

Search for R-Parity Violating Supersymmetry in
Final States with Four Leptons Including Taus
Using the ATLAS Detector

Suche nach R-Paritätsverletzender Supersymmetrie
in Ereignissen mit vier Lepton, einschließlich Taus,
mit dem ATLAS-Detektor



Masterarbeit der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Andrea Matic
geboren in München

München, den 22. Mai 2017

Gutachterin: Prof. Dr. Dorothee Schaile

Abstract

By increasing the LHC center-of-mass-energy from 8 TeV to 13 TeV in Run II the sensitivity to various supersymmetric models was enhanced enormously. The analysis in this work focuses on a group of signal models, in which the lightest supersymmetric particle (LSP) decays via an R -parity violating (RPV) process into two charged leptons and a neutrino. These signal models differ in their initially produced supersymmetric particles, which can be either charginos, neutralinos, gluinos or left handed sleptons and sneutrinos. A signal produced by these models is characterized by at least four charged leptons in the final state.

For the ICHEP conference in 2016 only events with electrons and muons were considered. Thereby only couplings of the LSP to light leptons could be studied. Now the analysis was extended to test another coupling, which describes the case that at least one of the two charged leptons in the LSP decay is a tau. For this purpose two new signal regions were introduced containing final states with one or two taus, respectively. The signal-to-background ratio in these regions was enhanced by applying a Z -veto and requiring the events to have a high effective mass. The Standard Model background for processes with prompt leptons was estimated fully with Monte Carlo simulation, while the backgrounds with non-prompt leptons were determined using dedicated control regions and a data-driven estimation method.

The full dataset of 2015 and 2016 corresponding to an integrated luminosity of 36.1 fb^{-1} was analyzed and background estimates derived. No significant excess of data over background predictions was found, thus exclusion limits on the various supersymmetric particles were calculated. For the case that the LSP couples only to light leptons, charginos and neutralinos with masses up to 1.4 TeV were excluded, as well as gluino masses up to 2.1 TeV and left handed sleptons and sneutrinos with masses up to 960 GeV. For the coupling scenario including taus exclusion limits on the particle masses were set up to 920 GeV for charginos and neutralinos, 1.6 TeV for gluinos and 620 GeV for left handed sleptons and sneutrinos.

Zusammenfassung

Mit der Erhöhung der LHC-Schwerpunktsenergie von 8 TeV auf 13 TeV in Run II konnte die Sensitivität auf verschiedene supersymmetrische Modelle erheblich gesteigert werden. Die Analyse in dieser Arbeit ist auf eine Gruppe von Signalmodellen ausgerichtet, in denen das leichteste supersymmetrische Teilchen (LSP) in einem R -Paritätsverletzenden (RPV) Prozess in zwei geladene Leptonen und ein Neutrino zerfällt. Die Signalmodelle unterscheiden sich jedoch hinsichtlich ihrer anfänglich produzierten Teilchen, bei denen es sich um Charginos, Neutralinos, Gluinos und linkshändige Sleptonen und Sneutrinos handeln kann. Ein Signal von diesen Modellen würde sich durch mindestens vier geladene Leptonen im Endzustand auszeichnen.

Da für die ICHEP-Konferenz 2016 nur Szenarien mit Elektronen und Myonen betrachtet wurden, konnte lediglich die Kopplung vom LSP an leichte Leptonen getestet werden. Nun wurde die Analyse erweitert, um auch solche Kopplungen zu studieren, die die Existenz von Taus im LSP-Zerfall erlauben. Aus diesem Grund wurden zwei neue Signalregionen eingeführt, die ein bzw. zwei Taus im Endzustand enthalten. Das Verhältnis von Signal zu Standardmodelluntergrund in diesen Regionen konnte deutlich gesteigert werden, indem ein Z -veto angewandt wurde und nur Ereignisse mit einer hohen effektiven Masse betrachtet wurden. Der Standardmodelluntergrund wurde für Prozesse mit echten Leptonen ausschließlich mit Monte-Carlo-Simulationen bestimmt, wohingegen für den Fake-Untergrund eine datenbasierte Abschätzung verwendet wurde, die auf ausgewählten Kontrollregionen beruht.

Der gesammelte Datensatz aus 2015 und 2016, der einer integrierten Luminosität von $\mathcal{L}dt = 36.1 \text{ fb}^{-1}$ entspricht, wurde analysiert und Untergrundabschätzungen durchgeführt. Die Ergebnisse zeigten keine signifikanten Abweichungen zu den Untergrunderwartungen, weshalb Ausschlussgrenzen für die Massen der supersymmetrischen Teilchen berechnet wurden. Für die Kopplung vom LSP an leichte Leptonen wurden Charginos und Neutralinos mit Massen von bis zu 1.4 TeV ausgeschlossen, ebenso wie Gluinomassen bis zu 2.1 TeV und linkshändige Sleptonen und Sneutrinos mit Massen bis zu 960 GeV. Für den Fall, dass das LSP auch an Taus koppelt, reichen die Ausschlussgrenzen bis 920 GeV für Charginos und Neutralinos, 1.6 TeV für Gluinos und 620 GeV für linkshändige Sleptonen und Sneutrinos.

Contents

1	Introduction	1
2	Theory	3
2.1	Standard Model of Particle Physics	3
2.1.1	Particle content	3
2.1.2	Mathematical description	3
2.1.3	Limitations of the Standard Model	5
2.2	Supersymmetry	7
2.2.1	Algebra	7
2.2.2	Particle content of the MSSM	8
2.2.3	Soft symmetry breaking	9
2.2.4	R - Parity	9
2.2.5	Benchmark models	9
2.2.6	RPV decay	10
3	The ATLAS experiment at the LHC	13
3.1	The LHC	13
3.2	The ATLAS detector	14
3.2.1	Coordinate system	14
3.2.2	Inner Detector	15
3.2.3	Calorimeters	15
3.2.4	Muon system	16
3.2.5	Trigger system	16
3.3	Particle identification	16
4	Data and Monte Carlo Simulation	19
4.1	Collected data	19
4.2	Monte Carlo simulation	19
4.2.1	Monte Carlo background samples	21
4.2.2	Monte Carlo signal samples	21

5	Object and event selection	23
5.1	Event cleaning	23
5.2	Object selection	24
5.3	Overlap removal	26
5.4	Trigger strategy	27
6	Signal region optimization	29
6.1	Z-veto	29
6.2	Discriminating variables	30
7	Background estimation	35
7.1	Irreducible background	35
7.2	Reducible background	35
7.2.1	Sources of fake leptons	36
7.2.2	Fake factor method	37
7.2.3	Systematic uncertainties	49
7.2.4	Validation	53
8	Results	55
8.1	Expected and observed yields	55
8.2	Statistical interpretation	56
8.2.1	Mathematical concepts	56
8.2.2	HistFitter setup	58
8.2.3	Model independent upper limits	58
8.2.4	Exclusion limits	59
9	Conclusion and outlook	67
A	MC samples	69
B	Region definitions	72
	Bibliography	73

Chapter 1

Introduction

With the discovery of the Higgs boson in 2012 [1, 2] all elementary particles of the Standard Model (SM) have been found. Apart from the particle content, also its predictions for the particle interactions have been confirmed with an amazing precision by many experiments. However, the SM is not a complete theory. There are several phenomena, which it cannot explain, for example: What is dark matter made of? Why is the Higgs mass stable, although it should suffer from enormous loop corrections? To answer these questions, a theory beyond the SM is required. One of the most promising candidates is Supersymmetry (SUSY). This theory introduces a symmetry between fermions and bosons, by which the number of elementary particles would be roughly doubled. The search for SUSY is intensively performed at CERN, the European Organization for Nuclear Research, where the Large Hadron Collider (LHC) is located. The LHC is the most powerful particle accelerator in the world, which brings beams of protons to almost the speed of light and collides them. In these high energetic reactions many other particles can be created, possibly also supersymmetric particles. As SUSY is a theory with many free parameters, there are plenty of different signatures that can be studied in the collisions.

The analysis in this work is specialized for SUSY processes leading to four or more leptons in the final state. This signature offers a high discrimination power against the large QCD multi-jet background. Furthermore, the event selection criteria do not contain any complicated requirements, which has the advantage of generality: The analysis is basically sensitive to new physics processes of any kind, provided that these are characterized by final states with multiple leptons. The models for which the analysis has been optimized belong to R -parity violating (RPV) processes. In RPV scenarios the lightest supersymmetric particle (LSP) is not stable, as proposed by many other SUSY models, but decays further into SM particles. In the case of this analysis the LSP is assumed to decay exclusively into leptons.

The four lepton analysis investigates data from the ATLAS detector and was already performed in Run I at a center of mass energy of $\sqrt{s} = 8$ TeV, in which a large parameter space of different RPV models was excluded [3]. In Run II, which started in June 2015 with an increased center of mass energy of $\sqrt{s} = 13$ TeV [4], the analysis was revived and first results were presented at the ICHEP conference in 2016. It was possible to exclude a large parameter space of a SUSY model, thereby exceeding the limits of Run I by approximately 400 GeV [5]. However, the ICHEP studies were based on only one benchmark model and on final states containing only light leptons. Thus only the coupling of the LSP to light leptons was tested. In the context of this work the analysis was extended to include further signal models and to test couplings, which allow the LSP to decay also to taus. For this purpose two new signal regions were introduced requiring the presence of at least one or two taus, respectively.

After an introduction to the theoretical and experimental backgrounds necessary for this

analysis, the optimization of two signal regions is presented, which provide the highest possible sensitivity to the RPV signals under study. Furthermore it will be explained, with which techniques the SM background processes in the signal regions are estimated. Finally, the results of the analysis for a dataset corresponding to an integrated luminosity of $\int \mathcal{L} dt = 36.1 \text{ fb}^{-1}$ will be shown and exclusion limits on masses of charginos, neutralinos, gluinos and sleptons in specific models are set.

Chapter 2

Theory

2.1 Standard Model of Particle Physics

The supersymmetric particles in the models considered in this analysis are assumed to decay almost instantaneously, so that the only particles, that can be directly detected, are already discovered particles from the Standard Model (SM) of particle physics. The SM is a theoretical framework which describes the fundamental particles and their electromagnetic, weak and strong interactions with great precision. This section gives an overview of the SM based on [6, 7, 8].

2.1.1 Particle content

Ordinary matter consists of fermions. These are particles of spin 1/2 and are further divided into leptons and quarks. Six quarks and six leptons exist ordered in three families, or generations. A quark family consists of an up-type and a down-type quark, while a lepton family contains a charged lepton together with a neutral lepton, the neutrino. The fermions of the SM with their electric charges and masses are listed in Table 2.1 and 2.2. Each fermion has an anti-particle with the same mass and spin, but opposite quantum numbers.

Interactions between fermions are mediated by force carriers, which are bosons of spin 1. The electromagnetic force is propagated by the massless photon γ coupling to particles with electric charge. Eight massless gluons g are the mediators of the strong force and couple to particles with color charge, i.e. quarks and gluons. The weak force is acting on particles with weak isospin and its propagators are three massive bosons: The electrically charged W^+ and W^- and the neutral Z . Furthermore, the SM contains a scalar Higgs boson H with spin 0. The SM bosons, their electric charges and masses are summarized in Table 2.3.

2.1.2 Mathematical description

To describe the particles and their interactions in the SM appropriately, it is necessary to have a mathematical framework which combines quantum mechanical and relativistic effects. This framework is called quantum field theory. In this theory a particle is not treated as a discrete mass point, but as a continuous system represented by a field $\phi(\vec{x}, t)$. It adopts the concept of classical mechanics, in which the equations of motions can be obtained by building a Lagrangian $L(q, \dot{q}, t)$ depending on the coordinates q_i and velocities \dot{q}_i of a mass point and the time t . In quantum field theory one defines a Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$ as a function

Generation	Quark	Electric charge [e]	Mass [MeV]
1	u	$+\frac{2}{3}$	2.2
	d	$-\frac{1}{3}$	4.7
2	c	$+\frac{2}{3}$	$1.27 \cdot 10^3$
	s	$-\frac{1}{3}$	96
3	t	$+\frac{2}{3}$	$173.2 \cdot 10^3$
	b	$-\frac{1}{3}$	$4.18 \cdot 10^3$

Table 2.1: The quarks of the SM. Masses are taken from [9]

Generation	Lepton	Electric charge [e]	Mass [MeV]
1	e	-1	0.511
	ν_e	0	$< 2 \text{ eV}$
2	μ	-1	105.7
	ν_μ	0	$< 2 \text{ eV}$
3	τ	-1	1776.9
	ν_τ	0	$< 2 \text{ eV}$

Table 2.2: The leptons of the SM. Masses are taken from [10].

Boson	Electric charge [e]	Mass [GeV]
γ	0	0
g	0	0
W^\pm	± 1	80.39
Z	0	91.188
H	0	125.09

Table 2.3: The bosons of the SM. Masses are taken from [11].

of the field and its derivatives with respect to the space-time coordinates $x^\mu = (t, x^1, x^2, x^3)$. The action S is defined by:

$$S = \int L dt = \int \mathcal{L}(\phi, \partial_\mu \phi) d^4x$$

Following the principle of least action, $\delta S = 0$, leads to the Euler-Lagrange equation:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$

by which the equations of motion of a mass point can be derived. If \mathcal{L} contains more than one field, then the Euler-Lagrange equation is applied separately to each field. The SM is based on three gauge symmetries. A gauge symmetry postulates, that \mathcal{L} is invariant under a local phase transformation $\phi \rightarrow \phi' = e^{i\theta(x^\mu)}\phi$ on the field. Quantum Chromodynamics (QCD) describes the interactions of colored particles and is represented by a $SU(3)$ symmetry group. Electromagnetic and weak interactions are described in a common symmetry group, $SU(2)_L \otimes U(1)_Y$, and are therefore summarized as *electroweak* interactions. The subscript L stands for the fact, that only left handed particles undergo weak interactions. Y refers to the hypercharge, which is defined by Gell-Mann-Nishijima relation:

$$Y = 2(Q - I_3)$$

This relation connects the charge Q and the third component of the weak isospin, I_3 , of a particle. The electroweak symmetry group, however, has the problem, that their interactions show no gauge invariance if a mass term is added into \mathcal{L} . As a consequence, W^\pm and Z would need to be massless. This issue can be solved by introducing the Higgs mechanism, through which the electroweak symmetry is spontaneously broken. Thereby W^\pm and Z become massive. Fermions acquire their mass by coupling to the scalar Higgs field, called Yukawa-coupling. Furthermore, the Higgs mechanism predicts the existence of an additional massive boson with spin 0: the Higgs boson.

2.1.3 Limitations of the Standard Model

Although the predictions of the SM have been confirmed to amazing precision by various experiments, the SM on its own is not a complete theory to describe all matter and interactions in the universe. In the following the deficiencies of the SM will be presented, which motivate the necessity for a theory beyond the SM (BSM).

Gravity

In the SM gravity is not included. However, at the Planck scale $M_P = (8\pi G_{Newton})^{-\frac{1}{2}} = 2.4 \cdot 10^{18}$ GeV the gravitational force gets comparable in strength to the forces of the SM [12]. Hence for energies greater than M_P a new theory is required that also accounts for quantum gravitational effects.

Dark matter

Astrophysical observations showed, that the universe is not solely composed of ordinary visible matter. Measurements of the space observatory Planck showed that ordinary matter amounts only to 4.9%. Most of the energy density in the universe consists of dark energy with a fraction of 68.3%. Dark energy is responsible for the accelerated expansion of the

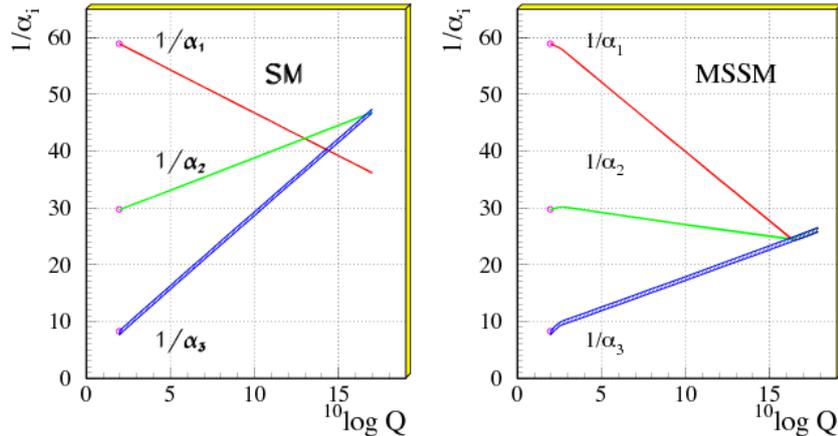


Figure 2.1: Evolution of the inverse gauge couplings as a function of the logarithm of the energy in the SM (left) and in the MSSM (right) [15].

universe. The remaining 26.8% of the energy density is in the form of dark matter [13]. The existence of dark matter was deduced from astronomical observations of galaxies. Their velocities were found to be so high that the gravitational force, generated by their visible matter alone, could not hold the galaxies together. However, the particle content of the SM does not contain a suitable candidate for dark matter. In general, a candidate for dark matter should be sufficiently massive and may not undergo electromagnetic and strong interactions.

Unification of gauge couplings

The concept of unified theories states that the weak, strong and electromagnetic interactions are three different aspects of one single interaction and that unification occurs at high energies. The strong and two electroweak couplings decrease with the logarithm of the energy, while one electroweak increases. In the SM the three coupling constants do not meet, but in a supersymmetric extension of the SM, the MSSM (which will be described in Section 3), the couplings are modified and an unification can happen at about 10^{16} GeV [14]. This is illustrated in Figure 2.1.

Hierarchy problem

The hierarchy problem [12] is related to the fact, that the weak energy scale M_W differs from the Planck scale M_P by 17 orders of magnitude. M_W is defined by the vacuum expectation value of the Higgs field, which is of the order of 10^2 GeV. The mass of the Higgs boson should be modified by quantum corrections caused by virtual loop diagrams of every particle that couples to the Higgs field. This includes SM as well as any BSM particles. Therefore arises the question, why the Higgs mass is still stable. This is referred to as fine tuning problem. A loop with a fermion f , shown in Figure 2.2, would give a correction to the squared Higgs mass m_H^2 by [12]:

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \dots \quad (2.1)$$

Here λ_f represents the coupling strength of the fermion to the Higgs boson. Λ_{UV} is a momentum cutoff scale which is used to regulate the loop integral and should be at least as high as the energy scale of new physics. If Λ_{UV} is taken to be M_P , then m_H is about 30 orders of magnitude larger than M_W . A solution to this problem can be obtained by including scalar

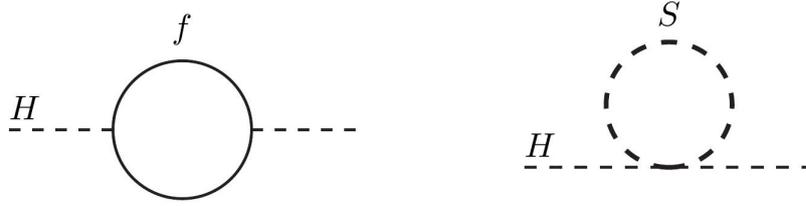


Figure 2.2: Radiative corrections to the Higgs mass caused by a fermion loop (left) and a scalar boson loop (right) [12].

bosons S into the theory, which also couple to the Higgs field. The virtual loop from a scalar, as illustrated in Figure 2.2, would result in a correction by [12]:

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} \Lambda_{UV}^2 + \dots \quad (2.2)$$

Hence, the quantum corrections from Equation 2.1 and Equation 2.2 would compensate each other, if each fermion is accompanied by two complex scalars with $|\lambda_f|^2 = \lambda_S$.

2.2 Supersymmetry

Supersymmetry (SUSY) is a theory which introduces a new symmetry between fermions and bosons. In the following the basic concepts of SUSY will be explained and the particle content of the Minimal Supersymmetric Standard Model (MSSM) will be presented, which is a supersymmetric extension of the SM with the smallest possible number of additional particles. Afterwards the properties of the benchmark models in this work will be discussed.

2.2.1 Algebra

The fundamental principle of SUSY [12] is that a fermionic state can be turned into a bosonic state, and vice versa, through the fermionic operator Q :

$$\begin{aligned} Q|\text{fermion}\rangle &= |\text{boson}\rangle \\ Q|\text{boson}\rangle &= |\text{fermion}\rangle \end{aligned}$$

Thereby the spin of the particle is changed by $\Delta s = 1/2$. Q^\dagger , the hermitian conjugate of Q , is also a generator of this symmetry transformation. The generators have a spin of $1/2$ and they must fulfill following anticommutation and commutation relations:

$$\{Q, Q^\dagger\} \propto P^\mu, \quad (2.3)$$

$$\{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0, \quad (2.4)$$

$$[P^\mu, Q] = [P^\mu, Q^\dagger] = 0, \quad (2.5)$$

with P_μ being the four-momentum generator of space-time translations. Equation 2.3 introduces a connection between SUSY and space-time translations, while Equation 2.5 indicates that supersymmetric transformations are independent of space-time positions. The SM particles together with their supersymmetric partners, shortly *superpartners*, are ordered in so called *supermultiplets*. As the squared-mass operator $-P^2$ commutes with Q , Q^\dagger (also a

consequence from Equation 2.5) and with all spacetime rotation and translation operators, all particles within the same supermultiplet have the same eigenvalues of $-P^2$. This means, that they have equal masses. Furthermore, Q and Q^\dagger commute with the generators of gauge transformations, too. As a consequence, particles from the same supermultiplet are also equal in all quantum numbers, except for the spin. A supermultiplet has the same number of fermionic as bosonic degrees of freedom:

$$n_F = n_B \quad (2.6)$$

The simplest supermultiplet satisfying this condition contains a Weyl fermion and a complex scalar field. A Weyl fermion is a fermion, which is massless and chiral, i.e. it has a left-handed and right-handed component. It has two helicity states, thus $n_F = 2$, while the real and imaginary part of the complex scalar field each yield $n_B = 1$. Such a constellation is called a *chiral* supermultiplet. Another combination, which is consistent with Equation 2.6, is the *gauge* supermultiplet: This is a vector boson with spin 1 along with a spin-1/2 Weyl fermion. Both have two possible helicity states and therefore $n_F = n_B = 2$. Moreover, the particles in the supermultiplets are massless at least until the gauge symmetry is spontaneously broken.

2.2.2 Particle content of the MSSM

In the MSSM, each particle of the SM is part of either of the two described supermultiplets. The fermions are members of the chiral group because only this multiplet allows a fermion to have different gauge transformation properties for its left-handed and right-handed component. Therefore each component has its own superpartner, which is a particle of spin 0. It obtains the same name as the corresponding fermion, but with a preceding “s”, for scalar. For example, the superpartner of an electron is called a selectron. The squarks and sleptons are denoted by a tilde above the fermions’ symbols, e.g. \tilde{e}_R stands for a right-handed selectron. It is important to note that the handedness refers only to the fermion’s helicity and not to the one of its superpartner.

As the Higgs boson has spin 0, it must also occupy a chiral supermultiplet. However, a supersymmetric theory requires an extended Higgs sector to prevent anomalies in the electroweak gauge symmetry and to supply the necessary Yukawa couplings. This new sector can consist of two supermultiplets of the form (H_u^+, H_u^0) and (H_d^0, H_d^-) . The superpartners are called higgsinos and are denoted by \tilde{H}_u^+ , \tilde{H}_u^0 , \tilde{H}_d^0 , \tilde{H}_d^- . The Higgs doublets have eight degrees of freedom in total. During the electroweak symmetry breaking three of them are used to give masses to the W and Z bosons, while the remaining ones result in five different Higgs bosons: Two CP even neutral h_0 and H^0 with h^0 being the lighter one, another neutral, but CP odd A^0 , and the charged H^+ and H^- .

The vector bosons of the SM reside in gauge supermultiplets together with their spin-1/2 superpartners, the gauginos. Again, the superpartners are indicated by a tilde above the SM symbols and they receive the same name as their counterparts with an additional “-ino” as suffix, e.g. the gluino is the superpartner of the gluon.

The MSSM predicts also the existence of so called neutralinos and charginos. They originate from effects of the electroweak symmetry breaking, by which the higgsinos and gauginos mix with each other. The neutral higgsinos \tilde{H}_u^0 and \tilde{H}_d^0 mix with the neutral bino \tilde{B} and wino \tilde{W}^0 , thereby producing the neutralinos $\tilde{\chi}_i^0$ ($i = 1, 2, 3, 4$). Moreover, the combination of the charged higgsinos \tilde{H}_u^+ and \tilde{H}_d^- with the winos \tilde{W}^+ and \tilde{W}^- generates the charginos $\tilde{\chi}_i^\pm$ ($i = 1, 2$). The charginos and neutralinos are labeled in ascending mass order, i.e. $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^\pm$ are the lightest of their sort.

2.2.3 Soft symmetry breaking

As stated above, the SM particles and their superpartners are supposed to have equal masses. However, no supersymmetric particles have been observed so far. This leads to the conclusion that, if they are realized in nature, they are heavier than their SM partners. Therefore SUSY must be a broken symmetry. In order to still guarantee a solution to the hierarchy problem, the symmetry breaking must be “soft” [12].

2.2.4 R - Parity

The theory of the MSSM allows processes, which violate the baryon (B) and lepton number (L). As a consequence, the proton could decay e.g. by $p^+ \rightarrow e^+ \pi^0$ with a very short lifetime. However, it has been experimentally confirmed that its lifetime exceeds 10^{32} years [12]. Such B and L violating processes can be eliminated by postulating the conservation of a new quantum number, the R -parity. It is defined by:

$$P_R = (-1)^{3(B-L)+2s} \quad (2.7)$$

where s refers to the spin of the particle. SM particles have even R -parity ($P_R = 1$), whereas supersymmetric particles have odd R -parity ($P_R = -1$). Imposing that P_R is multiplicatively conserved at each interaction vertex leads to following consequences:

- Supersymmetric particles can only be produced in even numbers, usually in pairs.
- A heavy supersymmetric particle always decays into an odd number of lighter supersymmetric particles, mostly just one.
- The lightest supersymmetric particle (LSP) must be stable.

It is also possible to avoid the experimental constraints from the proton decay while allowing for the existence of R -parity violating (RPV) scenarios [12]. This can be achieved if the Lagrangian conserves either B or L . Another option is to introduce alternative symmetries as the “baryon triality”, which is a quantum number relating B and the hypercharge Y , or to apply the concept of symmetry breaking also to R -parity.

2.2.5 Benchmark models

The analysis in this work targets various simplified models, which are shown in Figure 2.3. Instead of giving a full description of a supersymmetric model, a simplified model focuses on a certain decay chain. It considers only the particles occurring in the decay chain and the branching ratios of the depicted decays are assumed to be 100%.

Figure 2.3a shows the simplified model used in the ICHEP studies of 2016. In this model mass degenerate and wino-like charginos $\tilde{\chi}_1^+$ and $\tilde{\chi}_1^-$ are pair produced in the initial proton-proton collision. Each of the charginos decays into a W -boson and the LSP, which is a bino-like $\tilde{\chi}_1^0$. The LSP decays then via an RPV interaction to two charged leptons ℓ and a neutrino ν . The final state contains four charged leptons originating from the RPV decays and up to two more, depending on whether the W -bosons decay leptonically or hadronically.

The other simplified models were included into the analysis after the ICHEP conference. Also these models assume a bino-like $\tilde{\chi}_1^0$ LSP, which undergoes an RPV decay to two charged

leptons and one neutrino. However, they differ in the initially produced supersymmetric particles. In the model shown in Figure 2.3b a $\tilde{\chi}_1^\pm$ is produced together with a $\tilde{\chi}_2^0$, which decays then into the LSP and a Z -boson. The model in Figure 2.3c contains a gluino as next-to-lightest particle (NLSP). In its decay to the LSP two quarks are produced in addition. The last model studied in this analysis is shown Figure 2.3d, where a pair of left-handed sleptons is produced initially. These can be either charged sleptons $\tilde{\ell}_L$ or sneutrinos $\tilde{\nu}$, which decay to the LSP and a charged lepton / neutrino, respectively.

