same charge one of the W bosons will have decayed leptonically and the other hadronically, whereas if the charges are opposite sign both W decayed leptonically. Therefore is the decay of heavy quarks the main origin for the second μ^\pm , especially mesons with a b-quark.

When the μ^\pm 's in the final state have opposite sign the origin of the final state $\tau_{had-vis}$ is mainly from b-quarks. For the more interesting case of same sign μ^\pm 's, SS , 73% of the $\tau_{had-vis}$ are not matched to a true particle, but are likely to be faked by a jet from ISR/FSR instead of stemming from a W or a b.

Origin of μ^\pm	
OS	
Both μ^\pm from W^\pm	80%
1 μ^\pm from W^\pm and 1 μ^\pm from τ^\pm	17 %
1 μ^\pm from W^\pm and 1 μ^\pm from b	≈ 0.6 %
2 μ^\pm from τ	≈ 0.1 %
Others:	2 %
SS	
1 μ^\pm from W^\pm and 1 μ^\pm from a c-mesons	13.5 %
1 μ^\pm from W^\pm and 1 μ^\pm from a b-mesons	82.5 %
Other:	4 %

Table 6.9.1: Relative fractions for the origin of μ^\pm in $t\bar{t}$ events in the signal region.

⁴All cuts are applied except the b-jet veto and the P_T sum

6.9.2 W production in association with jets

Two of the three leptons in the signal region must come from other parts of the events and not the W decay. The origin of these leptons is investigated.

For $W \rightarrow l\nu$ the origin of the μ^\pm is shown in table 6.9.2. It is found that one μ^\pm always come from the W and the other from one of the heavy quarks.

None of the $\tau_{had-vis}$ can be matched to a true τ in $\Delta R < 0.4$, but rather to jets which may originate from all other parts of the event like ISR and FSR.

Origin of μ^\pm	
OS	
1 μ^\pm from W^\pm and 1 μ^\pm from b-quark	78 %
1 μ^\pm from τ^\pm and 1 μ^\pm from b/c meson	17 %
SS	
1 μ^\pm from W^\pm and 1 μ^\pm from b-meson	57 %
1 μ^\pm from W^\pm and 1 μ^\pm from c-mesons	43 %

Table 6.9.2: Investigation of the W+jets events

7

The background estimation methods

7.1 Background types

All contributions to the background are split into two parts: reducible background and irreducible background. Irreducible background contributions are defined by having the same final state as the signal with 3 good leptons, making it difficult to distinguish this from signal in the detector. Reducible background have at least one τ^\pm or one μ^\pm not stemming from a W or a Z boson. A data driven background estimation is used for estimating the reducible background. This is done since the second μ^\pm in W and t \bar{t} is from the tails of the isolation distributions which are not very well modelled in simulations. These processes also have limited statistics in the OSS region giving rise to a large uncertainty. Furthermore no simulations with sufficiently high statistics are available for the non EW backgrounds.

7.2 Irreducible background

The irreducible background, WZ and ZZ, will be determined by MC simulations. This is done since an estimate from data would not lead to better statistical precision and these backgrounds are well modelled in simulations since they contribute with three good leptons.

7.3 Reducible background

The reducible background involving μ^\pm stems from charged particles that manage to pass through the ATLAS detector and finally give a signal in the muon spectrometer. This can be:

1. A μ^\pm from a semi-leptonic decay of a heavy flavour quark (b or c quarks).

2. A μ^\pm from a pion or kaon from jets. It should be noticed that in this case the tracking part of the detector fails to find any kink in the track due to the mother particle.
3. A simple detector mistake, either a track found by random hits in the muon spectrometer from pions, kaons or any other charged particles. It rarely happens that these particles actually manage to reach the muon spectrometer causing a fake muon, so this source is negligible.

The energy deposit of the μ^\pm candidates, which is denoted as a fake μ^\pm will often be more spread and not as isolated as the energy from a μ^\pm from the hard process. A small fraction of these μ^\pm will however have an isolated energy contributing to the background in the signal region. These tails are not well described by simulation, and this supports the idea of using a data-driven method to estimate the reducible backgrounds [65].

The fake $\tau_{had-vis}$ stems normally from jet, and an important step is to investigate the nature of these jets. The properties of a jet depends on whether it originates from a quark or a gluon. Gluon-initiated jets will often give a wider signature in the detector, and are less likely to fake $\tau_{had-vis}$ compared to quark-initiated jets. Therefore it is important to have the same mixture of gluon-initiated and quark-initiated jets in the data-driven estimates as in the signal regions [22].

All this will be described in the following.

7.3.1 Definitions

To explain the fake-rate method it is useful to define three quantities: fake-rate, control region and sideband.

The procedure is to calculate fake-rates in the control-regions and then apply them to the side-bands, to obtain the number and kinematic distributions of the background events in the signal region, denoted as *OSS*.

Fake-rate :

The fake-rate is calculated from the number of signal leptons and jet-like leptons. Table 7.3.1 defines the signal leptons and the jet-like leptons. Thereby the jet-like leptons are considered close enough to the signal leptons to allow for a reasonable estimation from fake leptons to signal region using the fake-rate.

The muon fake-rate is a function of P_T and η and is defined as $FR(P_T, \eta)$.

$$FR_\mu(P_T, \eta) = \frac{N_{Signal-leptons}^{Data} - N_{Signal-leptons}^{MC,WZ+ZZ}}{N_{jet-like}^{Data} - N_{jet-like}^{MC,WZ+ZZ}} \quad (7.1)$$

$N_{Signal-leptons}^{Data}$ and $N_{jet-like}^{Data}$ are the number of signal leptons and jet-like leptons in data. $N_{Signal-leptons}^{MC,WZ+ZZ}$ and $N_{jet-like}^{MC,WZ+ZZ}$ are the numbers of true signal leptons and jet-like leptons in the WZ and ZZ MC simulations. This is done since these backgrounds are irreducible and as such are taken from simulations.

	Signal lepton	Jet-like leptons
Muons	$pt_{cone40}/pt < 0.08$ AND $et_{cone20}/pt < 0.08$	$0.08 < pt_{cone40}/pt < 5$ OR $0.08 < et_{cone20}/pt < 5$
Taus	Medium identification	BDT score > 0.3 Fail medium identification

Table 7.3.1: The definition of the signal lepton and jet-like leptons for both μ^\pm and $\tau_{had-vis}$.

The fake-rate for $\tau_{had-vis}$, $FR_\tau(P_T, \eta)$, is determined as a function of P_T , η and number of tracks, $FR_\tau(P_T, \eta, tracks)$. $FR_\tau(P_T, \eta, tracks)$ and $FR_\mu(P_T, \eta)$ will be abbreviated FR_τ and FR_μ .

From these fake-rates the probabilities for a fake muon and fake tau to pass selection can be found, these are denoted f_μ and f_τ :

$$f_{\mu/\tau} = \frac{FR_{\mu/\tau}}{1 + FR_{\mu/\tau}}$$

Control region :

FR_τ and FR_μ are measured in samples orthogonal to the signal region, but as close as possible to the signal region. When possible it will be investigated if there is the same mixture of fake jet types in these samples as found for the leptons in $W \rightarrow \mu\nu + jets$ and $t\bar{t}$ in section 6.9. The samples used are:

- A control region enriched in $Z \rightarrow \mu\mu + jets$:
This region is obtained by requiring a Z-peak of two μ^\pm with opposite charge.
- A control region that is enriched in multi-jets events.
This region is obtained by using a jet triggered sample.

The reason for investigating these two control regions is that the first region will mainly consist of quark-initiated jets and the latter of a mixture of quark-initiated and gluon-initiated jets. Thereby it is possible to investigate differences in FR_τ and FR_μ between these two kind of fake lepton sources and account for such differences in the final results.

Side-bands: The side-bands are the regions in which the obtained FR_τ and FR_μ are applied, thereby extrapolating the fake events into the signal region. These regions need to be as close as possible to the signal region. Several side-bands are considered for the estimation of the background, as an illustration two of the important side-bands are given here and shown in figure 7.3.1:

- 1 signal μ^\pm + 1 jet-like μ^\pm + 1 jet-like $\tau_{had-vis}$ candidate:
This region will mainly contain $W \rightarrow \mu\nu + jets$ and $t\bar{t}$, which are major contributors to the background as seen from the number of events in the signal region given in table 6.8.2.
- 2 jet-like μ^\pm + 1 jet-like $\tau_{had-vis}$ candidate:
This region consists mainly of multi-jets.

The irreducible background from WZ and ZZ must be subtracted from the side-bands before applying FR_τ and FR_μ to the jet-like lepton candidates.

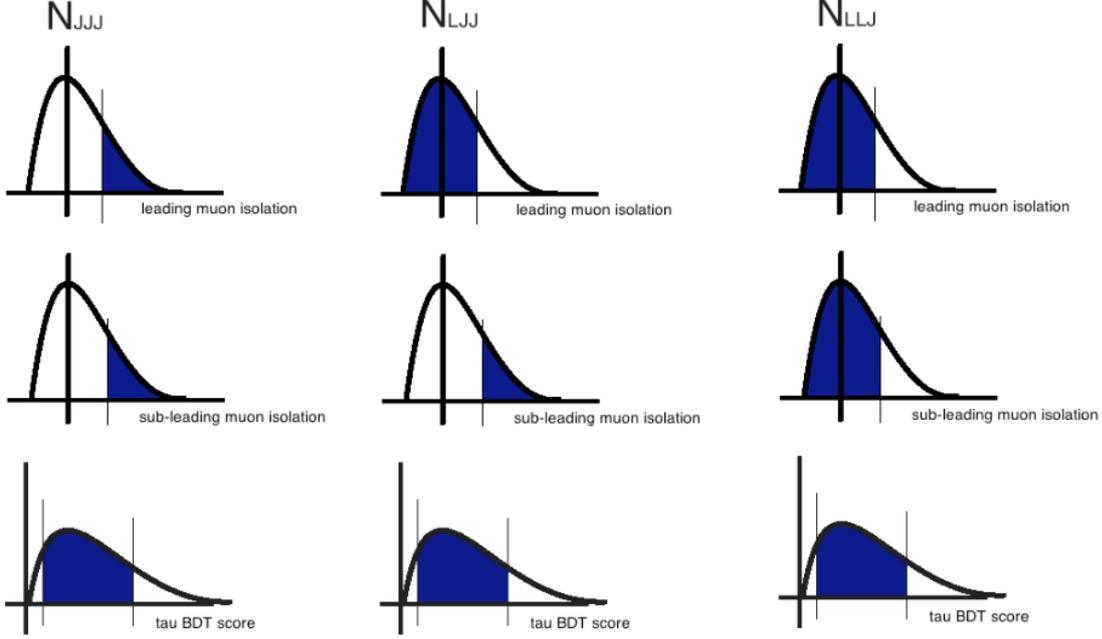


Figure 7.3.1: A cartoon showing the main idea behind the side-bands N_{JJJ} , N_{LJJ} and N_{LLL} . The shaded area is the selected part of each distribution.

7.3.2 Equations for the fake-rate estimation

The goal is now to derive the equations for applying the fake-rate to the side-bands to obtain the correct prediction for the data-driven background estimation in the signal region.

A useful notation is:

The index T and F: **T** is used as an index for a true isolated μ eg. from W or Z, or a true hadronic τ also from W or Z. **F** is used for non-isolated μ^\pm from eg $\pi^\pm \rightarrow \mu^\pm + \nu$ or b-decays. For $\tau_{had-vis}$ the **F** is almost always a jet.

The index J and L: **L** is used as an index for passing signal leptons and **J** is used as an index for passing jet-like cuts in table 7.3.1.

The reducible background in the signal region, N_{LLL} , is given by equation 7.2

$$\begin{aligned}
 N_{LLL} = & \epsilon_\mu^2 f_\tau N_{TTF} + \epsilon_\mu f_{sublead-\mu} \epsilon_\tau N_{TFT} + \epsilon_\mu f_{sublead-\mu} f_\tau N_{TFF} + f_{lead-\mu} \epsilon_\mu f_\tau N_{FTF} \\
 & + f_{lead-\mu} f_{sublead-\mu} \epsilon_\tau N_{FFT} + f_{lead-\mu} f_{sublead-\mu} f_\tau N_{FFF} \\
 & + f_{lead-\mu} \epsilon_\mu \epsilon_\tau N_{FTT}
 \end{aligned} \tag{7.2}$$

The first index denotes the leading μ , the second index denotes the sub-leading μ , and the third denotes the index of the τ . ϵ_τ is the efficiency for a true $\tau_{had-vis}$ to be selected as a signal tau. ϵ_μ is the efficiency for a true μ^\pm to be selected as a signal μ^\pm . It is very unlikely that a true μ^\pm will not be isolated and therefore this factor is set to 1.

Equation 7.2 continues to higher orders and an important step in this data-driven background estimation is to determine the necessary number of terms for equation 7.2. This will be discussed later in the thesis, but it is found that three objects cases are sufficient for the background estimation. This is therefore shown.

To solve equation 7.2 one needs to find independent equations to derive the terms N_{TTF} , N_{FFT} and so on. This is done by considering all the three terms in the side-bands. As for the signal region all side-bands can be expressed by the true terms, an example is given here for N_{LLJ} after subtraction of the di-boson events:

$$\begin{aligned} N_{LLJ} &= N_{TTF}(1 - f_\tau) + N_{FFT}f_\mu^2(1 - \epsilon_\tau) + N_{FFF}f_\mu^2(1 - f_\tau) \\ &+ N_{FTT}f_\mu(1 - \epsilon_\tau) \end{aligned} \quad (7.3)$$

The rest of the side-band equations are given in Appendix A. By solving this set of equations an expression for the data-driven background in the signal region is obtained:

$$\begin{aligned} N_{LLL} &= FR_\tau(N_{LLJ} - N_{LLJ}^{Diboson}) + FR_{\mu-sublead}(N_{LJL} - N_{LJL}^{Diboson}) \\ &+ FR_{\mu-lead}(N_{JLL} - N_{JLL}^{Diboson}) - FR_\tau FR_{\mu-sublead}(N_{LJJ} - N_{LJJ}^{Diboson}) \\ &- FR_{\mu-lead} FR_\tau(N_{JLJ} - N_{JLJ}^{Diboson}) - FR_\mu FR_{\mu-sublead}(N_{JLL} - N_{JLL}^{Diboson}) \\ &+ FR_\tau FR_\mu FR_{\mu-sublead}(N_{JJJ} - N_{JJJ}^{Diboson}) \end{aligned} \quad (7.4)$$

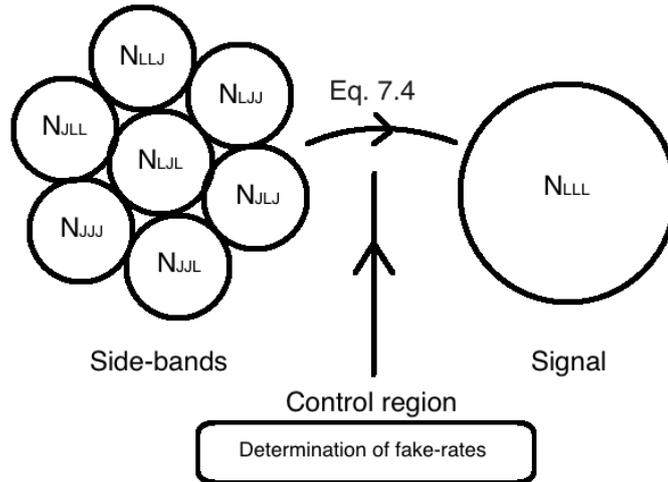


Figure 7.3.2: A simplified illustration of the data-driven background estimation

8

Determination of FR_μ and FR_τ

FR_μ and FR_τ are calculated in samples with different composition to test for changes dependent on the μ^\pm and τ origin. The samples used are enriched in multi-jets events and $Z \rightarrow ll + jets$ events. The origin of the leptons in the main background was found in section 6.9 and the goal now is to calculate FR_μ and FR_τ and check that the leptons used to calculate the fake-rates have similar origins.

8.1 Requirements for a sample enriched in $Z \rightarrow ll + jets$ events

The control region for the Z +jet is found by requiring the following of the μ^\pm :

- Two oppositely charged μ^\pm with mass 91 ± 10 GeV, which fulfill these criteria:
 - $P_T > 8$ GeV
 - Relative isolation in the calorimeter $(\frac{etcone20}{P_T}) < 0.08$
 - Relative isolation in the inner detector $(\frac{ptcone20}{P_T}) < 0.08$.¹
 - The impact parameter $|z_0| < 0.4$ mm

The μ^\pm which fulfill these criteria and thereby create a Z mass peak are shown in the plots in figure 8.1.1, where good agreement between data and MC simulation is evident. It is clear that these distributions consist of $Z \rightarrow \mu\mu + jets$ events, and thereby the desired region is obtained. The important samples for $Z \rightarrow \mu\mu + jets$ are the ones filtered to enrich light quarks, which is to be expected since these are the main part of the process.

¹ptcone20 is used here instead of ptcone40 to increase the available number of events for the fake-rate measurement

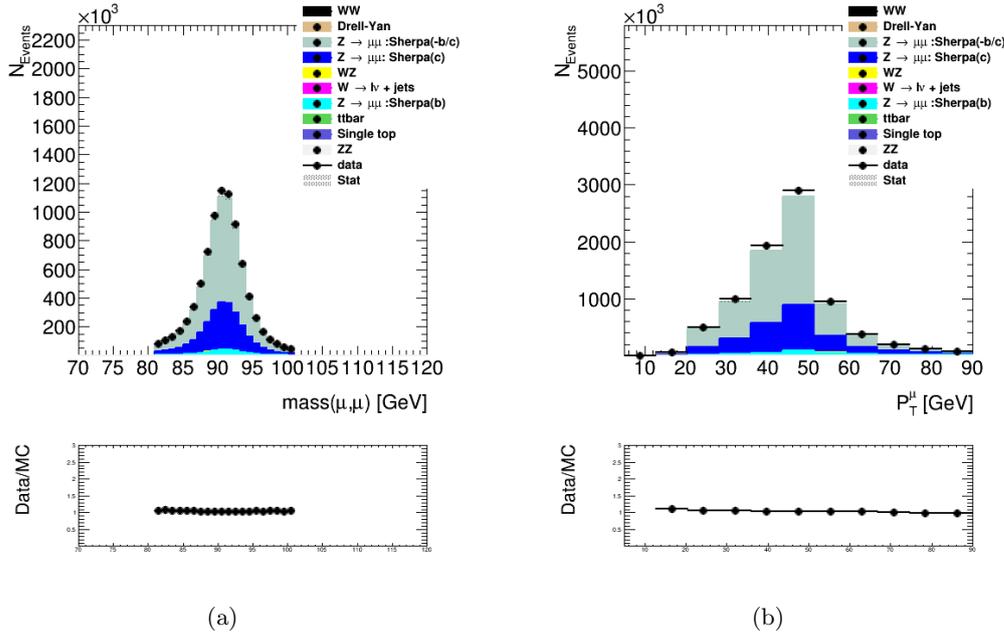


Figure 8.1.1: (a) The mass of the two μ^\pm . (b) P_T^μ distributions for the two μ^\pm .

8.1.1 FR_μ in the $Z \rightarrow ll + jets$ enriched sample

In the $Z \rightarrow ll + jets$ control region events with three μ^\pm 's are used to determine the FR_μ . It was found in section 6.9 that the origin of the μ^\pm in ttbar and $W \rightarrow \mu\nu + jets$ events are heavy quarks. Therefore the Sherpa MC samples listed in table 6.4.1 are used, which have higher statistics available for b-quarks and c-quarks.

To be orthogonal to the ZH signal a veto on other leptons in the event is enforced. The P_T^μ and $\frac{etcone20}{P_T}$ of the μ^\pm 's that fail this isolation are shown in 8.1.2(a) and 8.1.2(b) respectively. The μ^\pm 's which pass isolation are shown in 8.1.2(c) and 8.1.2(d). The $Z \rightarrow \mu\mu + jets$ events which pass the isolation requirement are mainly the events where a filter has been applied for b-quarks, whereas also the c-quark filter is important for μ^\pm 's failing isolation.

The origin of the μ^\pm are found in truth and stated in table 8.1.1. It is indicated that mesons containing a c-quark and a b-quark can both fail and pass isolation. Figure 8.1.2(c) and 8.1.2(d) show that the b-quark filter samples were the important ones for μ^\pm passing isolation.

In table 8.1.1 the percentage of μ^\pm passing isolation is given depending on the origin of the μ^\pm . These probabilities are of the same order, although with a bit higher probability for μ^\pm from b-quarks than c-quarks.

