DISSERTATION

The Search for Top-Squark Pair Production with the ATLAS Detector at $\sqrt{s} = 13$ TeV in the Fully Hadronic Final State

Philipp Mogg

Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

The Search for Top-Squark Pair Production with the ATLAS Detector at $\sqrt{s} = 13$ TeV in the Fully Hadronic Final State

Dissertation

zur Erlangung des Doktorgrades der
Fakultät für Mathematik und Physik der
ALBERT-LUDWIGS-UNIVERSITÄT
Freiburg im Breisgau

vorgelegt von
Philipp Mogg

Dezember 2019

Dekan:	Prof. Dr. Wolfgang Soergel
Betreuer der Arbeit:	Prof. Dr. Karl Jakobs, Dr. Christian Weiser
Erstgutachter:	Dr. Christian Weiser
Zweitgutachter:	Prof. Dr. Marc Schumann
Datum der mündlichen Prüfung:	21.02.2020
Prüfer:	Dr. Christian Weiser
	Prof. Dr. Gregor Herten
	Prof. Dr. Stefan Dittmaier

"Scientific revolutions don't change the universe. They change how humans interpret it."

—*The Science of Discworld: Judgement Day*, by Terry Pratchett († 2015)

Contents

1 Introduction

The understanding of the most fundamental objects and their behaviour and interactions has been a dream of many a scientist. Much progress in that regard was made in the twentieth century, starting with quantum mechanics and special relativity, and subsequently quantum field theory which combines the two, accompanied by progress in mathematics like the understanding of symmetries, all of which have fundamentally changed the way we interpret the universe. This lead to a long line of experimental discoveries in particle physics and continuously new theoretical interpretations to match those. The conclusion of this effort is today known as the Standard Model of Particle Physics, a theory that describes all experimentally observed particles and fundamental forces except gravity. The Standard Model has withstood all attempts to directly contradict it for a long time, and all its prediction have been proved true, with the discovery of the Higgs Boson by the ATLAS and CMS collaborations [1, 2] having delivered the last piece.

However, since the Standard Model does not describe gravity and cannot explain other cosmological findings, we must assume that the Standard Model is not a full description of the universe. Other hints, like extreme fine-tuning of constants in the theoretical description, lead to the interpretation of the Standard model as the low-energy representation of a more fundamental theory. For these reasons, the search for new physics at unprecedented energy scales was one of the main goal in the building of ATLAS and the Large Hadron Collider.

Supersymmetry is a prominent and well-studied theory which can address several of the shortcomings of the Standard Model. It postulates a symmetry between the two classes of particles – bosons, which have an integer spin, and fermions, which have a half-integer spin.

This would lead to a whole new family of particles, which could be at an energy scale that is reachable with the LHC; therefore a rich search program exists at ATLAS. This thesis describes a search for the supersymmetric partner of the top quark – the top squark. We are looking for events where a top-squark pair is produced in a proton-proton collision at a centre-of-mass energy of 13 TeV; the unstable top-squarks then decay in several steps into light- and heavy-flavour quarks and neutralinos – another type of hypothetical supersymmetric particle which leaves the detector without a trace – leading to a detector signature with a multitude of hadronic jets, including b-jets, and missing transverse momentum.
A similar search was already conducted with data taken during Run 1 of the LHC at $\sqrt{s} = 8\,\text{TeV}$ without finding any evidence for new physics [3]. The higher centre-of-mass energy and luminosity during LHC Run 2 and also improved detector performance and analysis techniques allow to greatly extend the sensitivity into areas that could not be excluded so far.
The theoretical foundations are discussed in Chpt. 2, including the Standard Model and the possibility to expand the model with supersymmetry. The experimental set-up that was used to produce data, which is the Large Hadron Collider and the ATLAS detector, is described in Chpt. 3. The methods to reconstruct physics objects from the raw detector data are described in Chpt. 4. A search for top squarks was conducted with data taken in 2015 and 2016 at $\sqrt{s} = 13\,\text{TeV}$ [4] with major contributions by the author; this is described in Chpt. 5, including the strategies to define search regions with a high potential purity of signal events, the methods to estimate Standard Model backgrounds, the evaluation of systematic uncertainties and finally the results and interpretation of the search. Only part of the Run-2 data was analysed in this context yet, a search with the full dataset is yet to follow this thesis; Chpt. 6 describes studies of new methods to improve the sensitivity that were conducted by the author. A conclusive summary is then given in Chpt. 7.

2 Theory background

This chapter describes the theory behind the physics discussed in this thesis. Our current best knowledge about elementary particle physics is the Standard Model, which is described in Sec. 2.1. Section 2.2 then describes Supersymmetry as a possible extension to the Standard Model, which has the potential to solve several of the shortcomings (discussed in Sec. 2.2.1) that the Standard Model has despite its undeniable usefulness.

2.1 The Standard Model

The Standard Model of particle physics (SM) describes our current understanding of all known fundamental particles and their interactions, with the exception of gravity. It is a quantum field theory (QFT), a framework that combines special relativity and quantum mechanics. The basic principles and notations of QFT are described in [5]. The SM emerged in the 1960s and 1970s as a result of the work of many physicists describing the strong interaction and the electroweak theory, which is a unified description of the electromagnetic and the weak interaction. The SM has since been extremely successful in making predictions and withstanding experimental tests and is still our best description of fundamental particles and interactions. The description given here is a summary of information that can be found in [5–7], if not stated otherwise.

2.1.1 Structure

The strong and electroweak interaction are described by the exchange of spin-1 particles (gauge bosons) which are quanta of gauge fielda. These bosons can carry mass as well as charges.
QFT is based on Lagrangian mechanics and makes use of Hamilton's principle. This means that a system behaves always in a way so that the action functional

$$\mathscr{S} = \int \mathrm{d}^4 x \, \mathscr{L}\left(\Phi, \partial_\mu \Phi\right) \tag{2.1}$$

becomes extremal. A system is therefore completely described by its Lagrangian density $\mathscr{L}$, in the following simply called Lagrangian, which is a function of the quantum field $\Phi(x)$ at a space-time point x, and its first derivative $\partial_\mu \Phi(x)$. The Euler-Lagrange equation gives the equations of motion:

$$\frac{\partial \mathscr{L}}{\partial \Phi} - \frac{\partial}{\partial x^\mu}\left[\frac{\partial \mathscr{L}}{\partial\left(\partial_\mu \Phi\right)}\right] = 0. \tag{2.2}$$

The SM is a gauge theory, which means that the Lagrangian must be invariant under a continuous group of local transformations (Lie group). According to the Noether Theorem, each symmetry (under global or local transformation) is always associated with a conserved quantity [8]. The SM is composed of the theory of Quantum Chromodynamics (QCD), which describes the strong interaction [9], and the Glashow-Salam-Weinberg theory of electroweak interaction [10,11]. QCD is described by the Lie group SU(3), with the colour charge C being the conserved quantity of this symmetry. The electroweak theory imposes a SU(2) $\otimes$ U(1) symmetry. The SU(2) group is associated with the conservation of the third component of the weak isospin T_3; the U(1) group leads to conservation of the weak hypercharge Y_W. The electric charge Q can be calculated from those two quantities by the Gell-Mann-Nishijima relation [12]:

$$Q = T_3 + \frac{1}{2} Y_W. \tag{2.3}$$

The group structure of the SM is thus:

$$\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y; \tag{2.4}$$

C and Y denote the conserved quantities, L indicates that the SU(2) transformation only applies to left-handed fermions.

2.1.2 Particle content

In order to formulate the Lagrangian for the SM, we need to know about the elementary particles it aims to describe. The SM is a phenomenological theory, meaning that the group structure and the particle content are based on observation. The knowledge about these particles is thus a result of many experiments that took place in more than a century[1].
Particles in the SM can be divided into two groups based on their spin: *Fermions* are particles with a half-integer spin number, all elementary fermions have the spin number $^1\!/_2$. *Bosons*, on the other hand, have an integer spin number[2].