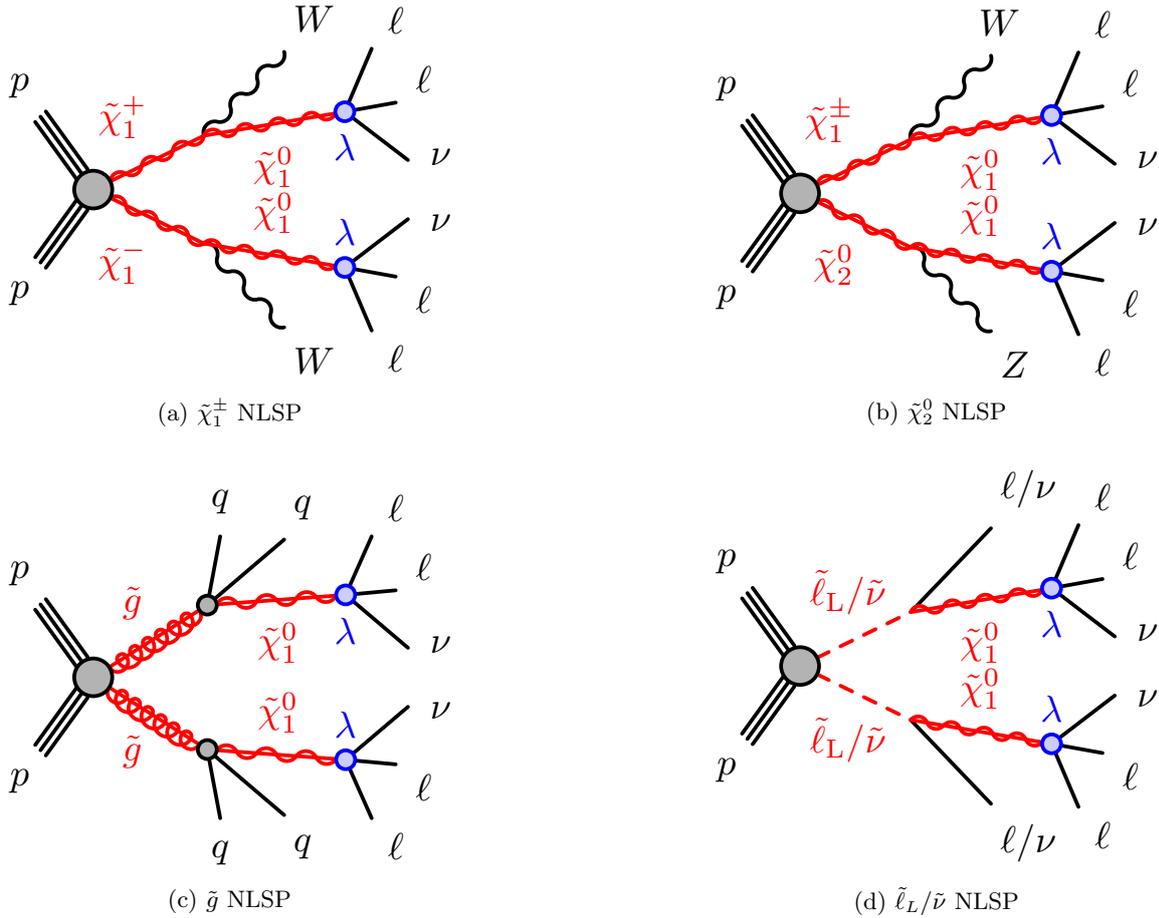


Figure 2.3: Benchmark models of the analysis.

2.2.6 RPV decay

In the simplified models presented above the LSP decay is mediated by the following L -violating superpotential term [16, 17]:

$$W_{LLE\bar{E}} = \frac{1}{2} \lambda_{ijk} L_i L_j \bar{E}_k \quad (2.8)$$

where $L_{i/j}$ are the $SU(2)_L$ doublet superfields and \bar{E}_k the right-handed $SU(2)_L$ singlet superfields. The indices i, j and k refer to the generation, while λ_{ijk} is a collection of nine Yukawa couplings with $\lambda_{ijk} = -\lambda_{jik}$. With this superpotential the LSP decays by:

$$\tilde{\chi}_1^0 \rightarrow \ell_k^\pm \ell_{i/j}^\mp \nu_{j/i} \quad (2.9)$$

The resulting lepton flavors depend on the indices of the associated λ_{ijk} couplings. In principle, the LSP can decay to any combination of charged lepton pairs. Table 2.4 shows two different decay modes with their corresponding branching ratios.

Sample	$\tilde{\chi}_1^0$ branching ratios		
$LL\bar{E}12k$	$e^+e^-\nu$ (1/4)	$e^\pm\mu^\mp\nu$ (1/2)	$\mu^+\mu^-\nu$ (1/4)
$LL\bar{E}i33$	$e^\pm\tau^\mp\nu$ (1/4)	$\tau^+\tau^-\nu$ (1/2)	$\mu^\pm\tau^\mp\nu$ (1/4)

Table 2.4: Decay modes and branching ratios for the LSP in the RPV models.

The naming of the samples is inspired by the couplings producing these final states. $LL\bar{E}12k$ scenarios ($k \in 1,2$) are based on λ_{121} and λ_{122} , thus the LSP can decay only into light charged leptons. In the $LL\bar{E}i33$ models the underlying couplings are λ_{133} and λ_{233} , so that at least one of the two charged leptons from the LSP decay is a tau.

For the ICHEP conference only final states were considered with at least four light leptons and no taus. Therefore only the λ_{12k} coupling was studied. These events are denoted by $4\ell 0\tau$, where $\ell = e, \mu$. For the simplified model with initial $\tilde{\chi}_1^+\tilde{\chi}_1^-$ production (see Figure 2.3a) it was possible to exclude $\tilde{\chi}_1^\pm$ masses up to 1.14 TeV [5].

To test the λ_{i33} coupling it is necessary to include also taus into the analysis. With a branching ratio of approximately 35% taus decay leptonically by:

$$\tau \rightarrow \ell \nu_\ell \nu_\tau \quad (2.10)$$

while in the other 65% they decay hadronically, i.e. into quarks:

$$\tau \rightarrow q\bar{q}'\nu_\tau \quad (2.11)$$

where \bar{q}' denotes an anti-quark of a different flavor than of the quark q . They are required to add up to unit charge. In the following only taus, which decay into quarks, are labeled as τ , while leptonically decaying taus are denoted by ℓ . The reason for this will be explained in Chapter 3. If all taus in a $LL\bar{E}i33$ model decay into leptons, one obtains $4\ell 0\tau$ final states as in the case of $LL\bar{E}12k$ models. However, to increase the sensitivity to $LL\bar{E}i33$ scenarios, one should consider also hadronically decaying taus. For this purpose the analysis is extended to include $3\ell 1\tau$ and $2\ell 2\tau$ final states. In the course of this work it will be described, how the $3\ell 1\tau$ and $2\ell 2\tau$ regions were integrated into the analysis.

Target	$N(e, \mu)$	$N(\tau)$
$4\ell 0\tau$	≥ 4	$= 0$
$3\ell 1\tau$	$= 3$	≥ 1
$2\ell 2\tau$	$= 2$	≥ 2

Table 2.5: Requirements on the lepton multiplicity for the different final states considered in the analysis.

An overview of the lepton requirements imposed on the different final states is shown in Table 2.5. For $4\ell 0\tau$ events one demands *at least* four light leptons, while for $3\ell 1\tau$ and $2\ell 2\tau$ events *exactly* three or two light leptons are required, respectively. This ensures, that the three regions are orthogonal to each other, i.e. that there is no overlap of events between these regions. That will play an important role in the statistical analysis of the results described in Chapter 8.

Chapter 3

The ATLAS experiment at the LHC

The European Organization for Nuclear Research (CERN), founded in 1954, is one of the most important institutions for probing the fundamental structure of matter. It is located at the border between France and Switzerland near Geneva and has by now 22 member states [18]. Currently, CERN's flagship project is the Large Hadron Collider (LHC), which is the largest particle accelerator in the world.

3.1 The LHC

The LHC is situated in a circular tunnel that was initially built for the Large Electron-Positron Collider (LEP). This tunnel has a circumference of 26.7 km and is located between 45 m and 170 m below the surface with an inclination of 1.4%. It consists of eight straight sections and eight arcs. Two counter-rotating beams of high energetic particles are used, which can be either protons or lead-ions. They are accelerated to almost the speed of light before they are brought to collision [19]. In 2010 the LHC started operation at a center-of-mass energy of $\sqrt{s} = 7$ TeV for proton-proton collisions, which was increased to $\sqrt{s} = 8$ TeV after two years. From 2013 until 2015 the LHC was shut down for maintenance and upgrade work, allowing it to restart operation with a notably higher center-of-mass energy of $\sqrt{s} = 13$ TeV [20]. Thereby it reached almost its design value of $\sqrt{s} = 14$ TeV.

The LHC is the last element of a large accelerator complex shown in Figure 3.1. The first element in the chain is a bottle of hydrogen gas. With the use of an electric field the electrons are stripped off the atoms, yielding protons. These are then accelerated by Linac 2 to an energy of 50 MeV. The second accelerator in the chain is the Proton Synchrotron Booster (PSB), through which the protons acquire an energy of 1.4 GeV. Afterwards they are transferred into the Proton Synchrotron (PS) and are accelerated to 25 GeV, followed by the Super Proton Synchrotron (SPS), where they reach an energy of 450 GeV. Finally, they are injected into the two beam pipes of the LHC [21, 22].

The proton beams are brought to collision at four interaction points, around which four big experiments are installed: The general purpose detectors ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid) and the two specialized experiments ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider beauty). At ALICE lead-ion collisions are analyzed with the aim to investigate the properties of the quark-gluon plasma. The LHCb experiment is specialized on B -physics to study the matter-antimatter asymmetry. There are also three smaller experiments at the LHC: TOTEM (TOTAl Elastic and diffractive cross section Measurement), LHCf (Large Hadron Collider forward) and MoEDAL (Monopole and Exotics Detector at the LHC) [22].

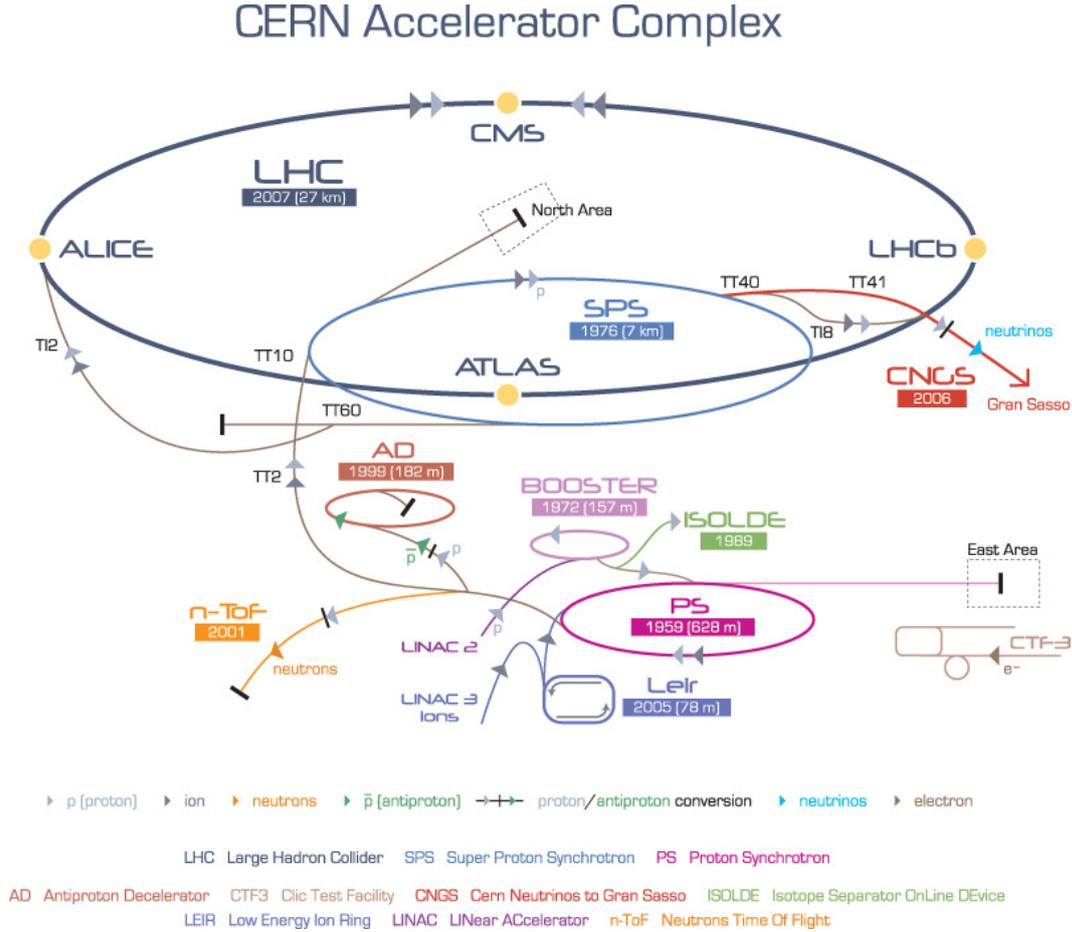


Figure 3.1: Overview of the LHC and its pre-accelerators [23].

3.2 The ATLAS detector

The ATLAS detector [24, 25] has a cylindrical geometry and measures 46 m in length and 25 m in diameter with a total mass of about 7000 tonnes. Its layout is sketched in Figure 3.2. A set of sub-detectors is arranged in layers around the beam line (barrel), while another two are ordered in disks perpendicular to the beam line (end-caps) making it possible to detect very forward or backward particles. This section will give an overview of the individual detector components and explain, how the different types of particles are reconstructed.

3.2.1 Coordinate system

To describe the position of the detected particles, one needs a suitable coordinate system. The interaction point of the beams is defined as the origin of the right handed coordinate system and the beam line is set as the z -axis. The positive x -axis is defined to point towards the center of the LHC ring, while the positive y -axis points upwards. Many event variables are measured in the $x - y$ plane, which is called the transverse plane. The azimuthal angle ϕ is the angle around the z -axis and the polar angle θ is measured from the z -axis. Commonly, θ is expressed through the pseudorapidity, given by $\eta = -\ln(\tan(\theta/2))$. For massive objects the rapidity $y = 1/2 \ln[(E + p_z)/(E - p_z)]$ is used instead of η or ϕ , where E is the particle's energy and p_z its momentum along the z -axis. The distance between two objects is defined by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

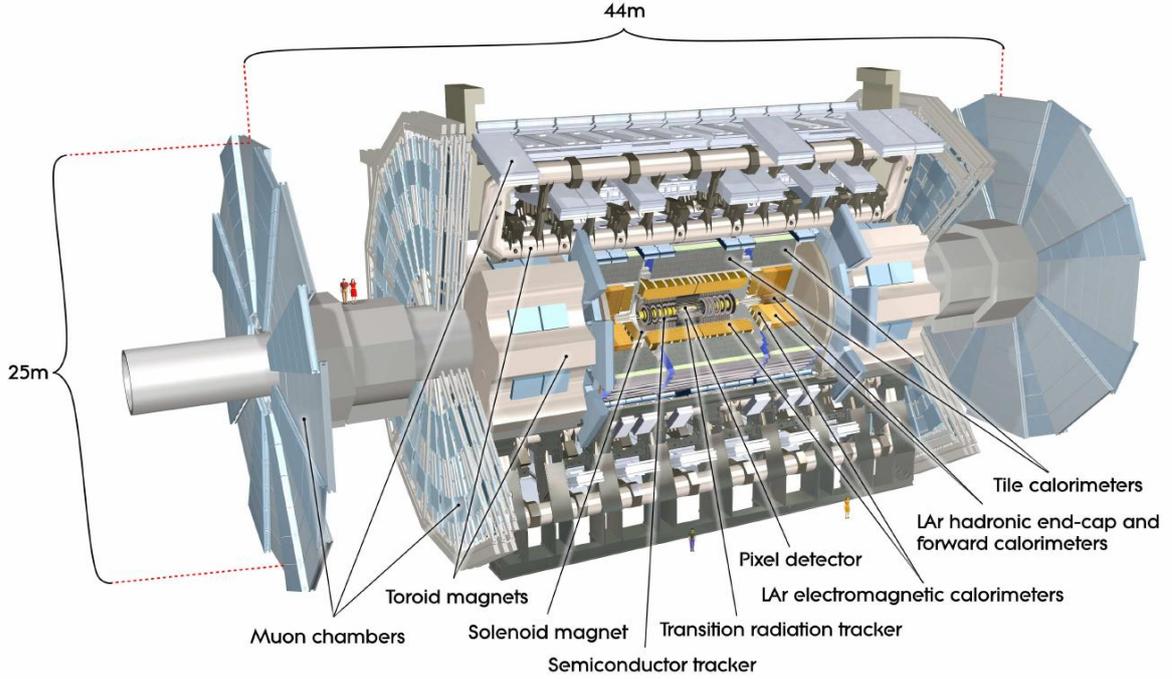


Figure 3.2: Sketch of the ATLAS detector

3.2.2 Inner Detector

The detector system closest to the beam line is the Inner Detector (ID), which is responsible for tracking charged particles. The ID is immersed in a magnetic field of 2 T, generated by a thin superconducting solenoid. Due to the Lorentz force, a charged particle traverses the ID in curved tracks. The charge and momentum of the particle can be inferred by the direction and strength of the curvature of the track.

The innermost part of the ID is the Insertable B-Layer (IBL). It was integrated into the detector during the long shutdown to improve the identification of hadrons containing b -quarks [26]. The IBL is surrounded by semiconductor pixel and silicon microstrip detectors, which provide a precise momentum and vertex measurement and cover the region $|\eta| < 2.5$. The outer part consists of straw-tube detectors. It is used for the identification of transition-radiation photons and is therefore named Transition Radiation Tracker (TRT).

3.2.3 Calorimeters

The ID is enclosed by a calorimeter system, consisting of the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL) and covering the range $|\eta| < 4.9$. They are responsible for the energy measurements, which work as follows: After passing through the ID, a particle is stopped in the calorimeter and deposits its energy there. The interaction with the calorimeter material leads to a shower of other particles [27]. The energy and position of this shower is then measured.

Electrons and photons are detected in the ECAL, which is a sampling calorimeter, i.e. it is composed of two materials ordered in alternate layers. Liquid-argon (LAr) layers are used to measure the energy deposit, while lead plates are used as absorbing material.

The ECAL is surrounded by the HCAL. This is also a sampling calorimeter designed for the

detection of strongly interacting particles. In its barrel region, called tile calorimeter, steel is used as absorber and scintillating tiles as active material. The end-cap region is composed of two parts: The hadronic end-cap calorimeter, with copper as absorbing medium, and the forward calorimeter, using copper and tungsten as absorber. Both calorimeters have a LAr active layer.

3.2.4 Muon system

Muons penetrate the ECAL as well as the HCAL, which allows a dedicated momentum measurement in an outer detector system. The muon system constitutes the outermost part of the ATLAS detector. For the muon reconstruction the same principle is applied as in the ID, namely the deflection of a particle's track due to a magnetic field. The magnetic field is produced by three air-core toroids (one barrel and two inserted end-caps) with an eight-fold azimuthal symmetry. The tracking proceeds through three stations of chambers. They are based on two different techniques: Monitored Drift Tubes (MDTs) measure the track in the range $|\eta| < 2$, whereas for $2 < |\eta| < 2.7$ this is done with Cathode Strip Chambers (CSCs). Besides for tracking, there are also muon chambers for triggering (see below). They use Resistive Plate Chambers (RPCs) in the barrel and Thin Gap Chambers (TGCs) in the end-cap regions.

3.2.5 Trigger system

At the LHC the proton bunches collide at a rate of approximately 30 MHz [28]. Additionally, each bunch crossing allows for several particle collisions. Due to limitations in bandwidth and storage capacities it is impossible to save all of these collisions to disk. Furthermore, most of the interesting processes are produced with a very small cross section, so that only a small fraction of collisions gives final states that are relevant for further analysis. Therefore it is important to minimize the recorded events without discarding the interesting processes. For this purpose the ATLAS detector uses a two-level trigger system.

The hardware-based first level trigger (L1) reduces the initial event rate to 100 kHz [28]. It evaluates coarse granularity information from the calorimeters and the muon system. By this it defines certain Regions-of-Interest (RoIs) in the detector within a decision time of $2.5 \mu\text{s}$ per event.

The RoIs are then transmitted to a software-based high level trigger (HLT), in which the full granularity detector information is analyzed by running suitable selection algorithms. With a processing time of approximately 200 ms, the HLT further reduces the event rate to about 1 kHz.

3.3 Particle identification

To determine which particles are produced in the proton-proton collisions, the information from all ATLAS sub-detectors needs to be analyzed and combined. In the following it will be explained, what the characteristic signatures of the individual particles are. These are also illustrated in Figure 3.3. However, it is important to note that thereby only particle candidates are reconstructed, because there is always a certain probability for a particle to be mis-identified.

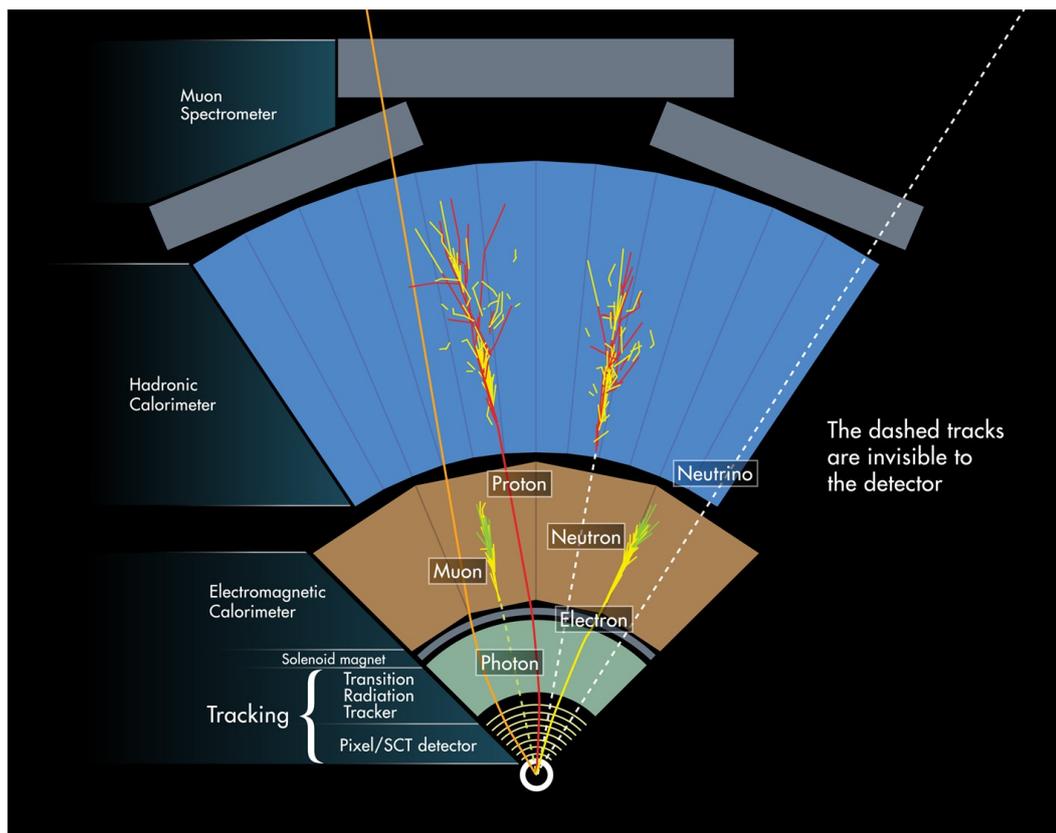


Figure 3.3: Illustration of a detector segment showing how particles are reconstructed at the ATLAS detector [29]. A dashed line indicates a particle, which is invisible to the particular detector component.

Electrons

Electrons leave a curved track in the ID and are stopped in the ECAL, where they induce a particle shower.

Photons

As photons do not carry an electric charge, they are invisible in the ID. They deposit their energy in the ECAL.

Muons

With a mean lifetime of 2.2×10^{-6} s muons can penetrate the full detector before they decay. They are tracked in the ID as well as the muon system and pass the ECAL and HCAL without causing a particle shower.

Quarks and gluons

Quarks and gluons cannot be observed as isolated objects because of color confinement, i.e. they form colorless combinations of quarks and antiquarks (hadrons). The only exception is the top quark, which decays before it can undergo hadronization. This collection of hadrons is called a jet. If a hadron is electrically charged, it leaves a track in the ID. It passes the ECAL and is then stopped in the HCAL, where it induces a shower of secondary particles. Furthermore, it is possible to “tag” heavy quarks, like b -quarks, in the jet by identifying displaced secondary vertices in the ID.

Taus

Due to a mean lifetime of 2.9×10^{-13} s most tau leptons decay before leaving the beam pipe, so that at ATLAS it is not attempted to measure taus from leptonic decays. Therefore the tau reconstruction concentrates on the hadronic decay modes, which occur at about 65% of the time. These decay modes produce mostly one or three charged pions, a tau neutrino and often additional neutral pions. The hadronic tau decays are classified according to the number of charged decay products, i.e. the number of tracks or *prongs* in the ID. The 1-prong decays are the most common, followed by the 3-prong decays [30]. The low track multiplicity is the main criterion to discriminate a tau jet from a quark or gluon initiated jet.

Neutrinos

As neutrinos interact only weakly, they penetrate the full detector without giving a signal. This results in a momentum imbalance in the transverse plane. Given that the initial protons have no transverse momentum, the p_T sum of all particles after the collision must be zero due to momentum conservation. This means that the contribution from the visible, reconstructed particles, each giving $\vec{p}_{T,i}^{\text{vis}}$, must be compensated by a missing transverse momentum \vec{p}_T^{miss} :

$$\sum_i \vec{p}_{T,i}^{\text{vis}} + \vec{p}_T^{\text{miss}} = 0 \quad (3.1)$$

Neglecting particle masses, one defines the missing transverse energy E_T^{miss} by:

$$E_T^{\text{miss}} := |\vec{p}_T^{\text{miss}}| = \left| - \sum \vec{p}_T^{\text{vis}} \right| \quad (3.2)$$

Therefore, measuring E_T^{miss} indicates the presence of neutrinos. Apart from neutrinos, E_T^{miss} can have also other sources, e.g. detector leakage or mis-measurements. Moreover, many BSM theories predict stable, weakly interacting particles, which would contribute to E_T^{miss} .

Chapter 4

Data and Monte Carlo Simulation

4.1 Collected data

At ATLAS the amount of collected data is expressed through the integrated luminosity L , which is connected to the number of events N by:

$$N = \sigma L = \sigma \int \mathcal{L} dt \quad (4.1)$$

with σ being the cross section for the studied process. In the last equation L was replaced by the time integral of the instantaneous luminosity \mathcal{L} . It is defined through several beam parameters:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \varepsilon_n \beta^*} F \quad (4.2)$$

where N_b is the number of protons per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic Lorentz factor, ε_n the normalized transverse beam emittance and β^* the beta function at the collision point. The geometric factor F is a measure for the luminosity reduction occurring due to the crossing angle at the interaction point [19].

The analysis in this work uses data, which was recorded in 2015 and 2016 at $\sqrt{s} = 13$ TeV. This corresponds to an integrated luminosity of 36.1 fb^{-1} .

4.2 Monte Carlo simulation

Particle collisions can be simulated with Monte Carlo (MC) generators. They are used for the estimation of signal and background processes and make it possible to compare theoretical predictions to data. The simulation of a specific process includes following parts [31, 32]:

- Hard scattering
- Parton showers
- Hadronization
- Underlying events
- Decays of unstable particles

The hard scattering constitutes the actual process of interest in a collision, i.e. in this step a high momentum transfer takes place or heavy objects are created. The probability distribution for a particular hard scatter is calculated from perturbation theory. Parton showers are the result of initial and final state radiation: Quarks and gluons can radiate off gluons, which itself can split into a quark-antiquark pair or radiate again gluons, thereby causing a shower. The simulation of these showers proceeds stepwise, starting from the hard process downwards to lower momentum scales. At a scale of the order of 1 GeV perturbation theory breaks down and the quarks and gluons start to hadronize. This simulation step makes use of non-perturbative hadronization models. Besides the hard scattering, a proton-proton collision allows also for more interactions, in which additional, soft (i.e. low energetic) particles are produced. These interactions are called underlying events or pile-up. The last step of the simulation deals with particle decays. As many of the particles produced in the collision are unstable resonances, it is necessary to model their subsequent decays down to particles, which can be observed in the detector.

To compare the number of data to the number of MC events N_{MC} , the latter must be scaled to the integrated luminosity $\int \mathcal{L} dt$ of the data samples and the predicted cross section σ_{MC} of the MC process. Furthermore, two additional correction factors need to be applied: the k-factor k and the filter efficiency ϵ_{filter} . The k-factor describes effects from next-to-leading order calculations, while the filter efficiency accounts for the fact, that the generator input can be subjected to filters to choose only useful particles. This results in the scaling factor f_{MC} :

$$f_{\text{MC}} = \frac{k \cdot \epsilon_{\text{filter}} \cdot \int \mathcal{L} dt \cdot \sigma_{\text{MC}}}{N_{\text{MC}}} \quad (4.3)$$

In the following it will be described, which MC generators are used in this analysis to simulate background and signal processes. Table 4.1 gives a summary of the various physics processes and their corresponding MC generators. A list of the MC samples with their dataset ID, cross section, k-factor and filter efficiency can be found in Appendix A.