FR_μ is found as a function of P_T^μ and η^μ with three bins in P_T^μ and three bins in η^μ . It is found that the distribution of the isolation in η^μ is very uniform making three bins sufficient. The obtained fake-rates are shown in figure 8.1.3. The results are shown with and without subtraction of the true WZ and ZZ, the importance of which is largest for the higher values

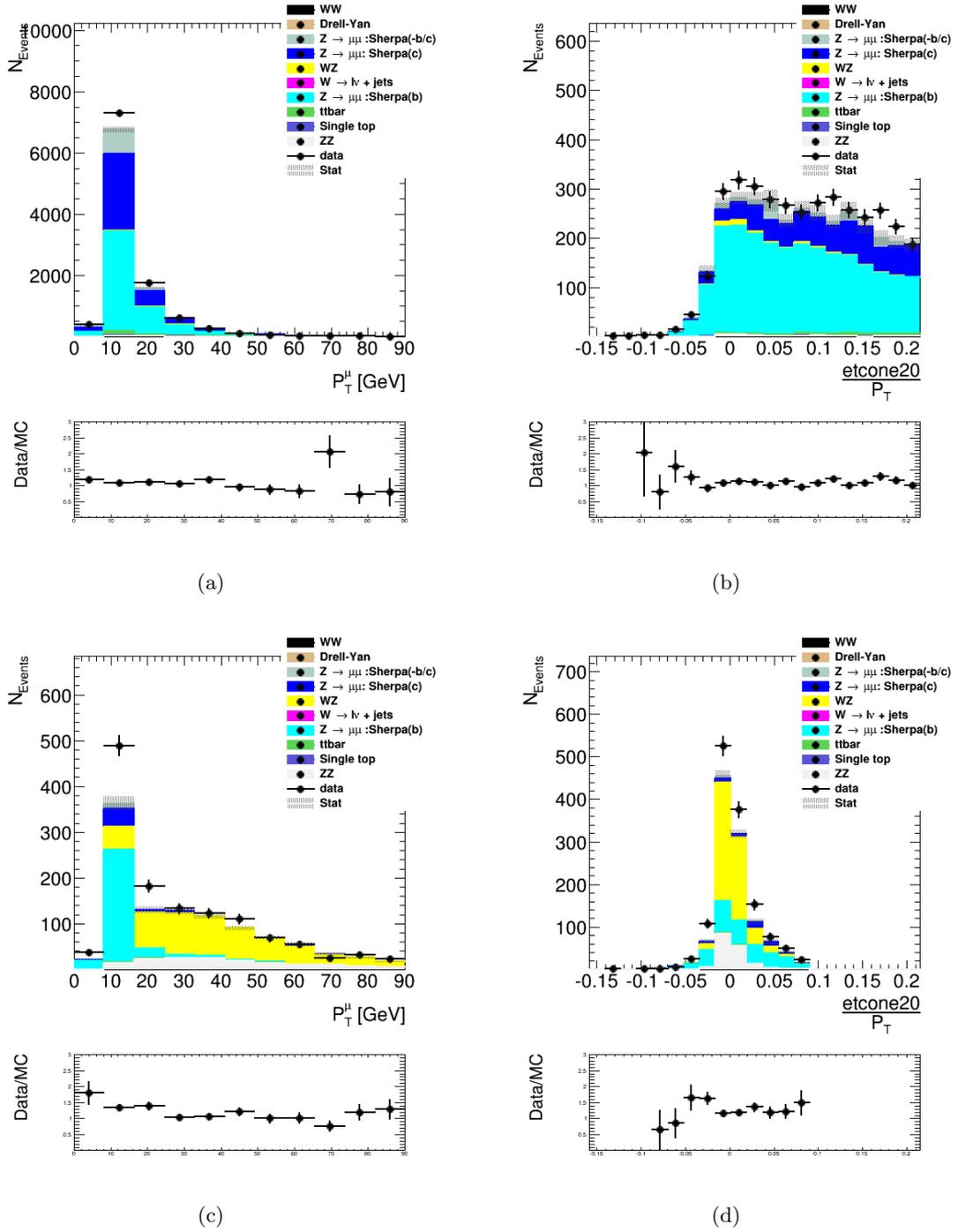
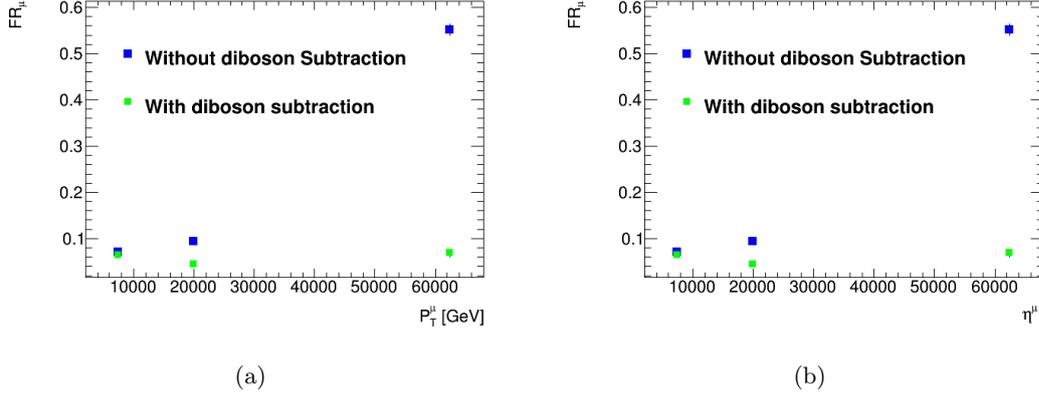


Figure 8.1.2: (a) and (b) show the P_T^μ and $\frac{\text{etcone20}}{P_T}$ of the μ^\pm failing isolation. (c) and (d) show the P_T^μ and $\frac{\text{etcone20}}{P_T}$ of the μ^\pm passing isolation.

of P_T^μ . This is due to the μ^\pm from these processes being more boosted and having a relative higher P_T . In table 8.1.2 FR_μ as a function of P_T^μ and η^μ is given.

Table 8.1.3 gives the variation of FR_μ with different triggers applied, the di-lepton trigger is only used when the single trigger does not fire. No significant trigger dependence is present.

Origin	All μ^\pm	Chance of μ^\pm passing isolation
Meson with a b-quark	42.6 %	58 %
Mesons with a c-quark	52.5 %	63 %
other	4.7 %	0 %

Table 8.1.1: The true origin of the third μ^\pm in the $Z \rightarrow ll + jets$ sampleFigure 8.1.3: (a) FR_μ as a function of P_T^μ . (a) FR_μ as a function of η^μ .

	$-2.5 < \eta^\mu < -1.2$	$-1.2 < \eta^\mu < -1.2$	$1.2 < \eta^\mu < 2.5$
$P_T^\mu < 15$ GeV	0.0846 ± 0.008	0.0465 ± 0.003	0.109 ± 0.009
15 GeV $< P_T^\mu < 25$ GeV	0.04 ± 0.01	0.0396 ± 0.005	0.077 ± 0.01
$P_T^\mu > 25$ GeV	0.01 ± 0.01	0.064 ± 0.01	0.15 ± 0.03

Table 8.1.2: FR_μ for the μ^\pm as a function of η^μ . and P_T^μ .

Trigger / P_T [GeV]	$8 < P_T^\mu < 15$	$15 < P_T^\mu < 25$	$P_T^\mu > 25$
EF_mu24i_tight OR EF_mu36_tight	0.0616 ± 0.003	0.0448 ± 0.004	0.0761 ± 0.008
EF_2mu13 OR EF_mu18_tight_mu8_EFFS	0.0898 ± 0.02	0.0403 ± 0.02	-

Table 8.1.3: FR_μ depending on the triggers. The di-muon triggers are used for events where the single triggers did not fire.

8.1.2 FR_τ in the $Z \rightarrow ll + jets$ enriched sample

In this section the fake-rate for the $\tau_{had-vis}$ candidates in the Z+jet enriched sample is found as defined in section 8.1. All $\tau_{had-vis}$ candidates with a identification BDT score above 0.3 are considered. The reason for the cut on the BDT score on 0.3 is inspired from [22], and will be treated in section 8.2.2

The plot in figure 8.1.4(a) shows the P_T distribution for the selected signal $\tau_{had-vis}$, and the

plot in figure 8.1.4(b) shows the P_T distribution for the jet-like $\tau_{had-vis}$ candidates. This region is totally dominated by $Z \rightarrow \mu\mu + jets$.

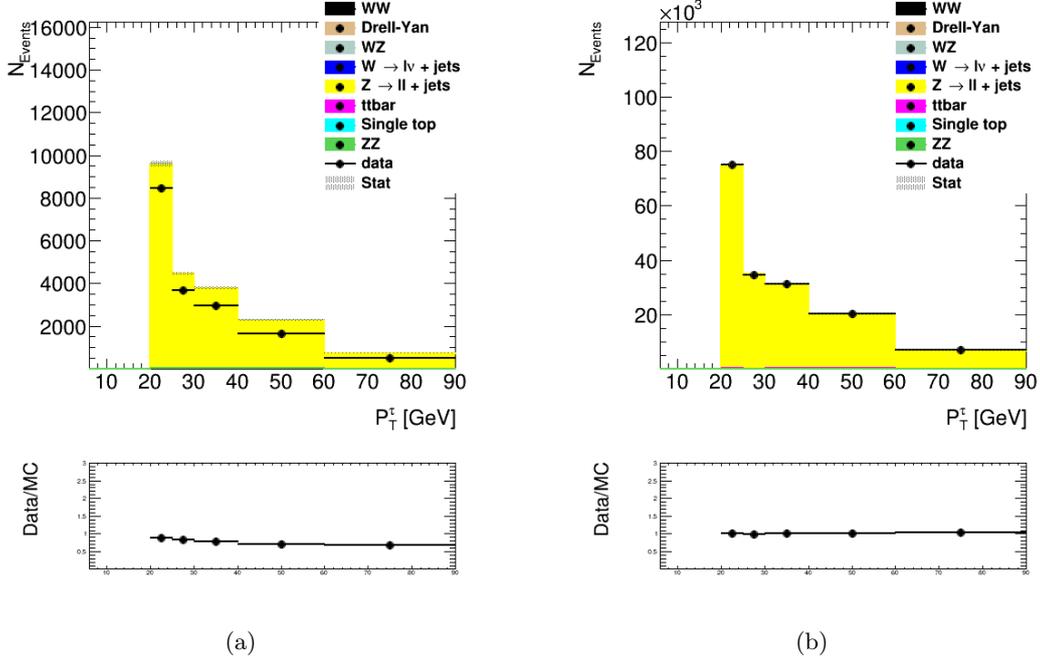


Figure 8.1.4: (a) P_T^τ for $\tau_{had-vis}$ passing the medium BDT, (b) P_T^τ distribution for $\tau_{had-vis}$ candidates with a BDT score above 0.3

The FR_τ is found for 1 prong and 3 prong separately as a function of P_T^τ and η^τ in Figure 8.1.5 and Table 8.1.4. For 3 prong $\tau_{had-vis}$ there is no variation of FR_τ with η , but for 1 prong taus there is a small dependence. This dependency is primarily for low P_T^τ , and the η^τ dependence is therefore done for $P_T^\tau < 30$ GeV. The relatively large bins for FR_τ are needed to keep the statistical uncertainty of FR_τ as low as possible.

1 prong				
[GeV]	[GeV]	[GeV]	[GeV]	[GeV]
$15 < P_T^\tau < 25$	$25 < P_T^\tau < 30$	$30 < P_T^\tau < 40$	$40 < P_T^\tau < 60$	$60 < P_T^\tau < 90$
$ \eta < 1.2 ; 0.103 \pm 0.001$	0.102 ± 0.002	0.0933 ± 0.002	0.078 ± 0.002	0.069 ± 0.003
$ \eta > 1.2 ; 0.128 \pm 0.002$	0.111 ± 0.003			
3 prong				
[GeV]	[GeV]	[GeV]	[GeV]	[GeV]
$P_T < 25$	$P_T < 30$	$30 < P_T < 40$	$40 < P_T < 50$	$50 < P_T < 90$
0.0653 ± 0.001	0.044 ± 0.001	0.0367 ± 0.001	0.0277 ± 0.001	0.0205 ± 0.002

Table 8.1.4: FR_τ as a function of P_T^τ in the $Z \rightarrow ll + jets$ sample. For 1 prong there a η splitting for $P_T^\tau < 30$ GeV is done

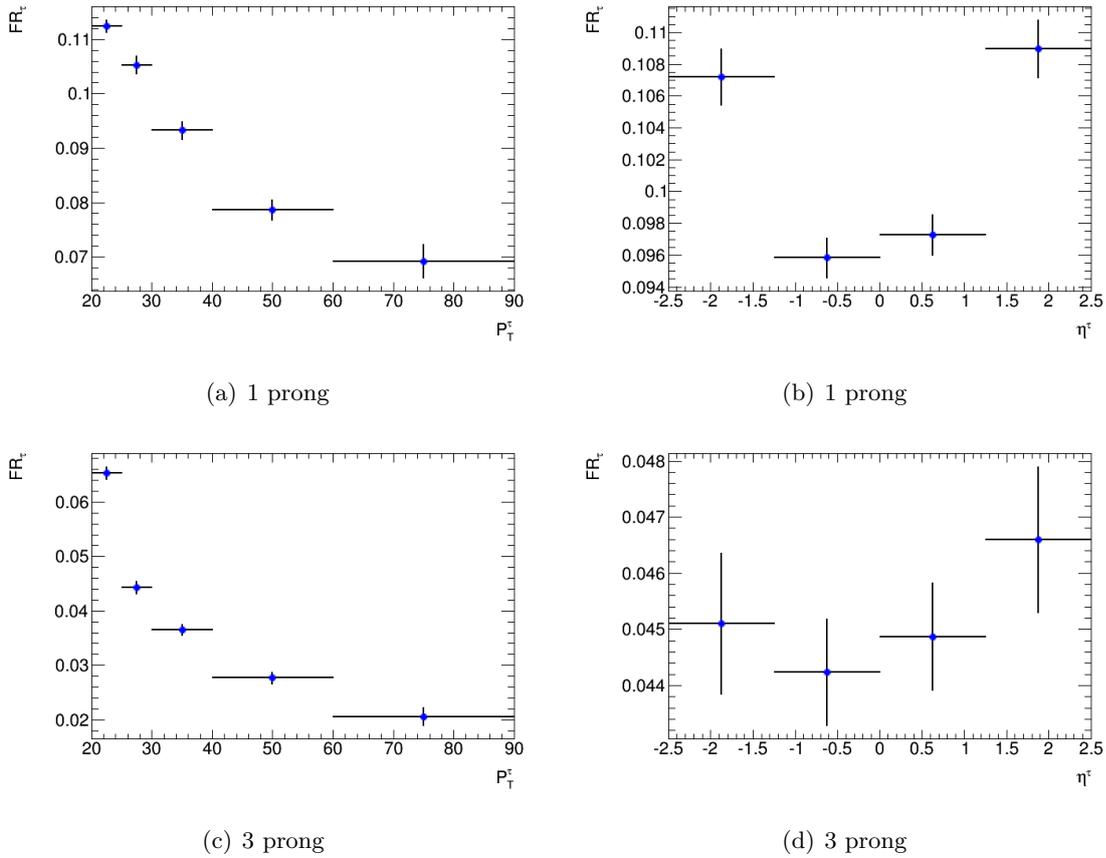


Figure 8.1.5: FR_τ as a function of P_T^τ (a) and η^τ (b) for 1 prong. FR_τ as a function of P_T^τ (c) and η^τ (d) for 3 prong. All plots are in the the $Z \rightarrow l + jets$ enriched sample.

8.2 Requirements for a sample enriched in multi-jets events

FR_μ and FR_τ are investigated in data triggered by jets and thus dominated by multi-jets jets. Other contributions in these samples are negligible due to the nature of these samples. When triggering on jets a pre-scaling is necessary to ensure a manageable amount of data stored to the disk. In ATLAS pre-scaling is done as a function of jet energy to include energetic jets as well as low energy jets. The factors use for the pre-scaling are shown in table 8.2.1.

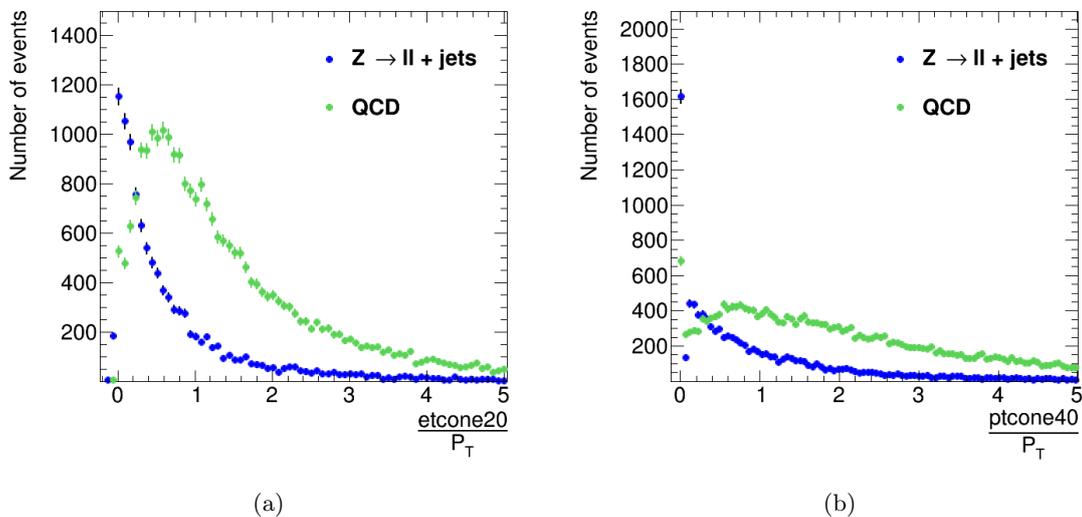
In this sample multi-jets are most likely a mixture of gluon-initiated and quark-initiated jets, where the former is expected to be wider and have more tracks than the latter. For samples enriched in $Z \rightarrow l + jets$ quark-initiated jets are the dominant part, whereas for multi-jets enriched samples it will be a mixture of the two. Whether or not this difference influences FR_μ and FR_τ is the point of interest.

minimum E_T^{jet} GeV	maximum E_T^{jet} GeV	Pre-scaling factor
0	80	50
80	140	20
200	300	10
300	400	5
400	10^9	1

Table 8.2.1: The table shows the pre-scaling used in ATLAS for multi-jets samples [66]

8.2.1 FR_μ in the multi-jets enriched sample

The jets in the multi-jets enriched sample have a more uniform distribution of isolation and more candidates are present at high $etcone20/p_T$ and $ptcone40/p_T$ values than in the $Z \rightarrow ll + jets$ events. The difference for the isolation cones are shown in Figure 8.2.1. These figures confirm that the μ^\pm candidates in multi-jets are another type of particles than in $Z \rightarrow ll + jets$ events. Based on this observation it can be expected that FR_μ calculated in this region differ from the FR_μ obtained in the $Z \rightarrow ll + jets$ sample .

Figure 8.2.1: The calorimeter (a) and tracking (b) isolation for the μ^\pm 's failing the isolation requirement.

To investigate this MC multi-jets samples generated with Pythia are used. These samples have a relatively low number of events and are therefore not used in the other parts of analysis where the muon trigger is needed. They can, however, be used here for an investigation of the truth origin of the μ^\pm

The primary result of the investigation is heavy quarks are the origin of 88.7 % of the μ^\pm 's compared to the 98% for $Z \rightarrow ll + jets$, as tabulated in table 8.1.1. This difference conse-

quently leads to different FR_μ .

Origin	All μ^\pm
Meson with a b-quark	53.4 %
Mesons with a c-quark	33.3 %
Others	13.3 %

Table 8.2.2: The true origin of the μ^\pm , which are used to determine the FR_μ in the multi-jets sample

Table 8.2.3 list the FR_μ 's obtained in the multi-jets region and in the $Z \rightarrow ll + jets$ region. The difference in FR_μ will be taken into account when transferring jet-like candidates from the side-bands into the signal region. If a side-band consists of mainly multi-jets jets the FR_μ found here will be used.

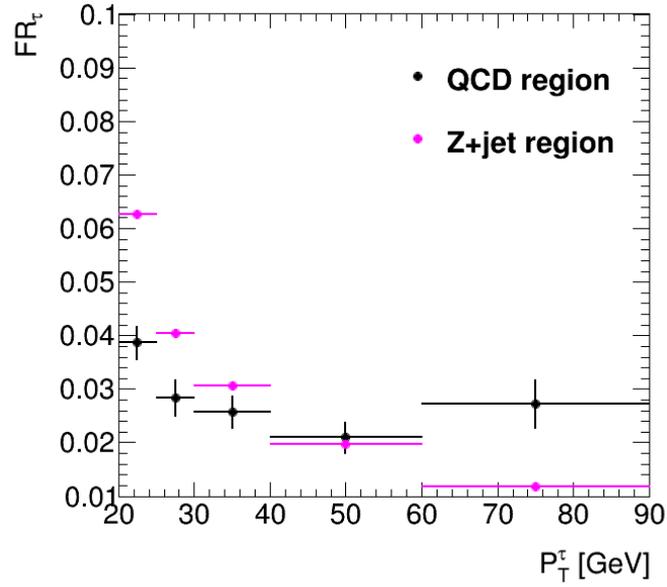
Sample	$8 \text{ GeV} < P_T^\mu < 15 \text{ GeV}$	$15 \text{ GeV} < P_T^\mu < 25 \text{ GeV}$	$P_T^\mu > 25 \text{ GeV}$
$Z \rightarrow \mu\mu + jets$	0.063 ± 0.003	0.044 ± 0.004	0.063 ± 0.01
multi-jets	0.021 ± 0.003	0.059 ± 0.006	0.111 ± 0.002

Table 8.2.3: The variation of FR_μ determined for a multi-jets enrich sample and a $Z \rightarrow \mu\mu$ enrich sample.