Fermions are the particles that make up ordinary matter. Elementary fermions can be further divided into two groups, *quarks* and *leptons*, with the main difference being that leptons do not interact via the strong interaction whereas quarks do. Quarks and leptons exist in three generations. Only the first generation[3] makes up normal matter (atoms), fermions of higher generations are unstable (with the exception of neutrinos). For quarks, each generation has an up-type (with an electric charge of $+^2\!/_3$ e) and a down-type (with an electric charge of $-^1\!/_3$ e) quark. Each quark also has a colour charge (three different colour charges – conventionally called red, green & blue – exist). The quarks and some of their properties are listed in Tab. 2.1. Nature usually presents us with bound states of quarks and/or their corresponding antiparticles. Bound states of one quark and one antiquark are called *mesons*,

[1]The first identified elementary particle was the electron, discovered by J.J. Thomson in 1897 [13]; the last SM particle to be discovered was the Higgs boson, discovered in 2012 by the ATLAS and CMS collaborations [1, 2].
[2]The Higgs Boson is the only known elementary spin-0 particle, all gauge bosons have a spin of 1.
[3]The generations are numbered according to the particle masses in ascending order.

Table 2.1: The three known generations of quarks and their properties (mass and electric charge) [14]

Generation	Name	Symbol	Q[e]	Mass
1	Up	u	$+2/3$	$2.2^{+0.5}_{-0.4}\,\mathrm{MeV}$
	Down	d	$-1/3$	$4.7^{+0.5}_{-0.3}\,\mathrm{MeV}$
2	Charm	c	$+2/3$	$1.275^{+0.025}_{-0.035}\,\mathrm{GeV}$
	Strange	s	$-1/3$	$95^{+9}_{-3}\,\mathrm{MeV}$
3	Top	t	$+2/3$	$173.0 \pm 0.4\,\mathrm{GeV}$
	Bottom	b	$-1/3$	$4.18^{+0.04}_{-0.03}\,\mathrm{GeV}$

with pions being prominent representatives of this group. Bound states of three quarks are called *baryons*, protons and neutrons are the best known among them. Bound states of quarks in general are called *hadrons*, they are all colour neutral from the outside (one colour and one anticolour charge cancel each other out, as well as three different colour charges).

For leptons, each generation consists of a charged lepton with $Q = -e$ and an electrically neutral neutrino. The leptons and some of their properties are listed in Tab. 2.2. The masses

Table 2.2: The three known generations of leptons and their properties (mass and electric charge) [14, 15]

Generation	Name	Symbol	Q[e]	Mass
1	Electron	e	-1	$510.9989461 \pm 0.0000031\,\mathrm{keV}$
	Electron neutrino	ν_e	0	$< 2\,\mathrm{eV}$
2	Muon	μ	-1	$105.6583745 \pm 0.0000024\,\mathrm{MeV}$
	Muon neutrino	ν_μ	0	$< 0.19\,\mathrm{MeV}$
3	Tau	τ	-1	$1.77686 \pm 0.00012\,\mathrm{GeV}$
	Tau neutrino	ν_τ	0	$< 18.2\,\mathrm{MeV}$

of neutrinos have not been measured yet, so only an upper limit can be given. In the SM, as it is described in this thesis, they are assumed to be zero[4].

On the side of elementary bosons, there are gluons, $W^{\pm}$ and Z^0 bosons, the photon and the Higgs boson. Gluons are the carriers of the strong interaction, they carry each a colour

[4]Massive neutrinos can be described as Dirac or Majorana neutrinos [16], but their nature is not known at this point, leaving them massless in the current SM.

and an anticolour charge, resulting in eight linearly independent gluons (“colour octet”). In principle one could form 9 linearly independent gluon states, but one of them would be a singlet state without colour charge and there is no evidence for this; this is a reason why the strong interaction is described by the SU(3), where the singlet state is forbidden, and not the U(3) group[5]. Gluons are assumed to be massless. The photon is the carrier of the electromagnetic interaction. It is also assumed to be massless and does not carry a colour or electric charge. The $W^{\pm}$ and Z^0 bosons are carriers of the weak interaction. They are both massive and the $W^{\pm}$ carries a positive or negative electric charge. The last particle of the SM is the Higgs boson, which was the last missing piece of the SM until discovered in 2012. It is not a gauge boson, but a quantum of the so-called Higgs field, explained in Sec. 2.1.4. The bosons and some of their properties are displayed in Tab. 2.3.

Table 2.3: The known bosons and their properties (mass and electric charge) [14]

Name	Symbol	el. charge [e]	Mass
Photon	γ	$< 10^{-35}$	$< 10^{-18}\,\mathrm{eV}$
W boson	$W^{\pm}$	± 1	$80.379 \pm 0.012\,\mathrm{GeV}$
Z boson	Z^0	0	$91.1876 \pm 0.0021\,\mathrm{GeV}$
Gluon	g	0	$< 0.19\,\mathrm{MeV}$
Higgs boson	H^0	0	$125.18 \pm 0.16\,\mathrm{GeV}$

2.1.3 The Standard Model Lagrangian

First we consider, how the Lagrangian of a freely propagating fermion would look like. A free massive fermion of spin $1/2$ must follow the Dirac equation

$$\left[i \sum_{\mu=0}^{3} \gamma^{\mu} \frac{\partial}{\partial x^{\mu}} - m \right] \psi = 0, \tag{2.5}$$

with ψ being a spinor with four components and γ^{μ} being hermitian anti-commutating 4×4 matrices with squares equal to the identity matrix. A common convention for these matrices

[5]The adjoint representation of a $\mathrm{SU}(n)$ has a dimensionality of $n^2 - 1$, for the $\mathrm{U}(n)$ group it is n^2.

are the Dirac matrices [17]:

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix}; \gamma^i = \begin{pmatrix} & \sigma^i \\ -\sigma^i & \end{pmatrix}, i \in \{1,2,3\}, \tag{2.6}$$

with $\mathbb{1}$ being the 2×2 identity matrix and σ^i the Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.7}$$

A Dirac field can therefore be described by the Lagrangian

$$\mathscr{L} = \bar{\psi} \left(i\partial_\mu \gamma^\mu - m \right) \psi, \tag{2.8}$$

with $\bar{\psi} = \psi^\dagger \gamma^0$. Enforcing Euler-Lagrange equation 2.2 then leads to Eq. 2.5.

This Lagrangian is invariant under global U(1) transformation $\psi \to e^{i\alpha}\psi$, with α being a real constant. However, a stronger condition can be imposed requiring invariance under local gauge transformation. This means that α is allowed to depend on space-time coordinates:

$$\psi \to e^{i\alpha(x)}\psi. \tag{2.9}$$

Equation 2.8 is not invariant under this local transformation because the derivative would introduce a factor $\partial_\mu \alpha(x)$, so the Lagrangian has to be modified. The goal is to modify the derivative such that the term $\partial_\mu \alpha(x)$ is cancelled. This is done by introducing a vector field A_μ that transforms as $A_\mu \to A_\mu + {}^1\!/_q \cdot \partial_\mu \alpha$. The modified derivative is then

$$\partial_\mu \to \partial_\mu - iqA_\mu. \tag{2.10}$$

A_μ is called a gauge field, q is a coupling constant. The Lagrangian now writes as

$$\mathscr{L} = \underbrace{i\bar{\psi}\partial_\mu\gamma^\mu\psi}_{\text{kinetic term of }\psi} - \underbrace{m\bar{\psi}\psi}_{\text{mass term of }\psi} - \underbrace{q\bar{\psi}\gamma^\mu A_\mu\psi}_{\text{interaction term}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{kinetic term of }A_\mu}, \tag{2.11}$$

with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ being the field strength tensor. This is so far only considering a U(1) symmetry. We can identify A_μ with the photon field and q with an electric charge of a particle and have obtained a quantum field theory that considers only the electromagnetic interaction (quantum electrodynamics, QED). For the full SM Lagrangian, we have to consider the full group structure.