Process	Generator(s)
ZZ, WZ, WW	SHERPA [33]
$VVV(WWW, WWZ, WZZ, ZZZ)$	SHERPA
Higgs ($t\bar{t}H$)	POWHEG [34] + PYTHIA 8 aMC@NLO + PYTHIA 8
$t\bar{t}Z, t\bar{t}W, t\bar{t}WW$	MADGRAPH 5_aMC@NLO [35] + PYTHIA 8 [36]
tWZ	aMC@NLO [37] + PYTHIA 8
$t\bar{t}\bar{t}, t\bar{t}t\bar{t}$	MADGRAPH 5_aMC@NLO + PYTHIA 8
$t\bar{t}$	POWHEG + PYTHIA 6 [38]
$Z + \text{jets}, W + \text{jets}$	MADGRAPH 5_aMC@NLO + PYTHIA 8
SUSY RPV signal	MADGRAPH 5_aMC@NLO + PYTHIA 8

Table 4.1: Overview of the MC generators used for the simulation of signal and background processes in the analysis.

4.2.1 Monte Carlo background samples

The generator SHERPA [33] is used for the modelling of diboson and triboson events, while $t\bar{t}$ and Higgs processes are simulated with POWHEG [34]. The only exception is $t\bar{t}H$, which is estimated with aMC@NLO. MADGRAPH 5 interfaced with aMC@NLO [35] is used to simulate the backgrounds coming from multitop events, Z +jets and $t\bar{t}$ along with a Z , W or WW . The contribution from tWZ processes is estimated with aMC@NLO [37]. Apart from the background processes themselves, also underlying events from the same or nearby bunch crossings need to be taken into account. These are generated with PYTHIA 8 [36] and are then overlaid to the hard scattering events. For the $t\bar{t}$ sample PYTHIA 6 [38] is used instead of PYTHIA 8. The passage of particles through the detector material is simulated with GEANT 4 [39] using the full ATLAS detector simulation (Fullsim)[40].

4.2.2 Monte Carlo signal samples

Signal samples are generated for RPV models satisfying $10 \text{ GeV} \leq m(\tilde{\chi}_1^0) \leq m(\text{NLSP}) - 10 \text{ GeV}$. This ensures that the decay products from the NLSP as well as the $\tilde{\chi}_1^0$ decay may be prompt. The RPV signal scenarios are simulated with MADGRAPH 5 interfaced with aMC@NLO. The detector response is modeled using the fast ATLAS simulation (AtFast-II) [40], which allows to speed up the MC processing compared to Fullsim. This is achieved by removing electromagnetic particles with low energy from the calorimeter and replacing them by pre-simulated showers. The other detector components are described by GEANT 4.

Chapter 5

Object and event selection

As discussed in Section 3.3, the reconstruction of the various particle types requires to combine information from the different detector components, so that the measured tracks and energy deposits are assigned to the different particles. However, this does not always proceed correctly: An object can be mis-identified as another one, e.g. a jet could falsely be interpreted as a tau. On the other hand, for example, it can happen that a real tau is not reconstructed at all. For this reason certain quality criteria are introduced, which help to identify the different particles as correctly as possible. In the case of leptons and b -jets one defines so called working points. These are sets of identification criteria providing certain reconstruction efficiencies. They are usually divided into the categories *loose*, *medium* and *tight*. Applying the tight quality criteria ensures a high purity of the reconstructed particles, i.e. a low mis-reconstruction rate. On the other hand, tighter quality criteria also increase the probability to falsely discard a particle, meaning that the reconstruction efficiency is lower and thus also the statistics. In each analysis it must be decided, which quality criteria give the best compromise between purity gain and statistical losses.

However, not all reconstructed particle candidates are also relevant for the analysis because the signal models under study predict certain properties for their event objects. For example, most analyses are interested in high energetic events, so that it would be generally useful to consider only events, in which the objects have quite high transverse momentum. Therefore additional selection criteria are applied on different variables to reject the events, which are not needed for the further analysis. These criteria can include e.g. restrictions on the number of certain particles or minimum values for their transverse momentum. In general, this selection is divided into several steps: The first step is a coarse preselection. Objects passing the preselection are denoted as *baseline* objects. Afterwards more sophisticated criteria are applied to make a final event selection. The remaining particles are then referred to as *signal* objects. Only the signal objects are considered as potential outcomes of a supersymmetric process. However, baseline objects are included into the analysis, too, as they are needed for the background estimation (see Chapter 7).

This section describes, which criteria are used in this analysis to identify the particles and select the interesting events.

5.1 Event cleaning

Before being considered for analysis, events must pass a cleaning process consisting of following parts:

- **Good Runs List**

The Good Runs List (GRL) is used to remove luminosity blocks (corresponding to approximately 1-2 minutes of data taking), which were affected by detector problems.

- **Removing additional detector problems**

While the GRL rejects entire luminosity blocks with poor detector performance, there can still remain single bad events in otherwise perfect luminosity blocks. Therefore event-level detector flags are introduced. They help to remove events that are considered bad due to problems in the LAr or tile calorimeter system. The detector flags and the GRL are applied only to data and leave Monte Carlo (MC) simulated events of course unaffected.

- **Trigger**

As the analysis is specialized on multileptonic final states, one can use lepton triggers to preselect interesting events. An event must fire at least one of the triggers described in Section 5.4.

- **Primary vertex**

Events are only kept if their reconstructed primary vertex has at least two tracks with $p_T > 400$ MeV. The primary vertex of an event is defined as the vertex with the highest $\sum p_T^2$ of associated tracks. This requirement helps to reduce effects from pile-up, i.e. underlying processes, which are not related to the hard collision.

- **Cosmic muon veto**

A veto is applied to events with a cosmic muon. A muon is said to originate from cosmic rays if $z_0 > 1$ mm and $d_0 > 0.2$ mm, where z_0 and d_0 are its longitudinal and transverse impact parameter with respect to the primary vertex. In a scattering experiment, the impact parameter is defined as the distance, by which the a particle would miss the scattering center, if it continued on its initial trajectory [6].

- **Bad muon veto**

An event is discarded if it contains a bad muon, that is a muon having $\sigma(q/p)/|q/p| > 0.2$ where q is the charge of the muon and p its momentum. $\sigma(q/p)$ is the uncertainty on the measured value of $|q/p|$.

- **Jet cleaning**

Jets arising from detector noise, cosmic rays or beam induced backgrounds are suppressed by applying the `VeryLooseBad` quality criteria described in [41]. These quality criteria include requirements on the signal pulse shape in the LAr calorimeters, energy ratios and track-based variables.

5.2 Object selection

Only if an event passes the event cleaning, the objects in this event can be considered as potential candidates for the analysis. The object selection consists of three phases: Preselection, overlap removal and signal selection. Objects satisfying the preselection criteria undergo an overlap removal (OR). The OR ensures that no double counting of objects takes place. Finally, if an object passes the OR, it is subjected to a last selection process picking out the signal objects. Baseline leptons, which pass the OR, but fail at least one of the signal criteria, are referred to as a loose leptons. These play an important role in the background estimation. In the following it will be explained, which preselection and signal criteria are applied to objects in this analysis. A summary of these criteria is shown in Table 5.1.

Electrons

Electron candidates must fulfill the very loose identification criteria as described in [42]. These criteria impose requirements e.g. on the energy deposits and shower widths in the ECAL or on the number of hits in the pixel and silicon microstrip detectors of the ID. Moreover, electrons are required to have $|\eta| < 2.47$ and $p_T > 7 \text{ GeV}$. This rather low p_T cut and the loose working point provide high statistics in the control regions (see Chapter 7) needed for the background estimation, while leaving the sensitivity to SUSY scenarios unaffected.

Signal electrons must satisfy the criteria of the medium working point. To suppress electrons originating from secondary vertices, signal electrons are required to have $d_0/\sigma(d_0) < 5$ and $z_0 \sin \theta < 0.5 \text{ mm}$, where $d_0/\sigma(d_0)$ is the uncertainty of d_0 in standard deviations. Furthermore, they must pass several p_T dependent isolation criteria. This guarantees better discrimination against semileptonic decays of hadrons and mis-identified jets.

Muons

Baseline muons must have $|\eta| < 2.7$ and $p_T > 5 \text{ GeV}$ and satisfy the medium identification requirements from [43]. Also here a low p_T threshold allows to increase statistics in the control regions without to affect the sensitivity on SUSY signals.

As in the case of electrons, signal muons are required to pass certain p_T dependent isolation requirements to reduce the contribution from semileptonic hadron decays and mis-identified jets. The contribution of muons from secondary vertices is reduced by demanding $d_0/\sigma(d_0) < 3$ and $z_0 \sin \theta < 0.5 \text{ mm}$.

Jets

Jet candidates are required to have $p_T > 20 \text{ GeV}$ and $|\eta| < 4.9$. They are reconstructed with the anti- k_t algorithm [44]. This is a sequential recombination algorithm, by which particles are iteratively combined starting from the ones with the closest distance to each other. The anti- k_t algorithm contains a radius parameter $R = 0.4$, which controls the extension of the jet. The calibration of jets follows the quality criteria described in [45], which includes corrections to the jet energy and resolution. In order to mitigate pile-up, jets with $p_T < 60 \text{ GeV}$ and $|\eta| < 2.4$ are required to have a significant fraction of tracks that originate from the primary vertex [46]. A variable, which gives a measure for this, is the jet vertex tagger (JVT). Its calculation is based on an algorithm, with which the ratio of the number of hard scatter jets to the number of hard scatter plus pile-up jets is computed in a training sample. A jet is classified as coming from pile-up if $\text{JVT} < 0.59$.

Signal jets are usually located in the central region of the detector and not in the forward region. Therefore a signal jet is defined as a non-pile-up jet satisfying a tighter pseudorapidity requirement of $|\eta| < 2.8$.

Taus

Tau identification at ATLAS is restricted to hadronic decays. The tau reconstruction algorithm uses jets as input, which have $p_T > 10 \text{ GeV}$ and $|\eta| < 2.47$. As above, the seed jets are reconstructed with the anti- k_T algorithm using $R = 0.4$. The algorithm analyzes track and calorimeter information within $\Delta R < 0.2$ of the seed jet to discriminate taus from gluon and quark initiated jets [47]. Tau candidates must have $p_T > 30 \text{ GeV}$ and either one or three prongs, i.e. tracks in the ID (see Section 3.3), with unit total charge. For baseline taus no working point is applied, again with the purpose of increasing statistics in the control regions.

The working point is then introduced as a signal requirement: Taus must pass the medium identification criteria, which are based on a boosted decision tree (BDT) algorithm [47].

Missing transverse energy

The missing transverse energy E_T^{miss} , defined in Equation 3.2, describes the energy imbalance in the transverse plane. For the reconstruction of E_T^{miss} , the \vec{p}_T values of all detector entities need to be summed. This includes all reconstructed and fully calibrated objects (electrons, muons, photons, taus and jets) as well tracks that originate from the primary vertex, but are not associated to any physics objects [48]. This allows E_T^{miss} to be almost independent of pile-up effects. The former contributions constitute the *hard term* and the latter the *soft term* in the calculation of E_T^{miss} , which is given by:

$$E_T^{\text{miss}} = | - (\vec{p}_T^e + \vec{p}_T^\mu + \vec{p}_T^\gamma + \vec{p}_T^{\text{jet}} + \vec{p}_T^\tau + \vec{p}_T^{\text{soft}}) | \quad (5.1)$$

Electron	Muon	Tau	Jet	Pile-up jet
Preselected				
$p_T > 7 \text{ GeV}$	$p_T > 5 \text{ GeV}$	$p_T > 30 \text{ GeV}$	$p_T > 20 \text{ GeV}$	$p_T < 60 \text{ GeV}$
$ \eta_{\text{cluster}} < 2.47$	$ \eta < 2.7$	$ \eta < 2.47$	$ \eta < 4.9$	$ \eta < 2.4$
VeryLoose	Medium			JVT < 0.59
				anti- k_t , $R = 0.4$
Overlap Removal see Subsection 5.3				
Signal				
Medium		Medium	$ \eta < 2.8$	
$d_0/\sigma(d_0) < 5$	$d_0/\sigma(d_0) < 3$		not pile-up jet	
$z_0 \sin \theta < 0.5 \text{ mm}$				
p_T dependent isolation				

Table 5.1: Preselected and signal object definition.

5.3 Overlap removal

The OR handles problems coming from the overlap of objects, e.g. what should be done when one signature in the detector is reconstructed as two distinct objects, or how to treat particles, that are very close to each other. In this analysis the OR consists of seven steps applied in the following order:

1. If a tau overlaps with an electron or muon within $\Delta R = 0.2$, then the tau is removed.
2. Any electron, which shares an ID track with a muon, is discarded.
3. If a jet and an electron overlap within $\Delta R = 0.2$, the jet is removed.
4. For an electron-jet overlap within $0.2 < \Delta R < 0.4$, the electron is discarded in order to suppress electrons arising from semileptonic decays of b or c hadrons.

5. A jet with less than two tracks is removed either if it overlaps with a muon within $\Delta R = 0.2$ or if the muon is matched to a track, which is associated to the jet.
6. To suppress muons from semileptonic decays of b or c hadrons, muons are discarded if they overlap with a jet within $\Delta R = 0.4$.
7. Jets within $\Delta R = 0.4$ of a tau, which fulfills the medium identification requirements, are discarded.

After OR an additional cleaning procedure is applied, which removes low-mass processes leading to leptons:

- An opposite sign pair (OS pair) of light leptons, i.e. e^+e^- , $e^\pm\mu^\mp$ or $\mu^+\mu^-$, is removed if the invariant mass of the OS pair m_{OS} is smaller than 4 GeV.
- The decay of an Υ -meson (consisting of $b\bar{b}$) can produce light lepton pairs of same flavor and opposite sign (SFOS pair). To veto this contribution, both members of an SFOS pair are removed if $8.4\text{ GeV} < m_{\text{SFOS}} < 10.4\text{ GeV}$.

5.4 Trigger strategy

Signal electrons and muons are used to trigger events containing four leptons. They must fire one of the single or double lepton triggers and pass the corresponding offline p_{T} thresholds shown in Table 5.2. The single lepton triggers have priority over the dilepton triggers, which are used only if the event fails the p_{T} requirement for the single lepton trigger. The offline p_{T} threshold is set to 1 GeV above the trigger p_{T} level. This ensures that the lepton triggering has an efficiency of basically 100%. The trigger thresholds were raised for data recording in 2015 compared to 2016 because of the increase in luminosity. To make data comparable to MC simulation of the trigger behavior, scale factors are applied on the MC events similarly as in Section 4.2.

Trigger	Offline p_{T} threshold [GeV]	
	2015	2016
Single isolated e	25	27
Single non-isolated e	61	61
Single isolated μ	21	25 or 27
Single non-isolated μ	41	41 or 51
Double e	13,13	18,18
Double μ	–	11,11 or 15,15
	19,9	21,9 or 23,9
Combined $e\mu$	8(e),25(μ)	8(e),25(μ)

Table 5.2: Overview of the triggers and offline p_{T} thresholds used in this analysis for 2015 and 2016. An “or” indicates that the trigger threshold has been raised during data taking.

Chapter 6

Signal region optimization

In order to be sensitive to a potential signal from RPV models, it is necessary to find suitable criteria to reduce the SM background while enhancing the signal. The collection of these criteria defines the so called signal regions (SR). For the $LL\bar{E}i33$ models two new signal regions are introduced targeting $3\ell 1\tau$ and $2\ell 2\tau$ final states. The optimization of these signal regions is performed using three reference models of the simplified model with initial $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production (see Figure 2.3a). These three models have $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^0$ masses lying just above the exclusion limits from Run I [3], i.e. models with these masses are not excluded by the Run I results and have still a potential of being discovered. The chosen models are: $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 100)$, $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 490)$ and $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (600, 300)$. This section describes, how the two signal regions are constructed based on the study of event variables and the statistical significance.

6.1 Z -veto

As the signal events contain at least two or three light leptons, with potentially more originating from the W -decays, one of the main SM background contribution are processes where a Z -boson decays into light leptons:

$$Z \rightarrow \ell^+ \ell^- \tag{6.1}$$

This is illustrated in Figure 6.1, which shows the invariant mass of an electron or muon pair for the SM background in $3\ell 1\tau$ and $2\ell 2\tau$ events. There is a clear peak around the Z -mass ($m_Z \approx 91\text{GeV}$) mainly caused by Z +jets and ZZ processes.

Therefore, an effective way to reduce this background source is to apply a Z -veto. The Z -veto discards events, in which the invariant mass of two light leptons $m(\ell^+ \ell^-)$ lies within a window of $\pm 10\text{GeV}$ around m_Z . Or equivalently, events must satisfy:

$$|m(\ell^+ \ell^-) - m_Z| > 10\text{ GeV} \tag{6.2}$$

To suppress radiative Z -decays, the condition from above is applied also to the invariant mass of three and four light leptons:

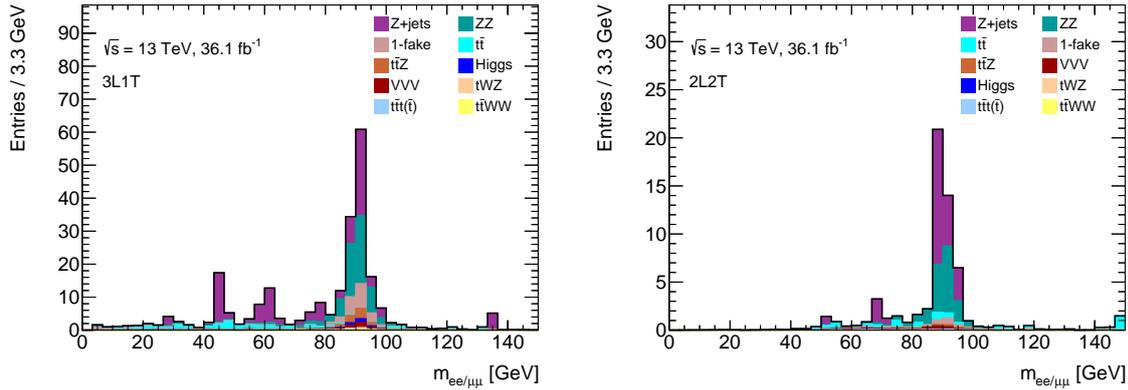


Figure 6.1: Invariant mass of ee or $\mu\mu$ pairs for SM background processes in $3l1\tau$ (left) and $2l2\tau$ events (right)

$$\begin{aligned} |m(\ell^+\ell^-\ell'^{\pm}) - m_Z| &> 10 \text{ GeV} \\ |m(\ell^+\ell^-\ell'^+\ell'^-) - m_Z| &> 10 \text{ GeV} \end{aligned} \quad (6.3)$$

where ℓ' does not necessarily need to have the same flavor as ℓ . As Z -bosons are not included in the $\tilde{\chi}_1^+\tilde{\chi}_1^-$ model, the signal is not expected to be affected by this requirement. This behavior is also expected for the models with a gluino or left handed slepton as NLSP (see Figure 2.3). Only the model with an initial $\tilde{\chi}_1^+\tilde{\chi}_2^0$ is assumed to be affected by this requirement as a Z -boson is produced in the decay of $\tilde{\chi}_2^0$ to $\tilde{\chi}_1^0$.

Therefore, for the three $\tilde{\chi}_1^+\tilde{\chi}_1^-$ models applying a Z -veto should result in a rather high signal-to-background ratio, while the opposite should be the case in events demanding the presence of a Z -boson. This can be seen in Figure 6.2 and Figure 6.3. The histograms show the distribution of H_T^{lep} for SM background and signal in $3l1\tau$ and $2l2\tau$ events, respectively, both for the cases with an explicit Z -requirement ($N_Z \geq 1$) as well as with a Z -veto. H_T^{lep} is defined as the sum of transverse momenta of all leptons:

$$H_T^{\text{lep}} = \sum_{\ell=e,\mu,\tau} p_T(\ell) \quad (6.4)$$

From the histograms it is clear, that the signal fraction is very low in events with at least one Z -boson, while with a Z -veto a decent signal-to-background ratio can be reached.

6.2 Discriminating variables

As can be seen in Figure 6.2 and Figure 6.3, the signal-to-background ratio is especially large for high values of H_T^{lep} . Hence it would be reasonable to include only high H_T^{lep} values into the signal region. However, cuts on other event variables could also be useful. This can be investigated in the same way as for H_T^{lep} , namely by plotting the distribution of a variable for SM background and signal without any requirements on the variable itself, but with all other cuts applied (in this case only the Z -veto). These types of plots are therefore called N-1 plots, where N denotes the number of discriminating variables of a signal region. However, it is complicated to determine an ideal value for a cut only by comparing the signal and background distributions in a histogram. Therefore it is required to have a measure for the

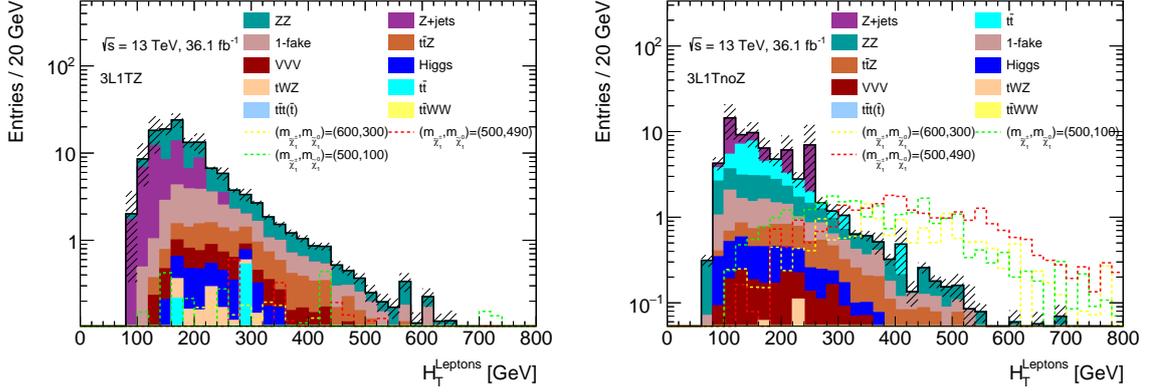


Figure 6.2: Distribution of H_T^{lep} for SM backgrounds and signal in $3l1\tau$ events with $N_Z \geq 1$ (left) and with an applied Z -veto (right).

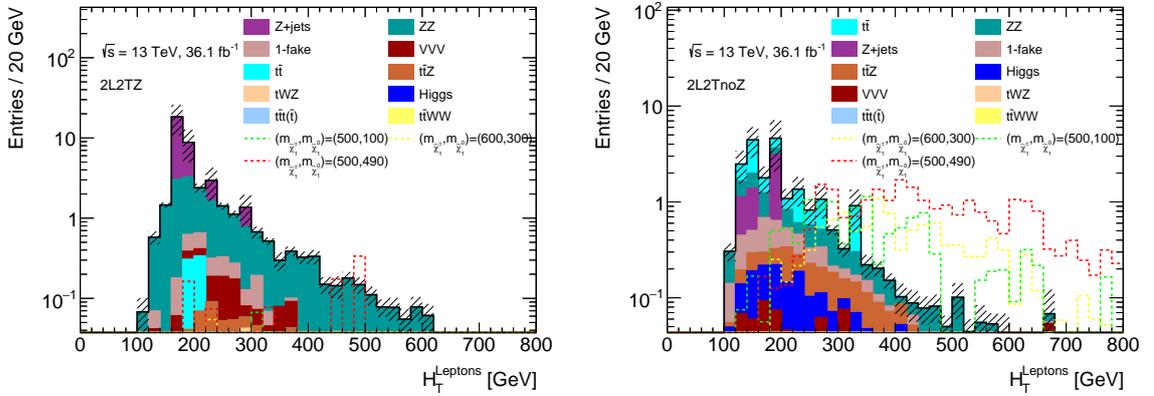


Figure 6.3: Distribution of H_T^{lep} for SM backgrounds and signal in $2l2\tau$ events with $N_Z \geq 1$ (left) and with an applied Z -veto (right).

discovery potential of a signal in a certain region. Then one can vary the applied cut on a variable and compare, how the discovery potential changes. The optimal cut should be chosen such that the discovery potential at this value is sufficiently high and still enough signal statistics is retained.

The discovery potential is determined by calculating the significance of a signal excess in a hypothesis test, in which the null hypothesis H_0 is the background-only and the alternative hypothesis H_1 the signal-plus-background assumption. The basis of the hypothesis test are number counting experiments for the background-only and signal-plus-background assumptions. Background events follow a Poisson probability distribution, which is given by:

$$P(n_i|b_i) = \frac{b_i^{n_i}}{n_i!} e^{-b_i} \quad (6.5)$$

Here $P(n_i|b_i)$ is the probability to observe n_i events in bin i when b_i background events are expected. The same holds also for the signal-plus-background model:

$$P(n_i|s_i, b_i) = \frac{(s_i + b_i)^{n_i}}{n_i!} e^{-(s_i + b_i)} \quad (6.6)$$

where s_i are the number of expected signal events in bin i . The outcome of the hypothesis test can be expressed by the p -value, which gives the probability to observe the same or a

more extreme result than the actual measured one, assuming the null hypothesis to be true. In this case it describes, how likely it is that an observed excess, potentially coming from the signal-plus-background model, is caused by background processes and their fluctuations. The p -value is equivalently formulated by the significance Z , which is the number of standard deviations a Gaussian variable would need to fluctuate in one direction to give the same p -value. The significance is obtained by:

$$Z = \Phi^{-1}(1 - p) \quad (6.7)$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian [49]. If a significance of 3σ is observed, one speaks about an evidence of excess. To claim a discovery, a significance of at least 5σ is required, which corresponds to a p -value of less than $2.9 \cdot 10^{-7}$. For sufficiently large background b and very low signal contamination, i.e. $s \ll b$, the significance can be approximated by:

$$Z = \frac{s}{\sqrt{b}} \quad (6.8)$$

In this analysis the significance is determined with a more accurate calculation taking into account also the uncertainty on background events. This calculation is already implemented in a function of the ROOT software framework [50]. The function used is [51]:

$$\text{RooStats::NumberCountingUtils::BinomialExpZ}(s, b, \Delta b_{\text{rel}}),$$

where s and b are the number of signal and background events and Δb_{rel} the relative uncertainty on the background. The function is based on the statistical evaluation of the Poisson distributed signal-plus-background model and of the relative background uncertainty. The latter is interpreted as originating due to an auxiliary observation that follows also a Poisson distribution from the background-only model. Δb_{rel} is given by the fraction of the absolute uncertainty over the background events:

$$\Delta b_{\text{rel}} = \frac{\Delta b}{b} = \sqrt{\left(\frac{\Delta b_{\text{stat}}}{b}\right)^2 + \left(\frac{\Delta b_{\text{syst}}}{b}\right)^2} \quad (6.9)$$

where in the last equation the background uncertainty was split into its statistical and systematic component. The statistical uncertainty is applied on an event-by-event basis, while for the systematic uncertainty an overall value of 30% is taken as approximation. This value is quite high compared to the uncertainty of 15%, which was obtained in the ICHEP search of 2016 [5]. As in this search taus were not included, a larger uncertainty is assumed in order to account for higher systematic uncertainties connected with the tau reconstruction.

To find the optimal threshold for a discriminating variable, the following method is used: Starting with no or only a small cut on the particular variable, the total number of signal and background events in the signal region is counted. For the resulting values of s, b and Δb_{rel} the significance is calculated. Afterwards the cut on the variable is increased and with the new s, b and Δb_{rel} the significance is determined. This is iteratively done over the whole variable range. Finally, the cut value is chosen such that the significance is maximal.

This optimization procedure is illustrated in Figure 6.4, which shows the N-1 plots for the effective mass m_{eff} in the first row. The effective mass is defined as the sum of transverse momenta of all objects and the missing transverse energy:

$$m_{\text{eff}} = \sum_{\ell=e,\mu,\tau} p_T(\ell) + \sum_j p_T(j) + E_T^{\text{miss}} \quad (6.10)$$

where the second term gives the p_T sum for jets. The initial total number of s and b corresponds to the integral of the signal and background entries in the N-1 plot. The cut on m_{eff} is increased bin per bin, so that the integrals must be calculated for each bin again to get the significance. Only histogram entries are included in the integral, which lie above the current cut value. These integrals constitute the cumulative distribution of the effective mass, shown in Figure 6.4 in the second row. The bottom panels show the resulting significance per bin. As the three reference signals have their significance maximum at different m_{eff} points, the final value for the cut should lie in between these maxima. For $3\ell 1\tau$ events a requirement of $m_{\text{eff}} > 700$ GeV is chosen, while for $2\ell 2\tau$ events a slightly looser cut of $m_{\text{eff}} > 650$ GeV is defined.