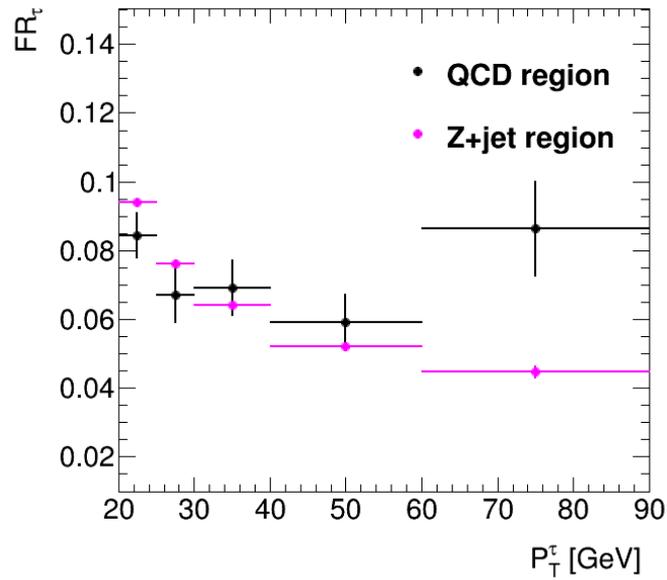
8.2.2 FR_τ in the multi-jets enriched sample

Fake-rates for $\tau_{had-vis}$ candidates, FR_τ , are calculated and further it is clarified why there a minimum cut value on the $\tau_{had-vis}$ identification BDT score cut.

The cut on the $\tau_{had-vis}$ candidate identification BDT score at 0.3 is motivated from [22]. This is a tag-and-probe analysis for $Z \rightarrow \tau\tau$, which shows that there are no true $\tau_{had-vis}$ with a BDT score below 0.3, but mainly multi-jets jets. The idea is the following: A jet can be either a quark-initiated jet or gluon-initiated jet. Gluon-initiated jets are expected to have a lower BDT score than quark-initiated jets, since they are wider and have more tracks. Removing candidates with a low BDT score therefore facilitates determination of similar FR_τ for gluon-initiated jets and quark-initiated jets. This assumption is tested by calculating FR_τ with or without the 0.3 BDT cut, as shown in Figure 8.2.2. FR_τ is more similar with the BDT cut in Figure 8.2.2(b) than without the BDT cut in figure 8.2.2(a). This behaviour is as expected since $Z \rightarrow ll + jets$ are mainly quark-initiated jets.



(a)



(b)

Figure 8.2.2: The fake-rate of taus calculated with and without (a) a cut on the BDT score (b).

9

The side-band regions of the data-driven background estimation method

To get a stable background estimation it is important to understand the composition of the side-bands and the statistics of the regions. From this a reasonable choice of side-bands can be made resulting in a robust background prediction for the signal region. The composition of the side-bands is also critical to ensure that the correct FR_μ and FR_τ are used.

In section 7.1 it was found how to estimated the reducible background in the signal region. For ease of reading these equations are repeated here. The equation for the truth objects are given in the equation:

$$\begin{aligned} N_{LLL} = & \epsilon_\mu^2 f_\tau N_{TTF} + \epsilon_\mu f_{sublead-\mu} \epsilon_\tau N_{TFT} + \epsilon_\mu f_{sublead-\mu} f_\tau N_{TFF} + f_{lead-\mu} \epsilon_\mu f_\tau N_{FTF} \\ & + f_{lead-\mu} f_{sublead-\mu} \epsilon_\tau N_{FFT} + f_{lead-\mu} f_{sublead-\mu} f_\tau N_{FFF} \\ & + f_{lead-\mu} \epsilon_\mu \epsilon_\tau N_{FTT} \end{aligned} \quad (7.2)$$

The truth terms are found by considering each side-band, which can also be expressed from the truth terms. The example of N_{LLJ} was used in section 7.1 :

$$\begin{aligned} N'_{LLJ} = & N_{TTF}(1 - f_\tau) + N_{FFT} f_\mu^2 (1 - \epsilon_\tau) + N_{FFF} f_\mu^2 (1 - f_\tau) \\ & + N_{FTT} f_\mu (1 - \epsilon_\tau) \end{aligned} \quad (7.3)$$

By combining all the equations for side-bands the equation determining the background will be as in equation 7.4:

$$\begin{aligned}
 N_{LLL} = & FR_{\tau}(N_{LLJ} - N_{LLJ}^{Diboson}) + FR_{\mu}(N_{LJL} - N_{LJL}^{Diboson}) + FR_{\mu}(N_{JLL} - N_{JLL}^{Diboson}) \\
 & - (N_{LJJ} - N_{LJJ}^{Diboson})FR_{\tau}FR_{\mu} - (N_{JLJ} - N_{JLJ}^{Diboson})FR_{\mu}FR_{\tau} \\
 & - (N_{JJL} - N_{JJL}^{Diboson})FR_{\mu}FR_{\mu} + (N_{JJJ} - N_{JJJ}^{Diboson})FR_{\tau}FR_{\mu}FR_{\mu} \quad (7.4)
 \end{aligned}$$

The first step is to test which side-bands are needed. The number of events in each three object side-bands are listed in table 9.1.1 and four object side-bands in table 9.1.2. (The four object side-bands consist of two μ^{\pm} 's and two $\tau_{had-vis}$) The number of events in table 9.1.2 are in most cases constitute about 4-5% of the similar three object side-bands in table 9.1.1. The four object side-bands statistics in table 9.1.2 are therefore found to be negligible, and only the three object side-bands should be considered. This choice will be evaluated in the validation test.

Sideband	N_{events}	$N_{events}^{Truthdi-boson}$	Total
N_{LLJ}	101	1.37	99.6
N_{LJJ}	1422	10.4	1412
N_{JLJ}	161	≈ 0	161
N_{LJL}	208	≈ 0	208
N_{JLL}	23	≈ 0.03	22.9
N_{JJJ}	2186	≈ 0	2186
N_{JJL}	97	≈ 0	97

Table 9.1.1: Number of events in three object side-bands

Sideband	N_{events}	$N_{events}^{Truthdi-boson}$	Total	%
N_{LLJJ}	2	0.01	1.98	1.9% of N_{LLJ}
N_{LJJJ}	54	0.43	53.6	3.8% of N_{LJJ}
N_{JLJJ}	7	≈ 0	7	4.3 % of N_{JLJ}
$N_{LJLJ}+N_{LJLJ}$	14	≈ 0	14	6.7% N_{LJL}
$N_{JLLJ}+N_{JLLJ}$	1	≈ 0	1	4.4 % of N_{JLL}
N_{JJJJ}	22	≈ 0	22	1 .1% of N_{JJJ}
$N_{JJJL}+N_{JJLJ}$	4	≈ 0	4	4 .1% of N_{JJL}

Table 9.1.2: Number of events in each side-band if there are four objects

In Appendix B the equations for each truth term is found. It should be noticed that these numbers are not used in the data-driven method, but merely provide confidence to the number of terms in equation 7.4. The goal is for the equation to be well behaved and robust against small fluctuations. From table 9.1.3 it is also possible to gather which terms are important for the background estimation. The number of events for each truth term is given in table 9.1.3 where both the systematic due to statistical and the fake-rate are taken into consideration.

Sideband	Number of events
N_{TTF}	17.8 ± 38.1
N_{TFF}	1310 ± 53.2
N_{FTF}	30.1 ± 31.1
N_{TFT}	276.1 ± 37.4
N_{FTT}	35.7 ± 16.6
N_{FFF}	2950 ± 210
N_{FFT}	-68.5 ± 150

Table 9.1.3: The number of events in each true term. The uncertainty is from variation of the fake-rate and the systematic uncertainty due to statistics

The composition of the side-bands is analysed with the help of simulated events, thereby getting a confirmation about the expected composition of each background:

The side-band N_{LLJ} in figure 9.1.1(a) :

In this side-band the main contributions are expected from $W \rightarrow \mu\nu + jets$, WZ and $t\bar{t}$.

The side-bands N_{LJJ} and N_{JLJ} in figure 9.1.1(c) and 9.1.2(a) :

Most of the background is again due to $W \rightarrow \mu\nu + jets$ and $t\bar{t}$. If the leading μ^\pm is a jet-like μ^\pm there are less events, as expected due to the single trigger isolation requirement. Furthermore there is an expected discrepancy between data and MC simulation in N_{JLJ} due to the lack of simulation of the non EW background or simulation problems to reproduce data.

The side-bands N_{LJL} and N_{JLL} in figure 9.1.2(c) and 9.1.3(a) :

In these side-bands not only $W \rightarrow \mu\nu + jets$ and $t\bar{t}$ are important, but also $Z \rightarrow \mu\mu + jets$. Again there are far less events if the leading μ^\pm is jet-like.

The side-bands N_{JJJ} and N_{JLJ} in figure 9.1.3(c) and 9.1.4(a) :

These two side-bands consist only of non EW background

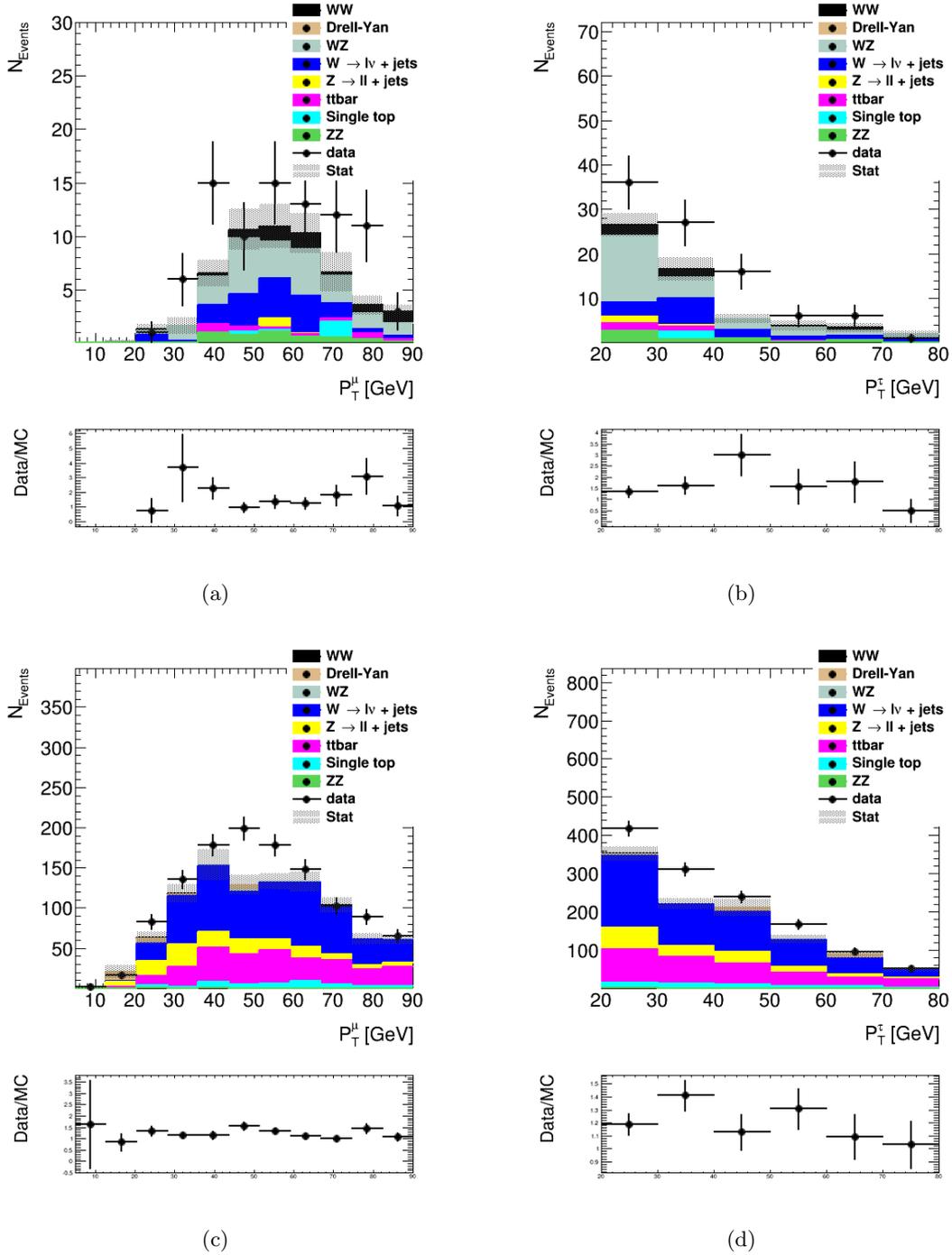


Figure 9.1.1: P_T^μ for the leading μ^\pm and P_T^τ in N_{LLJ} ((a), (d)), N_{LLJ} ((c),(b))

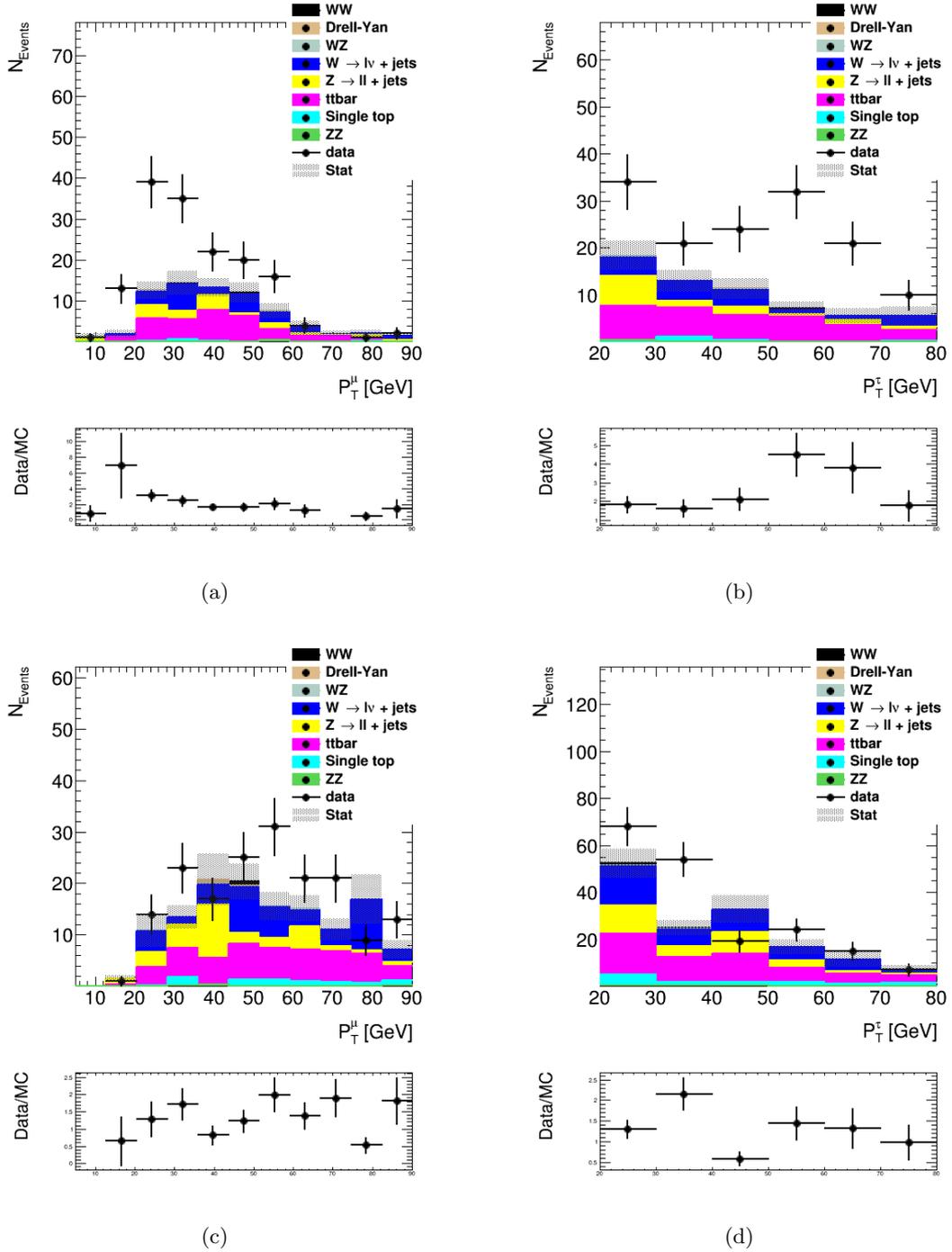
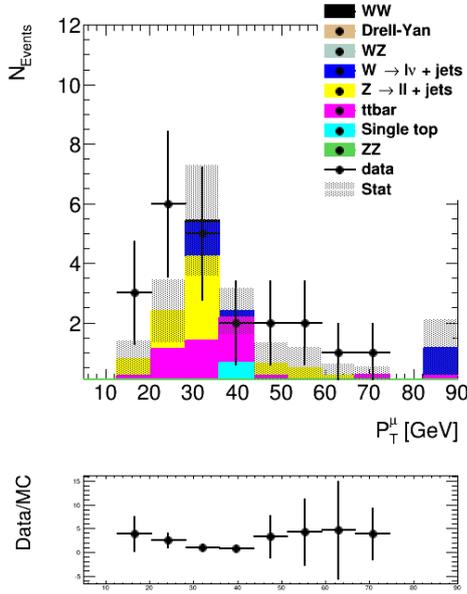
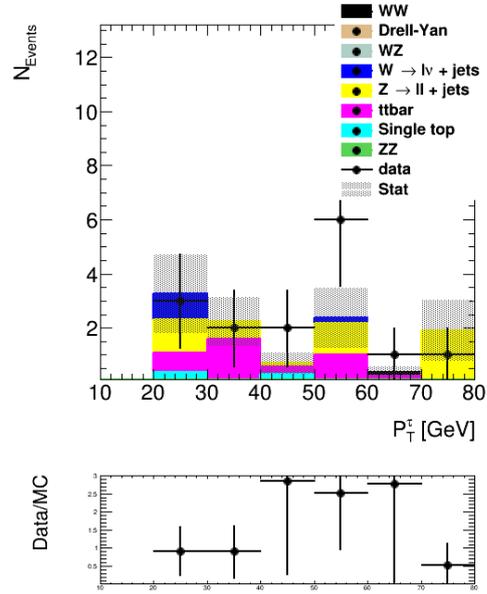


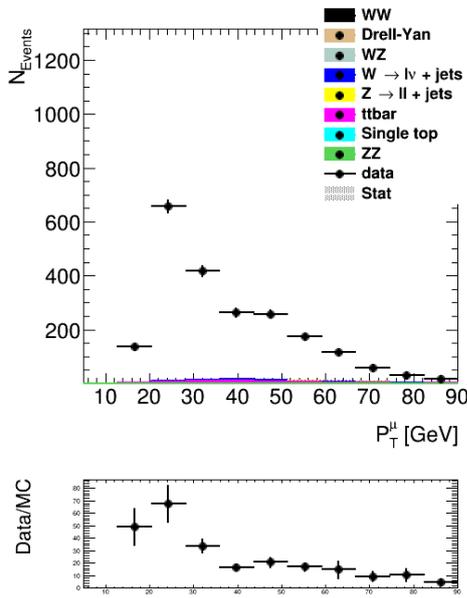
Figure 9.1.2: P_T^μ for the leading μ^\pm and P_T^τ in N_{JLJ} ((a),(b)) and N_{LJL} ((c),(d))



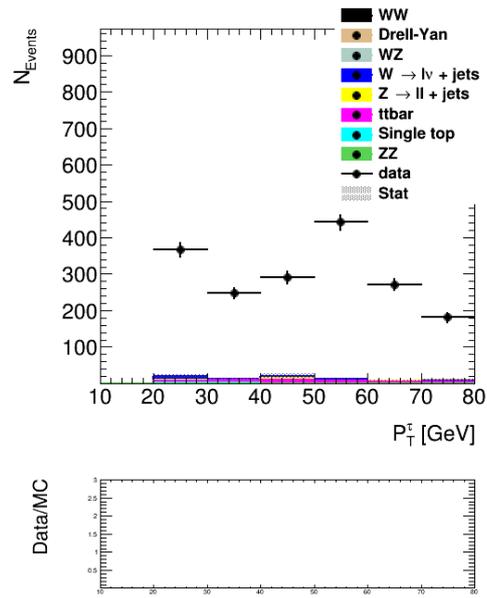
(a)



(b)



(c)



(d)

Figure 9.1.3: P_T^μ for the leading μ^\pm and P_T^τ in N_{JLL} ((a),(b)) N_{JJJ} ((c),(d))

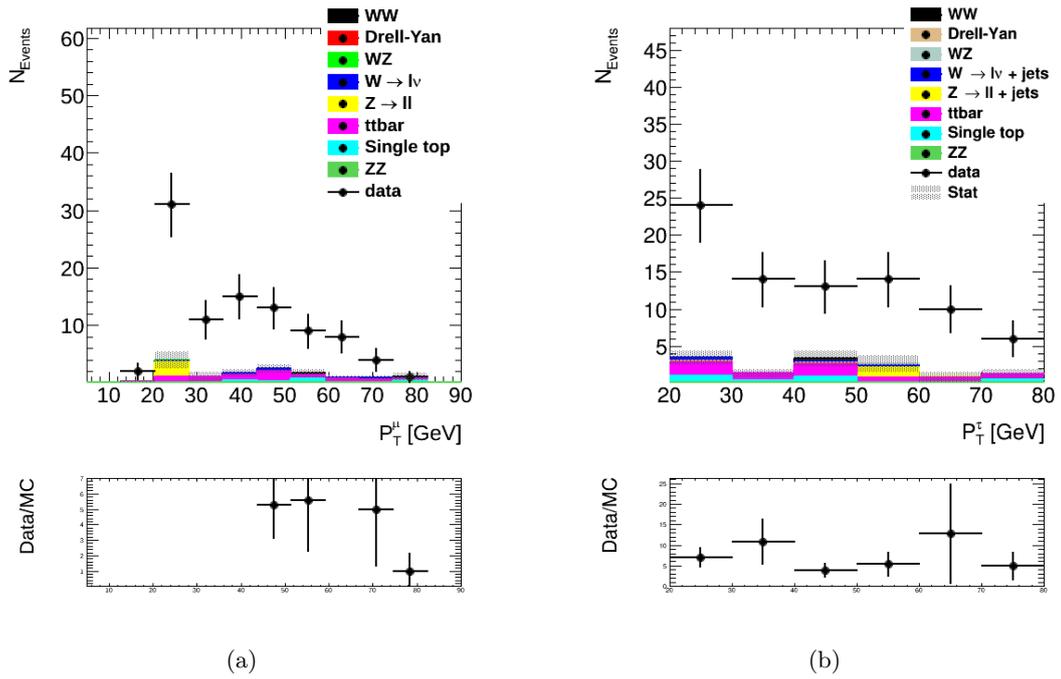


Figure 9.1.4: P_T^μ for the leading μ^\pm and P_T^τ in N_{JJL} ((a), (b))

Using FR_μ and FR_τ in the side-bands

From section 9 it is clear that the EW and non-EW background processes contribute differently to the side-bands. To scale events in the signal region a fake-rate correctly describing the candidates in the side-bands is needed. Thereby the correct number of reducible events in the signal region is predicted.