Now we look at the electroweak group structure $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$. With the same considerations as above we can introduce the covariant derivative

$$D_\mu^{\mathrm{EW}} = \partial_\mu - ig'\frac{Y_W}{2}B_\mu - ig\frac{\sigma_k}{2}W_\mu^k, \tag{2.12}$$

with B_μ, W_μ^k $(k = 1,2,3)$ being the gauge fields for the U(1) and SU(2) symmetry. g' and g are two different coupling constants, Y_W is the weak hypercharge and σ_k are the Pauli matrices, which are generators of the SU(2) group. The electroweak part of the Lagrangian is then

$$\mathscr{L}_{\mathrm{EW}} = i\bar{\psi}_f\gamma^\mu D_\mu^{\mathrm{EW}}\psi_f - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^k W^{\mu\nu,k}. \tag{2.13}$$

f runs over all the fermions. Left-handed fermions of the first generation take the form $l_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$ for leptons and $q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L$ for quarks, the other generations accordingly; this means that they are doublets in weak isospin, with $T = 1/2$ and the third component being $T_3 = +1/2$ for neutrinos and up-type quarks and $T_3 = -1/2$ for charged leptons and down-type quarks. The right handed fermions on the other hand have $T = 0$, which makes them singlets and take the form e_R, u_R and d_R. In the SM there are no right-handed neutrinos. $B_{\mu\nu}$ and $W_{\mu\nu}$ in Eq. 2.13 are field strength tensors and are defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.14}$$

$$W_{\mu\nu}^k = \partial_\mu W_\nu^k - \partial_\nu W_\mu^k + g\varepsilon_{klm}W_\mu^l W_\nu^m, \tag{2.15}$$

with ε_{klm} being the Levi-Civita symbol.

The strong interaction, which corresponds to the $\mathrm{SU}(3)_C$ group, can be treated in a similar way. The generators of the SU(3) group are the Gell-Mann matrices $\lambda_a/2$ [18]. The covariant derivative for the QCD sector thus becomes

$$D_\mu^{\mathrm{QCD}} = \partial_\mu - ig_S \frac{\lambda_a}{2} G_\mu^a, \tag{2.16}$$

with g_S being the strong coupling constant and G_μ^a ($a = 1, \ldots, 8$) the gluon gauge fields. With the field strength tensor

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_S f^{abc} G_\mu^b G_\nu^c, \tag{2.17}$$

f^{abc} being the structure constants of the SU(3) group, we can now write the full covariant derivative as

$$D_\mu = \partial_\mu - ig' \frac{Y_W}{2} B_\mu - ig \frac{\sigma_k}{2} W_{\mu,k} - ig_S \frac{\lambda_a}{2} G_\mu^a. \tag{2.18}$$

Considering everything discussed so far, the Lagrangian can be written as

$$\begin{aligned} \mathscr{L} &= \mathscr{L}_{\text{fermion}} \qquad + \qquad \mathscr{L}_{\text{gauge}} \\ &= i\bar{\psi}_f \gamma^\mu D_\mu \psi_f - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^k W^{\mu\nu,k} - \frac{1}{4} G_{\mu\nu}^a G^{\mu\nu,a}, \end{aligned} \tag{2.19}$$

where ψ_f runs through all left- and right-handed fermions. This would leave us so far with massless particles. Simply adding mass terms as in Eq. 2.11 of the form

$$m\bar{\psi}\psi = m(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) \tag{2.20}$$

would not result in a SU(2) invariant Lagrangian, since right-handed fermions are singlets, whereas left-handed fermions are doublets. Bosons with mass m would require additional terms in the Lagrangian of the form

$$\frac{m^2}{2} A_\mu A^\mu, \tag{2.21}$$

which would break gauge invariance. We know however from experiment that the $W^\pm$ and

Z^0 bosons actually have a mass, as well as fermions. This requires the introduction of the Brout-Englert-Higgs Mechanism.

2.1.4 The Brout-Englert-Higgs mechanism

This mechanism was proposed by three different groups [19–21]. The Brout-Englert-Higgs (BEH) mechanism introduces a new field, commonly known as Higgs field, which is an isospin doublet of two complex scalar fields:

$$\Phi = \begin{pmatrix} \Phi_1 + i\Phi_3 \\ \Phi_2 + i\Phi_4 \end{pmatrix} = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}. \tag{2.22}$$

Terms are added to the Lagrangian, that describe the interaction of Φ with the gauge fields:

$$\mathscr{L}_{\text{Higgs}} = \left(D^{\text{EW},\mu}\Phi\right)^\dagger \left(D^{\text{EW}}_\mu \Phi\right) - V(\Phi), \tag{2.23}$$

where the potential $V(\Phi)$ takes the form

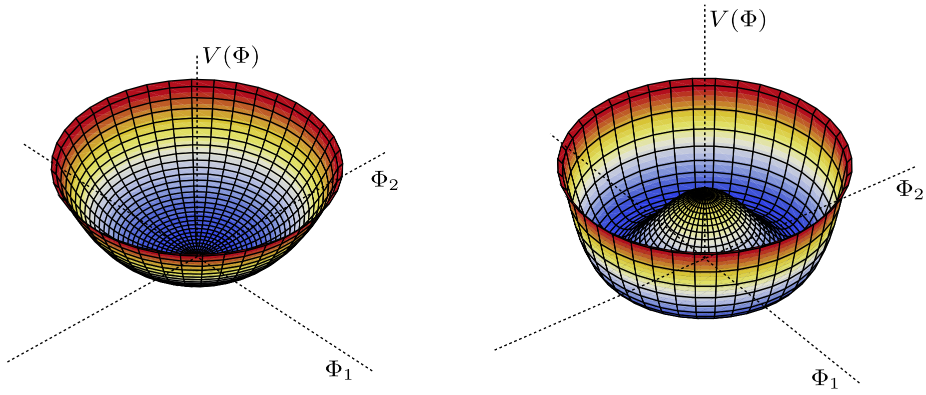
$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{2.24}$$

λ needs to be positive for a stable minimum. μ can in principle be a real or imaginary number, which leads to the two separate cases with μ^2 being negative or positive, with the crucial difference being the minimum. This is illustrated in Fig. 2.1 in the Φ_1-Φ_2 plane. With $\mu^2 < 0$, depicted in Fig. 2.1a, there is obviously just one distinct minimum at

$$\Phi = 0, \tag{2.25}$$

whereas with $\mu^2 > 0$ (the so-called “Mexican-hat potential” due to its appearance), depicted in Fig. 2.1b, the minimum calculates to

$$|\Phi|^2 = \frac{\mu^2}{2\lambda}. \tag{2.26}$$



(a) $\mu^2 < 0$ **(b)** $\mu^2 > 0$

Figure 2.1: The Higgs-Potential $V(\Phi)$ in the Φ_1-Φ_2 plane with different values for μ.