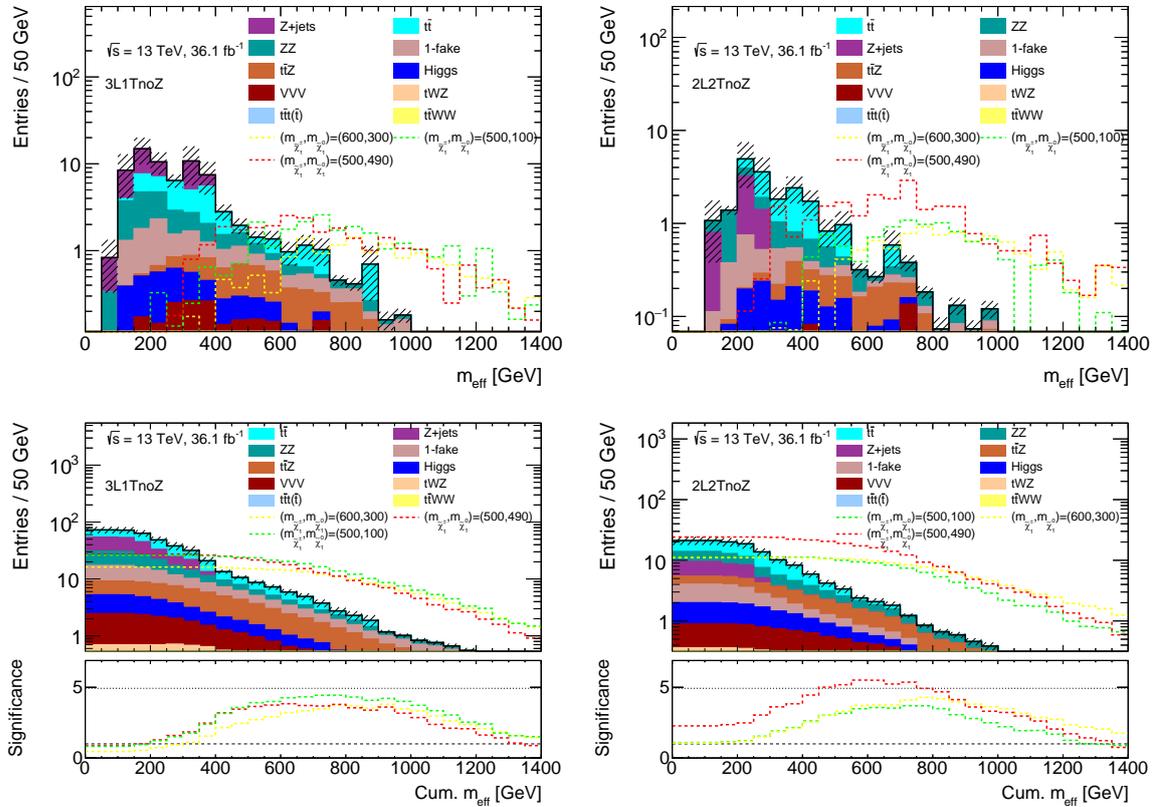


Figure 6.4: The top row shows the N-1 plots for the m_{eff} distribution of SM background and signal processes in $3\ell 1\tau$ (left) and $2\ell 2\tau$ events (right). The dashed lines represent the statistical uncertainty of the background. The corresponding cumulative distributions are shown in the histograms in the bottom row.

Apart from m_{eff} , also other variables were investigated as potential discriminating variables, for example H_T^{lep} or E_T^{miss} . However, the cut on m_{eff} turned out to be the most powerful. In many analyses it is common to use a set of several variables to define a signal region. This would require to study, how the significance changes under all possible combinations of cuts. In this case, however, this is not necessary because the m_{eff} cut alone already provides a sufficiently large discovery potential. As can be seen from the plots above, the significance is for all signal discovery models at least 3σ at the chosen cut value and can reach up to almost 6σ .

Furthermore, keeping the requirements on the signal regions simple has the big advantage of being sensitive to various new physics processes, which are characterized by high energetic multileptonic final states.

To summarize, two signal regions were constructed: One region targets $3\ell 1\tau$ events and is defined through a Z -veto and $m_{\text{eff}} > 700$ GeV. In the following this signal region will be referred to as SR1. The other one will be denoted by SR2, which is the region with $2\ell 2\tau$ final states, a Z -veto and $m_{\text{eff}} > 650$ GeV.

After defining the signal regions it is important to know, how sensitive they are to other signal models than the studied ones. Figure 6.5 shows the discovery significance for all models in the signal grid used in the analysis. One can see, that the 5σ level goes up to $\tilde{\chi}_1^\pm$ masses of approximately 550 GeV, while the 3σ contour includes points up to about 650 GeV.

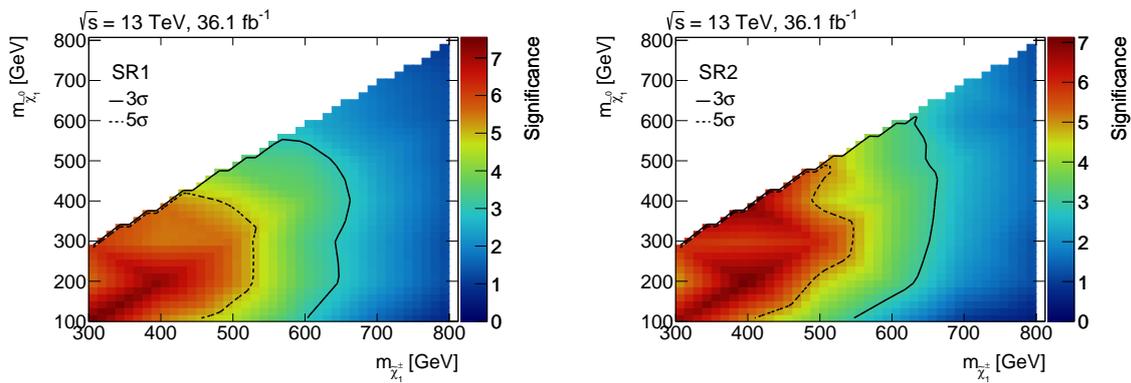


Figure 6.5: Discovery significance in the signal grid for SR1 (left) and SR2 (right).

The signal regions of the analysis are summarized in Table 6.1. Besides SR1 and SR2 it shows also the regions SR0A and SR0B, which are optimized for $4\ell 0\tau$ events [16, 17]. These signal regions allow to test the $LL\bar{E}12k$ as well as $LL\bar{E}i33$ models, in which the taus decay only leptonically.

	$N(e, \mu)$	$N(\tau)$	Z -boson	Selection
SR0A	≥ 4	$= 0$	veto	$m_{\text{eff}} > 600$ GeV
SR0B	≥ 4	$= 0$	veto	$m_{\text{eff}} > 1100$ GeV
SR1	$= 3$	≥ 1	veto	$m_{\text{eff}} > 700$ GeV
SR2	$= 2$	≥ 2	veto	$m_{\text{eff}} > 650$ GeV

Table 6.1: Signal region definitions.

Chapter 7

Background estimation

Using the signal region definitions shown in the previous chapter the SM backgrounds could be reduced, yet not completely eliminated. There remain several SM processes, which have the same signatures as an expected SUSY signal. Therefore it is important to have a good understanding of these processes and to accurately predict their contribution in the signal regions. The various backgrounds are grouped into two categories: The irreducible and reducible backgrounds. These two types of backgrounds are estimated in different ways and their predictions are finally tested in validation regions. In the following it will be explained, which properties these backgrounds have and which methods are used for their estimation.

7.1 Irreducible background

The irreducible background constitutes of processes leading to four “real” leptons. A lepton is classified as “real”, if it is prompt and genuinely isolated from other objects in the event. The dominating sources of the irreducible background are $t\bar{t}Z$ and ZZ processes, where the Z -boson and top quark decay into final states with leptons and are not rejected by the Z -veto. This can happen, when the Z -boson decays to taus, because the Z -veto is only applied to light leptons, or when the invariant mass m of the light leptons does not lie within the mass window $|m - m_Z| < 10 \text{ GeV}$. Other contributions to the irreducible background come from the production of three bosons (ZZZ , WZZ and WWZ), $t\bar{t}WW$, tWZ , Higgs and $t\bar{t}t\bar{t}/t\bar{t}tW$ (in the following summarized as multi-top background). The irreducible backgrounds are fully estimated from MC simulation, using the generators and samples listed in Section 4.2 and Appendix A.

7.2 Reducible background

The reducible background includes processes, where at least one “fake” lepton is produced. This is a lepton, which is non-prompt or not sufficiently isolated. Processes giving only one fake lepton are WZ , WWW and $t\bar{t}W$. The dominating contribution to the reducible background comes from process with two fake leptons, which are produced in $t\bar{t}$ and Z + jets events. Backgrounds yielding three or more fake leptons, e.g. W + jets, are very small and are therefore neglected in the analysis.

These processes yield one or two fake leptons for the following reasons: A W -boson produces only one charged lepton, when it decays leptonically by $W \rightarrow \ell\nu_\ell, \tau\nu_\tau$, while a Z -boson decay can result in two charged leptons. Therefore, a WZ process can give at most three charged

leptons. So if a WZ event passes the four-lepton requirement, at least one fake lepton must be involved. Similarly, WWW processes can give also at most three leptons. A top quark decays with a ratio of almost 100% by $t \rightarrow Wb$. The b quark is finally seen as a b -jet in the detector, so that the decay of a top quark can result at most in one charged lepton coming from the W decay. Therefore, also $t\bar{t}W$ processes cannot contain more than three real leptons in the final state. Analogously, $t\bar{t}$ events produce at most two real leptons. Hence, if a $t\bar{t}$ process passes the four-lepton requirement, there must be at least two fake leptons in the event. In order to obtain any charged leptons in Z +jets events, the Z -boson must decay leptonically. This gives again at most two charged leptons, so that also in this case two of the four leptons in the signal event are fakes.

7.2.1 Sources of fake leptons

Fake leptons can originate from different sources. To estimate the lepton contribution of each fake source, MC information about the lepton origins and types is used. For taus additional information regarding the partons inside the jet and the cone is evaluated. By combining these information one can infer the source of the fake leptons. The main sources are the following:

- **Heavy flavor (HF) leptons**

HF leptons arise from semileptonic decays of hadrons with a b or c quark. For $t\bar{t}$ events this is the dominant fake source.

- **Light flavor (LF) electrons and muons**

This type of fake leptons arises from other hadronic processes. In the case of electrons this is usually a misidentified hadron. LF muons can be produced by several processes, for example in-flight pion decays.

- **Conversion (CO) electrons**

Conversion is only a relevant fake source for electrons and is a minor effect for muons and taus. CO electrons originate from the interaction of photons with the detector material.

- **Quark jet (QJ) taus**

QJ taus are quark initiated jets being misidentified as a tau.

- **Gluon jet (GJ) taus**

GJ taus arise from the misidentification of gluon initiated jets. In contrast to QJ taus, GJ taus result only in few charged particles and have a lower fake rate.

- **Unknown (UN) sources**

If a lepton is neither real nor belonging to any of the fake sources above, one speaks about an unknown source. In such cases the combined MC information described above cannot be matched to any of the known sources because, for example, the lepton cannot be matched to any mother particle.

The composition of leptons in SR1 and SR2, as predicted from MC simulation, is shown in Figure 7.1. Electrons and muons are mostly real and the fake contributions are comparably small. For taus, however, the fake sources constitute a large fraction of the total number of taus, especially in SR1, and consist mostly of heavy flavor and quark jet taus. Therefore a correct modeling of the fake processes and a solid background estimation is required.

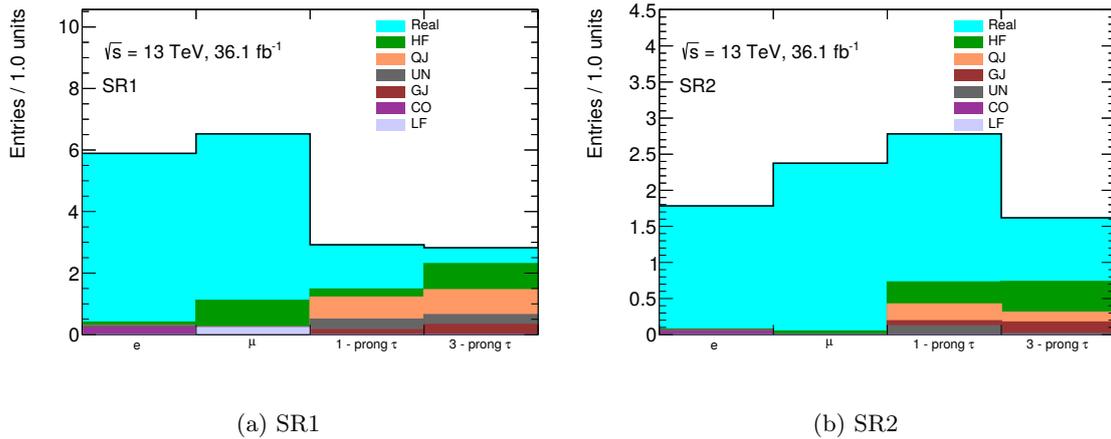


Figure 7.1: The lepton composition in SR1 (a) and SR2 (b) for the SM background (irreducible and reducible). The lepton origins are extracted from MC information.

7.2.2 Fake factor method

In this analysis the reducible background is estimated from data control regions (CR) using the so called fake-factor method [16]. The general approach of this method is to apply fake factors to data in control region events. The fake factor is a measure for the probabilities, that a fake lepton passes or fails signal criteria, and is derived from MC information. As in the case of signal regions, a control region is characterized by different selection criteria on the event variables. These criteria are defined such that the control regions are free of any expected SUSY signals. Furthermore, the control regions are required to have a large statistics of SM background processes in order to get reliable results for the background estimation.

Two types of control regions, CR1 and CR2, are defined for each signal region. The control regions have the same kinematic cuts as the corresponding signal region, but inverted selection criteria on the leptons:

- **SR:** In a signal region at least four signal leptons are required.
- **CR1:** CR1 demands exactly three signal leptons and at least one loose lepton. With CR1 the number of events with 1-fake and 2-fake leptons is estimated. Demanding “exactly” instead of “at least” three signal leptons ensures that the control region is orthogonal to the signal region, which means that an event cannot occur both in SR and CR1.
- **CR2:** In CR2 exactly two signal leptons and at least two loose leptons are required. CR2 is introduced to account for double counting of events with two fake leptons in CR1. Requiring “exactly” three signal leptons ensures again orthogonality to SR and also to CR1.

The fake factor is defined as the ratio of the number of signal and loose leptons:

$$F^{ij}(p_T, \eta) = \frac{N_{\text{signal}}^{ij}(p_T, \eta)}{N_{\text{loose}}^{ij}(p_T, \eta)} \quad (7.1)$$

F^{ij} depends on different parameters: The physics process, the fake lepton type and its fake source, p_T and η . Therefore the fake factor is determined for each fake source i of a lepton type l (electron, muon, 1-prong or 3-prong tau) and each MC process j separately as a function of p_T and η . For example, if the lepton l is a muon, then its possible fake sources i are muons from heavy flavor and light flavor processes. An overall fake factor F_w^l for each lepton type is then defined as the sum of these individual fake factors, weighted by the fractional contribution f of each SM process and a scale factor s :

$$F_w^l(p_T, \eta) = \sum_{i,j} (f^{ij}(p_T) \cdot s^i(p_T) \cdot F^{ij}(p_T, \eta)) \quad (7.2)$$

The scale factor is used to correct the fake factor for each fake source to data. Both for f and s only the p_T dependence is taken into account. After F_w^l has been measured for each control region, it is applied on data events in the respective control region. The number of reducible events in the signal region $N_{\text{SR}}^{\text{red}}$ is then estimated by:

$$N_{\text{SR}}^{\text{red}} = \left(N_{\text{CR1}}^{\text{data}} - N_{\text{CR1}}^{\text{irr}} \right) \cdot F_1 - \left(N_{\text{CR2}}^{\text{data}} - N_{\text{CR2}}^{\text{irr}} \right) \cdot F_1 F_2 \quad (7.3)$$

where $N_{\text{CRX}}^{\text{data}}$ and $N_{\text{CRX}}^{\text{irr}}$ are the number of data and irreducible events in control region CRX , respectively, and F_1 and F_2 are the weighted fake factors resulting if either the first or the second loose lepton (only for CR2) is used in the calculation. As data events in the control regions can also originate from irreducible processes, the MC predicted number of irreducible events $N_{\text{CRX}}^{\text{irr}}$ is subtracted from the number of data events $N_{\text{CRX}}^{\text{data}}$. The second term is subtracted from the first one in order to remove the double-counting of events with two fake leptons in the first term.

However, Equation 7.3 is only a schematic outline. Instead of multiplying an average fake factor with the total number of reducible events, it is actually applied event-by-event according to the lepton types of the loose leptons and their p_T and η . Furthermore, the estimation of $N_{\text{SR}}^{\text{red}}$ presented is only an approximate form. The precise calculation of $N_{\text{SR}}^{\text{red}}$ would require to consider also processes with three or more fake leptons, so that higher order terms would need to be included. However, as these processes are negligibly small, their effect is only taken into account by introducing an additional systematic uncertainty (see Section 7.2.3).

In the following a detailed discussion of the fake factor method will be presented. First it will be described, which control regions are used in this analysis. Afterwards the methods used to determine F_w^l , f and s will be explained.

Control regions

As stated above, exactly three signal and at least one loose lepton are required in CR1, while CR2 demands exactly two signal and at least two loose leptons. This gives different possibilities, how the leptons can be assigned as signal or loose to construct the control regions. For example for SR1, which targets $3\ell 1\tau$ final states, a possible CR1 can be constructed by demanding three signal light leptons and one loose tau. To distinguish between signal and loose leptons, the following notation is used: Signal light leptons and taus are denoted by “L” and “T”, respectively, whereas loose light leptons and taus are denoted by “l” and “t”. With this the control region from above can be indicated by CR1.LLLt.

However, not every combination must be considered for the analysis. This can be seen from Figure 7.2, which shows the number of signal fake leptons and their fake origins in $3\ell 1\tau$ and $2\ell 2\tau$ events with a Z -veto applied. For $3\ell 1\tau$, the fake events arise mostly from 2-fake

$\ell\tau$. Other contributions come from 1-fake ℓ , 1-fake τ and 2-fake $\ell\ell$. The latter one is very small and therefore it is not necessary to have a CR2 with two loose light leptons, i.e. for the analysis CR2.LTll is not needed. As the signal region definition requires at least one signal tau, the $3\ell1\tau$ events could also contain two fake taus. But clearly, these scenarios can be neglected. In $2\ell2\tau$, the fake events are overwhelmingly from 2-fake $\tau\tau$, followed by 1-fake τ . The minor contributions are from 1-fake ℓ and 2-fake $\ell\tau$, so that a CR1 with a loose light lepton is not required. This means that it is not necessary to include CR1.LTTl into the analysis. Furthermore, no $2\ell2\tau$ events originate from 2-fake $\ell\ell$. Therefore, for $2\ell2\tau$ a CR2 with two loose light leptons, i.e. CR2.TTll, is not needed, too.

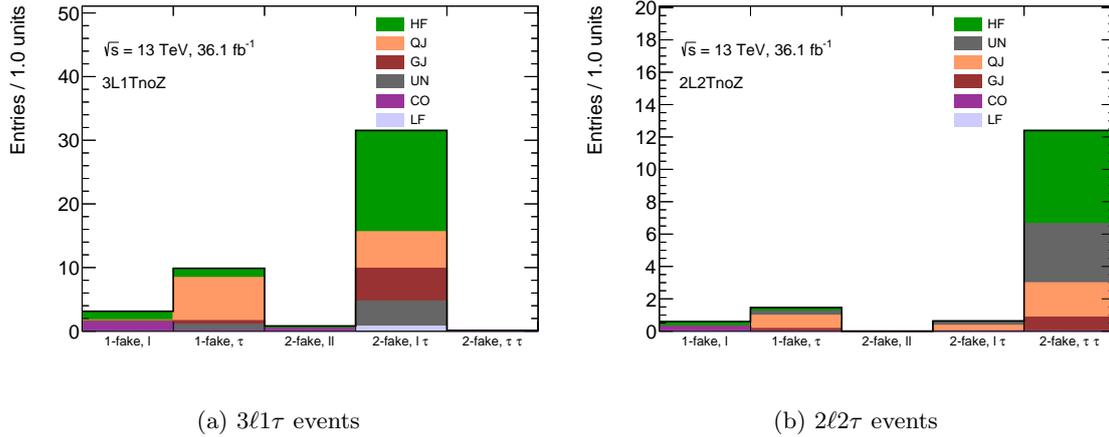


Figure 7.2: Multiplicity and composition of fake leptons for $3\ell1\tau$ (a) and $2\ell2\tau$ events (b) for the reducible background. In both regions a Z -veto is applied. The lepton origins are extracted from MC information.

Table 7.1 gives a summary of the resulting control regions in this analysis used to estimate the reducible background for $3\ell1\tau$ and $2\ell2\tau$ events. The control regions share the same kinematic cuts with their corresponding signal regions, i.e. the Z -veto and the high m_{eff} requirement. To estimate the number of reducible background events in the validation regions instead of in the signal regions, an inverted cut on m_{eff} is used (see Section 7.2.4). The control regions for $4\ell0\tau$ final states are listed in Appendix B.

Target	Control Region	$N(e, \mu)$	$N(e, \mu)$	$N(\tau)$	$N(\tau)$	Z -Boson	Selection
		signal	loose	signal	loose		
$3\ell1\tau$	CR1.LLLt	= 3	= 0	= 0	≥ 1	veto	$m_{\text{eff}} \geq 700$ GeV for SR
	CR1.LLTl	= 2	= 1	≥ 1	≥ 0		$m_{\text{eff}} < 700$ GeV for VR
	CR2.LLlt	= 2	= 1	= 0	≥ 1		
$2\ell2\tau$	CR1.LLTt	= 2	= 0	= 1	≥ 1	veto	$m_{\text{eff}} \geq 650$ GeV for SR
	CR2.LLtt	= 2	= 0	= 0	≥ 2		$m_{\text{eff}} < 650$ GeV for VR

Table 7.1: Definitions of the control regions.

Figure 7.3 shows the lepton composition in the full SM background, i.e. the reducible and irreducible background, for the three control regions targeting $3\ell1\tau$ final states. To increase statistics, no cut on m_{eff} is applied. One can see that in CR.LLLt and CR.LLlt 1-prong and 3-prong taus arise exclusively from fake sources. Only CR.LLTt contains a small fraction of

real taus. Real electrons constitute approximately one-third of the total number of electrons in CR_LLLt events, and about one-half in the other two control regions. For muons the contribution from real objects is clearly dominating in all control regions.

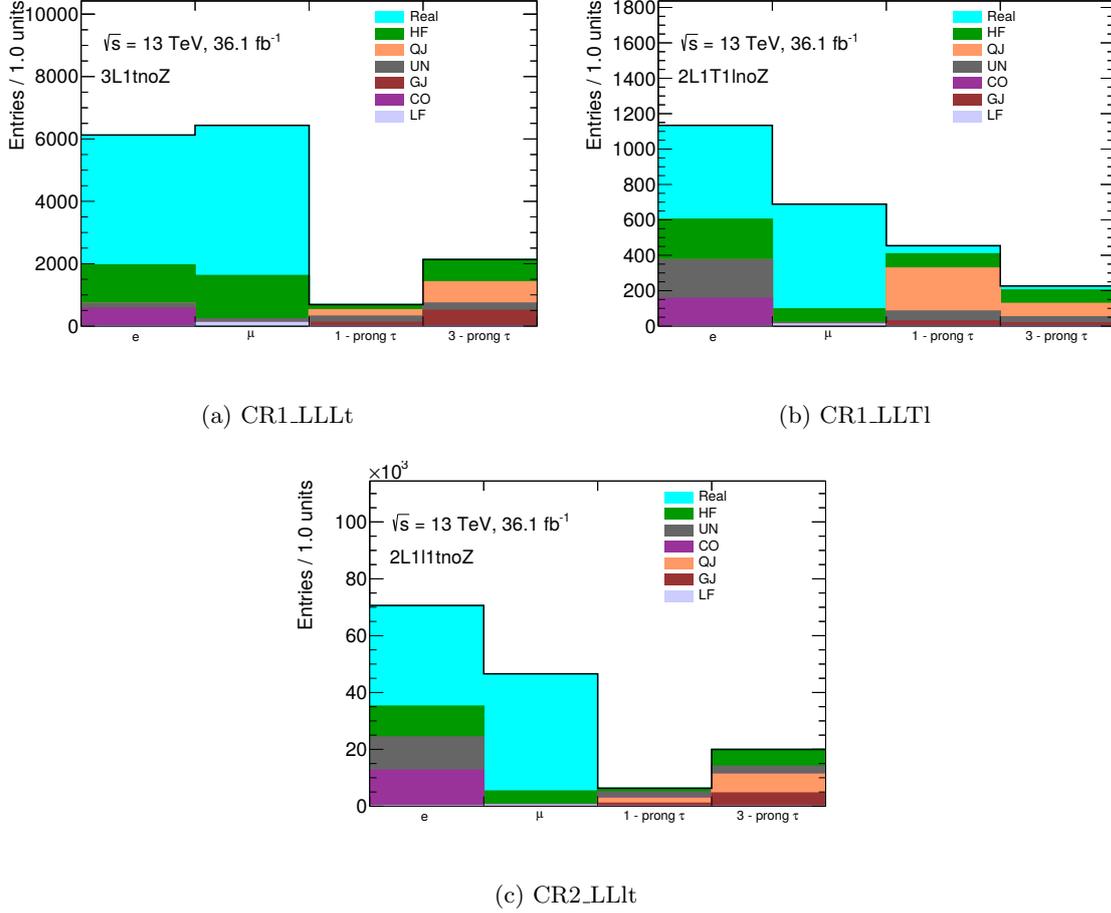


Figure 7.3: The lepton composition in $3\ell 1\tau$ control regions for the SM background (irreducible and reducible). In all regions a Z -veto is applied.

The lepton composition in the SM background for the two $2\ell 2\tau$ control regions is shown in Figure 7.4. Real taus occur only in CR_LLLt, while CR_LLLtI contains only fake taus. Light leptons originate almost completely from irreducible sources in both regions.

Fake factor calculation

To calculate the fake factor for a certain fake source and MC process events are selected, which contain exactly two real signal light leptons and a fake lepton. The latter can be either signal or loose. For the whole MC process the number of events with a fake signal lepton is counted, denoted by $N(2\ell_{\text{signal}}^{\text{real}} + 1\ell_{\text{signal}}^{\text{fake}})$ for a fake e or μ , and by $N(2\ell_{\text{signal}}^{\text{real}} + 1\tau_{\text{signal}}^{\text{fake}})$ for a fake 1-prong or 3-prong τ . Analogously, the number of events with a fake loose lepton is denoted by $N(2\ell_{\text{signal}}^{\text{real}} + 1\ell_{\text{loose}}^{\text{fake}})$ and $N(2\ell_{\text{signal}}^{\text{real}} + 1\tau_{\text{loose}}^{\text{fake}})$. The fake factor F is then given by:

$$F = \frac{N(2\ell_{\text{signal}}^{\text{real}} + 1\ell_{\text{signal}}^{\text{fake}})}{N(2\ell_{\text{signal}}^{\text{real}} + 1\ell_{\text{loose}}^{\text{fake}})} \quad (7.4)$$

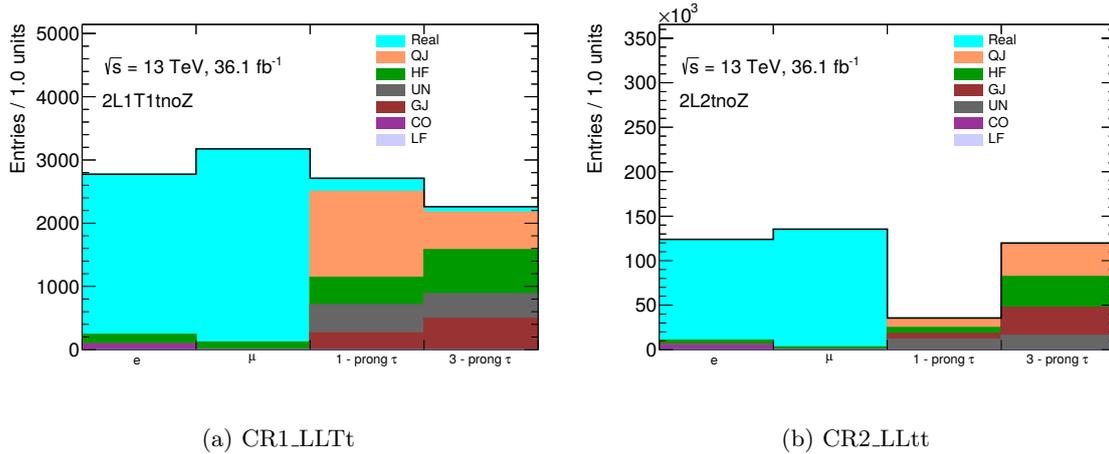


Figure 7.4: The lepton composition in $2\ell 2\tau$ control regions for the SM background (irreducible and reducible). In all regions a Z -veto is applied.

for loose light leptons. The same holds for taus:

$$F = \frac{N(2\ell_{\text{signal}}^{\text{real}} + 1\tau_{\text{signal}}^{\text{fake}})}{N(2\ell_{\text{signal}}^{\text{real}} + 1\tau_{\text{loose}}^{\text{fake}})} \quad (7.5)$$

The fake factor is stored bin-wise as a function of the p_T and η of the fake lepton, i.e. the calculation from above is performed in each $p_T - \eta$ bin. The binning is chosen such as to give approximately equal statistics in each bin.