FR_μ and FR_τ were calculated in two samples as described in section 7.3.1: One sample that is enriched in $Z \rightarrow \mu\mu + jets$ event applicable to EW background, and one sample that is enriched in multi-jets events applicable to the non-EW background.

From the composition it is now possible to determine if the fake-rates from the multi-jets enriched samples or the $Z \rightarrow \mu\mu + jets$ enriched samples should be used in each case. For FR_τ only a very small difference was found, and therefore the FR_τ found in multi-jets should only be used in pure multi-jets samples to keep the uncertainty low.

Figure 9.1.1, 9.1.2 and 9.1.3 show that the side-bands N_{LLJ} , N_{LJJ} , N_{LJL} , N_{JLL} consist mainly of EW background, therefore FR_μ and FR_τ from the $Z \rightarrow \mu\mu + jets$ enriched sample are used. The side-bands N_{JJJ} and N_{JJL} consist mainly of multi-jets and therefore the fake-rates from the multi-jets enriched sample are used. FR_τ has only a small variation between the two samples, but since N_{JJJ} and N_{JJL} is a pure multi-jets region the value from the multi-jets enriched sample is used.

The side-band N_{JLJ} is a mixture of EW background and non-EW background. For FR_τ only a very small difference is seen and therefore the FR_τ from the Z+jet enriched sample is used.

The non-EW background is around half of the events for $P_T^\mu < 40$ GeV, and therefore FR_μ is taken as an average between the two obtained FR_μ for $P_T^\mu < 40$ GeV and the uncertainty due to the fake-rate is found from the average of the upper(/lower) limit. Above 40 GeV the FR_μ from the $Z \rightarrow \mu\mu + jets$ enriched sample is used. A resume of which samples are used to determine FR_μ and FR_τ is given in table 9.1.4.

Side-band	FR_μ fom sample	FR_τ fom sample
$N_{LLJ}, N_{LJJ}, N_{LJL}, N_{JLL}$	$Z \rightarrow \mu\mu + jets$	$Z \rightarrow \mu\mu + jets$
N_{JLJ}	Mixture	$Z \rightarrow \mu\mu + jets$
N_{JJJ}, N_{JJL}	multi-jets	multi-jets

Table 9.1.4: The sample used to determine FR_μ and FR_τ for each side-band.

10

Validation of the data-driven background estimation method

Before estimating the background in the signal region it is necessary to validate the data-driven background estimation and ensure that the predicted background procedure is robust. This is done by verifying the soundness of the procedure on data sets not selected by the final selection for the signal regions.

10.1 Validation of the data-driven background estimation method in $W \rightarrow l\nu + jets$

The goal is to obtain a $W \rightarrow \mu\nu + jets$ region where one fake $\tau_{had-vis}$ side-band region can be tested. This done in a $W \rightarrow \mu\nu + jets$ region, since this is different from the $Z \rightarrow ll + jets$ region where the fake-rates were found. The following cuts are inspired by the ATLAS measurement of the $W \rightarrow l\nu + jets$ cross-section measurement in [67].

- **Trigger:** Passing the trigger EF_mu24i_tight.
- **Tau selection:** Exactly one $\tau_{had-vis}$ passing all tau selection criteria, but with no identification requirement.
- **Veto e^\pm and any bad jets**
- **Muon Selection:** One μ^\pm which pass the muon selection, with a looser isolation of 0.2 (The same cones as reference [67] are used: etcone30/pt, ptcone40/pt)

- **Transverse mass:** A lower threshold at the transverse mass between the μ^\pm and the ν : $M_T^{\mu,\nu} = \sqrt{2P_T^\mu P_T^\nu (1 - \cos(\phi^\mu - \phi^\nu))} = \sqrt{2P_T^\mu \cancel{E}_T (1 - \cos(\phi^\mu - \phi^{\cancel{E}_T}))} > 60 \text{ GeV}$. For the neutrino component \cancel{E}_T is used.
- **A threshold on \cancel{E}_T :** $\cancel{E}_T > 25 \text{ GeV}$
- **SS:** The μ^\pm and the $\tau_{had-vis}$ are required to have same sign of charge. Thereby eliminating events ending up in the signal region of the search.

From this a sample enriched in $W \rightarrow \mu\nu + jets$ is obtained as shown on figure 10.1.1(a) and figure 10.1.1(b).¹ The prediction for the situation with one $\tau_{had-vis}$ is done by the equation:

$$N_L = N_J \cdot FR_\tau$$

N_L is a $\tau_{had-vis}$ passing all the object selection, and N_J is the $\tau_{had-vis}$ candidates passing the jet-like selection from table 7.3.1.

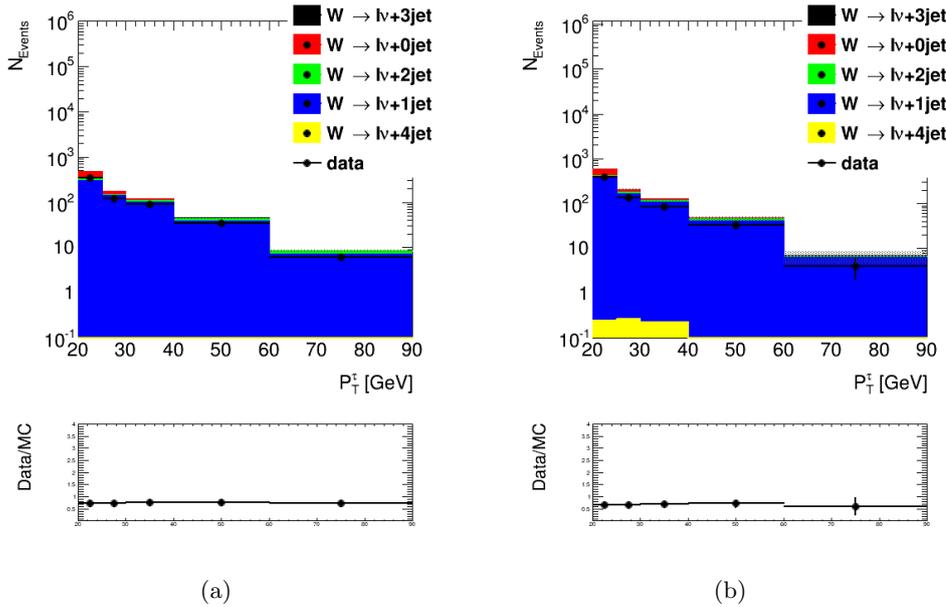


Figure 10.1.1: P_T^τ distributions for the $W \rightarrow \mu\nu + jets$ region for jet-like $\tau_{had-vis}$ multiplied by FR_τ (10.1.1(a)). P_T^τ distributions in the $W \rightarrow \mu\nu + jets$ region for signal $\tau_{had-vis}$ (10.1.1(b))

Figure 10.1.2(a) and 10.1.2(b) show P_T^τ and η^τ for the observed number and the predicted number of $\tau_{had-vis}$ in data only, the values are given in table 10.1.1. The agreement between observed and predicted values is very good, with a difference in total of $9.8 \pm 4.4\%$.

¹For these plots the generator Apgen is used for $W \rightarrow l\nu + jets$, since these samples have a filter applied according to the number of jets. In this test high statistics for $W \rightarrow \mu\nu + 1 \text{ jet}$ is desired and therefore these samples are used.

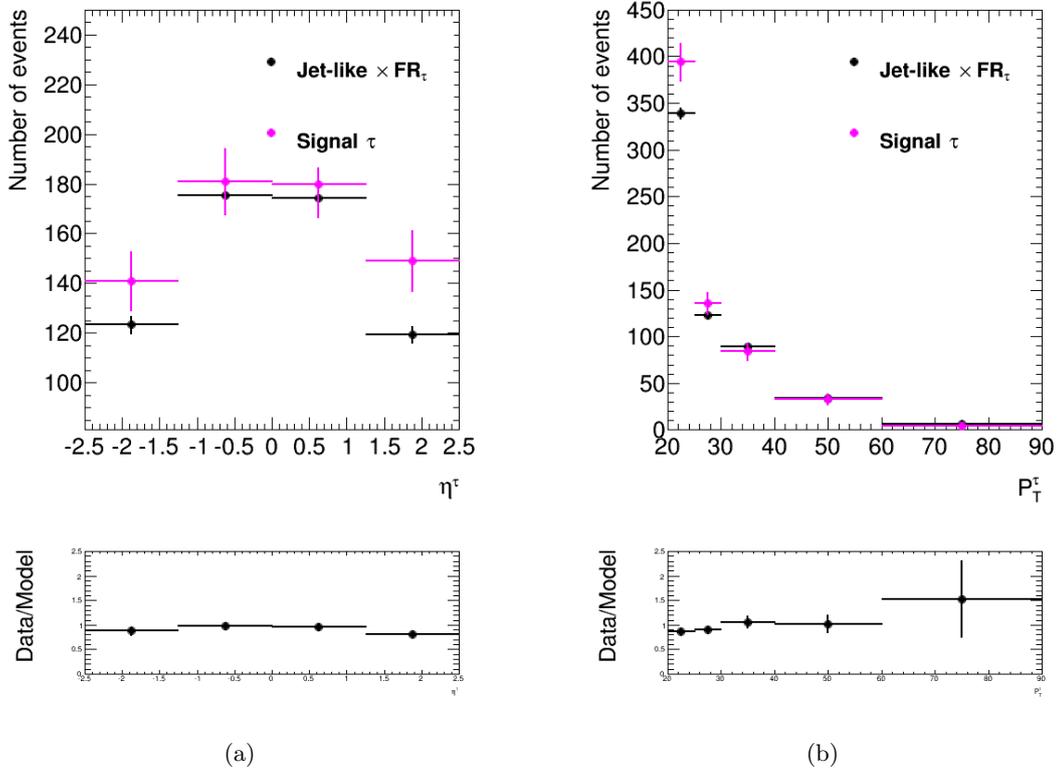


Figure 10.1.2: The predicted (a) and observed (b) number of events as a function of η^τ and P_T^τ .

Observed number of events	649 ± 25
Total number of expected events	590 ± 7.25
Difference between observed and expected	59.0 ± 26.4

Table 10.1.1: The table gives the number of expected and observed events in the W+jet control plot

10.2 Validation of the data-driven estimation method in a SSS control region

A control region is defined to test the data-driven estimation when two leptons are requested. The control region is strongly enriched in $W \rightarrow \mu\nu + jets$ and contains also a good fraction of $t\bar{t}$, and is orthogonal to the search due to a SSS requirement on the charge of the μ^\pm and $\tau_{had-vis}$.

The control region is defined with selection criteria very similar to those applied in section 10.1:

- **Trigger:** Passing the trigger `EF_mu24i_tight`.
- **Veto e^\pm and any bad jets**
- **Muon Selection:** At least one μ^\pm which pass the muon selection, with a looser isolation of 0.2 (The same cones as reference [67] are used: `etcone30/pt`, `ptcone40/pt`)
One more μ^\pm with the same sign of charge must be present, where the charge requirement is to kill $Z \rightarrow \mu\mu + jets$ contamination. No isolation requirement is applied to the additional μ^\pm .
- **Tau selection:** Exactly 1 $\tau_{had-vis}$ passing all tau selection, but no requirement of identification.
This $\tau_{had-vis}$ needs to have same charge as the leading μ^\pm , thus eliminating events ending up in the signal region of the search.
- **Transverse mass:** A lower threshold on the transverse mass between the μ^\pm and the ν : $M_T^{\mu,\nu} = \sqrt{2P_T^\mu P_T^\nu (1 - \cos(\phi^\mu - \phi^\nu))} = \sqrt{2P_T^\mu \cancel{E}_T (1 - \cos(\phi^\mu - \phi^{\cancel{E}}))} > 60000$.
For the neutrino component \cancel{E}_T is used.
- **A threshold on \cancel{E}_T :** $\cancel{E}_T > 25$ GeV

The region obtained from these criteria is shown in figure 10.2.1. The main components are $t\bar{t}$ and $W \rightarrow l\nu + jets$. It is found that $t\bar{t}$ is more important now than in section 10.1, since one more μ^\pm is now required. The data are not well described by simulation due to a contribution from non-EW backgrounds or mismodelling of simulations.

To obtain the predicted number of events the following equation is applied:

$$N_{LL} = N_{TF}f_\tau + N_{FT}\epsilon_\tau f_\mu + N_{FF}f_\mu f_\tau \quad (10.1)$$

The first index denote the μ and the second index the $\tau_{had-vis}$. By combining this equation with the expression for the side-bands N_{JL} , N_{LJ} and N_{LL} , an equation for the data-driven background with three terms is achieved:

$$N_{LL} = N_{JL}FR_\mu + N_{LJ}FR_\tau - N_{JJ}FR_\mu FR_\tau \quad (10.2)$$

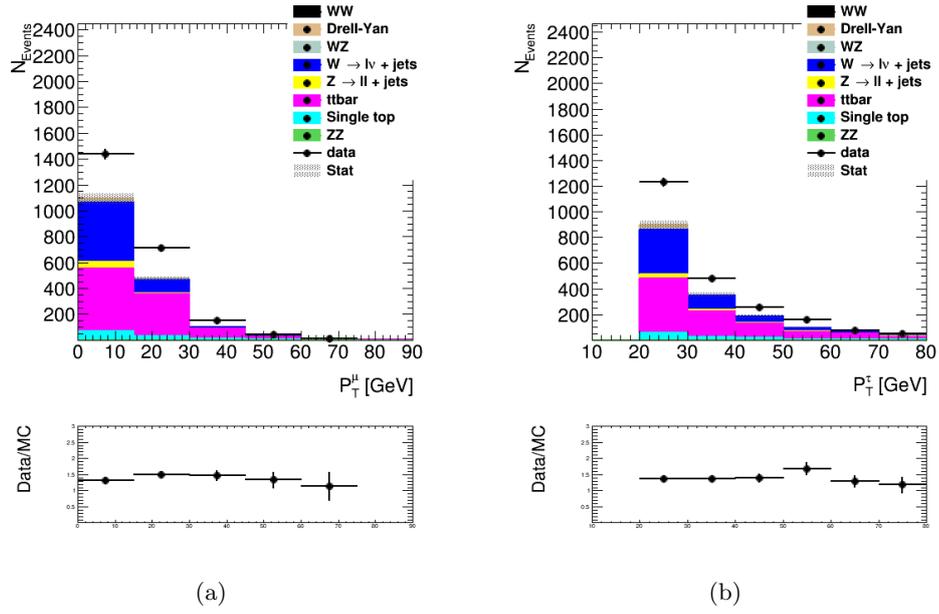


Figure 10.2.1: The SSS region validation

The number of events in each side-band mentioned in equation 10.2 are given in table 10.2.1. Also the number of event in side-bands with an additional $\tau_{had-vis}$ is listed in table 10.2.1. There is a high number of events in the two-object side-band, but for the 3 object case it is more limited. The N_{JLJ} case is 9.9 % of N_{JL} , but N_{JLJ} is also below the statistical uncertainty of N_{JL} . Therefore only the 2 object side-bands are used, which is similar to the procedure for the WH data-driven background estimation method.

2 object side-bands	N_{events}	3 object side-bands	N_{events}	%
N_{LJ}	88 ± 9.3	N_{LJJ}	1 ± 1	1.2 % of N_{LJ}
N_{JL}	121 ± 11.1	N_{JLJ}	12 ± 3.4	9.9 % of N_{JL}
N_{JJ}	2148 ± 46.3	N_{JJJ}	87 ± 9.3	4.1 % of N_{JJ}

Table 10.2.1: Number of events in the SSS side-band regions

Figure 10.2.2 gives the distribution of μ^\pm in the three side-bands. Clearly this region allows testing of a mixture of backgrounds close to that found in the signal region for the search. In this region there are very few events. It should however be noted than the whole region contains 2316 events, whereby it is possible to do the test even though the statistic is limited. The number of events in this test is given in table 10.2.2 and the distributions of the μ^\pm and $\tau_{had-vis}$ are shown in figure 10.2.3.