This means one gets infinite minima on a circle around $\Phi = 0$, the minimum is degenerate. One distinct minimum has to be chosen, but then this ground state is no longer invariant under U(1) transformation (rotation), which means that the U(1) symmetry is spontaneously broken. We can choose any minimum without loss of generality. Since the vacuum is not electrically charged, we can conveniently choose

$$\Phi_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}, \tag{2.27}$$

with the vacuum expectation value $v = \sqrt{\mu^2/\lambda}$. In this minimum, the Higgs Lagrangian (Eq. 2.23) now contains these terms:

$$\begin{aligned} \left|\left(-ig'\frac{Y_W}{2}B_\mu - ig\frac{\sigma_k}{2}W_{\mu,k}\right)\Phi_0\right|^2 = & \left(\frac{1}{2}vg\right)^2 W_\mu^+ W^{-,\mu} \\ & + \frac{1}{8}v^2\left(g^2 + g'^2\right) Z_\mu Z^\mu \\ & + 0\left(g^2 + g'^2\right) A_\mu A^\mu, \end{aligned} \tag{2.28}$$

where $W_\mu^\pm$, Z_μ and A_μ are linear combinations of the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ fields:

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp i W_\mu^2 \right), \tag{2.29}$$

$$Z_\mu = \cos(\theta_\mathrm{W}) W_\mu^3 - \sin(\theta_\mathrm{W}) B_\mu, \tag{2.30}$$

$$A_\mu = \sin(\theta_\mathrm{W}) W_\mu^3 + \cos(\theta_\mathrm{W}) B_\mu, \tag{2.31}$$

with the weak mixing angle (sometimes also called "Weinberg angle")

$$\theta_\mathrm{W} := \arctan \frac{g'}{g}. \tag{2.32}$$

We can see that the terms in Eq. 2.28 have the form of mass terms and thus obviate the need to plug in mass terms by hand. The resulting fields can be identified with the physical gauge bosons: $W_\mu^\pm$ are the $W^\pm$ bosons and have a mass of $m_W = 1/2 \cdot vg$, Z_μ is the Z^0 boson with the mass $m_Z = m_W / \cos\theta_\mathrm{W}$, and A_μ is the massless photon field. We can identify the electromagnetic coupling constant $q = e$ from Eq. 2.11 with $g \sin\theta_\mathrm{W}$.

Now the problem of massive gauge bosons is solved, but more happens when we have a closer look at the additional term in the Lagrangian introduced by the Higgs potential in Eq. 2.24. We can parametrise the field $\Phi(x)$ around the minimum:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.33}$$

The Lagrangian in Eq. 2.23 then writes as

$$\mathscr{L} = \frac{1}{2}(\partial_\mu h)^2 + \left[m_W^2 W_\mu^+ W^{-,\mu} + m_Z^2 Z_\mu Z^\mu \right] \left(1 + \frac{h}{v} \right)^2 - \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4. \tag{2.34}$$

So we have a field $h(x)$ with a mass term that results in the mass

$$m_H = \sqrt{2\lambda v^2}. \tag{2.35}$$

$h(x)$ hence describes a massive scalar particle that is called the Higgs boson. The other terms arising are the kinetic term $\frac{1}{2}(\partial_\mu h)^2$ and interaction terms, including interaction with the gauge bosons and self-interactions.

In order to describe the interaction of the Higgs field with fermions, further terms are introduced, that describe the Yukawa interaction:

$$\mathscr{L}_{\text{Yukawa}} = \underbrace{\lambda^e_{j,k}\bar{l}_{L,j}\Phi e_{R,k}}_{\mathscr{L}^e_{\text{Yukawa}}} + \underbrace{\lambda^d_{j,k}\bar{q}_{L,j}\Phi d_{R,k}}_{\mathscr{L}^d_{\text{Yukawa}}} + \underbrace{\lambda^u_{j,k}\bar{q}_{L,j}\Phi^c u_{R,k}}_{\mathscr{L}^u_{\text{Yukawa}}}. \tag{2.36}$$

The indices j and k run through the three generations. $\lambda_{j,k}$ are the Yukawa couplings and are 3×3 matrices. Φ^c is the Higgs charge conjugate $(\Phi^0, -\Phi^+)^\dagger$ and is needed to get the upper elements of the doublet. For leptons, the matrix can be diagonalised, leaving us with only the main-diagonal elements λ_e, λ_μ and λ_τ.

Choosing the minimum of the Higgs field, the leptonic part writes as

$$\mathscr{L}^e_{\text{Yukawa}} = \frac{\lambda^e_n v}{\sqrt{2}}\left(\bar{l}_{L,n}e_{R,n}\right) + \text{h.c.}, \tag{2.37}$$

resulting in lepton masses of

$$m_n = \frac{\lambda^e_n v}{\sqrt{2}}, \tag{2.38}$$

with $n = e, \mu, \tau$. Without right-handed neutrinos, left-handed neutrinos remain massless by this mechanism. The mass terms for quarks follow in a similar way. However, the matrices $\lambda^d_{j,l}$ and $\lambda^u_{j,l}$ are not diagonal in generation space. The masses of the quarks are obtained by diagonalizing the mass matrices

$$M^d = \frac{v}{\sqrt{2}}\lambda^d_{j,k}, M^u = \frac{v}{\sqrt{2}}\lambda^u_{j,k}. \tag{2.39}$$

This is done with four unitary matrices, $V^{u,d}_{L,R}$:

$$M^f_{\text{diag}} = V^f_L M^f V^{f\dagger}_R, f = u, d. \tag{2.40}$$

Since M^d and M^u are not diagonalisable simultaneously, the mass eigenstates are not the same as the interaction eigenstates. The rotation of basis is encoded in the Cabibbo–Kobayashi–Maskawa (CKM) matrix $V_{\mathrm{CKM}} = V_L^u V_R^{d\dagger}$:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.41}$$

d', s', b' represent the weak-interaction eigenstates, d, s, b are the mass eigenstates. The elements of the CKM matrix have been measured in different experiments and are known with various precision [14]:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97420 \pm 0.00021 & 0.2243 \pm 0.0005 & 0.00394 \pm 0.00036 \\ 0.218 \pm 0.004 & 0.997 \pm 0.017 & 0.0422 \pm 0.0008 \\ 0.0081 \pm 0.0005 & 0.0394 \pm 0.0023 & 1.019 \pm 0.025 \end{pmatrix}. \tag{2.42}$$

It should be noted that these experimental values do not consider theoretical restrictions. In order to conserve unitarity, $\left|V_{jk}\right| > 1$ is not possible for any of the elements.

Combining all our knowledge, with the terms from Eq. 2.19, 2.23 and 2.36, the full SM Lagrangian can now be written as

$$\mathscr{L}_{\mathrm{SM}} = \mathscr{L}_{\mathrm{fermion}} + \mathscr{L}_{\mathrm{gauge}} + \mathscr{L}_{\mathrm{Higgs}} + \mathscr{L}_{\mathrm{Yukawa}}. \tag{2.43}$$

2.2 Supersymmetric extensions

This thesis will focus on a search for supersymmetric particles, thus the understanding of the theory of Supersymmetry (SUSY) is vital. The concept of SUSY introduces a symmetry between fermions and bosons, which is not implemented in the SM. We introduce SUSY as an extension to the SM[6], meaning that all known particles and interactions are included, but an extended sector arises through the additional symmetry. This makes it a theory "beyond the Standard Model" (BSM). The information provided here is largely based on [22] and the lecture notes [23], which are based on the books [24, 25].

2.2.1 Motivation

While the SM is a very successful theory, it is clear that it cannot be a full description of the nature of the universe. The SM lacks a description of gravity. It fails to explain several empirical results, like the presence of dark matter and dark energy, or also the existence of neutrino masses (the latter can in principle be explained with a rather small extension). Also, the masses and mixing parameters, as well as the Higgs potential, are just set by hand without any explanation, which can be seen as unnatural. SUSY can give a direct explanation for some of these open questions. For some others SUSY can help indirectly. We will focus on the former here.