Figures 7.5–7.8 show the various fake factors for the dominant fake sources obtained from MC simulated $t\bar{t}$ events. For the other reducible backgrounds similar values for the fake factors are obtained. For electrons the heavy flavor fake factors are the highest in the lower p_T range of $p_T < 15$ GeV, where they reach up to ~ 0.27 . For $p_T > 15$ GeV the fake factors vary between 0 and 0.1. The fake factors for light flavor and conversion electrons have their maximum values at about 0.08 and 0.04, respectively. Also for muons the heavy flavor fake factor is relatively high for $p_T < 15$ GeV, which goes up to 0.7. For higher values of p_T it fluctuates between 0 and 0.2. For light flavor muons the low p_T fake factor is ~ 0.45 , while for the higher p_T region it is only ~ 0.1 .

The fake factors for taus are much smaller than for light leptons, and among the taus the fake factors are on average larger for 1-prong taus than for 3-prong taus. The general trend of obtaining decreasing values of the fake factor with higher p_T is also seen for taus, excluding some fluctuations like in Figure 7.8, where a high p_T bin yields the largest fake factor. For 1-prong taus, the quark jet fake factor has on average the largest values, which reach up to 0.1, while the gluon jet fake factors are the smallest and range from 0.018 to 0.035. In the case of 3-prong taus, the highest fake factor is observed for heavy flavor taus with a value of approximately 0.012. As for 1-prong taus, the gluon jet fake factors are in general the lowest with values varying between 0.0005 and 0.0025. Apart from the shown sources of fake taus, also the contribution from conversion taus is taken into account in the fake factor method.

The fake sources are usually related to non-prompt leptons with a low momentum. Therefore the fake factors are mostly higher for low p_T leptons. For example, the source of heavy flavor leptons is mostly the decay $b \rightarrow Wc$, where the W -boson decays leptonically. Because of the small mass difference between the b and the c quark the produced lepton has a low momentum. In the analysis electrons and muons must pass a p_T threshold of 5 GeV and

7 GeV, respectively, while taus are required to have a minimum p_T of 20 GeV (see Section 5.2). Therefore taus are less likely to originate from heavy flavor processes and consequently their heavy flavor fake factors are usually smaller than for electrons or muons.

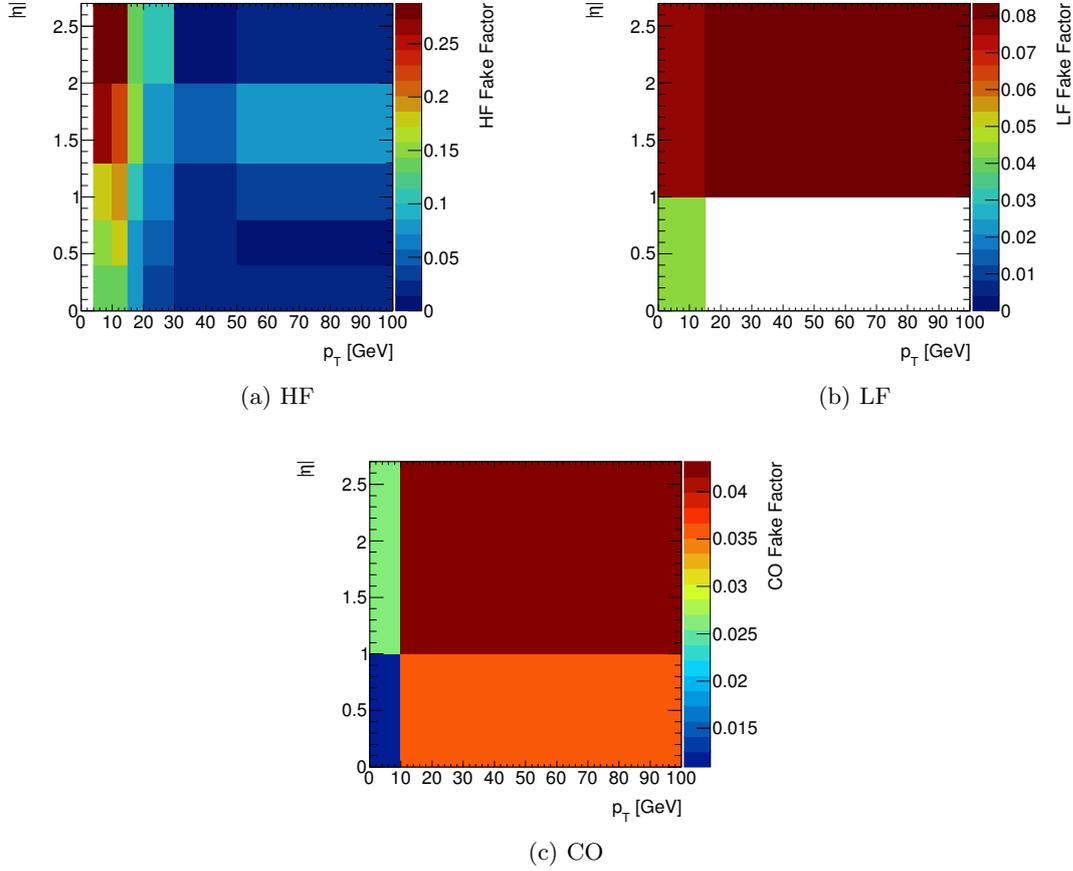


Figure 7.5: Fake factors obtained for electrons in $t\bar{t}$ events. An empty bin indicates, that there is no statistics in the particular $p_T - \eta$ bin.

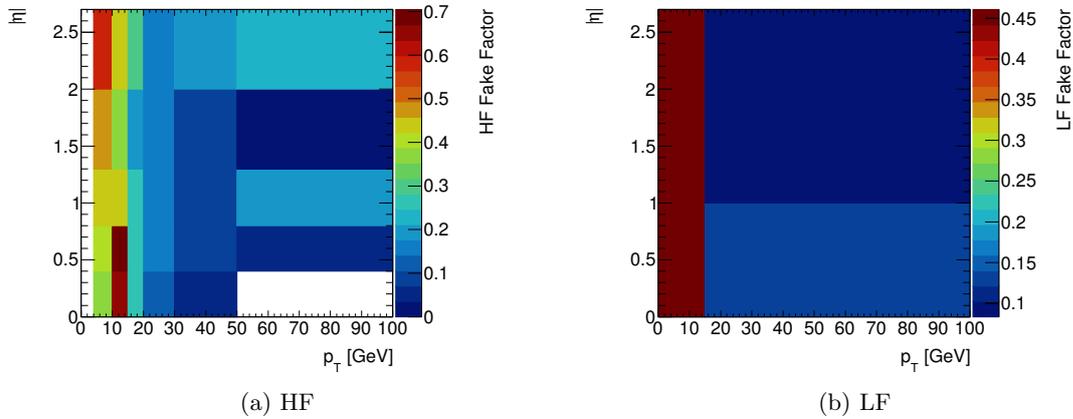


Figure 7.6: Fake factors obtained for muons in $t\bar{t}$ events. An empty bin indicates, that there is no statistics in the particular $p_T - \eta$ bin.

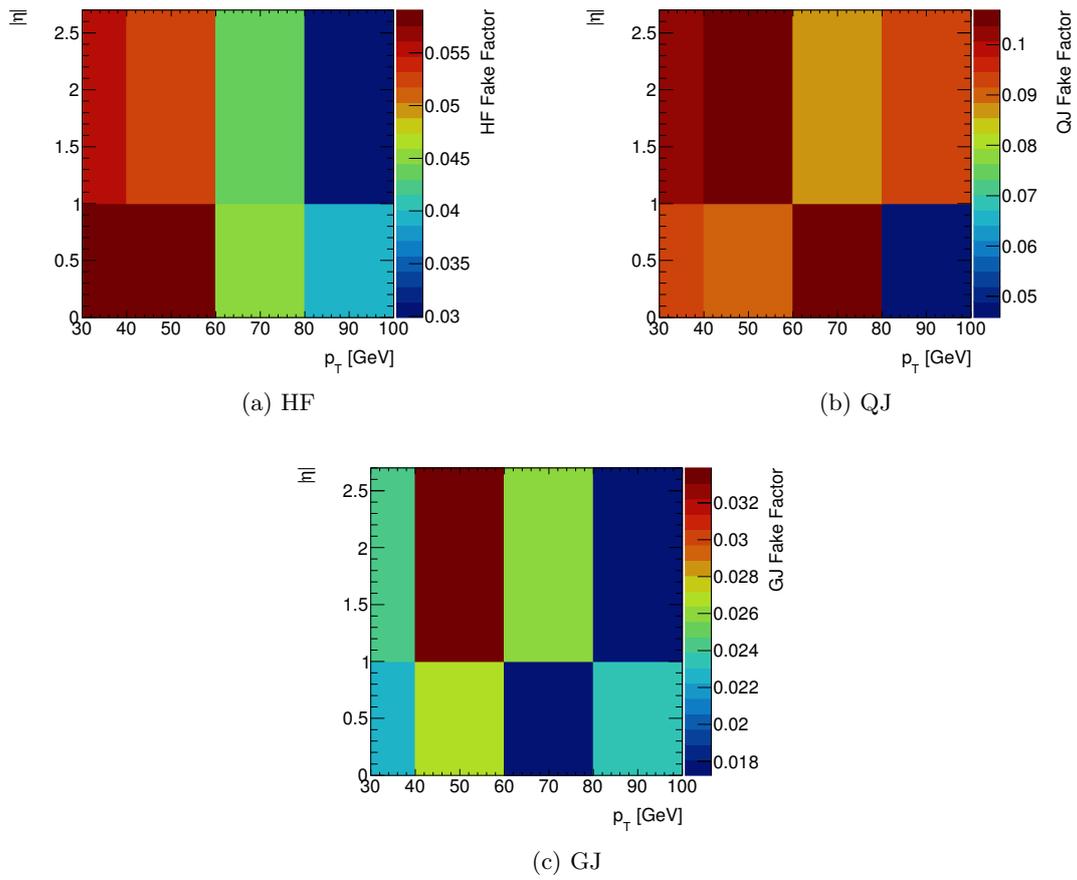


Figure 7.7: Fake factors obtained for 1-prong taus in $t\bar{t}$ events.

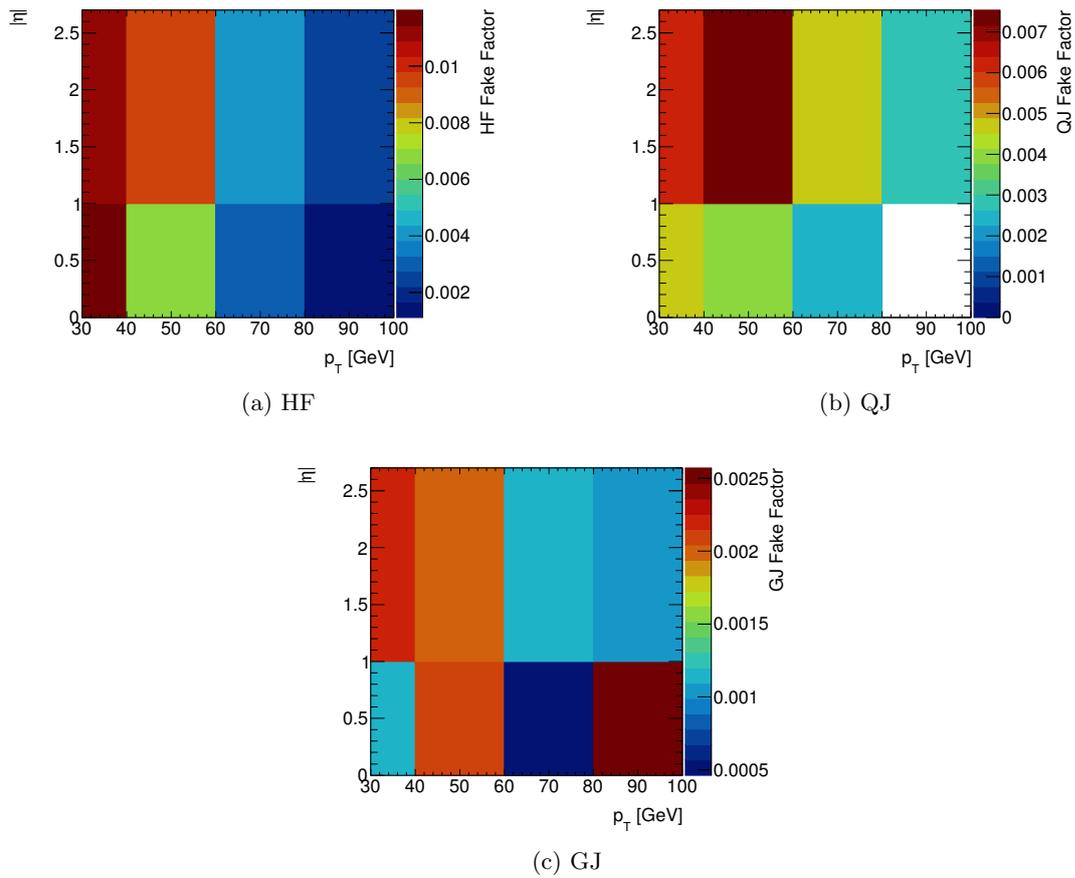


Figure 7.8: Fake factors obtained for 3-prong taus in $t\bar{t}$ events. An empty bin indicates, that there is no statistics in the particular $p_T - \eta$ bin.

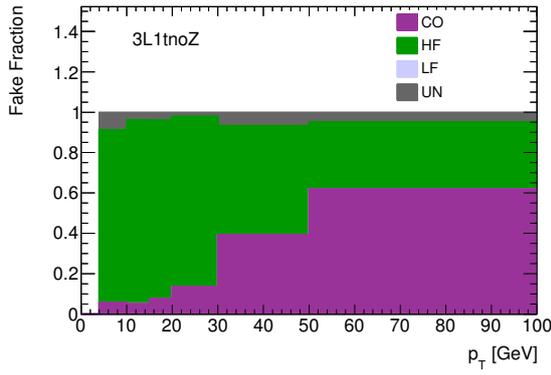
Fake fractions

The fake factors for each fake source are weighted according to their fractional contribution to the total number of fake leptons of one type. This is done for each reducible process. These weights are denoted as “fake fractions”. As they vary with the p_T of the fake leptons, they are also stored bin-wise.

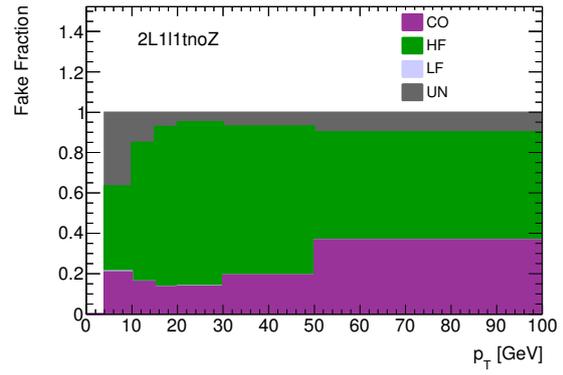
Figure 7.9 shows the fake fractions for each lepton type obtained from MC simulated $t\bar{t}$ events in the control regions which target $3\ell 1\tau$ final states. The plots for the $2\ell 2\tau$ events are shown in Figure 7.10. One can see that fake electrons with low p_T originate mostly from heavy flavor sources. With higher p_T the contribution from conversion processes increases, especially for $2\ell 2\tau$ events, in which it even becomes the dominant fake source. Furthermore, there is also a small contribution from unknown sources. In all control regions the light flavor fake fraction is zero in almost every bin. Fake muons arise overwhelmingly from heavy flavor sources. Only a small fraction of light flavor processes is seen. For 1-prong and 3-prong taus a similar behavior is observed in almost all regions. The largest contribution comes from heavy flavor processes with a fraction of about 50% for low p_T values, which increases to $\sim 70\%$. Simultaneously, the fraction of gluon jet and unknown processes is decreasing, while the quark jet fake fraction stays almost constantly at approximately 10% for all bins. In CR_LLt the fake fractions look slightly different. The gluon jet and unknown sources are approximately the same as in the other regions, but the contribution from heavy flavor sources is smaller and from quark jet sources larger.

The fake fractions for the other reducible processes were found to be roughly similar to the ones in $t\bar{t}$ events, but with a smaller fraction of heavy flavor leptons in processes without a top quark, for example in WZ . The contribution from heavy flavor processes is particularly large for top quark events due to the subsequent decay of the b quark, as explained in the previous section.

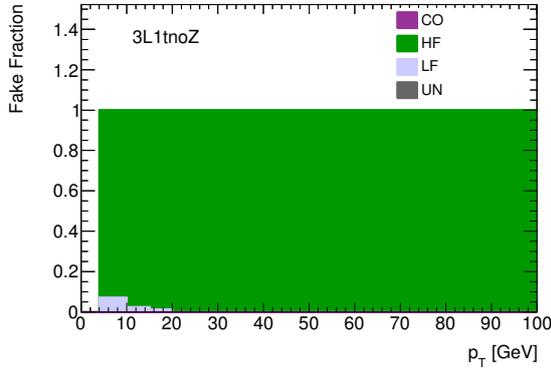
The fake fractions for one particular lepton in CR1 and CR2 are similar, e.g. CR1_LLLt and CR2_LLt for 3-prong taus. This is also the expected behavior, as CR1 and CR2 are identical in their selection criteria on kinematic variables and leptons, except that the requirement on one lepton is loosened in CR2 compared to CR1. On the other hand, the MC statistics is much larger in CR2 than in CR1 because loose leptons occur more frequently than signal leptons. Therefore, to increase statistics the fake fractions are measured only in CR2s. Furthermore, the fake fractions were found to depend only weakly on m_{eff} [16]. It is thus not necessary to apply the m_{eff} cuts from the signal region definitions in the fake fraction measurements.



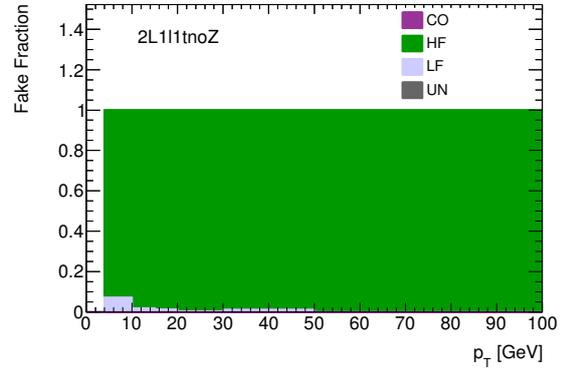
(a) Fake electrons in CR1_LLLt



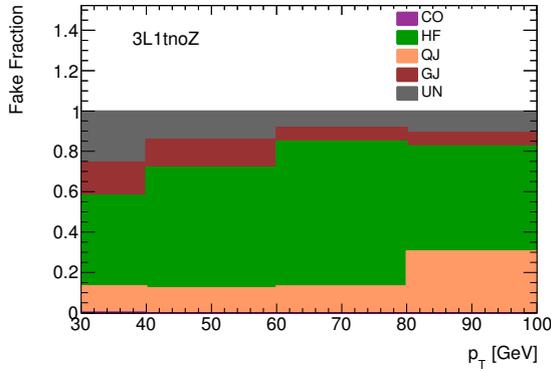
(b) Fake electrons in CR2_LLLt



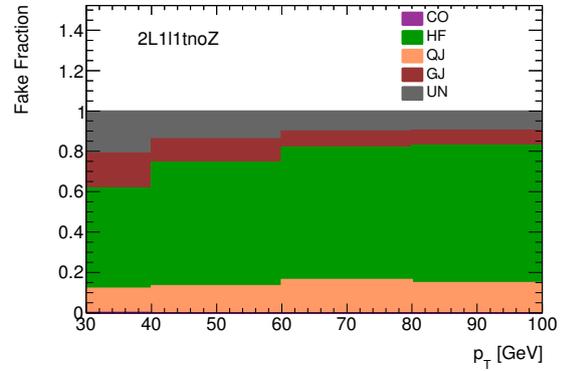
(c) Fake muons in CR1_LLLt



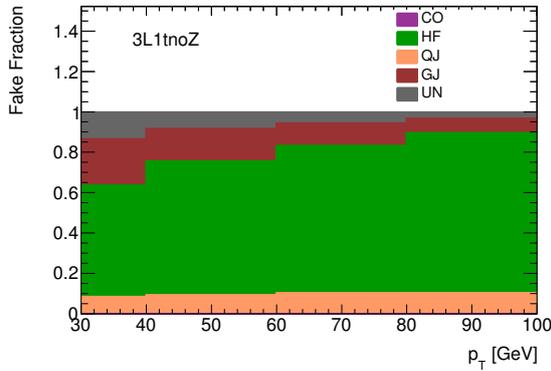
(d) Fake muons in CR2_LLLt



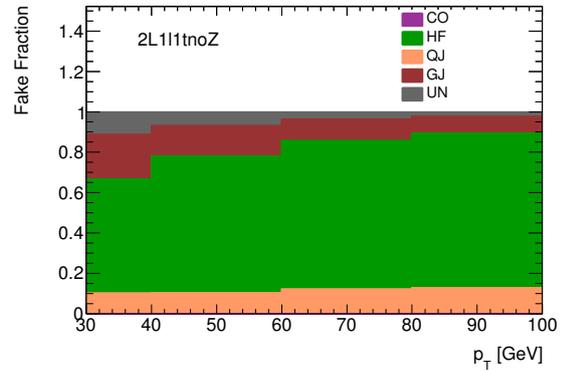
(e) Fake 1-prong taus in CR1_LLLt



(f) Fake 1-prong taus in CR2_LLLt

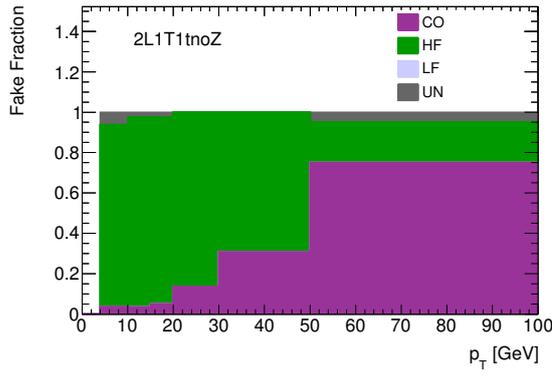


(g) Fake 3-prong taus in CR1_LLLt

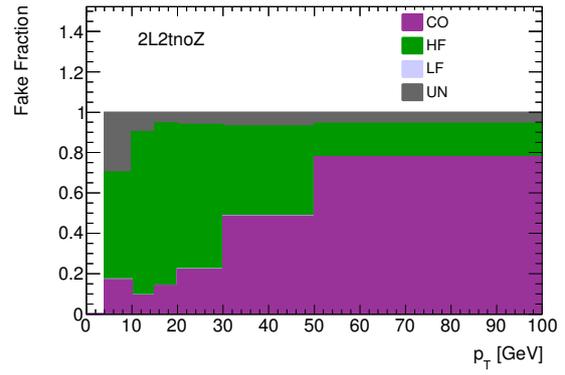


(h) Fake 3-prong taus in CR2_LLLt

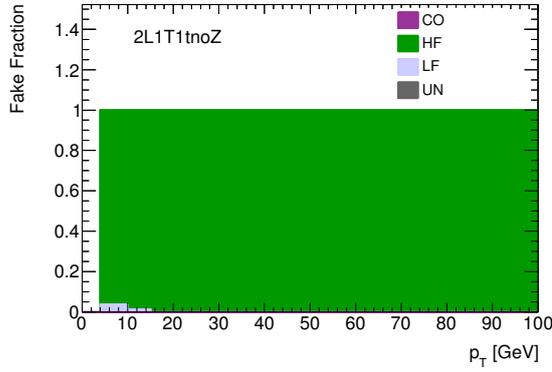
Figure 7.9: Fake fractions obtained in $t\bar{t}$ events with $3\ell 1\tau$ final states for the different lepton types. In all regions a Z -veto is applied. The CR1s are shown on the left and the CR2s on the right. CR1_LLLt is not shown due to poor statistics.



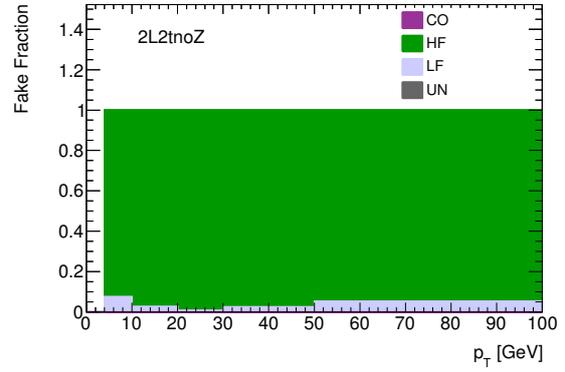
(a) Fake electrons in CR1_LLt



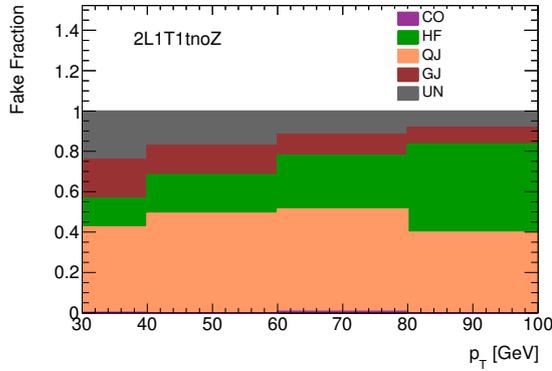
(b) Fake electrons in CR2_LLt



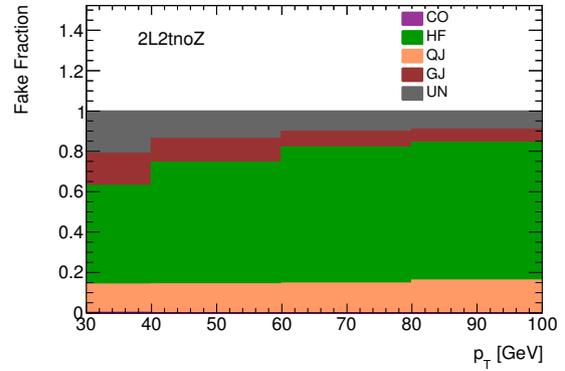
(c) Fake muons in CR1_LLt



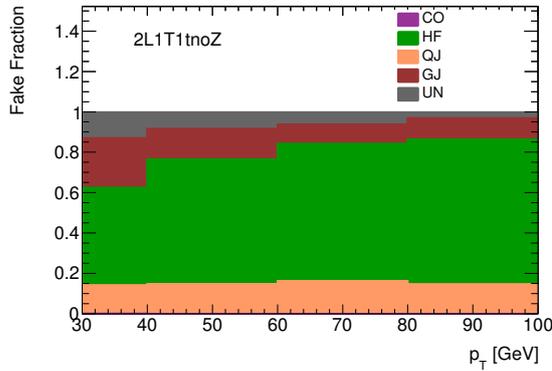
(d) Fake muons in CR2_LLt



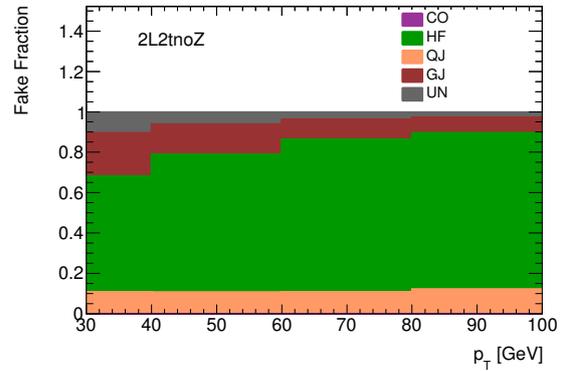
(e) Fake 1-prong taus in CR1_LLt



(f) Fake 1-prong taus in CR2_LLt



(g) Fake 3-prong taus in CR1_LLt



(h) Fake 3-prong taus in CR2_LLt

Figure 7.10: Fake fractions obtained in $t\bar{t}$ events with $2\ell 2\tau$ final states for the different lepton types. In all regions a Z -veto is applied. The CR1s are shown on the left and the CR2s on the right.