Observed number of events	5 ± 2.2
Total number of estimated number of events	5.23 ± 1.1

Table 10.2.2: The number of expected and observed events in the SSS test

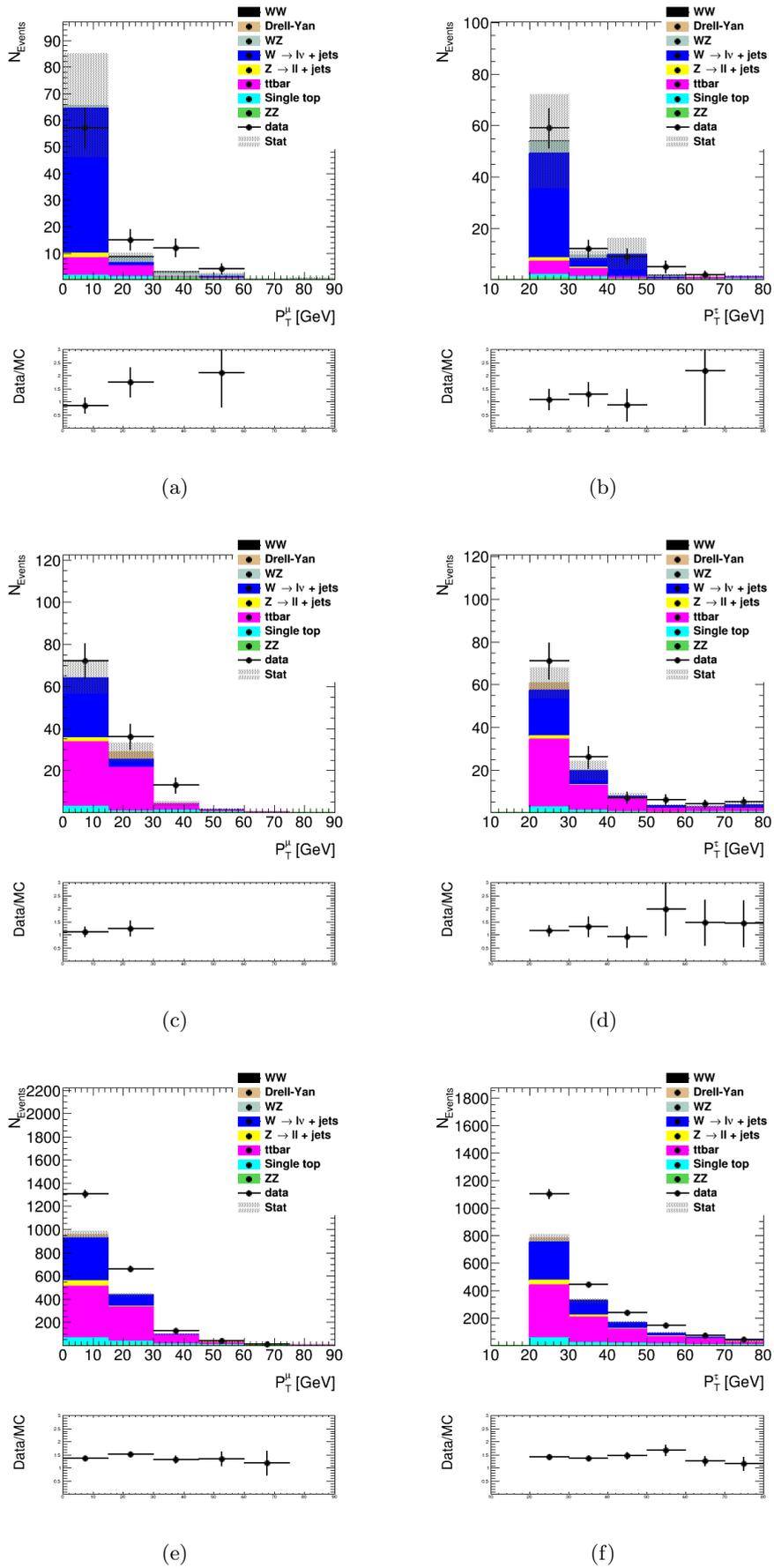


Figure 10.2.2: P_T^H and P_T^L distributions for the side-band regions for N_{LJ} ((a),(b)), N_{LJ} ((c),(d)) and N_{JJ} ((e),(f))

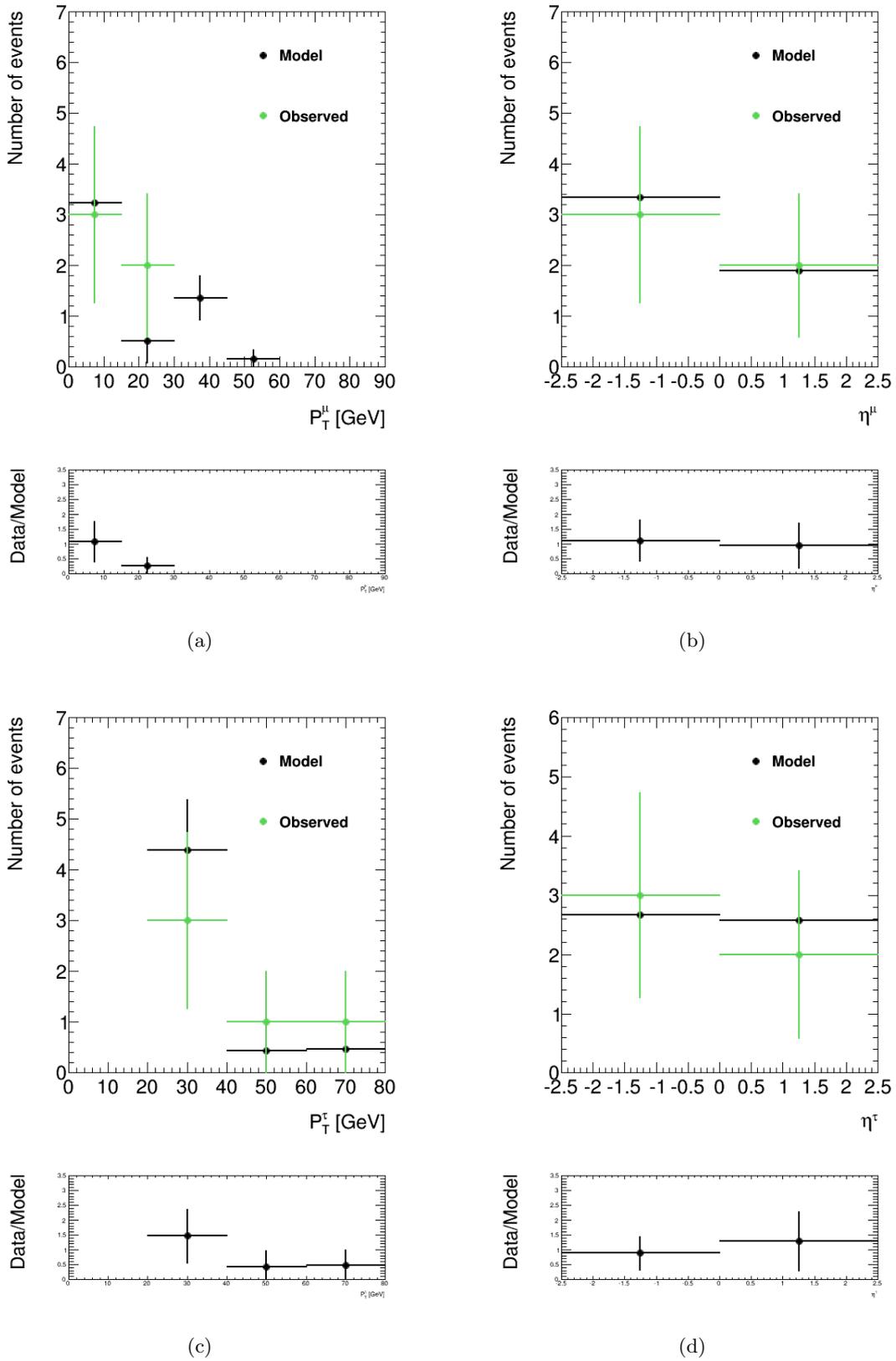


Figure 10.2.3: The distribution of the μ^\pm and $\tau_{had-vis}$ in the SSS test

10.3 Test of the data-driven background estimation method at a pre-selection stage

A test is then done at the pre-selection stage, where there is a high statistic. For the test of the pre-selection the same cuts are used as for the signal region *except* the *charge requirement of μ^\pm 's and $\tau_{had-vis}$* , and the *sum of the three objects P_T* . The same equation can be used as for the background in section 7. There is a small signal contamination in this region, but the relative magnitude is expected to be below 0.1 % and therefore deemed negligible.

The pre-selection region contains mainly $Z \rightarrow \mu\mu + jets$ as shown in figure 10.3.1. The test is applied to estimate the applicability of the method.

The predicted event numbers are plotted in figure 10.3.2 and tabulated in table 10.3.1. A good agreement is achieved by the observed and predicted number of events.

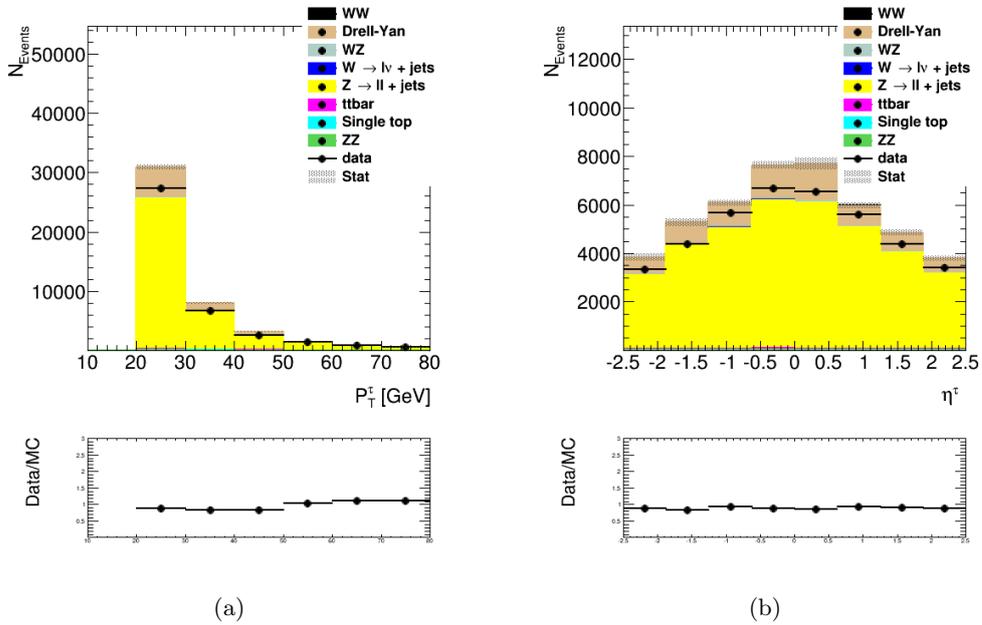


Figure 10.3.1: The distributions of the P_T^τ (d) and η^τ (e)

Observed number of events	$3.45 \times 10^4 \pm 188$
Total estimated number of events	$3.40 \times 10^4 \pm 56$
Fakes	$3.35 \times 10^4 \pm 54$
WZ,ZZ	446 ± 2

Table 10.3.1: The table gives the number of expected and observed events at the pre-selection stage

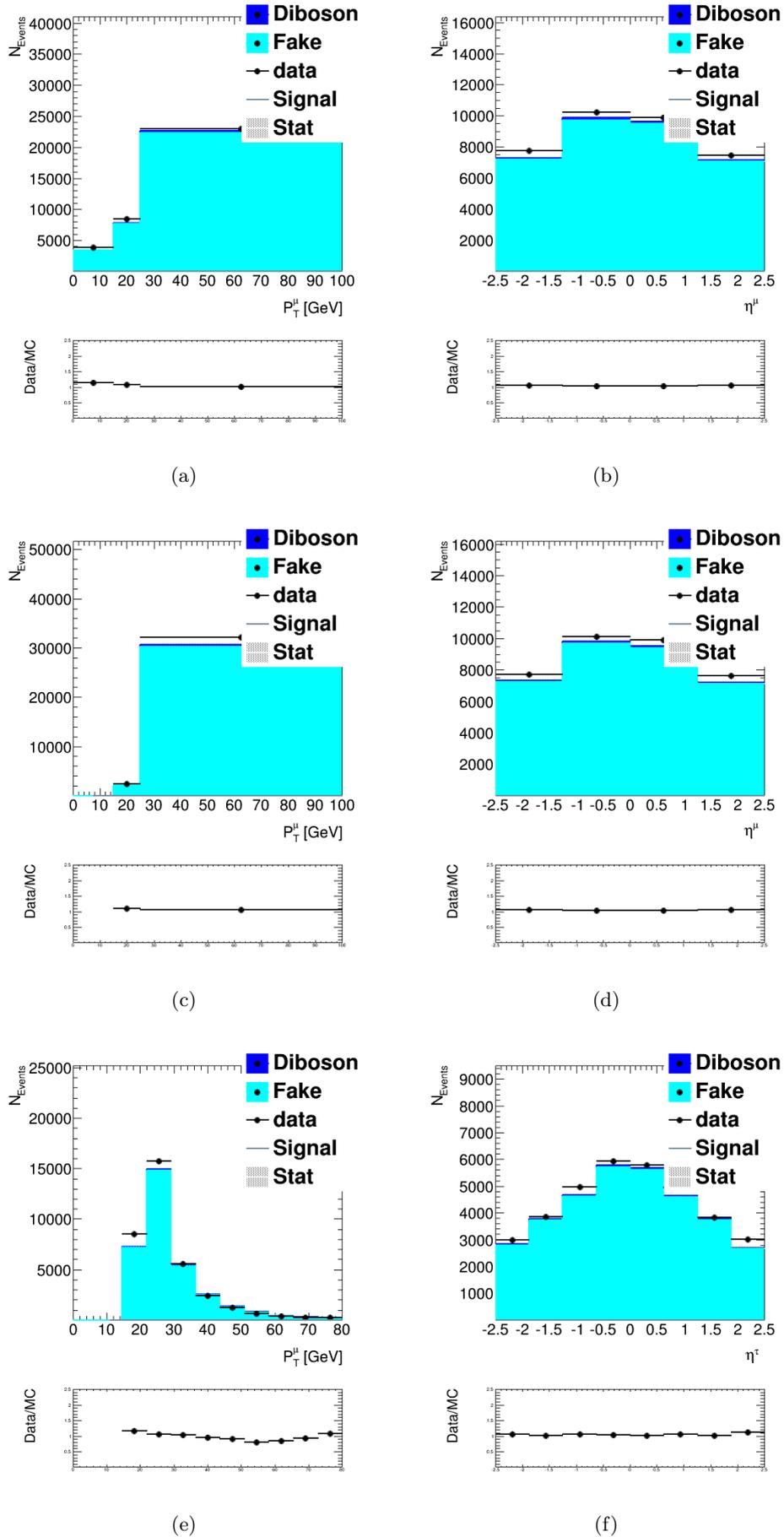


Figure 10.3.2: The distributions of η and P_T for the leading μ^{\pm} ((a) and (b)), sub-leading μ_{25}^{\pm} ((c) and (d)) and the $\tau_{had-vis}$ ((e) and (f))

10.4 Validation of FR_μ

A closure test is done on the FR_μ . This closure test has two purposes: First it checks whether the subtraction of μ^\pm from WZ and ZZ is done correctly, allowing the use of FR_μ and secondly it ensures that the kinematic distributions besides P_T^μ and η^μ are well described. The Z control region from section 8.1 is used where the Z is extracted.

The closure test is done by finding events with a third μ^\pm , and if this μ^\pm is jet-like it is multiplied by FR_μ . The true μ^\pm is from WZ and ZZ is subtracted before applying the FR_μ . The scaled jet-like P_T^μ and η^μ distributions are similar to the μ^\pm distributions attained when applying the request of a third isolated μ^\pm directly to this sample, showing that the FR_μ determination is satisfactory. These plots are shown in Figure 10.4.1 . Figure 10.4.1 shows that $\frac{d\sigma}{\sigma(d0)}$, ϕ^μ and z_0^μ are well described by the fake-rates.

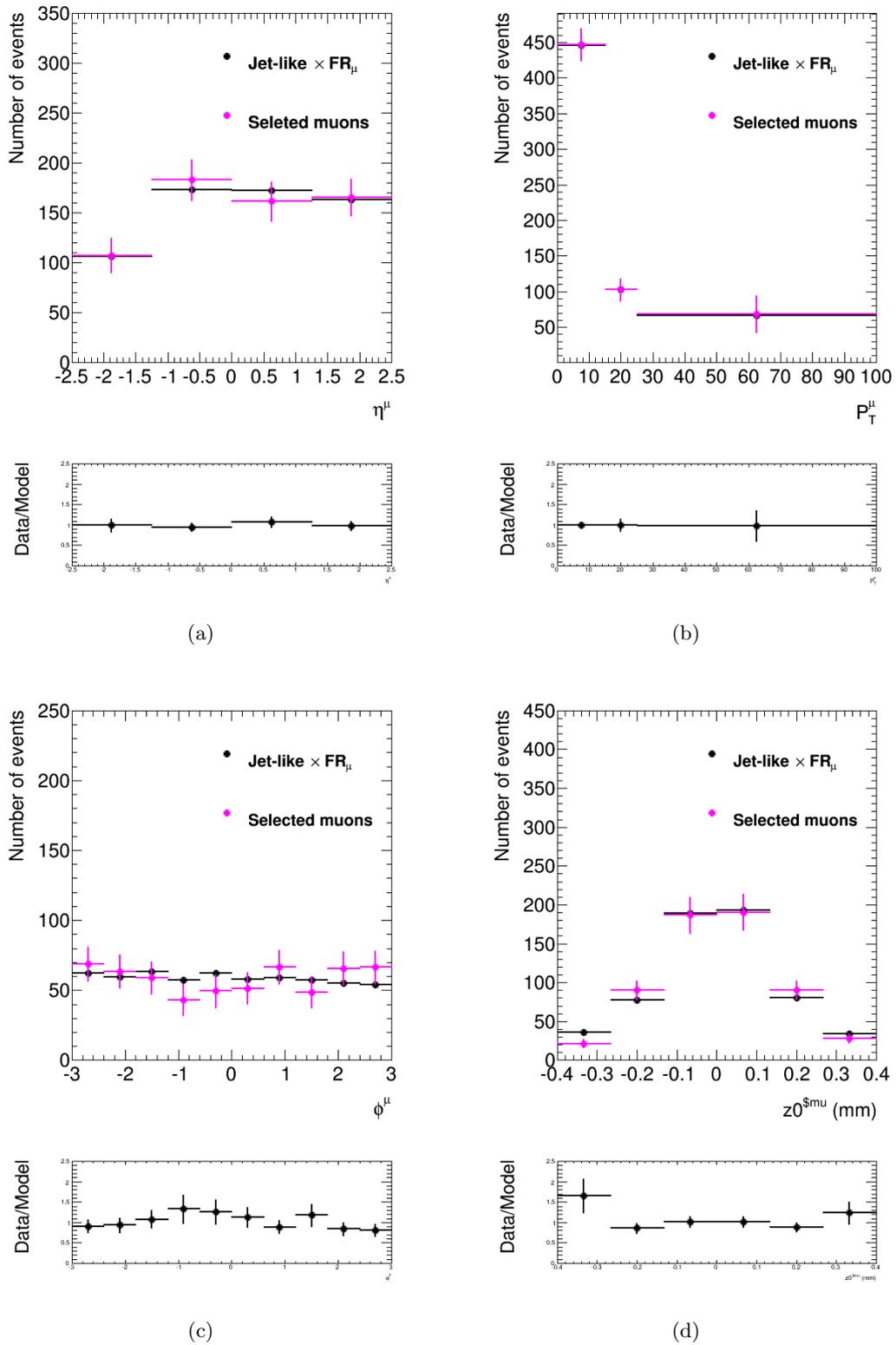


Figure 10.4.1: The closure tests done for of η^μ ((a)), P_T^μ ((b)), ϕ^μ ((c))an z_0^μ ((d)).

11

Results

11.1 Kinematic distributions in the signal region

Notice: Due to the fact that the analysis is internal in ATLAS and still blinded I will not report any data but only expectations.

The distributions in the signal region are shown in figure 11.1.1, figure 11.1.2 and figure 11.1.3. The numbers of events are given in table 11.1.1