The Hierarchy Problem

The masses of particles that we have discussed in the Lagrangian terms in Sec. 2.1.3 and 2.1.4 so far are only tree-level predictions and are not equivalent to the masses given in Sec. 2.1.2. In order to obtain predictions with higher precision, radiative corrections must be taken into account. These are divergent if we let the momentum go up to infinity. We can use the *momentum cut-off regularisation*, by applying a cut-off Λ as an upper limit for the momenta. This parameter Λ characterises the energy scale up to which the theory is valid. This scale should be at least as high as the scale where new physics is expected, which is the latest at the Planck scale ($M_P \approx 2.4 \cdot 10^{18}\,\text{GeV}$), where corrections from quantum gravity are

[6]SUSY as a concept can also be used in other fields than elementary particle physics, e.g. condensed matter physics or dynamical systems; this will not be discussed further here.

expected. For fermion masses this calculation results (in a simplified model that considers only one fermion and a scalar) in

$$\delta m_F \propto -\lambda_F^2 m_F \ln \frac{\Lambda^2}{m_F^2} + \ldots \tag{2.44}$$

Since there is only a logarithmic dependency on Λ, these terms remain small. For the mass of the scalar Higgs one gets

$$\delta m_S^2 \propto -\lambda_F^2 \left[\Lambda^2 - m_F^2 \ln \frac{\Lambda^2}{m_F^2} \right] + \ldots \tag{2.45}$$

Here the divergence is not logarithmic but quadratic, the logarithmic terms can be disregarded. For the SM this means that the physical Higgs mass becomes

$$m^2_{H_{\mathrm{SM}}^{\mathrm{phys}}} \approx m^2_{H_{\mathrm{SM}}} + \frac{c}{16\pi^2}\Lambda^2, \tag{2.46}$$

where H_{SM} is the Higgs mass from Eq. 2.35. The coefficient c depends on the dimensionless couplings. Assuming no new physics up to the Planck scale would lead to extremely high correction terms, but even when assuming that the SM is only valid up to the scale of grand unification (GUT), one still gets terms of $\delta m = \mathcal{O}(10^{16}\,\mathrm{GeV})$. This is not per se problematic, because the Higgs mass parameter $m_{H_{\mathrm{SM}}}$ can be fine-tuned to cancel out the correction terms exactly such that we get a physical Higgs mass at the electroweak scale. However, the extreme precision needed to achieve this and the high dependency of low-energy physics on parameters of a high-energy theory is often seen as unnatural [26, 27].

With SUSY this can be prevented: the quadratic terms are completely cancelled out by the supersymmetric partners (this is true, if the coupling strength of the SUSY particles to the Higgs is the same as of their SM partners), because fermion and boson loops contribute with opposite signs; then only logarithmic terms remain:

$$\delta m^2 \propto \frac{\lambda^2}{8\pi^2}(m_F^2 - m_\Phi^2) \ln \frac{\Lambda^2}{m_F^2}, \tag{2.47}$$

with Φ being the scalar partner of the fermion F. These terms remain relatively small, if m_Φ is not too high.

Cold Dark Matter

Dark Matter (DM) has first been suggested by F. Zwicky as a solution to observed nebularum velocities in the Coma cluster [28] and has since been backed by several observations like the rotation curves of spiral galaxies [29], the weak gravitational lensing effect in the Bullet cluster [30] and the cosmic microwave spectrum [31]. The ΛCDM model, which is very successful in describing the large-scale evolution of the universe, assumes a certain amount of cold dark matter (CDM). The best fit so far is based on Planck measurements and results in a CDM abundance of $\Omega_{\text{CDM}}h^2 = 0.1188 \pm 0.0010$ [31], which corresponds to $25.89 \pm 0.75\,\%$ of the total energy density of the universe[7]. 'Cold' means here that it moves slowly compared to the speed of light, in contrast to 'Hot Dark Matter', which describes highly relativistic particles. There is no strict classification into these categories, sometimes 'warm' is also used for possible DM particles with intermediate speed. The only possible DM candidates in the SM are neutrinos, but due to their low mass they are classified as hot DM and cannot explain the needed amount of CDM.

There is a variety of candidates for what could make up CDM, but a very popular one is a Weakly Interacting Massive Particle (WIMP), a particle which is stable (or has a very long lifetime compared to the age of the universe) and does not carry an electric or colour charge. Under certain conditions SUSY can provide such a candidate.

Unification of coupling constants

As already seen in Sec. 2.1, the electromagnetic and weak interaction can be described with a unified theory. However there are still two different coupling constants. Looking at

[7] Calculated with the reduced Hubble constant $h = H_0/(100\,\text{kms}^{-1}\text{Mpc}^{-1})$.

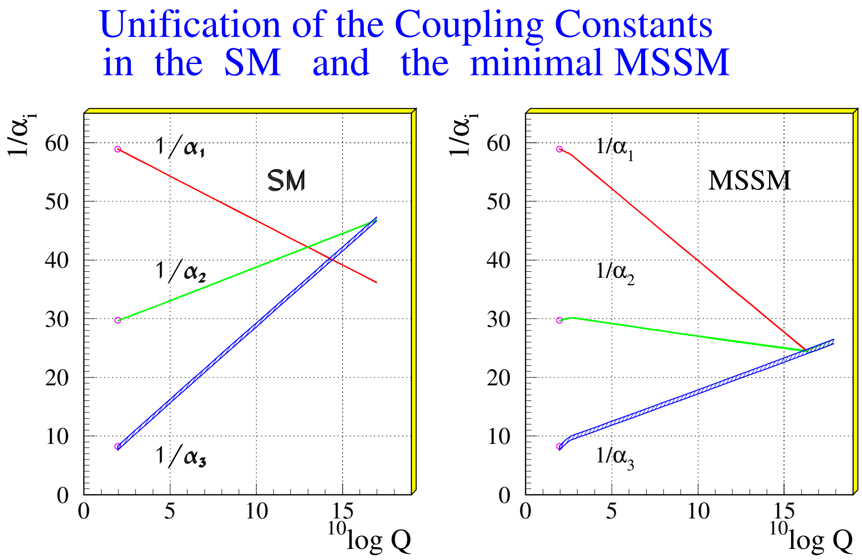


Figure 2.2: Running of the inverse $\mathrm{U}(1)_Y$ (α_1), $\mathrm{SU}(2)_L$ (α_2) and $\mathrm{SU}(3)_C$ (α_3) coupling constants versus the logarithm of the energy in GeV. In the SM case (left) a unification is not achieved. In the case of the Minimal Supersymmetric Standard Model (MSSM) with SUSY particles becoming relevant at 1 TeV the three couplings meet at a common point. [33]

the energy-dependent runnings of these two couplings and also the coupling of the strong interaction, one can suspect that there could be a unification of the three couplings at a high energy, meaning that they all meet at a common point. Such a unification is a desirable feature of a theory, and theories that achieve this are called *Grand Unified Theories* (GUTs). In the SM it can be shown that the three gauge couplings come close but do not actually meet at a common intersection. When including SUSY, further radiative corrections are added at a mass scale where SUSY particles become relevant. This leads to a unification, if the SUSY scale is at the order of $\sim 1\,\mathrm{TeV}$ [32]. This is depicted in Fig. 2.2.