Scale factors

As explained above, the fake factors are calculated only by using information from simulation. However, MC simulations do not always agree perfectly with data. Differences can arise for several reasons, for example due to mis-modeling in MC simulation, mis-measurements in the different detector components or poor understanding of data. Consequently, the fake factor obtained from simulation needs to be corrected to data by the application of a scale factor s . These scale factors are measured for each fake source separately in dedicated control regions, which are enriched in leptons of a given fake source. In these regions the fake factors are determined both for MC simulation as well as for data events. The scale factor is then defined as the ratio of the two fake factors:

$$s = \frac{F_{\text{data}}}{F_{\text{MC}}} \quad (7.6)$$

In the following it will be summarized, which control regions are used for the various fake sources and which scale factors are obtained (for details see [16]).

- **HF e, μ**

A $t\bar{t}$ dominated control region is used to measure the scale factors for heavy flavor electrons and muons. The scale factors were found to vary with the lepton p_T . The scale factor for muons decreases from ~ 1.04 to 0.73 as the p_T increases from 5 to 15 GeV. For electrons, the scale factor increases with higher p_T from about 1.16 to 1.35 .

- **CO e**

The scale factor for conversion electrons is determined from a control region enriched with the process $Z \rightarrow \mu\mu$. The photons, which initiate the conversion, originate from final state radiation of the muon pairs. The resulting scale factors have a small p_T dependence. They increase from approximately $1.2 - 1.6$ as the p_T increases from $7 - 25$ GeV.

- **QJ τ**

For fake taus arising from quark jet processes, the scale factors are measured in a $Z + \text{jets}$ dominated control sample separately for 1-prong and 3-prong taus. As before, the scale factors depend on the lepton p_T . As the tau p_T increases from $20 - 50$ GeV, the scale factors decrease from about $1.3 - 0.96$ for 1-prong taus, and from $\sim 1.4 - 1.2$ for 3-prong taus.

- **GJ, HF and CO τ**

The scale factors for fake taus coming from gluon jet, heavy flavor and conversion processes cannot be reliably measured from data because no suitable control region can be constructed enriched with these fake lepton types. Therefore they are assumed to be the same as the quark jet scale factors from above.

- **LF e, μ**

As the contribution from light flavor sources is very small, the scale factors for light flavor electrons and muons cannot be measured using data. For them an overall, p_T independent scale factor of 1.0 ± 0.25 is applied instead.

7.2.3 Systematic uncertainties

An essential point in the background estimation is to know, how accurate the obtained results are. Apart from statistical fluctuations, there are many systematic uncertainties that can increase the uncertainty on the result. In this section the systematic uncertainties will be presented, which are used in this analysis.

Theoretical uncertainties on MC processes

One category of systematic uncertainties are the theoretical uncertainties. These include the uncertainties on the production cross sections which are applied to the dominant irreducible backgrounds following the recommendations of the Physics Modeling Group (PMG) [52]. An uncertainty of 12% is assigned to the $t\bar{t}Z$, 6% to the ZZ and 20% to the triboson processes. Different uncertainties are used for the Higgs processes according to their production mechanisms. For vector-boson-fusion and associated production an uncertainty of 20% is assumed, while an uncertainty of 100% is applied to gluon-gluon-fusion and $t\bar{t}H$ processes. The cross section uncertainties for the signal processes are typically of the order of 10%.

In the analysis also the uncertainty on the acceptance is taken into account. The acceptance describes effects related to the cross section calculation in MC simulation. The cross section to produce a particular particle in a hadron collision depends on the parton distribution functions, describing which fraction of the total hadron momentum a single quark or gluon carries, and on the cross section of the hard process itself. The parton distribution functions and the cross section of the hard process both depend on factorization and renormalization scales, which are used to regulate divergences. The uncertainty on the acceptance is evaluated only for $t\bar{t}Z$ and ZZ processes, which are together with $t\bar{t}H$ the main irreducible backgrounds in the signal regions. For $t\bar{t}Z$ this is done by varying the scales. For ZZ also variations of the parton distribution functions are taken into account. The uncertainty on the acceptance has values ranging from 6 – 25%, depending on which regions are considered. For $t\bar{t}H$ the uncertainty on the acceptance is not calculated. Therefore the uncertainty on the production cross section was enlarged to account for this fact.

Furthermore, the MC predicted background yields can differ if a process is simulated with another generator. The uncertainty connected to the choice of generator is determined only for $t\bar{t}Z$ processes. For this purpose $t\bar{t}Z$ events from the generator MADGRAPH are compared with the events from SHERPA. This results in an uncertainty of $\sim 20\%$ [16, 17].

Experimental uncertainties on MC processes

Experimental uncertainties arise from detector inefficiencies, limitations in the measurement precision and reconstruction and identification uncertainties. Their effects are estimated by the Combined Performance (CP) Groups of ATLAS [53]. Additional scale factors, or *weights*, are introduced to correct MC simulation to data due to differences in the object reconstruction and identification. Systematic uncertainties are then applied to MC events to account for the uncertainties in this procedure and also for other sources of uncertainties. Systematic uncertainties regarding the jet reconstruction come e.g. from uncertainties on the jet energy resolution, the jet energy scale and the jet vertex tagger. For leptons there are uncertainties due to the the lepton energy scale and resolution, the lepton identification and the isolation. Other systematic uncertainties consider uncertainties on the trigger efficiencies and the soft term in E_T^{miss} . The analysis includes also uncertainties on the luminosity and the pile-up simulation.

Systematic uncertainties on the reducible background

The dominant systematic uncertainty on the reducible background is caused by the statistical uncertainty of data in the different CR1s and CR2s. Apart from data, also the statistical uncertainty of the MC events must be taken into account because the fake factors and fake fractions, used in the background estimation, are calculated from simulation. Another group of systematic uncertainties are the uncertainties on the scale factors. Furthermore, an additional uncertainty is introduced, which accounts for the neglected terms in the fake factor method describing the contribution from processes with three or more fake leptons.

Obtained systematic uncertainties in the signal regions

Table 7.2 and Table 7.3 show a breakdown of the systematic uncertainties in SR1 and SR2, respectively. Listed are the total number of background uncertainty events for each uncertainty as well as the relative uncertainty with respect to the total background. The uncertainties can be correlated between each other and therefore do not necessarily add up quadratically to the total background uncertainty. Uncertainties with a contribution of less than 0.1% are not shown in the tables.

The statistical uncertainties are clearly the dominant uncertainties, followed by the uncertainty on the $t\bar{t}H$ cross section. The latter contribution is so large because of the assumed uncertainty of 100% on this process. Also for the other two dominant irreducible processes ZZ and $t\bar{t}Z$ the uncertainties on the cross sections constitute a large fraction of the background uncertainty. Another quite large contribution comes from the neglected terms in the fake factor method. Furthermore, also the tau uncertainties regarding their energy scales and reconstruction efficiencies contribute a great part to the total background uncertainty, while the uncertainties on the jet, electron and muon reconstruction are on average smaller.

Uncertainty of channel	SR1
Total background expectation	4.20
Total statistical ($\sqrt{N_{\text{exp}}}$)	± 2.05
Total background systematic	± 1.14 [27.20%]
Statistical	± 1.02 [24.4%]
ttH_Theo_CrossSection	± 0.46 [11.0%]
ttZ_Theo_CrossSection	± 0.11 [2.5%]
Tau_Eff_InnerDetector	± 0.10 [2.4%]
ZZ_Theo_CrossSection	± 0.07 [1.8%]
Tau_Smearing_EnergyScale_Insitu	± 0.06 [1.3%]
Reducible_NeglectedTerms	± 0.05 [1.2%]
Tau_Eff_Reconstruction	± 0.05 [1.2%]
JetEnergyScale_NuisanceParameters_Group_1	± 0.05 [1.1%]
Electron_Eff_InnerDetector	± 0.04 [0.85%]
VVV_Theo_CrossSection	± 0.03 [0.76%]
Muon_Eff_Systematic	± 0.02 [0.52%]
Tau_Eff_ElectronOverlapRemoval	± 0.02 [0.44%]
JetEnergyResolution	± 0.02 [0.43%]
Tau_Smearing_EnergyScale_Detector	± 0.02 [0.36%]
Electron_Eff_Isolation	± 0.01 [0.36%]
PileUp_Reweighting	± 0.01 [0.32%]
tWZ_Theo_CrossSection	± 0.01 [0.32%]
Tau_Smearing_EnergyScale_Model	± 0.01 [0.30%]
Electron_Eff_Reconstruction	± 0.01 [0.28%]
JetEnergyScale_NuisanceParameters_Group_3	± 0.01 [0.23%]
JetEnergyScale_EtaIntercalibration_NonClosure	± 0.01 [0.22%]
Reducible_HF_LightLepton	± 0.01 [0.21%]
JetEnergyScale_NuisanceParameters_Group_2	± 0.01 [0.20%]
Muon_Isolation_Systematical	± 0.01 [0.19%]
Electron_EnergyMomentumScale	± 0.01 [0.14%]
Muon_MuonSpectrometer	± 0.01 [0.12%]
ElectronTrigger_ScaleFactor	± 0.01 [0.12%]
Tau_Eff_InnerDetector_HighPt	± 0.00 [0.11%]
Muon_Eff_Statistical	± 0.00 [0.11%]
Reducible_HF_Tau	± 0.00 [0.11%]
MissingTransverseEnergy_SoftTracks_Scale	± 0.00 [0.10%]

Table 7.2: Breakdown of the dominant systematic uncertainties on the SM background estimates in SR1. The individual uncertainties can be correlated. The percentages indicate the relative size of the uncertainties with respect to the total expected background.

Uncertainty of channel	SR2
Total background expectation	2.46
Total statistical ($\sqrt{N_{\text{exp}}}$)	± 1.57
Total background systematic	± 0.60 [24.56%]
Statistical	± 0.42 [16.9%]
ttH_Theo_CrossSection	± 0.29 [11.8%]
Reducible_NeglectedTerms	± 0.24 [9.8%]
Tau_Eff_InnerDetector	± 0.14 [5.5%]
ttZ_Theo_CrossSection	± 0.07 [3.0%]
Tau_Smearing_EnergyScale_Insitu	± 0.07 [2.9%]
Tau_Eff_Reconstruction	± 0.07 [2.7%]
ZZ_Theo_CrossSection	± 0.07 [2.7%]
JetEnergyScale_NuisanceParameters_Group_1	± 0.06 [2.6%]
PileUp_Reweighting	± 0.04 [1.7%]
VVV_Theo_CrossSection	± 0.04 [1.5%]
Reducible_ScaleFactors	± 0.03 [1.2%]
Tau_Eff_ElectronOverlapRemoval	± 0.03 [1.1%]
JetEnergyResolution	± 0.02 [0.86%]
Tau_Smearing_EnergyScale_Detector	± 0.02 [0.82%]
Muon_Eff_Systematical	± 0.01 [0.60%]
Electron_Eff_InnerDetector	± 0.01 [0.52%]
JetEnergyScale_NuisanceParameters_Group_2	± 0.01 [0.49%]
JetEnergyScale_EtaIntercalibration_NonClosure	± 0.01 [0.48%]
Tau_Smearing_EnergyScale_Model	± 0.01 [0.46%]
JetEnergyScale_NuisanceParameters_Group_3	± 0.01 [0.43%]
tWZ_Theo_CrossSection	± 0.01 [0.42%]
Muon_InnerDetector	± 0.01 [0.39%]
Tau_Eff_InnerDetector_HighPt	± 0.01 [0.25%]
Electron_Eff_Isolation	± 0.01 [0.24%]
Electron_Eff_Reconstruction	± 0.00 [0.16%]
Muon_Isolation_Systematical	± 0.00 [0.14%]
MuonTrigger_ScaleFactor	± 0.00 [0.14%]
Muon_Eff_Statistical	± 0.00 [0.11%]
Reducible_QJ_Tau	± 0.00 [0.10%]

Table 7.3: Breakdown of the dominant systematic uncertainties on the SM background estimates in SR2. The individual uncertainties can be correlated. The percentages indicate the relative size of the uncertainties with respect to the total expected background.

7.2.4 Validation

After including the systematic uncertainties into the background estimation, the modeling for the irreducible and reducible backgrounds are tested in validation regions. These are required to have the same background composition as the signal regions, but must not overlap with them. For this reason the same lepton selection and a Z -veto is applied in the validation regions, but the cut on m_{eff} is inverted. The two regions used to validate the background estimation in SR1 and SR2 are denoted by VR1 and VR2. Their definitions are summarized in Table 7.4. The veto on loose light leptons is also applied to signal region events and ensures, that there is no overlap between the signal and validation regions with taus and the control regions for $4\ell 0\tau$ events (see Appendix B).

Validation Region	$N(e, \mu)$	$N(e, \mu)$	$N(\tau)$	$N(\tau)$	Z -boson	Selection
	signal	loose	signal	loose		
VR1	= 3	= 0	≥ 1	≥ 0	veto	$m_{\text{eff}} < 700 \text{ GeV}$
VR2	= 2	= 0	≥ 2	≥ 0	veto	$m_{\text{eff}} < 650 \text{ GeV}$

Table 7.4: Definitions of the validation regions.

The predicted yields for the irreducible and reducible backgrounds are compared to data yields in the validation regions. If the background was estimated correctly, then the expected and observed values should agree within the uncertainties. The resulting yields for an integrated luminosity of $\int \mathcal{L} dt = 36.1 \text{ fb}^{-1}$ are shown in Table 7.5.

Sample	VR1	VR2
Observed	114	29
SM Total	80 ± 6	26 ± 5
ZZ	8.1 ± 1.1	3.0 ± 0.6
$t\bar{t}Z$	1.87 ± 0.30	0.64 ± 0.14
Higgs	3.2 ± 2.9	1.2 ± 1.1
VVV	1.18 ± 0.28	0.27 ± 0.08
tWZ	0.23 ± 0.06	0.14 ± 0.04
MultiTop	0.014 ± 0.003	0.001 ± 0.000
$t\bar{t}WW$	0.050 ± 0.008	0.000 ± 0.000
Reducible	65 ± 4	20 ± 4

Table 7.5: Expected and observed yields for 36.1 fb^{-1} in the validation regions. Uncertainties are statistical and systematic.

For VR2 a good agreement between the expected and observed yields is seen. However, in VR1 a data excess of approximately 30 events is observed, which lies above the uncertainty range of the predicted yields and corresponds to a significance of $\sim 2.9\sigma$. Several studies were performed to investigate this issue [16] and it was found, that the excess is limited to $e\mu\mu\tau$ and $e\mu\mu\tau$ events, while no excess is seen in $ee\tau\tau$ and $\mu\mu\mu\tau$ events. Furthermore, it was localized to be in the region $m_{\text{eff}} = 200 - 300 \text{ GeV}$, where a deviation from a typical

m_{eff} distribution is observed. No indication for background mis-modeling was found within the available statistics. Therefore it was concluded that the excess is caused by statistical fluctuations.

Chapter 8

Results

When all important parameters of the analysis are well understood, the data is compared to the SM background expectation in the signal regions. An excess of data over predicted background events can be a sign for new physics. However, if no significant deviation is observed, one can exclude a subspace of the investigated models. The obtained results need then to be interpreted with respect to their discovery or exclusion significance.

In this chapter, first the observed yields will be presented and afterwards the relevant statistical methods will explained. Finally the statistical interpretation of the results will be shown.

8.1 Expected and observed yields

Table 8.1 shows the expected and observed yields in each signal region for an integrated luminosity of $\int \mathcal{L} dt = 36.1 \text{ fb}^{-1}$. The uncertainties on the SM background include statistical as well as systematic uncertainties. In SR0A, SR0B and SR2 the observed data is consistent with the SM expectations, while SR1 is the only region in which a deviation from SM expectations is seen.

Sample	SR0A	SR0B	SR1	SR2
Observed	13	2	8	2
SM Total	10.7 ± 1.9	1.44 ± 0.25	4.2 ± 1.1	2.5 ± 0.6
ZZ	2.6 ± 0.7	0.36 ± 0.10	0.32 ± 0.12	0.30 ± 0.10
$t\bar{t}Z$	3.0 ± 0.8	0.56 ± 0.16	0.58 ± 0.19	0.32 ± 0.11
Higgs	1.2 ± 1.2	0.14 ± 0.13	0.5 ± 0.5	0.30 ± 0.30
VVV	0.80 ± 0.19	0.22 ± 0.05	0.16 ± 0.05	0.19 ± 0.07
tWZ	0.25 ± 0.07	0.044 ± 0.015	0.066 ± 0.026	0.051 ± 0.017
MultiTop	0.174 ± 0.020	0.076 ± 0.009	0.044 ± 0.012	0.003 ± 0.001
$t\bar{t}WW$	0.101 ± 0.016	0.044 ± 0.009	0.066 ± 0.017	0.006 ± 0.002
Reducible [54]	2.44 ± 0.29	0.000 ± 0.031	2.5 ± 0.6	1.29 ± 0.33

Table 8.1: Expected and observed yields for 36.1 fb^{-1} in the signal regions. Statistical as well as systematic uncertainties are included.

8.2 Statistical interpretation

A statistical analysis is essentially based on calculating probabilities. However, there are different ways how the meaning of probability can be interpreted. One of the most common interpretations is the frequentist statistics. In the frequentist statistics the probability to observe a particular result in an experiment is defined as its relative frequency in the limit $N \rightarrow \infty$, where N is the number of experiments. These experiments are required to be repeatable and must be performed under identical conditions [55].

In this analysis all results are interpreted with frequentist methods, which will be presented in the following.

8.2.1 Mathematical concepts

As in Chapter 6, hypothesis tests are performed to quantify the outcome of an experiment in terms of its discovery or exclusion significance. These hypothesis tests are based on the evaluation of a likelihood. The likelihood includes the parameters of interest, for example the rate of a signal process. The concept of likelihood can be explained with an event counting experiment, in which a variable x is measured in a signal region and the resulting data is summarized in a histogram. Its number of entries are denoted by $\mathbf{n} = n_1, n_2, \dots$. The expectation value for bin i to have n_i entries is calculated by:

$$E[n_i] = \mu s_i + b_i \quad (8.1)$$

where s_i and b_i are the number of expected signal and background events in bin i , respectively, and μ the strength of the signal process, which is zero for the background-only hypothesis and one for the signal-plus background model [49]. The signal strength is constrained to $\mu \geq 0$ because in this experiment the presence or absence of a signal cannot result in less events than predicted from the background-only model. However, a negative μ can be a valid assumption for signals allowing for interference effects, as for example in neutrino oscillation. The signal and background events s_i and b_i are obtained by:

$$s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(x, \boldsymbol{\theta}_s) \quad (8.2)$$

$$b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(x, \boldsymbol{\theta}_b) \quad (8.3)$$

where s_{tot} and b_{tot} are the total number of signal and background events and $f_s(x, \boldsymbol{\theta}_s)$ and $f_b(x, \boldsymbol{\theta}_b)$ are the corresponding probability density functions (PDFs) for the variable x [49]. The shapes of the PDFs are characterized by the nuisance parameters $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_b$, describing the effect of systematic uncertainties. In the following all nuisance parameters are summarized by $\boldsymbol{\theta} = (\boldsymbol{\theta}_s, \boldsymbol{\theta}_b)$. They are constraint by auxiliary measurements and treated as adjustable parameters like the signal strength μ .

In a statistical test one is interested in the probability to see a particular outcome of an experiment assuming a certain hypothesis to be true. As the experiment can depend on different variables, the total probability of obtaining a specific result is given by the product of the individual probabilities for each bin [55], which is called the likelihood. In this case the likelihood refers to the total outcome of the event counting experiment and it is given by [56]:

$$L(\mathbf{n}, \boldsymbol{\theta}^0 | \mu, \boldsymbol{\theta}) = \mathcal{P}_{\text{SR}} \times \mathcal{C}_{\text{sys}} = \prod_{i \in \text{SR}} P(n_i | \lambda_i(\mu, \boldsymbol{\theta})) \times \mathcal{C}_{\text{sys}}(\boldsymbol{\theta}^0, \boldsymbol{\theta}) \quad (8.4)$$

Equation 8.4 describes the likelihood to obtain \mathbf{n} histogram entries under a hypothesis that predicts a certain value for μ . It is defined as the product of the probability \mathcal{P}_{SR} to see \mathbf{n} events in the signal region and of the constraint terms \mathcal{C}_{sys} corresponding to the auxiliary measurements. \mathcal{P}_{SR} is given by the product of the Poisson distributions $P(n_i | \lambda_i(\mu, \boldsymbol{\theta}))$ for each bin i in the signal region. The expectation value of each Poissonian is denoted by $\lambda_i(\mu, \boldsymbol{\theta})$, which is a function of μ and $\boldsymbol{\theta}$. \mathcal{C}_{sys} is the product of the probability distributions for each θ_i in $\boldsymbol{\theta}$. These are typically modeled by Gaussians. The central values of the auxiliary measurements are denoted by $\boldsymbol{\theta}^0$. In the statistical tests $\boldsymbol{\theta}$ is varied around $\boldsymbol{\theta}^0$. The likelihood can contain also other factors, depending on the type of statistical test. For example, if information from the control regions is included additionally, then the likelihood would be extended by a factor \mathcal{P}_{CR} . The value of μ is tested by using a profile likelihood ratio $\lambda(\mu)$. This is defined by:

$$\lambda(\mu) = \frac{L(\mu, \hat{\boldsymbol{\theta}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \quad (8.5)$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ maximize the likelihood, while $\hat{\boldsymbol{\theta}}$ maximizes the likelihood for a specific fixed value of μ [49]. The profile likelihood ratio is broadened by the presence of nuisance parameters, which represents the loss of information about μ caused by systematic uncertainties. It can have values in the range $0 \leq \lambda(\mu) \leq 1$, where higher values of $\lambda(\mu)$ correspond to better agreement between data and μ . An equivalent description of $\lambda(\mu)$ is given by the test statistic q_μ :

$$q_\mu = -2 \ln \lambda(\mu) \quad (8.6)$$

Opposed to $\lambda(\mu)$, higher values of q_μ yield less agreement between data and μ . Depending on the type of statistical test, e.g. the calculation of the discovery or exclusion significance, different test statistics for the signal strength are applied. The level of agreement or disagreement between data and the tested μ is quantified with the p -value p_μ , which gives the probability of obtaining the observed or a more extreme value for q_μ . It is defined by:

$$p_\mu = \int_{q_{\mu, \text{obs}}}^{\infty} f(q_\mu | \mu, \boldsymbol{\theta}) dq_\mu \quad (8.7)$$

where $q_{\mu, \text{obs}}$ is the observed value for q_μ and $f(q_\mu | \mu, \boldsymbol{\theta})$ is the PDF describing q_μ for a certain μ and $\boldsymbol{\theta}$. The p -value can be transformed into the significance Z by Equation 6.7.

To determine $f(q_\mu | \mu, \boldsymbol{\theta})$ pseudo experiments must be thrown. However, this procedure usually needs large computing power and is highly time consuming. Therefore it is common to use instead the so called asymptotic approximation, which is based on Wilks' theorem [49]. This states that $f(q_\mu | \mu, \boldsymbol{\theta})$ can be exactly calculated in the limit of a large statistics sample. In this analysis all statistical tests are performed with asymptotic formulae and the Asimov dataset. This is a special artificial dataset that can be used to determine the standard deviation of $\hat{\mu}$ [49]. In the Asimov dataset all data is set equal to their expectation values.

The crucial part in a hypothesis test is the decision, when the null hypothesis H_0 should be accepted or rejected. Ideally, the probability of falsely rejecting H_0 should be kept small. For this purpose a significance level α is defined. If a statistical test gives a p -value below α ,

then H_0 is accepted. Usually a significance level of 5% is chosen, i.e. $p_\mu \leq 0.05 = \alpha$. This can be equivalently expressed by the confidence level (CL), which is defined as $1 - \alpha$. In this analysis all results are evaluated at a CL of 95%.

Following the method above, a model is excluded if $p_{s+b} < 5\%$, which corresponds to $CL_{s+b} = 95\%$. However, this can be problematic if the sensitivity to the signal processes is very low, i.e. when the number of signal-plus-background events is almost equal to the number of background-only events ($s + b \approx b$). Such signal models might be falsely excluded. To avoid this, the CL_s method [57] is used instead. The CL_s value is defined by:

$$CL_s = \frac{p_{s+b}}{1 - p_b} \quad (8.8)$$

If the test statistics of $s + b$ and b are close to each other, then p_b is large and the denominator therefore small. Through this CL_s is larger than CL_{s+b} . On the other hand, the CL_s method is still valid if $s + b$ and b are sufficiently different. In this case p_b is small and the denominator is ~ 1 , so that $CL_s \approx CL_{s+b}$. Therefore it is more practical to use the CL_s method, according to which a model is excluded if $CL_s < 5\%$. For the statistical tests in this analysis only the CL_s method is used.

8.2.2 HistFitter setup

The statistical analysis on the obtained results is performed with the software framework HistFitter [56], which uses the software packages HistFactory, RooStats and RooFit. These packages are again based on ROOT.

In this analysis a rather simple fit configuration is used for the statistical tests, taking into account only the signal regions and no control and validation regions. Usually, better results are observed by including also control regions into the analysis. However, the fit configuration is based on the configuration for the ICHEP studies of 2016, in which a dataset corresponding to an integrated luminosity of $\int \mathcal{L} dt = 13.1 \text{ fb}^{-1}$ was analyzed. This gave rather small statistics in signal and control regions. On the other hand, the background estimation at this time performed sufficiently well, so that no control regions were included into the fit strategy. However, with increasing luminosity the analysis would definitely profit from control regions, so that they should be introduced into the fit configuration at some later point, when even more statistics than 36.1 fb^{-1} is available.

Besides the signal regions, the fit configuration uses as input the observed data, the estimated background processes and the different RPV signals presented in Section 2.2.5. The statistical tests are performed for each signal model separately. All systematic uncertainties described in Section 7.2.3 are applied on the signal and background processes.

With the HistFitter framework a variety of statistical tests can be performed. In this analysis it is used to determine model independent upper limits and model dependent exclusion limits.

8.2.3 Model independent upper limits

Model independent upper limits allow to quantify the significance of an observed deviation of data over SM expectations without any assumptions on the signal model. It can thus be studied if the deviation is potentially a sign of new physics. To claim a discovery a significance of 5σ is required.

Using the data and background expectation as shown in Table 8.2, an upper limit on the number of allowed signal events is calculated using a hypothesis test. The H_0 hypothesis

is here the background-only assumption and H_1 the signal-plus-background assumption. It is called upper limit because this is the highest value of μ for which p_μ is greater than 5%. To determine the expected upper limit on signal events S_{exp}^{95} , the likelihood in the statistical test is evaluated for $s = 0$. The upper limit on the visible cross section of any signal model is obtained by dividing S_{obs}^{95} by the integrated luminosity. The visible cross section is given by the product of acceptance, reconstruction efficiency and production cross section and is therefore denoted by $\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$.

Table 8.2 shows the resulting $\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$, S_{obs}^{95} and S_{exp}^{95} . The table shows also the different CL_b values for the background-only hypothesis, the p -value and the corresponding significance Z . The highest Z is seen for SR1. However, even though twice as many events were observed than predicted in this region, the obtained Z of 1.43σ gives no significant indication for any new physics processes.

Sample	SR0A	SR0B	SR1	SR2
Observed	13	2	8	2
SM Total	10.7 ± 1.9	1.44 ± 0.25	4.2 ± 1.1	2.5 ± 0.6
$\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$ [fb]	0.32	0.13	0.29	0.13
S_{obs}^{95}	11.5	4.6	10.3	4.9
S_{exp}^{95}	$8.8^{+4.1}_{-1.6}$	$4.0^{+2.3}_{-1.3}$	$6.2^{+3.1}_{-1.2}$	$5.2^{+2.5}_{-1.9}$
CL_b	0.71	0.63	0.88	0.41
$p(s = 0)$	0.27	0.34	0.08	0.50
Z	0.60	0.43	1.43	0

Table 8.2: Model independent limits in the different signal regions: Shown are the observed upper limit on the visible cross section $\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$, the observed and expected upper limit on the number of signal events S_{obs}^{95} and S_{exp}^{95} , the CL_b value for the background-only hypothesis, the p -value and the corresponding significance Z to the p -value.