Fakes	11.1 ± 1.2
WZ,ZZ	5.4 ± 1.1
Signal	1.13 ± 0.04

Table 11.1.1: The number of expected events in *OSS*

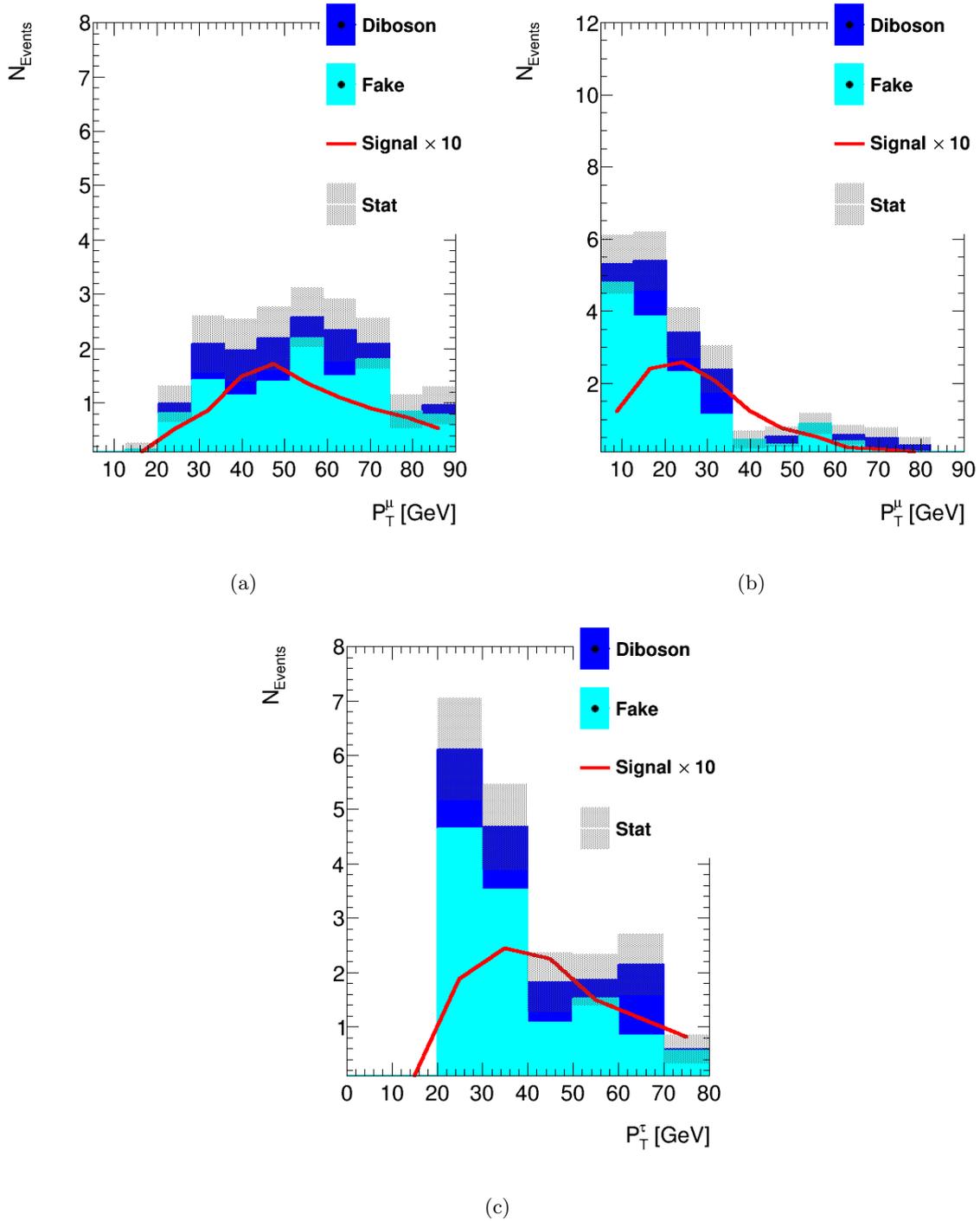
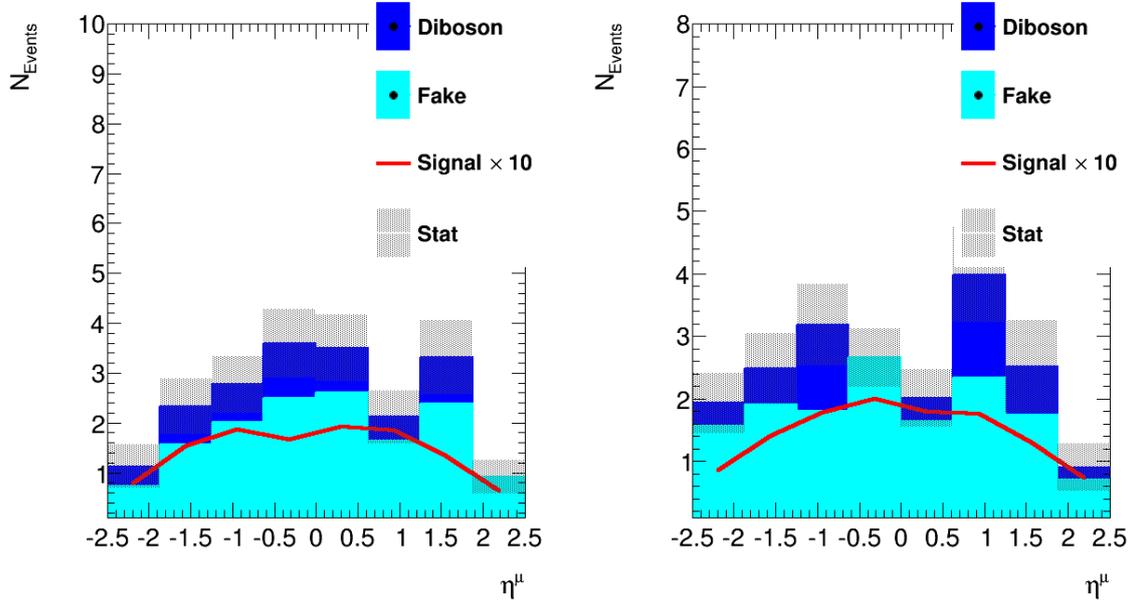
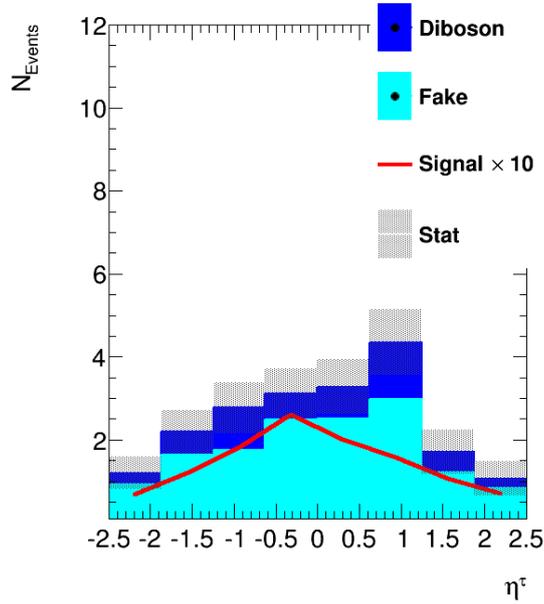


Figure 11.1.1: (a) and (b) shows P_T^μ of the leading and sub-leading μ^\pm . (c) gives P_T^\pm



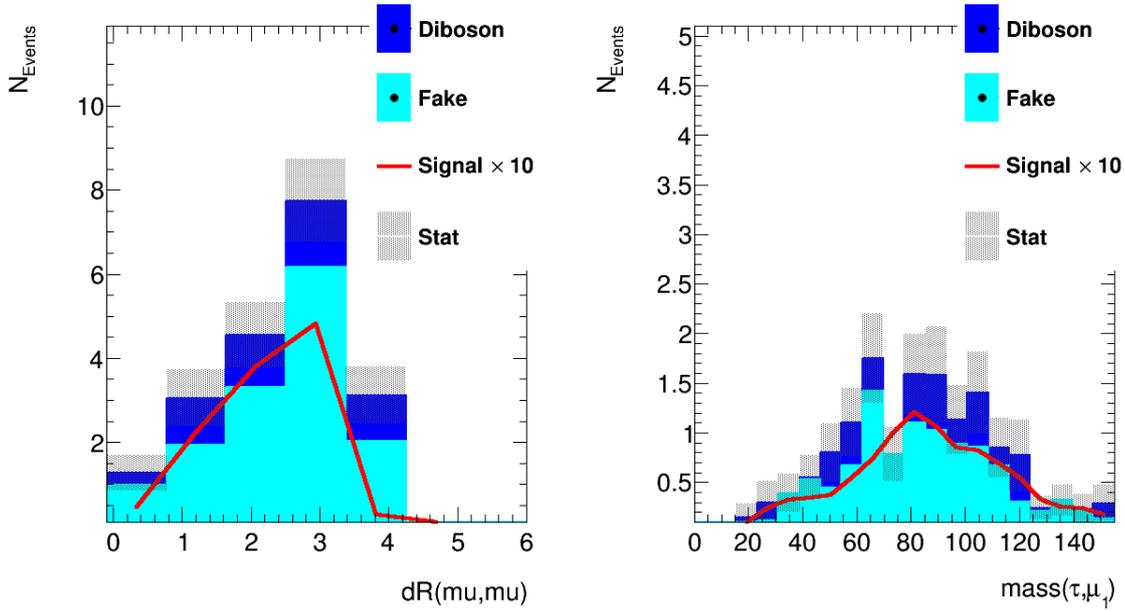
(a)

(b)



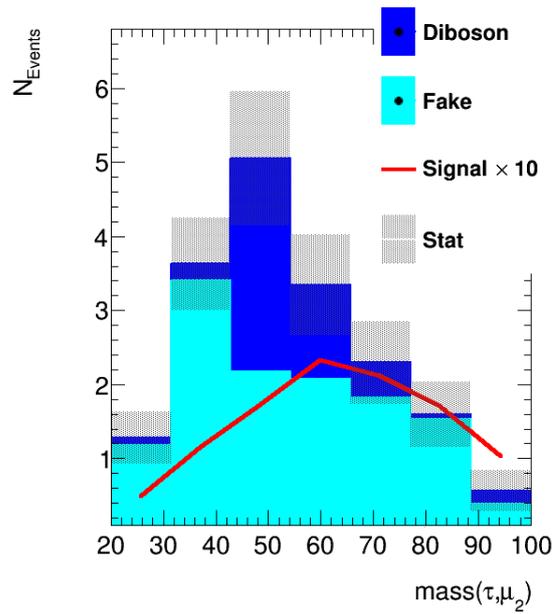
(c)

 Figure 11.1.2: (a) and (b) shows η_T^μ of the leading and sub-leading μ^\pm . (c) gives η_T^τ



(a)

(b)



(c)

Figure 11.1.3: (a) shows ΔR between the two muons, (a) shows mass of the tau and the leading muon and (c) the mass mass of the tau and the sub-leading muon

11.2 Systematic Uncertainties

The main cause of uncertainty in this analysis have proven to be statistical limitations. Even though this is the main contribution, other sources do also give rise to systematic uncertainties. The uncertainties are divided into three cases: uncertainty on the data-driven background estimation, experimental uncertainties, and theoretical uncertainties.

11.2.1 Uncertainty on the data-driven background estimation

The reducible background contributions are determined from FR_μ and FR_τ and these fake-rates are the source of uncertainty for this kind of background.

Uncertainty due to FR_μ : The FR_μ is varied for each μ^\pm according to the uncertainty at a given η^μ and P_T^μ as obtained in section 8.1.2 and 8.2.3. When the average of the two FR_μ is used the average of the upper/lower limit of the two FR_μ is used. The difference in the yield between the upper and lower limit is found and the maximum difference quantify the uncertainty. The uncertainty, stated in table 11.2.1, is the maximal relative difference between the obtained background yields, after subtraction of the di-boson events.

Uncertainty due to FR_τ : The uncertainty due to FR_τ is found similarly as for FR_μ : FR_τ is varied for each $\tau_{had-vis}$ according to the given η^τ and P_T^τ . For the FR_τ found in the QCD samples the FR_τ is used and the uncertainty from table 8.1.4 and figure 8.2.2. Table 11.2.1 gives the maximal relative difference in the background yield by varying the FR_τ .

Uncertainty due to the fake-rate method As seen in figure 10.2.1 there is a disagreement in the first bin between the model and the observed number of events. This disagreement is assigned as an uncertainty, resulting in a 14% disagreement. To be conservative this is used in the whole P_T range.

11.2.2 Experimental uncertainty on simulations

The experimental uncertainties are the consequences of pile up and other effects affecting simulation and data differently, causing the the scale-factors to have an uncertainty. Such

Uncertainty source	Data [%]
FR_μ	8.6
FR_τ	2.5
Method	14

Table 11.2.1: The uncertainty due to the FR_μ and FR_τ .

uncertainty is considered for simulation events which are truth matched to the given process is used. The uncertainties for this kind of background are:

Luminosity The luminosity uncertainty is 2.8% [47].

Pile-up One of the main corrections to MC simulations is pile-up as described in section 2.4.1. The profile of simulation events is scaled to have the same structure as data, and the correction factor is varied within its uncertainty.

Trigger The efficiency of the trigger is different for simulated events than data requiring compensation by a scale-factor. The scale-factor is varied and the main difference from the nominal number is used as the uncertainty.

Uncertainty on $\tau_{had-vis}$:

There are two kinds of systematic uncertainties for the $\tau_{had-vis}$:

- The efficiency differ between MC simulation and data for $\tau_{had-vis}$ identification and reconstruction. The uncertainty on the efficiency scale factor is 2-3 % for 1 prong and 3-5 % for 3 prong.
Likewise a correction is needed in the simulation for the tau the energy scale to be like data. A correction is found by a $Z \rightarrow \tau\tau$ tag and probe analysis [68]. The uncertainty on the tau energy scale is 2-4%, and is varied accordingly.

Uncertainty on μ^\pm ; Systematic uncertainties for μ^\pm is generally lower than for the $\tau_{had-vis}$. It is however still important to consider it:

- Also the μ^\pm have a efficiency scale factor for identification, isolation and reconstruction as described in section 2.5. These efficiency scale-factors are varied within their uncertainty. For simulated events to match the data energy resolution a μ^\pm momentum smearing is done. The resolution of this smearing depends on the P_T and vary between 0.7% for low P_T to 4% for 100 GeV [69].

bjet The b-jet tagging efficiency has been measured for ttbar events in data with a total uncertainty of about $\pm 2\%$ for jets with a transverse momentum up to 100 GeV. The efficiency scale factor is varied influencing the final signal yields [70][71]

11.2.3 Theoretical uncertainties

The theoretical uncertainties are:

QCD scale One of the main theoretical uncertainties is the choice of renormalization scale, which corrects for the ultraviolet divergences. The actual scale is unknown, but it is known that a Higgs particle is produced. Therefore this scale is varied by a factor of 2 around m_H value in MCFM, which is a generator describing the hard process

Uncertainty source	Diboson (%)	signal (%)
Luminosity	2.8	2.8
Pile-up	0.54	0.36
μ^\pm reconstruction, ID and isolation	1.1	2.7
Trigger	1.1	1.1
Muons smearing	1.4	<0.1
$\tau_{had-vis}^\pm$ reconstruction and ID	<0.1	2.6
$\tau_{had-vis}^\pm$ energy scale	6.1	0.84
B-tagging	0.80	0.20

Table 11.2.2: Overview of all uncertainties on di-boson and signal simulated samples

to NLO[40]. The factorization scale corrects for the infrared divergences, and is also varied by a factor of 2 around m_H value. The uncertainty extracted from this is the maximum found over the entire mass range in [72] which is 1.1% for WH. These numbers are provided by the LHC Higgs Cross Section Working Group.[72] On the di-boson background an overall uncertainty due this effect is 4%.

Parton distribution function The available PDF are found by different groups and have small differences. In MCFM the PDF was varied and in [72] found a maximum difference of 2.4 % in the entire mass range. On the di-boson background an overall uncertainty due this effect is 5%.

Branching ratio of the $\tau_{had-vis}$ The branching ratio of the $H \rightarrow \tau\tau$ is found in [72]. The uncertainty is 5.6 % .

EW correction: The signal samples Powheg+Pythia samples are simulating the QCD at NLO and the EW at LO. The simulation done NLO for EW can be done via the generator HAWK. [73] By comparing the P_T of the Higgs for the difference due to this higher order can be found. This technique was used for the $gg \rightarrow H$, since the change in the P_T^{Higgs} is very important for the quark loop in this production. The effect is 0.8% on $m_h = 125$ GeV

11.3 Expected limits

In order to properly evaluate the results the invariant mass histogram of the sub-leading μ^\pm and the $\tau_{had-vis}$ needs to be examined more closely. The expected number of events in each bin n_i can be written as

$$E[n_i] = \mu s_i + b_i \quad (11.1)$$

Where s_i and b_i are the number of signal and background events in bin n_j respectively. μ is the signal strength and a $\mu = 1$ means that the observed number of events follow the signal

hypothesis whereas $\mu = 0$ means the background fully describe the situation. s_i and b_i can be written as

$$s_i = s_{tot} \int_{bini} f_s(x, \theta_s) dx \quad \wedge \quad b_i = b_{tot} \int_{bini} f_b(x, \theta_b) dx \quad (11.2)$$

s_{tot} and b_{tot} are the total number of signal and background events. $f_s(x, \theta_s)$ and $f_b(x, \theta_b)$ are the probability density functions (pdfs) for signal and background events (with a variable x). θ_s and θ_b are the nuisance parameteres. There is a nuisance parameter for each uncertainty eg. listed in table 11.2.2 . [74] [75] Then it is possible to write the likelyhood function $L(\mu, \theta)$:

$$L(\mu, \theta) = \prod_{i=1}^N \frac{(\mu s_i + b_i)^{n_j}}{n_j!} e^{-\mu s_i + b_i} \quad (11.3)$$

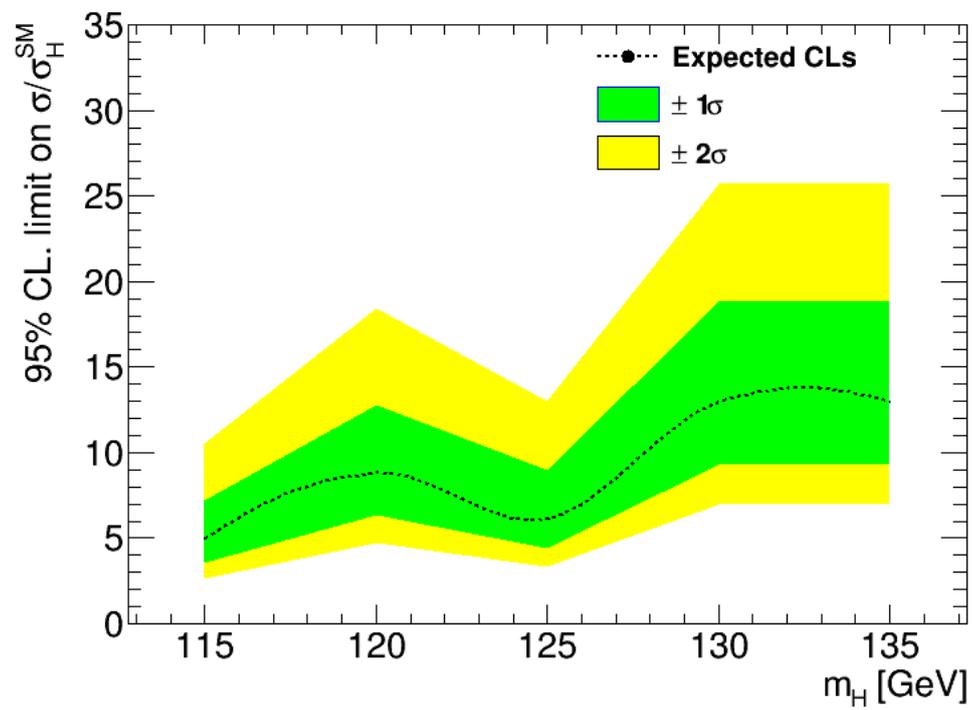
To test the likely-hood one can investigate the ratio of likely-hoods:

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})} \quad (11.4)$$

$\hat{\theta}$ is the θ value which gives the maximum L for a specific μ . (conditional) $\hat{\mu}$ and $\hat{\theta}$ give the maximum L for a non-specific μ . (Unconditional) It is then useful to introduce the p-value

$$p_u = \int_{t_u, obs}^{\inf} f(t_\mu | \mu) dt_\mu \quad (11.5)$$

Where $t_\mu = -2 \ln \lambda(\mu)$. A p-value on 0.05 corresponds to a 95% confidence level.[74] [75] Figure 11.3.1 shows the expected limit plot as a function of the different Higgs masses.

Figure 11.3.1: Expected limit depending on the m_H hypothesis

12

Conclusion

This thesis presents an improvement of tau identification and a search for the SM Higgs produced in association with a vector boson.

The tau identification for hadronic decaying taus is done by using a recently available algorithm which is able to reconstruct π^0 's. From this π^0 reconstruction three new variables are defined, which have then been used to separate jets from taus. The variables are found to improve tau identification in the full P_T range. Furthermore it allows for the lowering of the tau identification threshold from 15 GeV to 20 GeV. These two achievements have been used in ATLAS to give a higher signal yield in several tau searches.

The search for the SM Higgs produced in association with a vector boson is performed via the final state with two muons and one hadronic decaying tau. A data-driven background estimation for the reducible background and simulation for estimating the irreducible background.

This method has been used due to various reasons. One important point is that the simulations do not provide the necessary statistical information in the signal region. Another reason is that the tail of the isolation distributions in the simulation are not well modelled. A robust background estimation is instead obtained by the combination of simulation and data-driven background. This background estimation method has been proven to foresee the expected number of data in various tests for events close to the signal region in terms of background estimation. Of special importance is a validation test in a region where all three final objects have the same-sign, and it is found to nicely foresee the observed number of events. This method and the muon fake-rates determined in this work will be used for the ATLAS final result in this production channel, to be unblinded and published later this year.

Bibliography

- [1] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys.Lett. **12** (1964) 132–133.
- [2] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156–1163. <http://link.aps.org/doi/10.1103/PhysRev.145.1156>.
- [3] G. S. Guralnik et al., *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587. <http://link.aps.org/doi/10.1103/PhysRevLett.13.585>.
- [4] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323. <http://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [5] ATLAS Collaboration Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Tech. Rep. CERN-PH-EP-2013-102, CERN, Geneva, Juli, 2013.
- [6] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266. <http://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [7] M. E. Peskin et al., *An Introduction To Quantum Field Theory (Frontiers in Physics), Year = 1995*. ISBN 9780201503975.
- [8] M. E. Peskin et al., *Hight P_T Particle Physics at Hadron Colliders*. Cambridge University Press, 2005.
- [9] I. van Vulpen, *The Standard Model Higgs Boson Part of the Lecture Particle Physics II , (Lecture notes)*, 2014.
- [10] A. Pich, *The Standard model of electroweak interactions*, arXiv:0705.4264 [hep-ph].
- [11] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics*, Phys. Rev. D **86** (2012) 010001. <http://link.aps.org/doi/10.1103/PhysRevD.86.010001>.
- [12] CMS Collaboration Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].
- [13] ATLAS Collaboration Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

- [14] M.Kado, *Higgs Physics in ATLAS (talk)*, 2014.
- [15] M. Trottier-McDonald, *Evidence for the Standard Model Higgs Boson Decaying to One Semi-hadronically Decaying Tau and One Leptonically Decaying Tau at ATLAS*, 2014.
- [16] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3.Higgs Properties*; Tech. Rep. CERN-2013-004, CERN, Geneva, Jul, 2013.
- [17] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*; Tech. Rep. CERN-2012-002, CERN, Geneva, Jan, 2012.
- [18] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*; Tech. Rep. CERN-2011-002, CERN, Geneva, Jan, 2011.
- [19] ATLAS Collaboration Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys.Lett. B **726** (2014) 032, [arXiv:1307.1427](#) [hep-ph] [hep-ph].
- [20] C. Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, [arXiv:1207.7235](#) [hep-ex].
- [21] ATLAS Collaboration Collaboration, *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb⁻¹ of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, March, 2014.
- [22] ATLAS Collaboration Collaboration, *Performance of the Reconstruction and Identification of Hadronic τ Decays in ATLAS with 2011 Data*, Tech. Rep. ATLAS-CONF-2012-142, CERN, Geneva, Oct, 2011.
- [23] ATLAS Collaboration Collaboration, *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2013-108, CERN, Geneva, Nov, 2013.
- [24] S. Dittmaier and M. Schumacher, *The Higgs Boson in the Standard Model, From LEP to LHC: Expectations, Searches, and Discovery of a Candidate*, Prog.Part.Nucl.Phys **70** (2013) 082, [arXiv:1211.4828](#) [hep-ph] [hep-ph].
- [25] The CMS Collaboration, *Search for the standard model Higgs boson produced in association with W and Z bosons in pp collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. CERN-PH-EP-2012-253, CERN, Geneva, Sep, 2012.
- [26] <http://cern.ch>, 2014.
- [27] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [28] ALICE Collaboration Collaboration, B. Alessandro et al., *ALICE: Physics performance report, volume II*, J.Phys. **G32** (2006) 1295–2040.