2.2.2 General structure

Supersymmetry is what we obtain, when we extend our theory with one or more spin-1/2 generators. This was not thought to be possible because of the Coleman-Mandula theorem, which states that the restriction for combining space-time and internal symmetries is, that the conserved quantities – apart from generators of the Poincaré group – must be Lorentz scalars [34]. This theorem considers only bosonic generators. The Haag–Łopuszański–Sohnius theorem generalized later, that spin-$1/2$ operators are a possible extension (spin-$3/2$ or higher is ruled out) [35]. This Lie superalgebra is then the maximal allowed algebra that is consistent with relativistic QFT. Such a SUSY operator Q would work like this, when applied:

$$Q\,|\mathrm{Boson}\rangle \propto |\mathrm{Fermion}\rangle \quad \text{and} \quad Q\,|\mathrm{Fermion}\rangle \propto |\mathrm{Boson}\rangle\,. \tag{2.48}$$

In the simplest case (which is the only case we will discuss here), there is exactly one such operator. This would mean, that every particle has a supersymmetric partner with a spin difference of $1/2$. Going through the particles of the SM one finds that there are no particles that can be such partners of each other. Introducing the operator Q to the SM thus means introducing one new particle with spin $s - 1/2$ for each SM particle with spin s. Together they form a supermultiplet. The quantum numbers of particles are all the same as for their superpartners except for the spin.

The SM fermions and their partners – denoted with a preceding 's' (for 'scalar'), e.g. the partner of the electron is called 'selectron' – form chiral supermultiplets, which contain a Weyl spinor and a complex scalar field. Their general form is

$$\Phi(y,\theta) = \varphi(y) + \sqrt{2}\theta\psi(y) + (\theta\theta)F(y). \tag{2.49}$$

$y^\mu := x^\mu + i\theta\sigma^\mu\theta^\dagger$ is a complex coordinate in superspace; x^μ are the four space-time coordinates, the four fermionic coordinates are arranged in the two Weyl spinors θ, $\theta^\dagger$. The complex scalar field φ describes the sleptons and squarks, the left-handed Weyl spinor field ψ describes the leptons and quarks. The complex scalar field F is an auxiliary field[8]. The Higgs bosons and their partners are also contained in chiral superfields. Antichiral fields are written as

$$\bar{\Phi}(\bar{y},\theta^\dagger) = \varphi^*(\bar{y}) + \sqrt{2}\theta^\dagger\psi^\dagger(\bar{y}) + (\theta^\dagger\theta^\dagger)F^*(\bar{y}), \tag{2.50}$$

[8] Auxiliary fields do not have a kinetic term; they are mere book-keeping devices to make sure that the algebra closes off-shell.

with $\bar{y}^\mu := x^\mu - i\theta\sigma^\mu\theta^\dagger = (y^\mu)^*$. σ^μ is the four-vector of Pauli matrices: $\sigma^\mu = \begin{pmatrix} \mathbb{1} & \sigma^1 & \sigma^2 & \sigma^3 \end{pmatrix}$ and equivalently $\tilde{\sigma}^\mu = \begin{pmatrix} \mathbb{1} & -\sigma^1 & -\sigma^2 & -\sigma^3 \end{pmatrix}$. The Lagrangian of such a superfield can be written as

$$\begin{aligned} \mathscr{L} =& (\partial_\mu \varphi^*)(\partial^\mu \varphi) - \frac{i}{2}(\psi^\dagger \tilde{\sigma}^\mu \partial_\mu \psi + \psi \sigma^\mu \partial_\mu \psi^\dagger) \\ &+ \frac{m}{2}(\psi\psi + \psi^\dagger\psi^\dagger) + g(\varphi\psi\psi + \varphi^*\psi^\dagger\psi^\dagger \\ &+ F^*F - (m\varphi + g\varphi^2)F - (m\varphi^* + g\varphi^{*2})F^*. \end{aligned} \tag{2.51}$$

It can also be written without auxiliary fields:

$$\begin{aligned} \mathscr{L} =& (\partial_\mu \varphi^*)(\partial^\mu \varphi) - m^2|\varphi|^2 \\ &- \frac{i}{2}(\psi^\dagger \tilde{\sigma}^\mu \partial_\mu \psi + \psi \sigma^\mu \partial_\mu \psi^\dagger) + \frac{m}{2}(\psi\psi + \psi^\dagger\psi^\dagger) \\ &+ g(\varphi\psi\psi + \varphi^*\psi^\dagger\psi^\dagger) - mg|\varphi|^2(\varphi + \varphi^*) - g^2|\varphi|^4. \end{aligned} \tag{2.52}$$

The first line describes the free bosonic field $\varphi(x)$ with mass m, the second line the free fermionic field $\psi(x)$ with the same mass m. The last line contains the interaction terms; the first two terms are the Yukawa terms with coupling constant g, the last two describe the self-couplings of the bosonic field.

The Lagrangian of the chiral superfields can be divided into the kinetic term and the superpotential

$$\mathscr{L} = \mathscr{L}_{\text{kin}} - [W(\Psi)|_{\theta\theta, y\to x} + \text{h.c.}]; \tag{2.53}$$

the superpotential $W(\Psi)$ is a holomorphic function that contains all the mass and interaction terms. $\theta\theta$ behind the vertical line means that this is the highest order component (in contrast to vector superfields with the highest order $(\theta\theta)(\theta^\dagger\theta^\dagger)$); $y \to x$ means that we transform into the x coordinate by replacing $y = x + i\theta\sigma\theta^\dagger$. We can write the potential as

$$W(\Phi) = \frac{m}{2}\Phi^2 + \frac{g}{3}\Phi^3. \tag{2.54}$$

Developing it as a function of φ and ψ we get

$$W(\Phi)|_{\theta\theta} + \text{h.c.} = (m\varphi + g\varphi^2)F - \frac{1}{2}(m + 2g\varphi + 2g\varphi)(\psi\psi), \tag{2.55}$$

which results in the Lagrangian

$$\mathscr{L}_W = \frac{m}{2}(\psi\psi + \psi^\dagger\psi^\dagger) + g(\varphi\psi\psi + \varphi^*\psi^\dagger\psi^\dagger) - (m\varphi + g\varphi^2)F - (m\varphi^* + g\varphi^{*2})F^*. \tag{2.56}$$

The gauge bosons and their partners – denoted by a succeeding "-ino" – are contained in vector multiplets (a spin-1 field and a Weyl spinor). They are usually expressed with the help of auxiliary fields; they can be reduced by chosen a convenient gauge. Here they are given in the Wess-Zumino gauge as a function of space-time x^μ and the Weyl spinors θ and $\theta^\dagger$:

$$V_{\mathbf{WZ}}(x,\theta,\theta^\dagger) = (\theta\sigma^\mu\theta^\dagger)A_\mu + i(\theta\theta)(\theta^\dagger\lambda^\dagger) - i(\theta^\dagger\theta^\dagger)(\theta\lambda) + \frac{1}{2}(\theta\theta)(\theta^\dagger\theta^\dagger)D \tag{2.57}$$

D,λ and A_μ are the component fields, D is a real scalar auxiliary field, λ a Weyl spinor and A_μ a vector field. The Lagrangian then writes as

$$\mathscr{L}_V = -\frac{i}{2}(\lambda\sigma^\mu\partial_\mu\lambda^\dagger + \lambda^\dagger\tilde{\sigma}^\mu\partial_\mu\lambda) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}D^2 \tag{2.58}$$

Using the covariant derivative $D_\mu = \partial_\mu - {}^i\!/\!{}_2 \cdot qA_\mu$, the kinetic term (in its gauge invariant form) is

$$\begin{aligned}\mathscr{L}_{\text{kin}} =& (D_\mu\varphi)^*(D^\mu\varphi) - \frac{i}{2}(\psi^\dagger\tilde{\sigma}^\mu D_\mu\psi + \psi\sigma^\mu D^*_\mu\psi^\dagger) \\ &+ |F|^2 + \frac{i}{\sqrt{2}}q[\varphi^*(\psi\lambda) - \varphi(\psi^\dagger\lambda^\dagger)] + \frac{1}{2}qD|\varphi|^2\end{aligned} \tag{2.59}$$

It is not intrinsic to SUSY, but if we use Supersymmetry to describe the nature of elementary particle, we need an additional part to the Lagrangian to break the symmetry. This is because if there was no such breaking mechanism the supersymmetric particles would all have the same mass as their partners. Since this is known from experiment not to be the case, SUSY can only be introduced as a broken symmetry. Such SUSY-breaking terms should only lead to logarithmic divergences (*Soft SUSY breaking* [36]) and have lots of free parameters in their most general form, without assuming a specific mechanism. Such a SUSY-breaking Lagrangian would be of the form

$$\begin{aligned}
\mathscr{L}_{\text{soft}} = &-\frac{1}{2}M_i\bar{\lambda}_i\lambda_i && \text{for the gaugino masses} \\
&-m_{\tilde{f}}^2|\tilde{f}|^2+\dots && \text{for the sfermion and Higgs masses} \\
&-W_2(\varphi)-W_3(\varphi)+\text{h.c.} && \text{bilinear/trilinear terms and tadpole terms}
\end{aligned} \tag{2.60}$$

The breaking might happen in a "hidden sector" and be communicated via messenger particles. Examples for specific breaking mechanisms are the gravity mediated, the gauge mediated and the anomaly mediated SUSY breaking.