8.2.4 Exclusion limits

If no significant excess of data over SM background expectation is observed, one can set exclusion limits on the parameter space of the signal models. In this analysis the models contain only two free parameters, i.e. the NLSP mass and the LSP mass. By calculating the exclusion significance for each grid point of a signal model exclusion contours can be derived. This is similar to the procedure explained in Chapter 6, by which the discovery contours from Figure 6.5 were obtained. However, in this case the hypotheses are inverted: The signal-plus-background model is taken as H_0 , while H_1 represents the background-only scenario.

Figures 8.1 and 8.2 show the resulting exclusion contours for the RPV model with initial $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production (see Figure 2.3a) in SR1 and SR2, respectively. The x-axis represents the mass of the $\tilde{\chi}_1^0$ LSP and the y-axis the $\tilde{\chi}_1^\pm$ mass. Points below $m_{\tilde{\chi}_1^0} = 450$ GeV are not considered because these were already excluded in Run I. The solid red line displays the observed limit, while the two red dotted lines indicate which limits are observed when the production cross sections of the signal processes are varied up or down by one standard deviation. The gray dashed line shows the expected limit, which is calculated by setting $s = 0$ in the likelihood of the exclusion fit. The yellow band is its $\pm 1\sigma$ uncertainty.

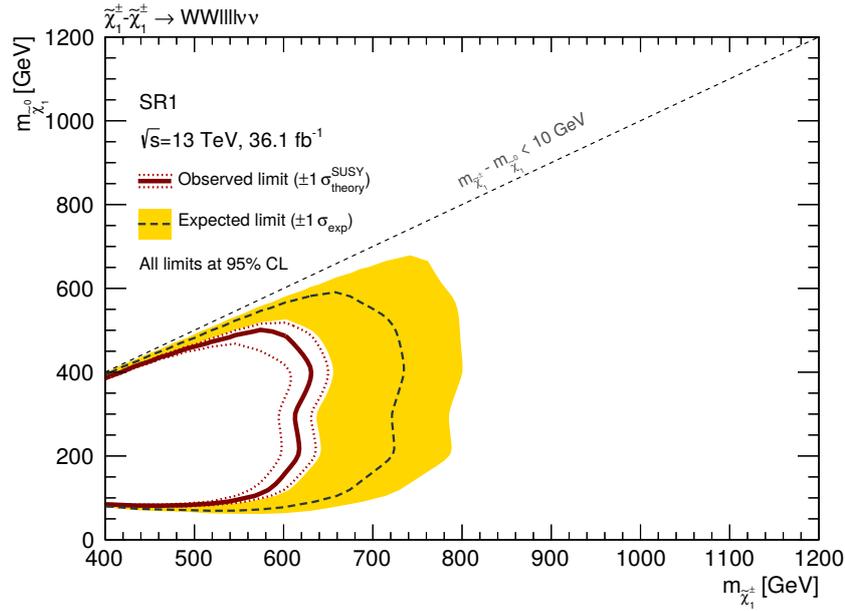


Figure 8.1: 95% CL exclusion limits for signal models with initial $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production in SR1 for $LL\bar{E}i33$ models. The solid red line shows the observed limit and the dotted red lines the observed limits obtained when the cross sections of the SUSY processes are varied by $\pm 1\sigma$. The gray dashed line indicates the expected limit and the yellow band its $\pm 1\sigma$ uncertainty.

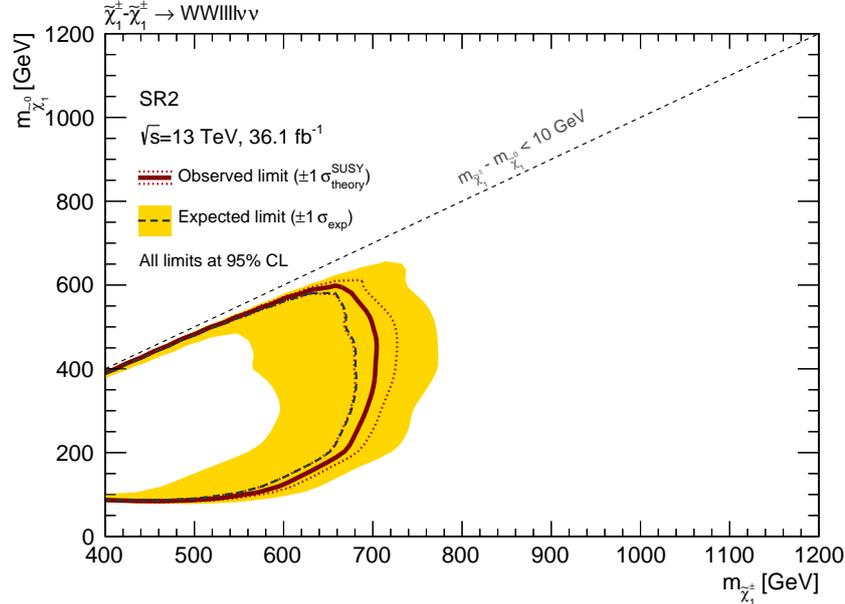


Figure 8.2: 95% CL exclusion limits for signal models with initial $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production in SR2 for $LL\bar{E}i33$ models. The solid red line shows the observed limit and the dotted red lines the observed limits obtained when the cross sections of the SUSY processes are varied by $\pm 1\sigma$. The gray dashed line indicates the expected limit and the yellow band its $\pm 1\sigma$ uncertainty.

By comparing the two contours it is seen that SR1 is expected to have a higher exclusion power. The expected limit reaches to chargino masses up to ~ 700 GeV, which exceeds the expected limit of SR2 by about 50 GeV. As in SR1 a deviation of data over SM expectations is seen, the observed limit is weaker than the expected. In SR2 almost the same number of events is seen as expected, so that the observed and expected limit are close to each other.

One can increase the exclusion sensitivity for a signal model by combining the results from different signal regions. Depending on how the regions are related to each other, different methods must be applied. If the regions are statistically independent, i.e. do not share events, then the exclusion fit can be performed simultaneously for all regions. However, this is not possible if the signal regions overlap. This is the case e.g. for SR0A and SR0B, which differ only by a cut on m_{eff} , so that SR0A is a subset of SR0B. These can be combined by comparing their significance for each grid point and using each time only the value of the region with the best expected exclusion.

SR1 and SR2 are orthogonal because they differ from each other in their lepton requirements. Therefore they can be combined during the fitting process. Furthermore, a signal from $LL\bar{E}i33$ models can also appear in $4\ell 0\tau$ final states provided that all taus decay leptonically. Hence the exclusion sensitivity on $LL\bar{E}i33$ signals can be further increased by including also a signal region with four light leptons into the fit. For this purpose SR0B is chosen because it was found to have a better exclusion power than SR0A for the whole grid.

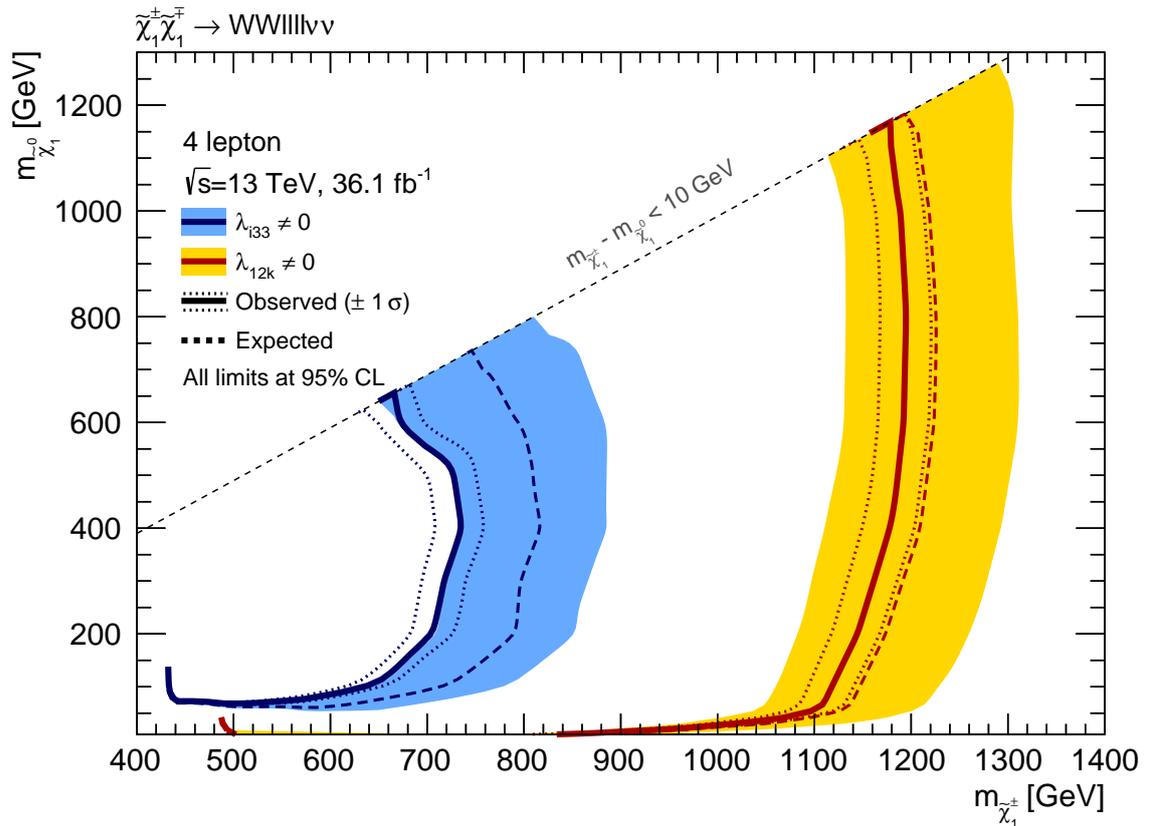


Figure 8.3: 95% CL exclusion limits for signal models with initial $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production. The blue contour refers to the $LL\bar{E}i33$ models and the red-yellow contour to the $LL\bar{E}12k$ models. The solid lines show the observed limits and the dotted lines the observed limits obtained by varying the signal cross sections by $\pm 1\sigma$. The dashed lines indicate the expected limits and the colored bands their $\pm 1\sigma$ uncertainties.

The result of this combination is shown in Figure 8.3, in which the blue contour refers to the $LL\bar{E}i33$ models. Charginos with masses up to 700 GeV are excluded by this. The yellow-red contour shows the obtained exclusion limits for the $LL\bar{E}12k$ models. For these models only SR0B is used in the fit. The limits reach up to $m_{\tilde{\chi}_1^\pm} \approx 1.2$ TeV. Figure 8.4 shows the exclusion limits for the model with initial $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ production (see Figure 2.3b). For both couplings the observed limits are ~ 100 GeV higher than for the $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ model.

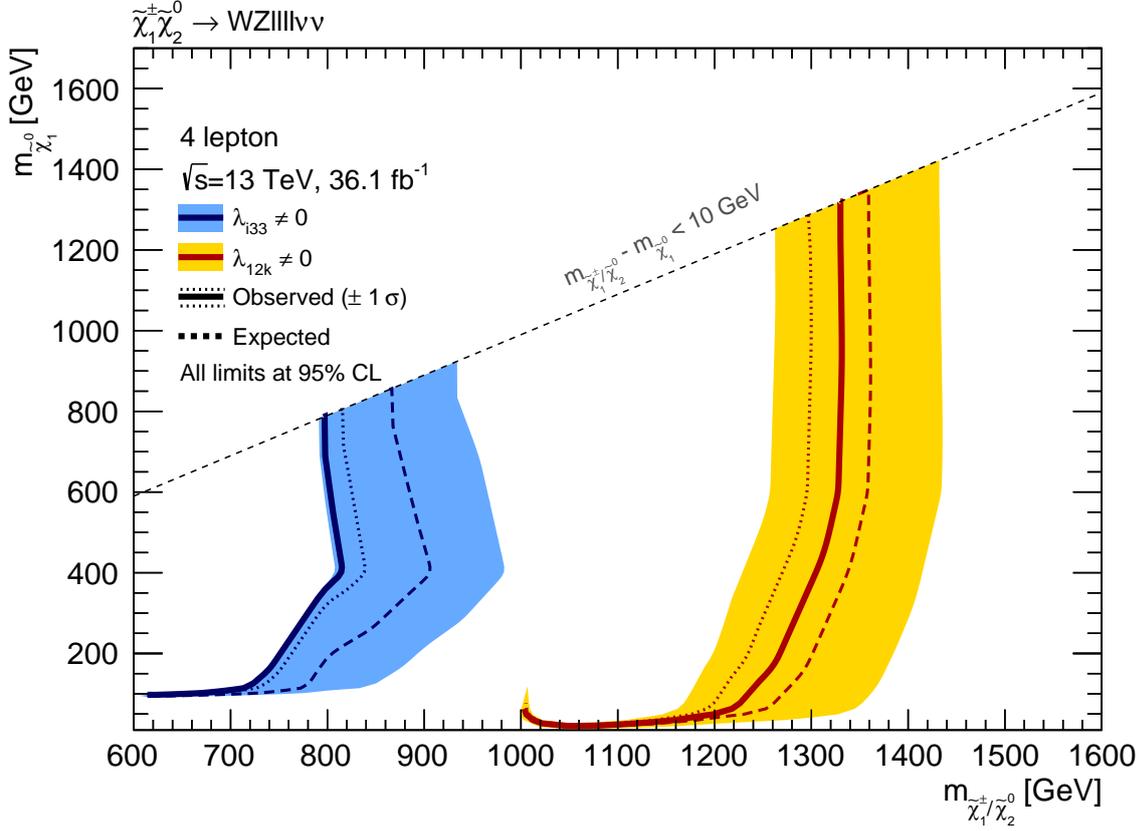
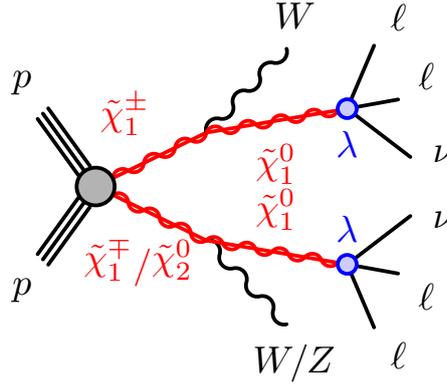
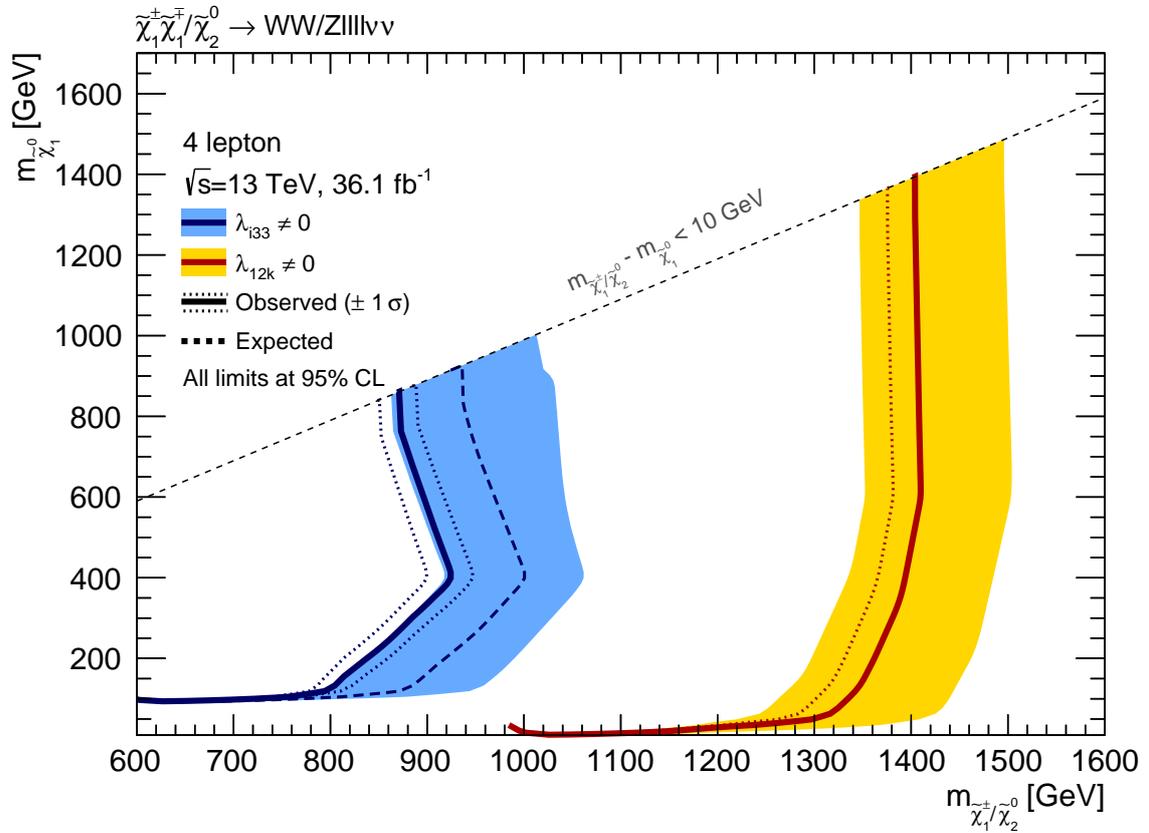


Figure 8.4: 95% CL exclusion limits for signal models with initial $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ production. The blue contour refers to the $LL\bar{E}i33$ models and the red-yellow contour to the $LL\bar{E}12k$ models. The solid lines show the observed limits and the dotted lines the observed limits obtained by varying the signal cross sections by $\pm 1\sigma$. The dashed lines indicate the expected limits and the colored bands their $\pm 1\sigma$ uncertainties.

One can construct from these two simplified models a combined model by merging their processes. In general this new model would describe scenarios, in which a $\tilde{\chi}_1^\pm$ is produced together with either a $\tilde{\chi}_1^\mp$ or a $\tilde{\chi}_2^0$. The diagram for this process is shown in Figure 8.5. The corresponding limits can be seen in Figure 8.6. For the $LL\bar{E}i33$ models $\tilde{\chi}_1^\pm/\tilde{\chi}_2^0$ masses of up to 920 GeV are excluded, while for $LL\bar{E}12k$ the observed limits reach up to 1.4 TeV.

The limits for the simplified model with initial gluino pair production (see Figure 2.3c) are shown in Figure 8.7. Gluino masses up to 1.6 TeV / 2.1 TeV are excluded for $LL\bar{E}i33$ / $LL\bar{E}12k$ scenarios, respectively. Finally, the limits for the model with left handed sleptons and sneutrinos (see Figure 2.3d) is shown in Figure 8.8, where $\tilde{\ell}_L/\tilde{\nu}$ masses are excluded up to 620 GeV for $LL\bar{E}i33$ and up to 960 GeV for $LL\bar{E}12k$ models.

Figure 8.5: Simplified model with initial $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp / \tilde{\chi}_2^0$ production.Figure 8.6: 95% CL exclusion limits for signal models with initial $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp / \tilde{\chi}_2^0$ production. The blue contour refers to the $LL\bar{E}i33$ models and the red-yellow contour to the $LL\bar{E}12k$ models. The solid lines show the observed limits and the dotted lines the observed limits obtained by varying the signal cross sections by $\pm 1\sigma$. The dashed lines indicate the expected limits and the colored bands their $\pm 1\sigma$ uncertainties.

Except for the gluino case, all contours for the $LL\bar{E}i33$ models showed no sensitivity for low $\tilde{\chi}_1^0$ masses. In this region the mass splitting between the NLSP and LSP is large. This leads to boosted decay products, which are difficult to reconstruct as distinct objects. That effect can also be seen for the $LL\bar{E}12k$ models, where the limits for models with large mass splittings are on average ~ 100 GeV lower compared to models with smaller mass splittings.

To improve this issue, the isolation criteria used in this analysis need to be optimized.

With the presented results the exclusion limits from Run I could be exceeded. In Run I $LL\bar{E}i33$ and $LL\bar{E}12k$ scenarios were studied for the modes with initial $\tilde{\chi}_1^+ \tilde{\chi}_1^-$, $\tilde{g}\tilde{g}$ and $\tilde{\ell}_L \tilde{\ell}_L / \tilde{\nu} \tilde{\nu}$ production, too. In the case of $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production, the observed limits are higher than the corresponding Run I limits by ~ 150 GeV and ~ 400 GeV for $LL\bar{E}i33$ and $LL\bar{E}12k$ models, respectively. For $\tilde{g}\tilde{g}$ production the exclusion limits of Run I are exceeded by approximately 600 GeV for both couplings. Finally, in the model with an initial $\tilde{\ell}_L \tilde{\ell}_L / \tilde{\nu} \tilde{\nu}$ pair the limits are exceeded by ~ 200 GeV and ~ 400 GeV for $LL\bar{E}i33$ and $LL\bar{E}12k$ models, respectively.

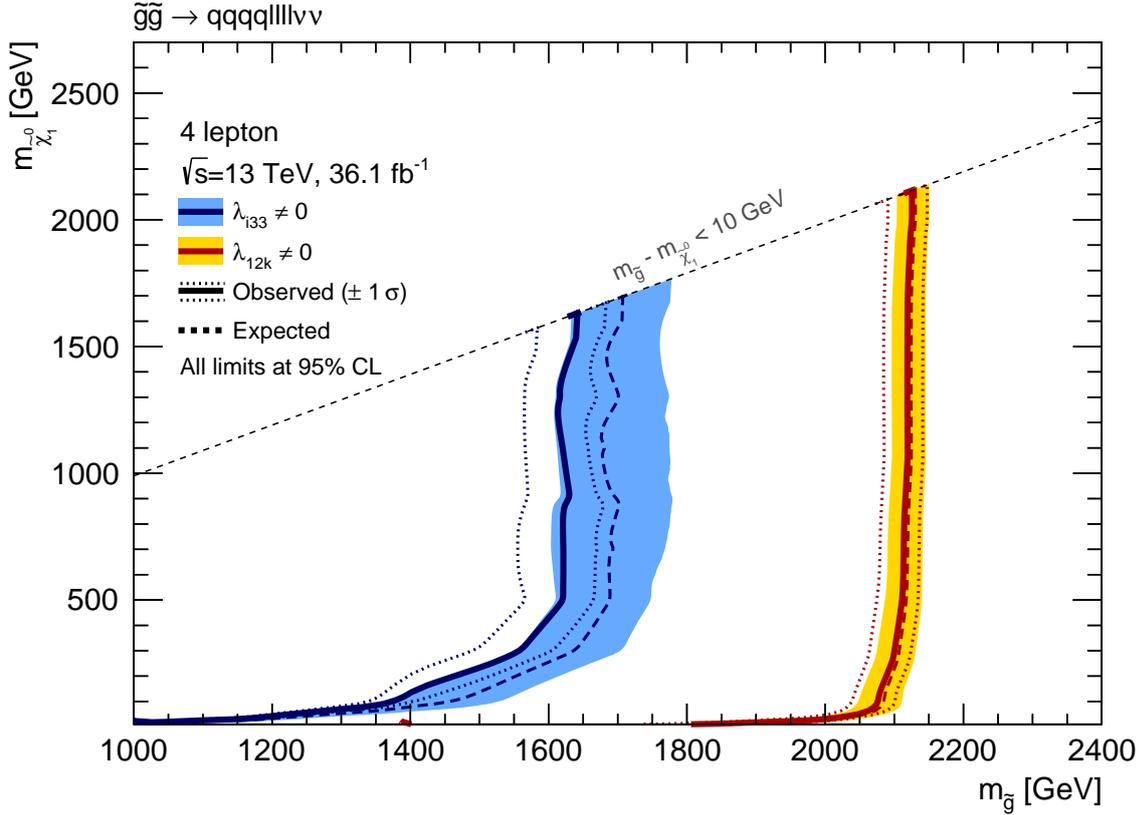


Figure 8.7: 95% CL exclusion limits for signal models with initial $\tilde{g}\tilde{g}$ production. The blue contour refers to the $LL\bar{E}i33$ models and the red-yellow contour to the $LL\bar{E}12k$ models. The solid lines show the observed limits and the dotted lines the observed limits obtained by varying the signal cross sections by $\pm 1\sigma$. The dashed lines indicate the expected limits and the colored bands their $\pm 1\sigma$ uncertainties.

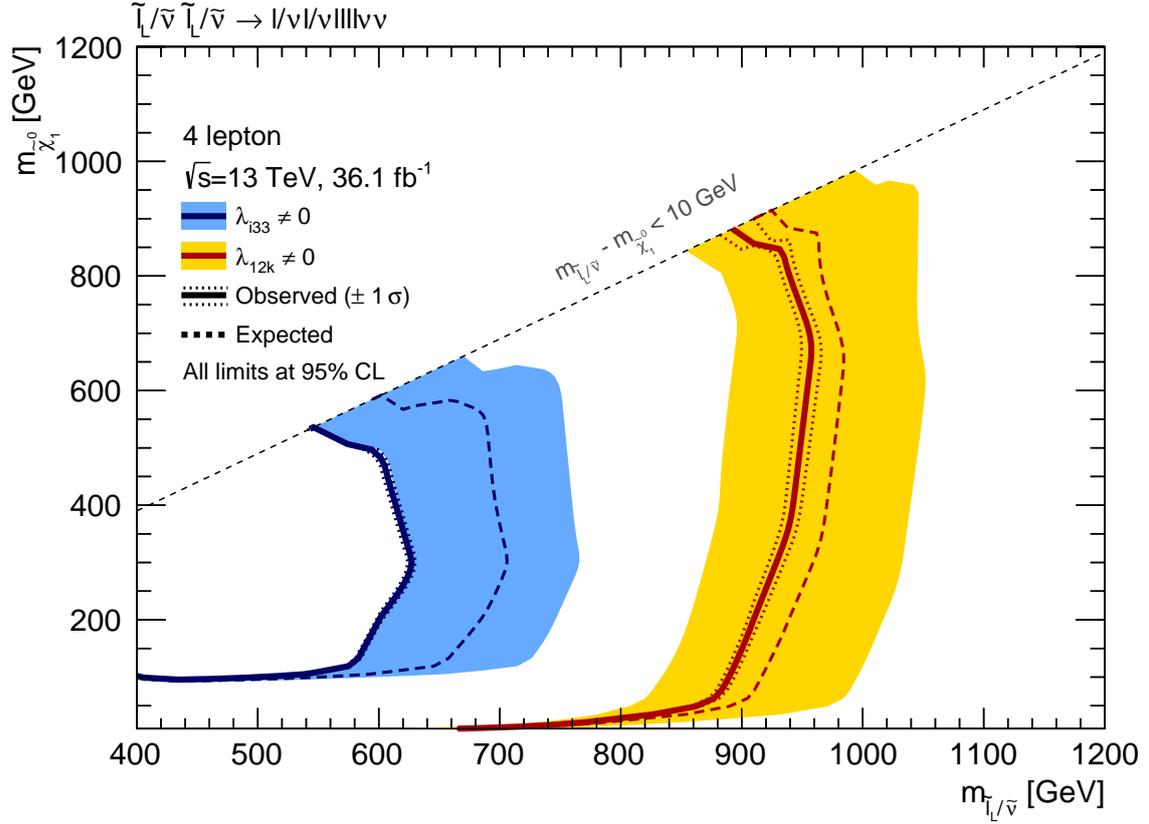


Figure 8.8: 95% CL exclusion limits for signal models with initial $\tilde{\ell}_L \tilde{\ell}_L / \tilde{\nu} \tilde{\nu}$ production. The blue contour refers to the $LL\bar{E}i33$ models and the red-yellow contour to the $LL\bar{E}12k$ models. The solid lines show the observed limits and the dotted lines the observed limits obtained by varying the signal cross sections by $\pm 1\sigma$. The dashed lines indicate the expected limits and the colored bands their $\pm 1\sigma$ uncertainties.

Chapter 9

Conclusion and outlook

An analysis was presented targeting SUSY RPV models using multileptonic final states. The different steps of the analysis follow closely the methods of the ICHEP studies from 2016 [5], in which one benchmark model was studied. In this model the lightest supersymmetric particle (LSP) decays via an R -parity violating (RPV) process into two charged leptons and a neutrino. The initially produced SUSY particles are assumed to be a $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ pair. For the ICHEP conference only four-lepton events with electrons and muons were studied. Thereby only models could be investigated, in which the LSP couples exclusively to light leptons (referred to as $LL\bar{E}12k$ models). The ICHEP studies were extended by three additional RPV signals with initial pair production of $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$, $\tilde{g}\tilde{g}$ and $\tilde{\ell}_L \tilde{\ell}_L / \tilde{\nu}\tilde{\nu}$. Furthermore, another coupling scenario was included into the analysis, which allows the LSP to decay also to taus (referred to as $LL\bar{E}i33$ models).