-
- [29] CMS Collaboration Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [30] LHCb Collaboration Collaboration, J. Alves, A. Augusto et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [31] LHCf Collaboration Collaboration, O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.
- [32] TOTEM Collaboration Collaboration, G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.
- [33] M. A.D. et al., *Parton distributions for the LHC*, Eur.Phys.J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].
- [34] A. Bhatti et al., *Jet Physics at the Tevatron*, Ann.Rev.Nucl.Part.Sci. **60** (2010) 267–297, arXiv:1002.1708 [hep-ex].
- [35] T. Sjostrand, S. Mrenna, and P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Comm **178** (2008) 028, arXiv:hep-ph/0710.3820.
- [36] <https://sherpa.hepforge.org/trac/wiki/SherpaManual>.
- [37] M. L. Mangano, F. Piccinini, A. D. Polosa, M. Moretti, and R. Pittau, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, Journal of High Energy Physics **7** (2003) 1, hep-ph/0206293.
- [38] G. Corcella et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, JHEP **01** (2001) 010, arXiv:hep-ph/0011363.
- [39] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].
- [40] J. M. Campbell et al., *MCFM for the Tevatron and the LHC*, Tech. Rep. arXiv:1007.3492. FERMILAB-CONF-10-244-T, Jul, 2010. Comments: Talk presented by R.K Ellis at Loops and Legs in Quantum Field Theory 2010, Woerlitz, Germany, April 25-30, 2010, (6 pages and 4 figures).
- [41] N. Davidson et al., *PHOTOS Interface in C++: Technical and Physics Documentation*, arXiv:1011.0937 [hep-ph].
- [42] *The tau decay library TAUOLA: Version 2.4*, Comput.Phys.Comm. **76** (1993) 361–380.
- [43] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303.
- [44] <http://atlas.ch>, 2014.

- [45] I. Vaynman, *ATLAS E-M Calorimeter Resolution and Neural Network Based Particle Classification (talk)*, 2014.
- [46] R. Voss et al., *The CERN Large Hadron Collider, accelerator and experiments vol. 1, JINST 3*,.
- [47] ATLAS Collaboration Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** no. arXiv:1302.4393. CERN-PH-EP-2013-026, (2013) 2518. 40 p.
- [48] V. Balagura, *Notes on van der Meer Scan for Absolute Luminosity Measurement*, Nucl.Instrum.Meth. **A654** (2011) 634–638, arXiv:1103.1129 [physics.ins-det].
- [49] ATLAS Collaboration Collaboration, *Measurement of the production cross section for W^- bosons in association with jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys.Lett. **B698** (2011) 325–345, arXiv:1012.5382 [hep-ex].
- [50] ATLAS Collaboration Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur.Phys.J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].
- [51] ATLAS Collaboration Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur.Phys.J. **C74** (2014) 2941, arXiv:1404.2240 [hep-ex].
- [52] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008.
- [53] Cacciari et al., *The Anti- $k(t)$ jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].
- [54] ATLAS Collaboration Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*,.
- [55] “J. Beringer et al. (Particle Data Group), Phys. Rev. D86, 010001 (2012).”.
- [56] M. Trottier-McDonald, *Identifying hadronic tau decays using calorimeter topological clusters at atlas*, 2010.
- [57] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008.
- [58] ATLAS Collaboration Collaboration, *Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2013-064, CERN, Geneva, Jul, 2013.

-
- [59] M. L., “https://www.uvic.ca/science/physics/vispa/research/projects/atlas/Lefebvre_Split2010.pdf, LHC Days in split, 2010.”.
- [60] A. H. others, *TMVA - Toolkit for Multivariate Data Analysis*, ArXiv Physics e-prints (2007), physics/0703039.
- [61] H. Voss, *(Talk) Multivariate Data Analysis and Machine Learning in High Energy Physics (IV)*, 2009.
- [62] ATLAS Collaboration Collaboration, *(Under preparation) Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment at sqrt(8) TeV*,.
- [63] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of hard double-parton interactions in $W(\rightarrow l\nu) + 2$ jet events at $\sqrt{s}=7$ TeV with the ATLAS detector*, New J.Phys. **15** (2013) 033038, arXiv:1301.6872 [hep-ex].
- [64] ATLAS Collaboration Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2014-004, CERN, Geneva, Feb, 2014.
- [65] ATLAS Collaboration Collaboration, *Measurement of total ZZ production cross section in the four-lepton channel in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2013-020,, CERN, Geneva, Nov, 2013.
- [66] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PerformanceDPD#DESD_CALJEThttps://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PerformanceDPD#DESD_CALJET.
- [67] ATLAS Collaboration Collaboration, *Measurement of the $W \rightarrow l\nu$ and $Z/\gamma^* \rightarrow \ell\ell$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **1012** (2010) 060, arXiv:1010.2130 [hep-ex].
- [68] ATLAS Collaboration Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC in 2011*, Tech. Rep. ATLAS-CONF-2012-054, CERN, Geneva, Jun, 2012.
- [69] ATLAS Collaboration Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, arXiv:1407.3935 [hep-ex].
- [70] ATLAS Collaboration Collaboration, *Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2014-046, CERN, Geneva, Feb, 2014.

- [71] ATLAS Collaboration Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011.
- [72] <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt8TeV>.
- [73] <http://omnibus.unifreiburg.de/sd565/programs/hawk/hawk.html>.
- [74] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, European Physical Journal C **71** (2011) 1554, [arXiv:1007.1727](https://arxiv.org/abs/1007.1727) [physics.data-an].
- [75] ATLAS Collaboration Collaboration, *Expected Performance of the ATLAS Experiment - Statistical Combination of Several Important Standard Model Higgs Boson Search Channels*, [arXiv:0901.0512](https://arxiv.org/abs/0901.0512) [hep-ex].

Appendices



Side-bands equation

Using the notation, N'_{XXX} , of the side-band after subtraction of the di-boson events:

$$N'_{XXX} = N_{XXX} - N_{XXX}^{Diboson} \quad (\text{A.1})$$

For the case of three objects the equations of the side-bands are given here:

$$\begin{aligned} N'_{LLJ} &= -N_{TTF}(f_\tau - 1) - N_{FFT}f_\mu^2(\epsilon_\tau - 1) - N_{FFF}f_\mu^2(f_\tau - 1) - N_{FTT}f_\mu(\epsilon_\tau - 1) \\ &\quad - N_{FTF}f_\mu(f_\tau - 1) - N_{TFT}f_\mu(\epsilon_\tau - 1) - N_{TFF}f_\mu(f_\tau - 1) \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} N'_{LJJ} &= -N_{TFT}\epsilon_\tau(f_\mu - 1) - N_{TFF}f_\tau(f_\mu - 1) - N_{FFT}\epsilon_\tau f_\mu(f_\mu - 1) \\ &\quad - N_{FFF}f_\tau f_\mu(f_\mu - 1) \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} N'_{JLL} &= -N_{FTT}\epsilon_\tau(f_\mu - 1) - N_{FTF}f_\tau(f_{\mu,2} - 1) - N_{FFT}\epsilon_\tau f_\mu(f_\mu - 1) \\ &\quad - N_{FFF}f_\tau f_\mu(f_\mu - 1) \end{aligned} \quad (\text{A.4})$$

$$\begin{aligned} N'_{LJJ} &= N_{TFT}(\epsilon_\tau - 1)(f_\mu - 1) + N_{TFF}(f_\tau - 1)(f_\mu - 1) \\ &\quad + N_{FFT}f_\mu(\epsilon_\tau - 1)(f_\mu - 1) + N_{FFF}f_\mu(f_\tau - 1)(f_\mu - 1) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} N'_{JLL} &= N_{FTT}(\epsilon_\tau - 1)(f_\mu - 1) + N_{FTF}(f_\tau - 1)(f_\mu - 1) \\ &\quad + N_{FFT}f_\mu(\epsilon_\tau - 1)(f_\mu - 1) + N_{FFF}f_\mu(f_\tau - 1)(f_\mu - 1) \end{aligned} \quad (\text{A.6})$$

$$N'_{JLL} = N_{FFT}\epsilon_\tau(f_\mu - 1)(f_\mu - 1) + N_{FFF}f_\tau(f_\mu - 1)(f_\mu - 1) \quad (\text{A.7})$$

$$N'_{JJJ} = -N_{FFT}(\epsilon_\tau - 1)(f_\mu - 1)(f_\mu - 1) - N_{FFF}(f_\tau - 1)(f_\mu - 1)(f_\mu - 1) \quad (\text{A.8})$$

B

Truth parts contamination

The data-driven background estimation was found in equation 7.2, and due to the importance repeated here with the necessary terms.

$$\begin{aligned} N_{LLL} = & N_{TTF}\epsilon_\mu\epsilon_\mu f_\tau + N_{TFT}\epsilon_\mu f_\mu\epsilon_\tau + N_{TFF}\epsilon_\mu f_\mu f_\tau + N_{FTF}f_\mu f_\tau\epsilon_\mu + N_{FFT}f_\mu f_\mu\epsilon_\tau \\ & + N_{FFF}f_\mu f_\mu f_\tau + N_{FTT}f_\mu\epsilon_\mu\epsilon_\tau \end{aligned} \quad (\text{B.1})$$

Each of these true terms are an important piece of the background estimation, and therefore it is investigated which side-band are important in each term.

For the following is introduced the notation for the side-bands after subtraction of the diboson events:

$$N'_{XXX} = N_{XXX} - N_{XXX}^{Diboson} \quad (\text{B.2})$$

For each true term are produced two figures: One figure shows the contribution of a given side-band to the truth term, and the other figure shows the absolute values. The distributions show P_T^μ of the leading μ^\pm . From these plots it is possible to compare the size of all contributions independent of sign. The equation for each true term and the reference to the correct figure is given in B.0.1

The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

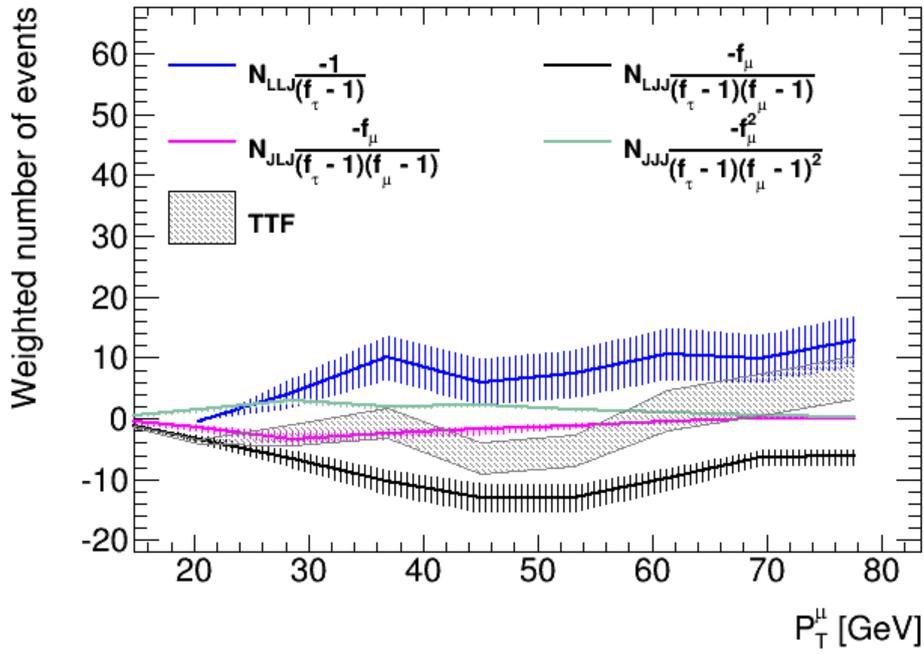
In all of the above the side-bands N_{JJJ} and N_{JJL} contribute to the truth terms, which is critical since these only consist of QCD jets and the applied fake-rate have a high uncertainty.

True term	Equation	Figure
N_{TTF}	$= N'_{LLJ} \frac{1}{f_\tau - 1} - N'_{JLJ} \frac{f_\mu}{(f_\tau - 1)(f_\mu - 1)}$ $- N'_{LJJ} \frac{f_\mu}{(f_\tau - 1)(f_\mu - 1)} - N'_{JJJ} \frac{f_\mu^2}{(f_\tau - 1)(f_\mu - 1)^2}$	B.0.1(a) B.0.1(b)
N_{TFT}	$= N'_{LJJ} \frac{f_\tau}{(\epsilon_\tau - f_\tau)(f_\mu - 1)} + N'_{LJL} \frac{f_\tau - 1}{(\epsilon_\tau - f_\tau)(f_\mu - 1)}$ $- N'_{JLJ} \frac{(f_\mu - f_\tau f_\mu)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$ $+ N'_{JJJ} \frac{f_\tau f_\mu}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$	B.0.2(a) B.0.2(b)
N_{FTT}	$= N'_{JLJ} \frac{f_\tau}{(\epsilon_\tau - f_\tau)(f_\mu - 1)} + N'_{JLL} \frac{(f_\tau - 1)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)}$ $- N'_{JLJ} \frac{(f_\mu - f_\tau f_\mu)}{((\epsilon_\tau - f_\tau)(f_\mu - 1)^2)}$ $+ N'_{JJJ} \frac{f_\tau f_\mu}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$	B.0.3(a) B.0.3(b)
N_{TFF}	$= -N'_{LJJ} \frac{\epsilon_\tau}{(\epsilon_\tau - f_\tau)(f_\mu - 1)} - N'_{LJL} \frac{(\epsilon_\tau - 1)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)}$ $- N'_{JLJ} \frac{f_\mu(\epsilon_\tau - 1)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$ $- N'_{JJJ} \frac{\epsilon_\tau f_\mu}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$	B.0.4(a) B.0.4(b)
N_{FTF}	$= -N'_{JLJ} \frac{\epsilon_\tau}{((\epsilon_\tau - f_\tau)(f_\mu - 1))} - N'_{JLL} \frac{(\epsilon_\tau - 1)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)}$ $- N'_{JLJ} \frac{f_\mu(\epsilon_\tau - 1)}{((\epsilon_\tau - f_\tau)(f_\mu - 1)^2)}$ $- N'_{JJJ} \frac{\epsilon_\tau f_\mu}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$	B.0.5(a) B.0.5(b)
N_{FFT}	$= N'_{JJJ} \frac{-f_\tau}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2} - N'_{JLJ} \frac{(f_\tau - 1)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$	B.0.6(a) B.0.6(b)
N_{FFF}	$= N'_{JJJ} \frac{\epsilon_\tau}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2} + N'_{JLJ} \frac{(\epsilon_\tau - 1)}{(\epsilon_\tau - f_\tau)(f_\mu - 1)^2}$	B.0.7(a) B.0.7(b)

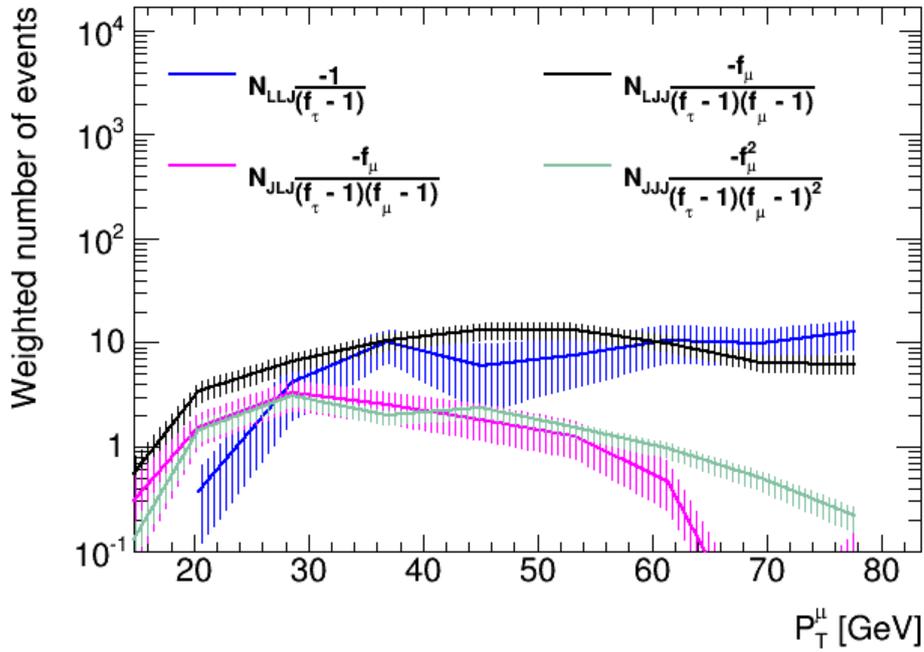
Table B.0.1: The different true terms expressed from the side-bands

In most cases, however the contribution from N_{JJJ} is small, and often more or less cancel by N_{JLL} .

Table 9.1.3 gives the total number of events found in each side-band where both the systematic due to statistical and the fake-rate are taken into consideration. It is found from the integrals of the previous plots. The numbers in the table shows that in the N_{FFF} and N_{FFT} term have a very high uncertainty, since are the important (and only) contributions N_{JJJ} and N_{JLL} . According to equation B.1 these terms will be multiplied by two or tree fake-rates and therefore the contribution will be small in the signal region.

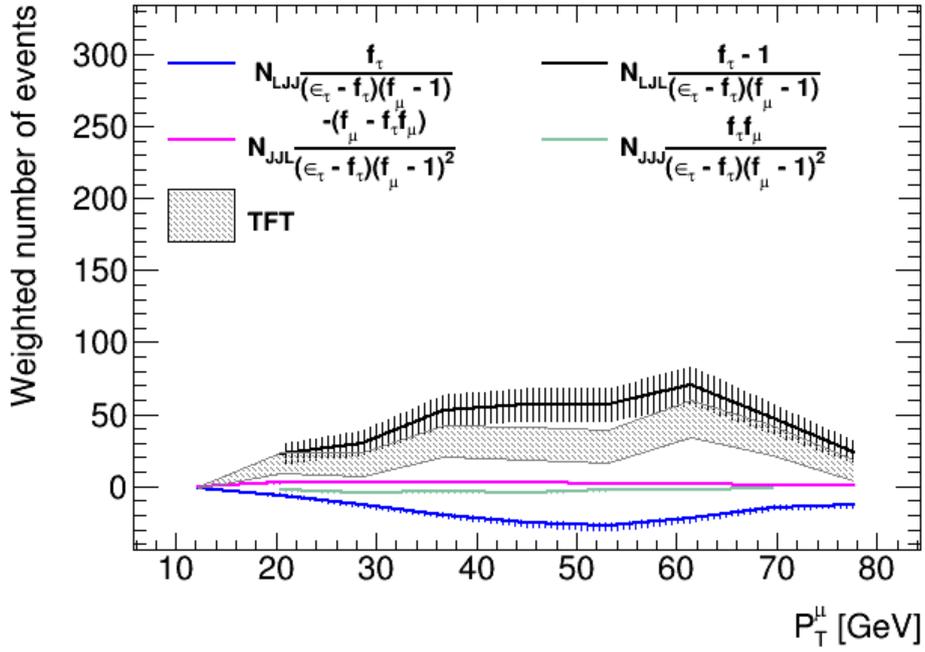


(a)

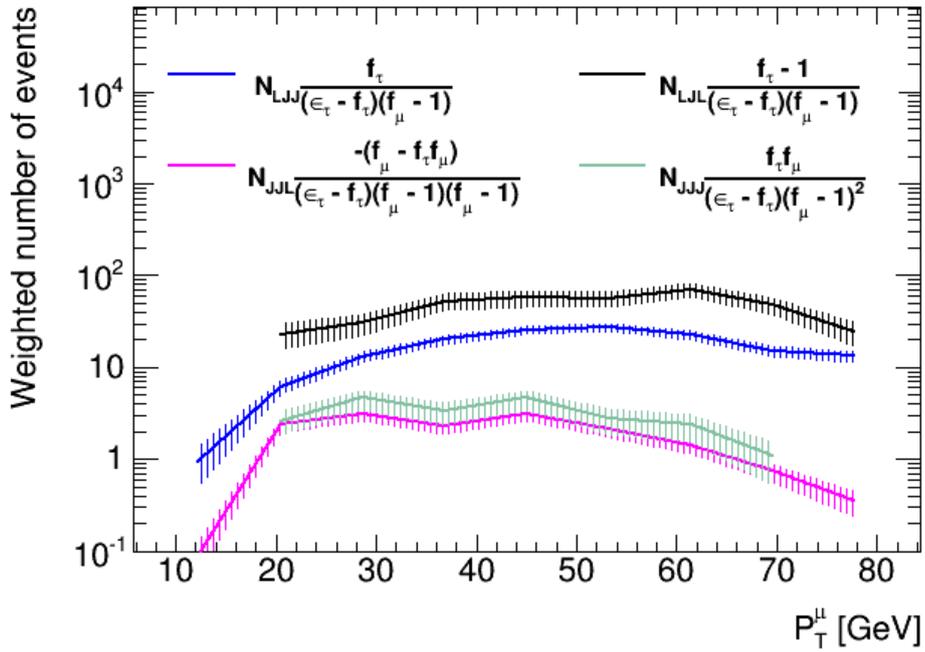


(b)