2.2.3 The MSSM

The Minimal Supersymmetric Standard Model (MSSM) [37,38] is the supersymmetric model that describes all known particles and interactions of the current Standard Model and extends it with as few additional particles as possible. The gauge groups are the same as in the SM. All SM particles are described within superfields (see Sec. 2.2.2), which basically doubles the particle content. The Higgs sector needs special attention: since the superpotential is holomorphic, a right-handed field like the charge conjugate in Eq. 2.36, which is needed in the SM to give mass to up-type particles, is not allowed in the superpotential. In order to give mass to both up- and down-type fermions, an extended Higgs sector is needed. The minimal solution, which is used in the MSSM, are two left-handed superfields, with a hypercharge of $Y=-1$ and $Y=+1$, respectively. This doubles the degrees of freedom in the Higgs sector, so together with the doubling done by the superfields the degrees of freedom in the Higgs sector are quadrupled with respect to the SM, in contrast to everything else, where they are just doubled. The ratio of the vacuum expectation values of the two Higgs doublets is quantified by the angle β:

$$\tan\beta = \frac{\langle H_u^0\rangle}{\langle H_d^0\rangle}. \tag{2.61}$$

All superfields in the MSSM are summarized in Tab. 2.4.

Table 2.4: The superfields in the MSSM and their behaviour under gauge transformation (e.g. **1** means singlet, **2** means doublet, ...). For the $\mathrm{U}(1)_Y$ group the quantum number (hypercharge Y_W) is given. For the quark and lepton sector, only the first generation is shown, the others are equivalent.

Superfield	$\mathrm{SU}(3)_C$	$\mathrm{SU}(2)_L$	$\mathrm{U}(1)_Y$	Content
$\hat{L} = \begin{pmatrix} \hat{\nu}_{eL} \\ \hat{e}_L \end{pmatrix}$	**1**	**2**	-1	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \tilde{\nu}_{eL} \\ \tilde{e}_L \end{pmatrix}$
$\hat{\bar{E}}$	**1**	**1**	2	$e_R^\dagger, \tilde{e}_R^*$
$\hat{Q} = \begin{pmatrix} \hat{u}_L \\ \hat{d}_L \end{pmatrix}$	**3**	**2**	$\frac{1}{3}$	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} \tilde{u}_L \\ \tilde{d}_L \end{pmatrix}$
$\hat{\bar{U}}$	**3**	**1**	$-\frac{4}{3}$	$u_R^\dagger, \tilde{u}_R^*$
$\hat{\bar{D}}$	**3**	**1**	$\frac{2}{3}$	$d_R^\dagger, \tilde{d}_R^*$
$\hat{H}_d = \begin{pmatrix} \hat{H}_1^0 \\ \hat{H}_1^- \end{pmatrix}$	**1**	**2**	-1	$\begin{pmatrix} h_1^0 \\ h_1^- \end{pmatrix}, \begin{pmatrix} \tilde{h}_1^0 \\ \tilde{h}_1^- \end{pmatrix}$
$\hat{H}_u = \begin{pmatrix} \hat{H}_2^+ \\ \hat{H}_2^0 \end{pmatrix}$	**1**	**2**	1	$\begin{pmatrix} h_2^+ \\ h_2^0 \end{pmatrix}, \begin{pmatrix} \tilde{h}_2^+ \\ \tilde{h}_2^0 \end{pmatrix}$
$\hat{G}^j$	**8**	**1**	0	$G_j, \tilde{g}_j$
$\hat{W}^k$	**1**	**3**	0	$W_k, \tilde{w}_k$
$\hat{B}$	**1**	**1**	0	$B, \tilde{b}$

The Lagrangian of the MSSM can be divided into

$$\mathscr{L}_{\mathrm{MSSM}} = \mathscr{L}_{\mathrm{gauge}} + \mathscr{L}_{\mathrm{matter}} + \mathscr{L}_{\mathrm{D}} + \mathscr{L}_{\mathrm{W}} + \mathscr{L}_{\mathrm{soft}}. \tag{2.62}$$

All D parts from Eq. 2.59 and 2.58 were moved to $\mathscr{L}_{\mathrm{D}}$, all F parts to the superpotential part $\mathscr{L}_{\mathrm{W}}$. The full gauge part is

$$\begin{aligned} \mathscr{L}_{\mathrm{gauge}} = & -\frac{1}{4} G^{j\mu\nu} G^b_{\mu\nu} - \frac{1}{4} W^{k\mu\nu} W^a_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} \\ & + \mathrm{Tr}\left[\tilde{g}^\dagger i D^\mu \gamma_\mu \tilde{g}\right] + \mathrm{Tr}\left[\tilde{w}^\dagger i D^\mu \gamma_\mu \tilde{w}\right] + \tilde{b}^\dagger i \partial^\mu \gamma_\mu \tilde{b}; \end{aligned} \tag{2.63}$$

$j = 1 \ldots 8$, $k = 1 \ldots 3$. The gauginos are here in the form of Dirac spinors. The matter part is

$$\mathscr{L}_{\mathrm{matter}} = \sum_k \bar{\psi}_k i D^\mu \gamma_\mu \psi_k + \sum_l \left|D_\mu \phi_l\right|^2 + i \sum_{k,l,m} \frac{g_{V^m}}{\sqrt{2}} [\bar{\psi}_{Lk} T^m \tilde{V}^m \phi_l - \tilde{V}^{\dagger m} T^m \psi_{Lk} \phi_l^*]; \tag{2.64}$$

k runs over the fermions f and higgsinos $\tilde{h}_i$, l over the sfermions $\tilde{f}$ and Higgs bosons h_i; V^m are the vector bosons with the corresponding coupling constant g_{V^m} and generator of the gauge group T^m. ψ and $\tilde{\nu}$ are 4-component spinors.