For this purpose two new signal regions are introduced containing four-lepton final states with taus. These two signal regions are characterized by $3\ell 1\tau$ and $2\ell 2\tau$ events. Three reference models from the $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ simplified model are used for the optimization of the signal regions. In these three models the chargino and neutralino have masses, which lie slightly above the exclusion limits from Run I. The chosen reference models are: $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 100)$, $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 490)$ and $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (600, 300)$. By applying a Z -veto the SM background is vastly reduced while leaving the signal contribution almost unaffected. The signal regions are further optimized by demanding high m_{eff} values. The highest discovery significance is obtained with a requirement of $m_{\text{eff}} \geq 700 \text{ GeV} / 650 \text{ GeV}$ for $3\ell 1\tau / 2\ell 2\tau$ events, respectively.

The SM background contribution is estimated in two different ways: The irreducible background is estimated fully from MC simulation, while the reducible background is estimated from data control regions using the fake factor method. In the course of this work the essential steps of the fake factor method were presented: Calculating the fake factor for each fake source and SM process, determining the fake fractions for each lepton and SM process and scaling the MC predicted fake factors to data.

Results were shown for an integrated luminosity of $\int \mathcal{L} dt = 36.1 \text{ fb}^{-1}$. The data is found to be consistent with SM background expectations within 1.4σ . The largest deviation is seen in SR1, which targets $3\ell 1\tau$ final states. Exclusion limits are set on the various RPV scenarios under study. In the case of an initially produced $\tilde{\chi}_1^\pm / \tilde{\chi}_2^0$ the exclusion limits reach up to $\tilde{\chi}_1^\pm / \tilde{\chi}_2^0$ masses of 920 GeV for $LL\bar{E}i33$ models and up to 1.4 TeV for $LL\bar{E}12k$ models. Gluino masses are excluded up to 2.1 TeV / 1.6 TeV for $LL\bar{E}12k / LL\bar{E}i33$ scenarios, respectively. Finally, $\tilde{\ell}_L / \tilde{\nu}$ masses are excluded up to 960 GeV for $LL\bar{E}12k$ models, and up to 620 GeV for $LL\bar{E}i33$ models. In all cases the exclusion limits from Run I are exceeded: The smallest deviation from the Run I results is seen for $LL\bar{E}i33$ models with $\tilde{\chi}_1^+ \tilde{\chi}_1^-$ production, in which

the limits of Run I are exceeded by ~ 150 GeV, while the largest enhancement is obtained for $\tilde{g}\tilde{g}$ models, where the exclusion limits are for both couplings ~ 600 GeV larger than in Run I.

Although a large parameter subset of different RPV models could be excluded, there are many possibilities to further extend and improve the presented analysis. First of all, one should rerun the analysis as soon as more data is available because the current integrated luminosity yields very low statistics in the signal regions. With more statistics one can investigate, if the deviation in SR1 is really only a statistical fluctuation or if it increases. Apart from this, also the sensitivity in the signal grids is expected to increase with more data. Another important point is to find suitable methods, which enhance the discrimination of boosted objects. Thereby one can gain sensitivity to models with a high mass splitting between the next-to-lightest supersymmetric particle (NLSP) and the LSP. This is particularly relevant for boosted taus. Furthermore, one can also study the impact of the analysis on other RPV models, for example the right handed analog to the $\tilde{\ell}_L\tilde{\ell}_L$ model. In general, the analysis is very simple in its selection criteria, which allows to be sensitive to a variety of models with high m_{eff} and events with multiple leptons. Therefore the analysis is not restricted to RPV signals, but could be also tested on R -parity conserving models with a similar signature as already successfully done in Run I.

Appendix A

MC samples

In the following the SM and SUSY signal samples used in this analysis are listed. Table A.1 shows the SM cross sections, k-factors and filter efficiencies for the irreducible backgrounds and Table A.2 for the reducible backgrounds.

SUSY Signal

The signal samples have the form:

```
mc15_13TeV.*.MGPy8EG_A14N_C1C1_*_*_LLEi33
```

where the first * must be replaced by the dataset ID, the second * by the $\tilde{\chi}_1^\pm$ mass and the third * by the $\tilde{\chi}_1^0$ mass of the particular grid points.

Sample	σ [pb]	k	ϵ
ZZ			
mc15_13TeV.363490.Sherpa_221_NNPDF30NNLO_III	1.2557	1.0	1.0
mc15_13TeV.361075.Sherpa_CT10_ggllllNoHiggs	0.011826	1.67	1.
For a cross-check			
mc15_13TeV.361603.PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_ZZllllmll4	1.2568	1.	1.
t\bar{t}Z			
mc15_13TeV.410111.MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np0	0.0096235	1.51	1.
mc15_13TeV.410112.MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np1	0.017344	1.51	1.
mc15_13TeV.410113.MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np0	0.0096462	1.51	1.
mc15_13TeV.410114.MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np1	0.017361	1.51	1.
mc15_13TeV.410115.MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np0	0.0098874	1.51	1.
mc15_13TeV.410116.MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np1	0.017790	1.51	1.
For the evaluation of systematics			
mc15_13TeV.410202.MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np0_scaleUp	as 410111		
mc15_13TeV.410203.MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np0_scaleDn	as 410111		
mc15_13TeV.410204.MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np1_scaleUp	as 410112		
mc15_13TeV.410205.MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np1_scaleDn	as 410112		
mc15_13TeV.410206.MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np0_scaleUp	as 410113		
mc15_13TeV.410207.MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np0_scaleDn	as 410113		
mc15_13TeV.410208.MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np1_scaleUp	as 410114		
mc15_13TeV.410209.MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np1_scaleDn	as 410114		
mc15_13TeV.410210.MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np0_scaleUp	as 410115		
mc15_13TeV.410211.MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np0_scaleDn	as 410115		
mc15_13TeV.410212.MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np1_scaleUp	as 410116		
mc15_13TeV.410213.MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np1_scaleDn	as 410116		
t\bar{t}WW			
mc15_13TeV.410081.MadGraphPythia8EvtGen_A14NNPDF23_ttbbarWW	0.0080975	1.2231	1.
VVV			
mc15_13TeV.407311.Sherpa_221_NNPDF30NNLO_6l0v_EW6	0.00010235	1.	1.
mc15_13TeV.407312.Sherpa_221_NNPDF30NNLO_5l1v_EW6	0.00056766	1.	1.
mc15_13TeV.407313.Sherpa_221_NNPDF30NNLO_4l2v_EW6	0.0043684	1.	1.
mc15_13TeV.407314.Sherpa_221_NNPDF30NNLO_3l3v_EW6	0.015846	1.	1.
mc15_13TeV.361072.Sherpa_CT10_IIIjj_EW6	0.031496	0.91	1.
t\bar{t}t\bar{t}/t\bar{t} (includes t\bar{t}W)			
mc15_13TeV.410080.MadGraphPythia8EvtGen_A14NNPDF23_4topSM	0.0091622	1.0042	1.
mc15_13TeV.304014.MadGraphPythia8EvtGen_A14NNPDF23_3top_SM	0.00164	1.	1.
tWZ			
mc15_13TeV.410215.aMcAtNloPythia8EvtGen_A14_NNPDF23LO_260000_tWZDR	0.015558	1.	1.
Higgs			
VH			
mc15_13TeV.341421.PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_eveWWlvlv	0.00133518	0.987551064706	1.0
mc15_13TeV.341423.PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_mumuWWlvlv	0.00133477	0.987852755431	1.0
mc15_13TeV.341427.PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_tauvtauWWlvlv	0.00119259	1.1056182392	1.0
mc15_13TeV.341429.PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_eveWWlvlv	0.00192652	1.07134893874	1.0
mc15_13TeV.341431.PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_mumuWWlvlv	0.00192652	1.07134893874	1.0
mc15_13TeV.341435.PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_tauvtauWWlvlv	0.00192666	1.07127335768	1.0
mc15_13TeV.341447.Pythia8EvtGen_A14NNPDF23LO_WH125_Zllgam	0.00017831	1.20353707588	1.0
mc15_13TeV.341449.PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_eeWWlvlv_VpT	0.024531	1.	1.
mc15_13TeV.341451.PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_mumuWWlvlv_VpT	0.024536	1.	1.
mc15_13TeV.341453.PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_tauvtauWWlvlv_VpT	0.0249	1.	1.
mc15_13TeV.341947.Pythia8EvtGen_A14NNPDF23LO_ZH125_ZZ4l	0.015767	1.45604363544	0.010256
mc15_13TeV.341964.Pythia8EvtGen_A14NNPDF23LO_WH125_ZZ4l	0.00029076	1.27767114215	1.0
VBF			
mc15_13TeV.341488.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_VBFH125_ZZ4lep	0.00103101	0.978615629651	1.0
ggF			
mc15_13TeV.344376.PowhegPythia8EvtGen_CT10_AZNLO_ggZH125_eeWWlvlv	0.00183	1.	1.
mc15_13TeV.344378.PowhegPythia8EvtGen_CT10_AZNLO_ggZH125_mumuWWlvlv	0.00183	1.	1.
mc15_13TeV.344380.PowhegPythia8EvtGen_CT10_AZNLO_ggZH125_tauvtauWWlvlv	0.00183	1.	1.
mc15_13TeV.341471.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_ggH125_ZZ4lep	0.0081266	1.45488273486	1.0
ttH			
mc15_13TeV.343365.aMcAtNloPythia8EvtGen_A14_NNPDF23_NNPDF30ME_ttH125_dilep	0.048237	1.	1.
mc15_13TeV.343366.aMcAtNloPythia8EvtGen_A14_NNPDF23_NNPDF30ME_ttH125_semilep	0.200556788	1.	1.
mc15_13TeV.343367.aMcAtNloPythia8EvtGen_A14_NNPDF23_NNPDF30ME_ttH125_allhad	0.20890322	1.	1.

Table A.1: Irreducible background samples

Sample	σ [pb]	k	ϵ
$t\bar{t}$			
mc15_13TeV.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad	696.11	1.1949	0.543
For a cross-check			
mc15_13TeV.410009.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_dil	696.12	1.1949	0.1053
mc15_13TeV.410500.PowhegPythia8EvtGen_A14_ttbar_hdamp172p5_nonallhad	730.28	1.139	0.543
Z + jets			
mc15_13TeV.361500.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_Np0	1401.6	1.232	1.
mc15_13TeV.361501.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_Np1	211.99	1.232	1.
mc15_13TeV.361502.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_Np2	67.305	1.232	1.
mc15_13TeV.361503.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_Np3	18.679	1.232	1.
mc15_13TeV.361504.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_Np4	7.2910	1.232	1.
mc15_13TeV.361505.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_Np0	1402.0	1.232	1.
mc15_13TeV.361506.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_Np1	211.95	1.232	1.
mc15_13TeV.361507.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_Np2	67.353	1.232	1.
mc15_13TeV.361508.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_Np3	18.633	1.232	1.
mc15_13TeV.361509.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_Np4	7.3013	1.232	1.
mc15_13TeV.361510.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_Np0	1397.8	1.232	1.
mc15_13TeV.361511.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_Np1	211.40	1.232	1.
mc15_13TeV.361512.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_Np2	67.176	1.232	1.
mc15_13TeV.361513.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_Np3	18.609	1.232	1.
mc15_13TeV.361514.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_Np4	7.2749	1.232	1.
mc15_13TeV.361628.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_lowMll_Np0	2677.1	1.232	1.
mc15_13TeV.361629.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_lowMll_Np1	44.988	1.232	1.
mc15_13TeV.361630.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_lowMll_Np2	29.292	1.232	1.
mc15_13TeV.361631.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_lowMll_Np3	6.117	1.232	1.
mc15_13TeV.361632.MadGraphPythia8EvtGen_A14NNPDF23LO_Zee_lowMll_Np4	2.2091	1.232	1.
mc15_13TeV.361633.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_lowMll_Np0	2677.1	1.232	1.
mc15_13TeV.361634.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_lowMll_Np1	44.98	1.232	1.
mc15_13TeV.361635.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_lowMll_Np2	29.302	1.232	1.
mc15_13TeV.361636.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_lowMll_Np3	6.1088	1.232	1.
mc15_13TeV.361637.MadGraphPythia8EvtGen_A14NNPDF23LO_Zmumu_lowMll_Np4	2.2231	1.232	1.
mc15_13TeV.361638.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_lowMll_Np0	2412.4	1.232	1.
mc15_13TeV.361639.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_lowMll_Np1	46.019	1.232	1.
mc15_13TeV.361640.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_lowMll_Np2	28.852	1.232	1.
mc15_13TeV.361641.MadGraphPythia8EvtGen_A14NNPDF23LO_Ztautau_lowMll_Np3	6.0513	1.232	1.
W + jets			
mc15_13TeV.361520.MadGraphPythia8EvtGen_A14NNPDF23LO_Wenu_Np0	13939.	1.20185	1.
mc15_13TeV.361521.MadGraphPythia8EvtGen_A14NNPDF23LO_Wenu_Np1	1894.0	1.20185	1.
mc15_13TeV.361522.MadGraphPythia8EvtGen_A14NNPDF23LO_Wenu_Np2	642.66	1.20185	1.
mc15_13TeV.361523.MadGraphPythia8EvtGen_A14NNPDF23LO_Wenu_Np3	179.18	1.20185	1.
mc15_13TeV.361524.MadGraphPythia8EvtGen_A14NNPDF23LO_Wenu_Np4	70.785	1.20185	1.
mc15_13TeV.361525.MadGraphPythia8EvtGen_A14NNPDF23LO_Wmunu_Np0	13935.	1.20185	1.
mc15_13TeV.361526.MadGraphPythia8EvtGen_A14NNPDF23LO_Wmunu_Np1	1893.3	1.20185	1.
mc15_13TeV.361527.MadGraphPythia8EvtGen_A14NNPDF23LO_Wmunu_Np2	642.70	1.20185	1.
mc15_13TeV.361528.MadGraphPythia8EvtGen_A14NNPDF23LO_Wmunu_Np3	179.19	1.20185	1.
mc15_13TeV.361529.MadGraphPythia8EvtGen_A14NNPDF23LO_Wmunu_Np4	70.761	1.20185	1.
mc15_13TeV.361530.MadGraphPythia8EvtGen_A14NNPDF23LO_Wtaunu_Np0	13920.	1.20185	1.
mc15_13TeV.361531.MadGraphPythia8EvtGen_A14NNPDF23LO_Wtaunu_Np1	1891.9	1.20185	1.
mc15_13TeV.361532.MadGraphPythia8EvtGen_A14NNPDF23LO_Wtaunu_Np2	641.87	1.20185	1.
mc15_13TeV.361533.MadGraphPythia8EvtGen_A14NNPDF23LO_Wtaunu_Np3	179.21	1.20185	1.
mc15_13TeV.361534.MadGraphPythia8EvtGen_A14NNPDF23LO_Wtaunu_Np4	71.012	1.20185	1.
WZ, WW			
mc15_13TeV.363491.Sherpa_221_NNPDF30NNLO_llv	4.5877	1.	1.
mc15_13TeV.363492.Sherpa_221_NNPDF30NNLO_llvv	12.465	1.	1.
mc15_13TeV.361069.Sherpa_CT10_llvjj_ss_EW4	0.025765	0.91	1.
mc15_13TeV.361070.Sherpa_CT10_llvjj_ss_EW6	0.043375	0.91	1.
mc15_13TeV.361071.Sherpa_CT10_llvjj_EW6	0.042287	0.91	1.
mc15_13TeV.361072.Sherpa_CT10_llljj_EW6	0.031496	0.91	1.
mc15_13TeV.361077.Sherpa_CT10_ggllvv	0.85492	0.91	1.
mc15_13TeV.363356.Sherpa_221_NNPDF30NNLO_ZqqZll	15.563	1.	0.13961
mc15_13TeV.363358.Sherpa_221_NNPDF30NNLO_WqqZll	3.437	1.	1.
$t\bar{t}W$			
mc15_13TeV.410066.MadGraphPythia8EvtGen_A14NNPDF23LO_ttW_Np0	0.17656	1.32	1.
mc15_13TeV.410067.MadGraphPythia8EvtGen_A14NNPDF23LO_ttW_Np1	0.14062	1.32	1.
mc15_13TeV.410068.MadGraphPythia8EvtGen_A14NNPDF23LO_ttW_Np2	0.13680	1.32	1.

Table A.2: Reducible background samples

Appendix B

Region definitions

Table B.1 shows the definition of the various signal, control and validation regions used in the analysis. The control and validation regions are defined such as that they are orthogonal to the signal regions.

Sample	$N(e, \mu)$ signal	$N(e, \mu)$ loose	$N(\tau)$ signal	$N(\tau)$ loose	Z-boson	Selection
SR0A	≥ 4	≥ 0	$= 0$	≥ 0	veto	$m_{\text{eff}} > 600 \text{ GeV}$
SR0B	≥ 4	≥ 0	$= 0$	≥ 0	veto	$m_{\text{eff}} > 1100 \text{ GeV}$
VR0	≥ 4	≥ 0	$= 0$	≥ 0	veto	$m_{\text{eff}} < 600 \text{ GeV}$
CR1_LLLl	$= 3$	≥ 1	$= 0$	≥ 0	veto	as SR/VR
CR2_LLll	$= 2$	≥ 2	$= 0$	≥ 0	veto	as SR/VR
SR1	$= 3$	$= 0$	≥ 1	≥ 0	veto	$m_{\text{eff}} > 700 \text{ GeV}$
VR1	$= 3$	$= 0$	≥ 1	≥ 0	veto	$m_{\text{eff}} < 700 \text{ GeV}$
CR1_LLLt	$= 3$	$= 0$	$= 0$	≥ 1	veto	as SR/VR
CR1_LLt1	$= 2$	$= 1$	≥ 1	≥ 0	veto	as SR/VR
CR2_LLlt	$= 2$	$= 1$	$= 0$	≥ 1	veto	as SR/VR
SR2	$= 2$	$= 0$	≥ 2	≥ 0	veto	$m_{\text{eff}} > 650 \text{ GeV}$
VR2	$= 2$	$= 0$	≥ 2	≥ 0	veto	$m_{\text{eff}} < 650 \text{ GeV}$
CR1_LLt1	$= 2$	$= 0$	$= 1$	≥ 1	veto	as SR/VR
CR2_LLtt	$= 2$	$= 0$	$= 0$	≥ 2	veto	as SR/VR

Table B.1: Signal, control and validation region definitions.

Bibliography

- [1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett. B716 (2012) 1-29*, 2012. arXiv:1207.7214 [hep-ex].
- [2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B 716 (2012) 30*, 2012. arXiv:1207.7235 [hep-ex].
- [3] ATLAS Collaboration. Search for supersymmetry in events with four or more leptons in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector. *Phys. Rev. D. 90, 052001 (2014)*, 2014. arXiv:1405.5086 [hep-ex].
- [4] CERN Homepage. <https://home.cern/about/updates/2015/06/lhc-experiments-back-business-record-energy>, Accessed: 10-05-2017.
- [5] ATLAS Collaboration. Search for supersymmetry in events with four or more leptons in $\sqrt{s} = 13$ TeV pp collisions using 13.3 fb⁻¹ of ATLAS data. *ATLAS-CONF-2016-075*, 2016. <https://cds.cern.ch/record/2206245>.
- [6] D. Griffiths. *Introduction to Elementary Particles*. WILEY-VCH, 2008.
- [7] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [8] D. Schaile. Lecture Notes on "Advanced Particle Physics", 2015. LMU München.
- [9] C. Patrignani et al. (Particle Data Group). Quark Summary Table. *Chin. Phys. C, 40, 100001 (2016)*, 2016. <http://pdg.lbl.gov/2016/tables/rpp2016-sum-quarks.pdf>.
- [10] C. Patrignani et al. (Particle Data Group). Lepton Summary Table. *Chin. Phys. C, 40, 100001 (2016)*, 2016. <http://pdg.lbl.gov/2016/tables/rpp2016-sum-leptons.pdf>.
- [11] C. Patrignani et al. (Particle Data Group). Boson Summary Table. *Chin. Phys. C, 40, 100001 (2016)*, 2016. <http://pdg.lbl.gov/2016/tables/rpp2016-sum-gauge-higgs-bosons.pdf>.
- [12] S. P. Martin. A Supersymmetry Primer. DOI: 10.1142/9789812839657-0001, 1998. arXiv:hep-ph/9709356.
- [13] Planck Mission Cardiff University. <http://planck.cf.ac.uk/results/cosmic-microwave-background>. Accessed: 10-05-2017.
- [14] K. Müller. Einführung supersymmetrie. <http://www.physik.uzh.ch/~kmueller/text/vorlesung/susy.pdf>, 2002. Universität Zürich.
- [15] A. V. Gladyshev and D. I. Kazakov. IS (Low Energy) SUSY STILL ALIVE? *Lectures at the European School of High-Energy Physics, June 2012*, 2012. arXiv:1212.2548 [hep-ph].

- [16] ATLAS Collaboration. Search for supersymmetry in events with four or more leptons in 36.1 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ pp collisions with the ATLAS detector. *ATL-COM-PHYS-2017-160*, 2016. Internal Note.
- [17] ATLAS Collaboration. Search for supersymmetry in events with four or more leptons in $\sqrt{s} = 13 \text{ TeV}$ pp collisions with ATLAS. 2017. To be published.
- [18] CERN Homepage. <https://home.cern/about>. Accessed: 10-04-2017.
- [19] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation Vol.3 no.8*, 2008. <http://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08001/pdf>.
- [20] CERN Media and Press. <http://press.cern/backgrounders/facts-figures>. Accessed: 10-04-2017.
- [21] CERN Homepage. <https://home.cern/about/accelerators>, Accessed: 10-04-2017.
- [22] LHC the Guide. <http://cds.cern.ch/record/2255762/files/CERN-Brochure-2017-002-Eng.pdf>, 2017.
- [23] Picture of LHC accelerator complex. <https://bigscience.web.cern.ch/bigscience/en/lhc/lhc2.html>. Accessed: 10-04-2017.
- [24] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation Vol.3 no.8*, 2008. <http://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08003/pdf>.
- [25] ATLAS Collaboration. ATLAS Detector and Physics Performance: Technical Design Report, Volume 1. *ATLAS-TDR-14 ; CERN-LHCC-99-014*, 1999. <https://cds.cern.ch/record/391176/files/ce-0317330.pdf>.
- [26] CERN Homepage - Run II Update. <http://home.cern/about/updates/2014/05/new-subdetector-atlas>. Accessed: 10-04-2017.
- [27] D. Green. *High p_T Physics at Hadron Colliders*. Cambridge University Press, 2005.
- [28] A. R. Martinez. The Run-2 ATLAS Trigger System. *ATL-DAQ-PROC-2016-003*, 2016. <https://cds.cern.ch/record/2133909>.
- [29] J. Pequeno and P. Schaffner. *An computer generated image representing how ATLAS detects particles*. 2013. <https://cds.cern.ch/record/1505342>.
- [30] ATLAS Collaboration. Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS. *ATLAS-CONF-2011-152*, 2011. <https://cdsweb.cern.ch/record/1398195>.
- [31] M. H. Seymour and M. Marx. Monte Carlo Event Generators. *Lectures given at 69th Scottish Universities Summer School in Physics 2012 (SUSSP 69)*, 2013. arXiv:1304.6677 [hep-ph].
- [32] A. Buckley et al. General-purpose event generators for LHC physics. *DOI: 10.1016/j.physrep.2011.03.005*, 2011. arXiv:1101.2599 [hep-ph].
- [33] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP 02 (2009) 007*, 2008. arXiv:0811.4622 [hep-ph].
- [34] S. Alioli et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP 06 (2010) 043*, 2010. arXiv: 1002.2581 [hep-ph].

- [35] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP07(2014)079*, 2014. arXiv:1405.0301 [hep-ph].
- [36] T. Sjöstrand. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun. 191 (2015) 159*, 2014. arXiv:1410.3012 [hep-ph].
- [37] S. Frixione and B.R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP 0206:029,2002*, 2002. arXiv:hep-ph/0204244.
- [38] S. Mrenna T. Sjöstrand and P. Skands. PYTHIA 6.4 Physics and Manual. *JHEP 0605:026,2006*, 2006. arXiv:hep-ph/0603175.
- [39] S. Agostinelli et al. GEANT4 – a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment; Vol. 506, pages 250-303*, 2003.
- [40] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur.Phys.J.C70:823-874,2010*, 2010. arXiv:1005.4568 [physics.ins-det].
- [41] ATLAS Collaboration. Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector. *ATLAS-COM-CONF-2015-024*, 2015. <https://cds.cern.ch/record/2016323>.
- [42] ATLAS Collaboration. Electron identification measurements in ATLAS using $\sqrt{s} = 13$ TeV data with 50 ns bunch spacing. *ATL-PHYS-PUB-2015-041*, 2015. <https://cds.cern.ch/record/2048202>.
- [43] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C (2016) 76:292*, 2016. arXiv:1603.05598 [hep-ex].
- [44] G. P. Salam M. Cacciari and G. Soyez. The anti- k_t jet clustering algorithm. *JHEP 0804:063,2008*, 2008. arXiv:0802.1189 [hep-ph].
- [45] ATLAS Collaboration. Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. *ATL-PHYS-PUB-2015-015*, 2015. <https://cds.cern.ch/record/2037613>.
- [46] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J. C (2016) 76:581*, 2016. arXiv:1510.03823 [hep-ex].
- [47] ATLAS Collaboration. Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC. *ATL-PHYS-PUB-2015-045*, 2015. <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045/>.
- [48] ATLAS Collaboration. Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s} = 13$ TeV. *ATL-PHYS-PUB-2015-027*, 2015. <http://cds.cern.ch/record/2037904>.
- [49] E. Gross G. Cowan, K. Cranmer and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur.Phys.J.C71:1554,2011*, 2010. arXiv:1007.1727 [physics.data-an].
- [50] ROOT Software Framework. <https://root.cern.ch/>. Accessed: 12-05-2017.

- [51] RooStats::NumberCountingUtils::BinomialExpZ. https://root.cern.ch/root/html526/RooStats__NumberCountingUtils.html#RooStats__NumberCountingUtils:BinomialExpZ. Accessed: 12-05-2017.
- [52] Recommendations for systematic uncertainties on modelling of mc processes from the physics modelling group. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MC15SystematicUncertainties>. Accessed: 02-05-2017.
- [53] Overview and links to the ATLAS CP groups. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasPhysics>. Accessed: 02-05-2017.
- [54] Numbers calculated by the 4-lepton team using contributions by J. Junggeburth, M. Klein and the author.
- [55] R. J. Barlow. *A Guide to the Use of Statistical Methods in the Physical Sciences*. WILEY-VCH, 1989.
- [56] M. Baak et al. HistFitter software framework for statistical data analysis. *Eur. Phys. J. C (2015) 75: 153.*, 2015. arXiv:1410.1280 [hep-ex].
- [57] J. Lorenz. *Search for strongly interacting supersymmetric particles decaying to final states with an isolated lepton with the ATLAS detector at the LHC*. PhD thesis, LMU München, 2014.

Danksagung

An dieser Stelle möchte ich mich bei allen Leuten bedanken, die mich bei der Anfertigung dieser Arbeit unterstützt und mir geholfen haben:

- Prof. Dorothee Schaile, die mir die Möglichkeit gegeben hat, an ihrem Lehrstuhl zu arbeiten.
- Dr. Jeanette Lorenz für die ausgezeichnete Betreuung und Unterstützung im letzten Jahr wie auch für das Korrekturlesen dieser Arbeit.
- Frau Grimm-Zeidler für ihre Hilfe bei allen organisatorischen Angelegenheiten.
- Johannes Junggeburch dafür, dass er sich stets für meine Fragen Zeit genommen und mir gerne auch alles mehrfach erklärt hat.
- Christina Potter for helping me with any questions in the analysis and her support during the last year.
- Matt Klein for his help with analysis questions.
- In general the whole 4-lepton team. It was great working with you together.
- Allen Mitgliedern des Lehrstuhls für die angenehme Arbeitsatmosphäre, die Hilfsbereitschaft und die vielen Kaffeerunden, PubQuiz-Abende und vieles mehr. Es hat sehr viel Spaß gemacht in dieser Gruppe zu arbeiten.
- Meinen Eltern und meiner Schwester, die mich immer unterstützt haben und in jeder Situation für mich da waren.
- Bernhard Flierl, der immer für mich da war und mich unterstützt, ermuntert und aufgebaut hat, besonders während der stressigen Zeit des Zusammenschreibens.

Selbständigkeitserklärung

Ich versichere hiermit, die vorliegende Arbeit mit dem Titel

**Search for R-Parity Violating Supersymmetry in Final States with Four
Leptons Including Taus Using the ATLAS Detector**

selbständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Andrea Matic

München, den 22. Mai 2017