Figure B.0.1: The N_{TTF} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

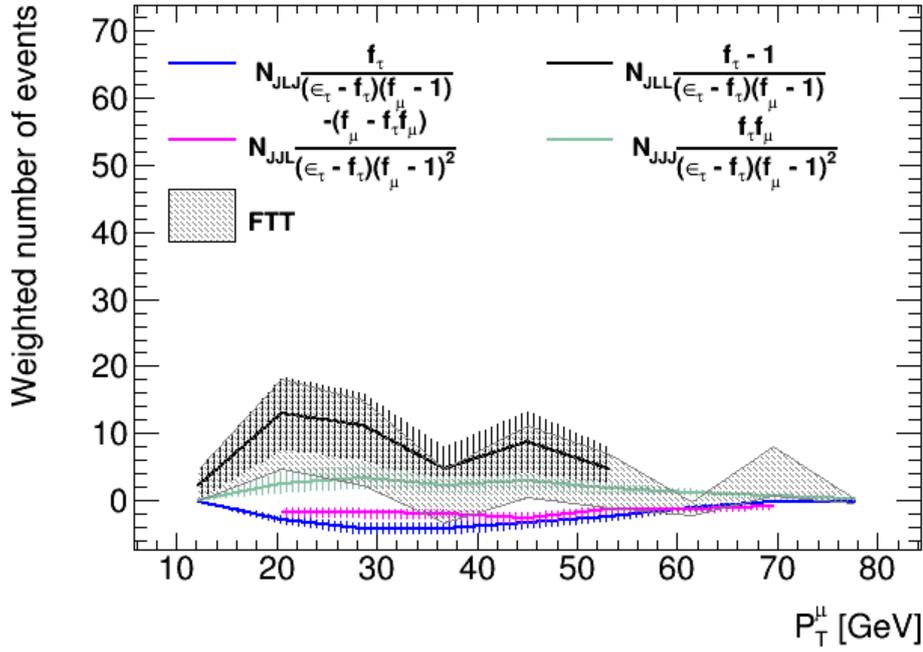


(a)

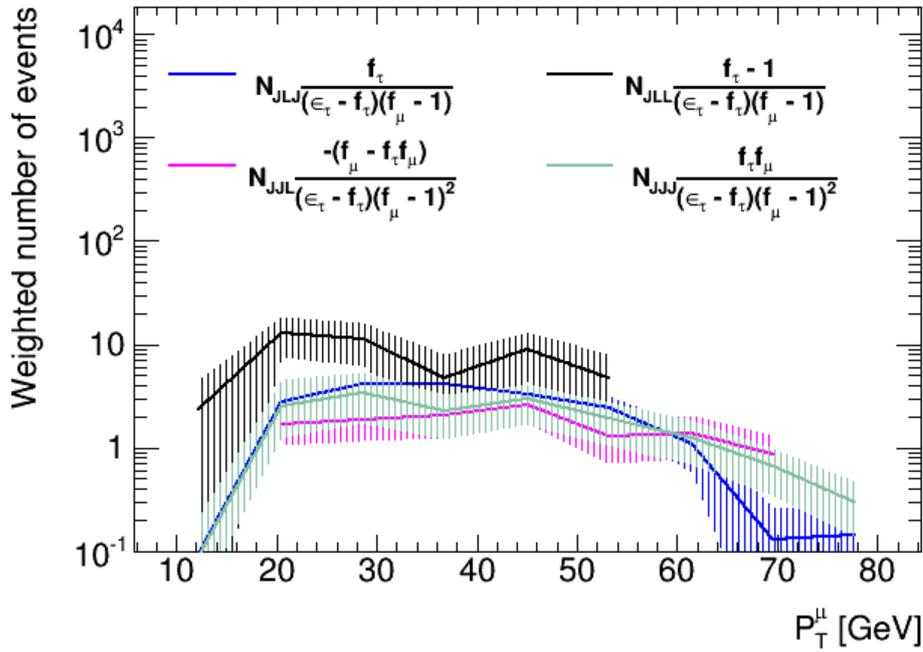


(b)

Figure B.0.2: The distribution of the leading μ^\pm for the N_{TFT} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

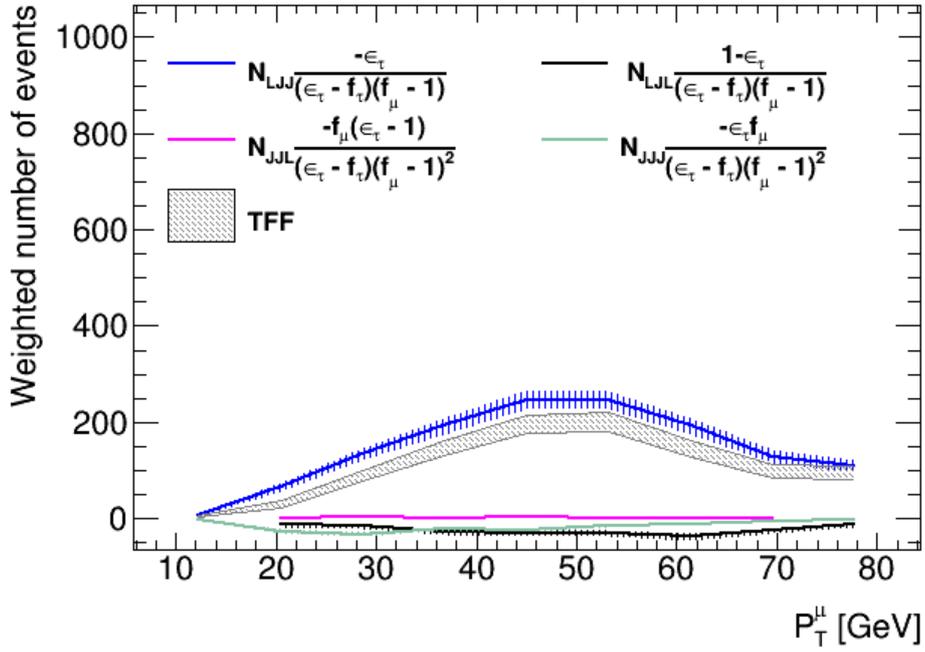


(a)

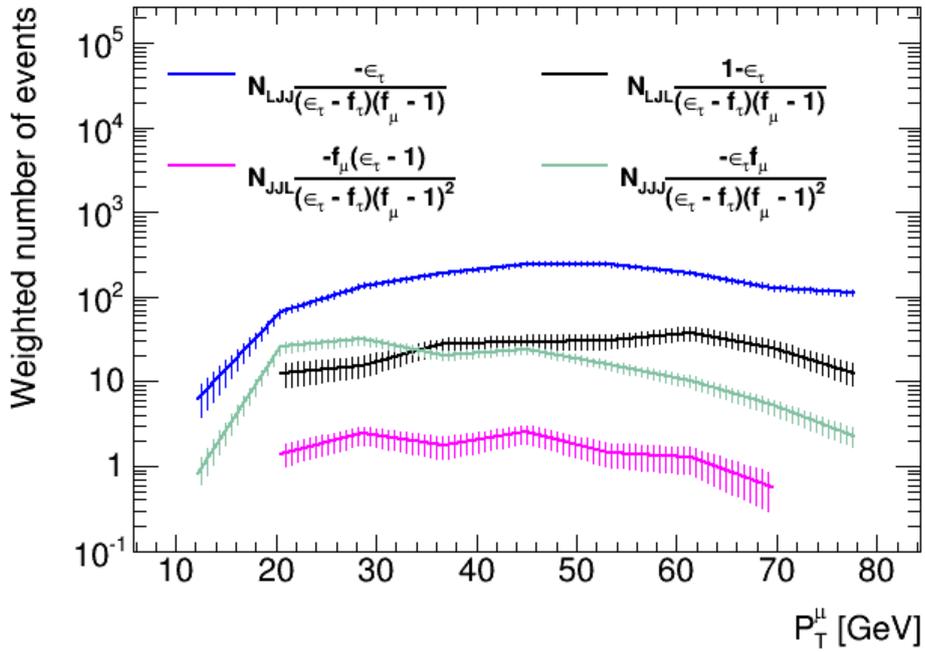


(b)

Figure B.0.3: The distribution of the leading μ^\pm for the N_{FTT} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

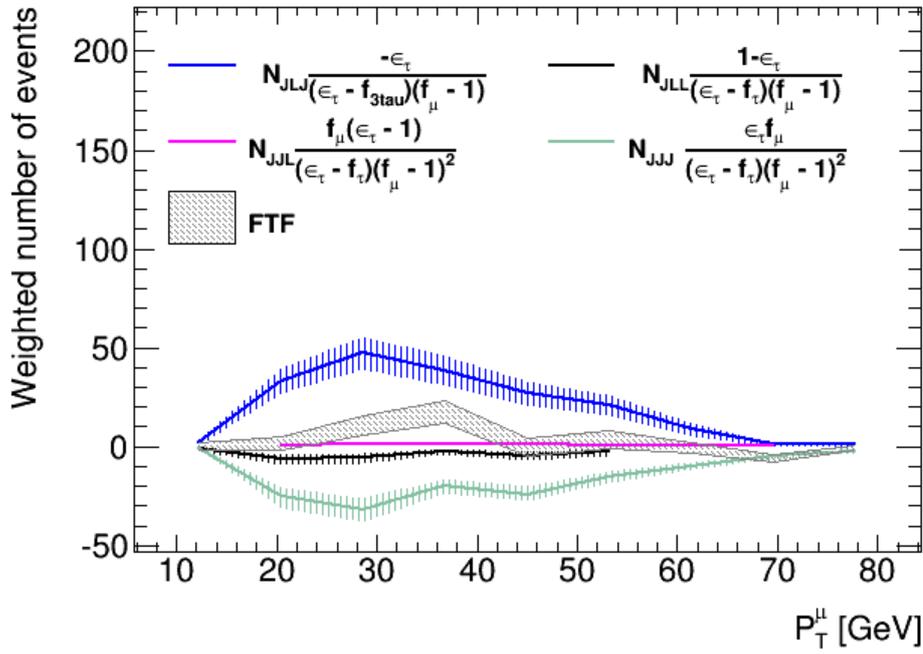


(a)

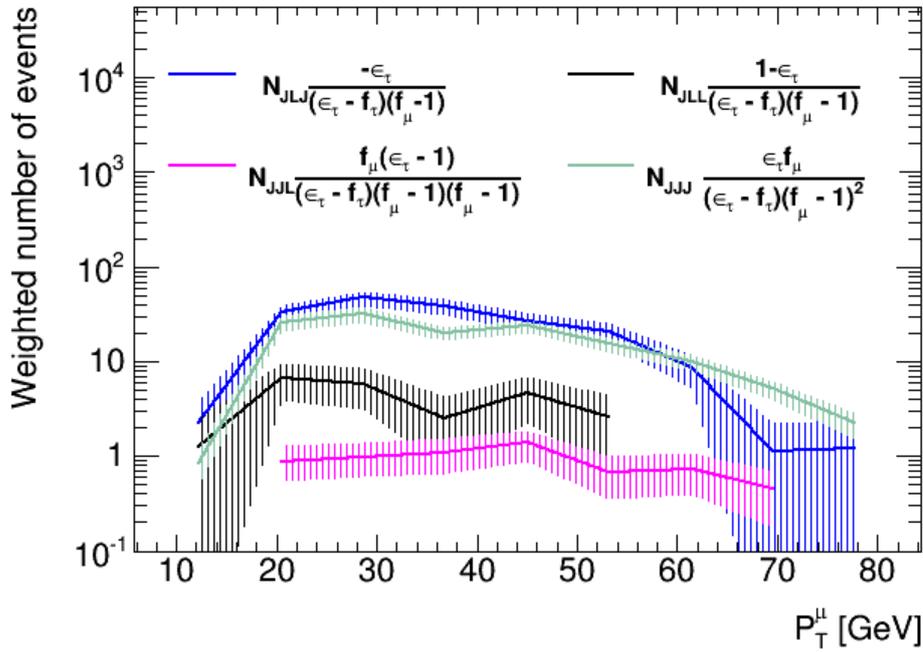


(b)

Figure B.0.4: The distribution of the leading μ^\pm for the N_{TFF} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

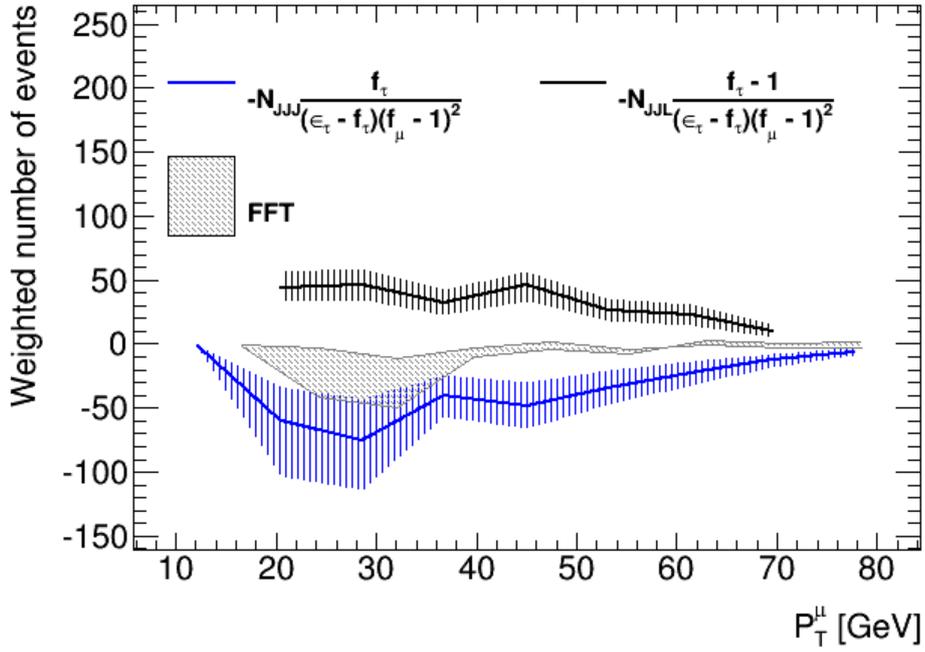


(a)

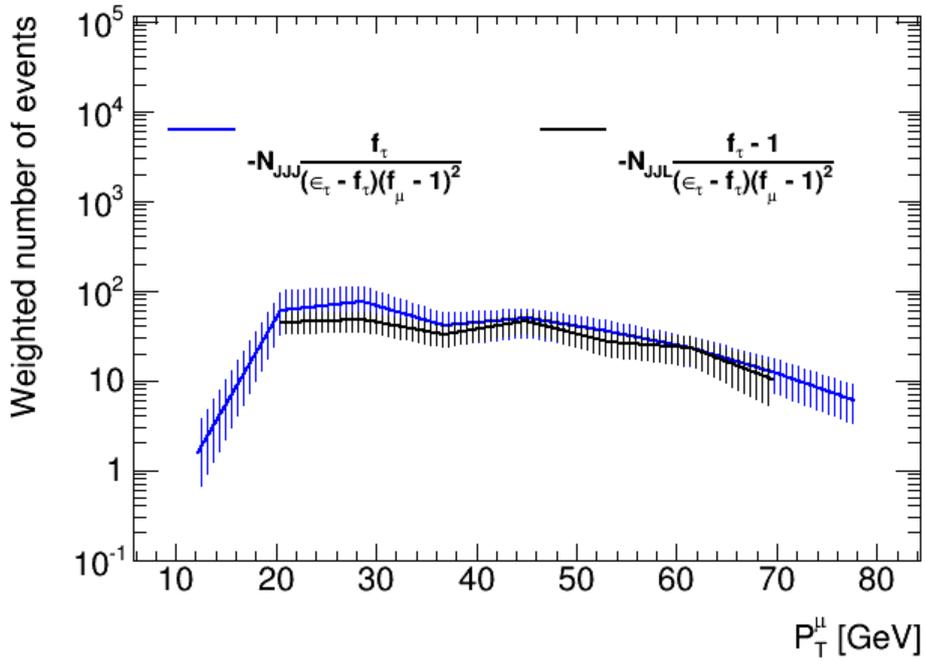


(b)

Figure B.0.5: The distribution of the leading μ^\pm for the N_{FTF} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

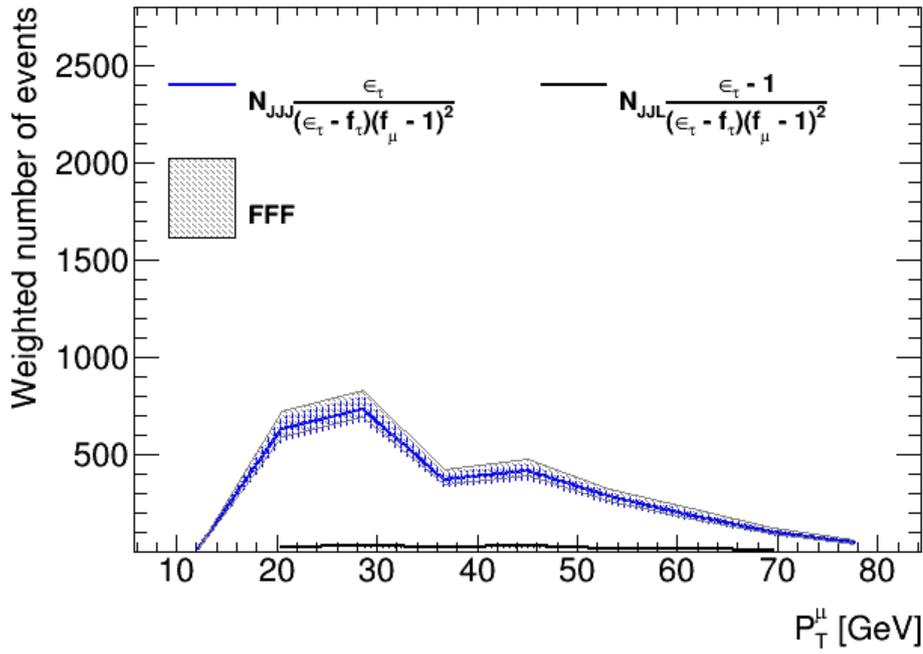


(a)

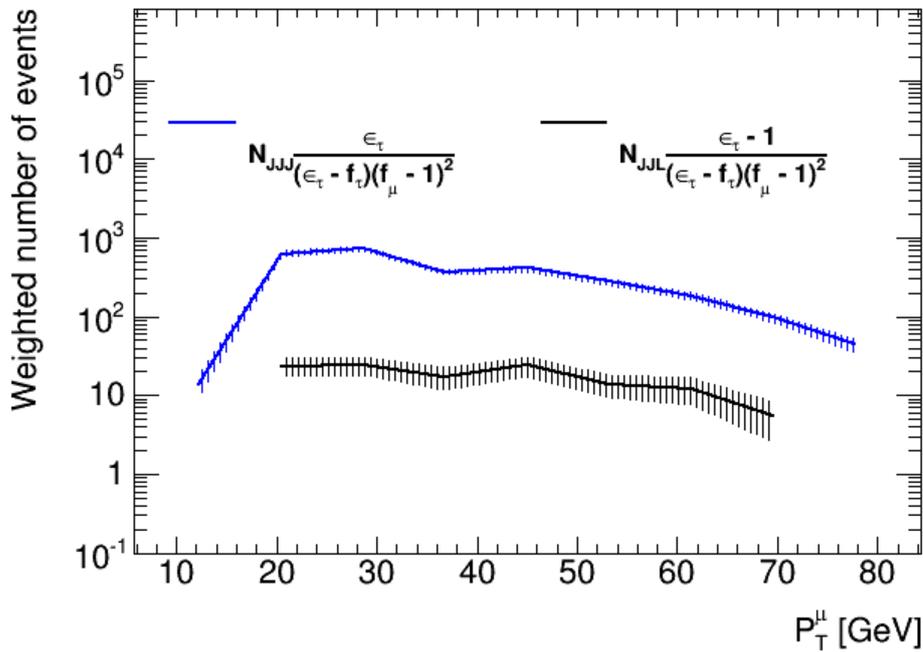


(b)

Figure B.0.6: The distribution of the leading μ^\pm for the N_{FFT} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.



(a)



(b)

Figure B.0.7: The distribution of the leading μ^\pm for the N_{FFF} term: The uncertainty of the distributions are found by varying the fake-rates between a lower and an upper limit. The shaded area in the figures signify the uncertainty from this variation and the systematic uncertainty due to statistics.