The D part is

$$\mathscr{L}_{\mathrm{D}} = -\frac{1}{2}\sum_m \left|D_m^V\right|^2 \quad \text{, with} \quad D_m^V = -g_{V^m}\sum_l \phi_l^* T^m \phi_l. \tag{2.65}$$

The superpotential part of the Lagrangian is

$$\mathscr{L}_{\mathrm{W}} = -\sum_l \left|\frac{\partial W}{\partial \phi_l}\right|^2 - \left(\frac{1}{2}\sum_{kl} \bar{\Psi}_{kL} \frac{\partial^2 W}{\partial \phi_k \partial \phi_l} \psi_{lL} + \text{h.c.}\right). \tag{2.66}$$

The superpotential W can be divided into to parts $W = W_R + W_{\not{R}}$. The first part is

$$W_R = -\sum_{i,j=1}^{2} \varepsilon_{ij}\mu \hat{H}_1^i \hat{H}_2^j + \sum_{i,j=1}^{2}\sum_{r,s=1}^{2} \varepsilon_{ij}[\lambda_{rs}^e \hat{H}_1^i \hat{L}_{jr} \hat{\bar{E}}_s + \lambda_{rs}^d \hat{H}_1^i \hat{Q}_{jr} \hat{\bar{D}}_s + \lambda_{rs}^u \hat{H}_2^i \hat{Q}_{jr} \hat{\bar{U}}_s], \tag{2.67}$$

ε_{ij} is the Levi-Civita symbol in two dimensions. i and j are the $SU(2)$-doublet indices. r and s are generation numbers, λ_{rs} are the corresponding elements of the Yukawa coupling matrix. The second part is

$$W_{\not{R}} = \sum_{i,j=1}^{2}\sum_{r=1}^{3} \mu_r H_u^i L_{jr} + \sum_{i,j=1}^{2}\sum_{r,s,t=1}^{3} \varepsilon_{ij}[\lambda_{rst} \hat{L}_{ir} \hat{L}_{js} \hat{\bar{E}}_t + \lambda'_{rst} \hat{L}_{ir} \hat{Q}_{js} \hat{\bar{D}}_t] + \lambda''_{rst} \hat{\bar{U}}_r \hat{\bar{D}}_s \hat{\bar{D}}_t \tag{2.68}$$

This second part introduces terms that would lead to processes which change the lepton and baryon number, e.g. the decay of protons. Such processes have not been observed in nature so far and can thus be assumed to either not exist or have a high mean life time compared to the age of the universe [39,40]. This part of the Lagrangian can be set to zero by postulating conservation of R-parity [41]. While R-parity violating models are not strictly forbidden, only constrained, in this thesis only R-parity conserving SUSY is considered. R-parity is defined as

$$P_R = (-1)^{3(B-L)+2s}, \tag{2.69}$$

B is the baryon number, L the lepton number and s the spin number of a particle. This means that all SM particles have $P_R = 1$ and all their superpartners have $P_R = -1$. This has some major impacts on SUSY phenomenology: It means that SUSY particles can only be produced in pairs. Also it means that in the decay of a SUSY particle an odd number of SUSY particles must be produced, resulting in a stable lightest supersymmetric particle (LSP). The LSP – if it does not carry any charges – is a candidate for CDM [42, 43], giving us another motivation to look for R-parity conserving SUSY.

Finally, the soft SUSY breaking part in its most general form is

$$\begin{aligned}
\mathscr{L}_{\text{soft}} = &-\frac{1}{2}\sum_{k=1}^{3} M_k \tilde{\lambda}_k^a \tilde{\lambda}_k^a - m_1^2|H_1|^2 - m_2^2|H_2|^2 + B\mu\varepsilon_{ij}(H_1^i H_2^j + \text{h.c.}) \\
&+ \sum_{r,s=1}^{3}\Big[-M_{\tilde{q}_{rs}}^2(\tilde{u}_{Lr}^\dagger \tilde{u}_{Ls} + \tilde{d}_{Lr}^\dagger \tilde{d}_{Ls}) - M_{\tilde{u}_{rs}}^2 \tilde{u}_{Rr}^\dagger \tilde{u}_{Rs} - M_{\tilde{d}_{rs}}^2 \tilde{d}_{Rr}^\dagger \tilde{d}_{Rs} + (\text{lept.}) \\
&+ \varepsilon_{ij}\Big(a_{Drs} H_1^i \tilde{q}_r^j \tilde{d}_{Rs}^\dagger + a_{Urs} H_2^i \tilde{q}_r^j \tilde{u}_{Rs}^\dagger + (\text{lept.}) + \text{h.c.}\Big)\Big].
\end{aligned} \tag{2.70}$$

"lept." stands for the equivalent leptonic terms. M_k are three real gaugino-mass parameters; $\tilde{\lambda}_k^a$ stand for the gauginos $\tilde{B}$, $\tilde{W}^a$ and $\tilde{g}^a$; m_1 and m_2 are mass terms for the Higgs doublets; $M_{\tilde{Q}}, \dots$ are 3×3 matrices in generation space that lead to non-vanishing off-diagonal terms in the mass matrices; the complex $a_D, \dots$ are trilinear coupling matrices leading to flavour-violating processes.

In this most general form, the MSSM has more than 100 free parameters. They can be reduced, when assuming a specific SUSY breaking mechanism or taking experimental constraints into account.

Neutralinos, Charginos and Stops

In the MSSM there are four colour- and electric-neutral spin-1/2 sparticles – the bino $\tilde{B}$, a wino $\tilde{W}^0$ and the higgsinos $\tilde{H}_d^0$ and $\tilde{H}_u^0$ – which have all the same quantum numbers, which means that the mass eigenstates – called *neutralinos* – can be a mix of the gauge eigenstates. The mass matrix gets contributions from the superpotential, from the SUSY-breaking gaugino masses and from electroweak symmetry breaking. In the basis of $(\tilde{B}, \tilde{W}^0, \tilde{H}_d^0, \tilde{H}_u^0)$ it is written as:

$$M_{\tilde{\chi}^0} = \begin{pmatrix} M_1 & 0 & -c_\beta s_W m_Z & s_\beta s_W m_Z \\ 0 & M_2 & c_\beta c_W m_Z & -s_\beta c_W m_Z \\ -c_\beta s_W m_Z & c_\beta c_W m_Z & 0 & -\mu \\ s_\beta s_W m_Z & -s_\beta c_W m_Z & -\mu & 0 \end{pmatrix}, \qquad (2.71)$$

using the short forms $s_\beta = \sin\beta$, $c_\beta = \cos\beta$, $s_W = \sin\theta_W$, $c_W = \cos\theta_W$. The mass terms – which are rather long – can be obtained by diagonalisation. For the lightest neutralino, which is usually assumed to be the LSP, one gets the effect that it is gaugino-like for $|\mu| \gg M_1, M_2$ and higgsino-like for $|\mu| \ll M_1, M_2$ (which leads also to $m_{\tilde{\chi}_1^0} \sim m_{\tilde{\chi}_2^0}$).

Similarly we get *charginos* from the mixing of the charged higgsinos and winos. The mixing matrix in the basis $(\tilde{W}^+, \tilde{H}_u^+, \tilde{W}^-, \tilde{H}_d^-)$ is:

$$M_{\tilde{\chi}^\pm} = \begin{pmatrix} 0 & 0 & M_2 & \sqrt{2} c_\beta m_W \\ 0 & 0 & \sqrt{2} s_\beta m_W & \mu \\ M_2 & \sqrt{2} s_\beta m_W & 0 & 0 \\ \sqrt{2} s_\beta m_W & \mu & 0 & 0 \end{pmatrix}. \qquad (2.72)$$

Since this matrix can be divided into two equivalent 2×2 matrices (for positive and negative charge) and all other elements are 0, the diagonalisation is simpler than for neutralinos and one gets

$$m^2_{\chi_1^\pm} = \frac{1}{2}\left[|M_2|^2 + |\mu|^2 + 2m_W^2 - \sqrt{(|M_2|^2 + |\mu|^2 + 2m_W^2)^2 - 4\left|\mu M_2 - m_W^2 \sin 2\beta\right|^2}\right], \qquad (2.73)$$

$$m^2_{\chi_2^\pm} = \frac{1}{2}\left[|M_2|^2 + |\mu|^2 + 2m_W^2 + \sqrt{(|M_2|^2 + |\mu|^2 + 2m_W^2)^2 - 4\left|\mu M_2 - m_W^2 \sin 2\beta\right|^2}\right]. \qquad (2.74)$$

In the limit $|\mu| \gg M_2$ one gets a gaugino-like $\chi_1^\pm$ and a higgsino-like $\chi_2^\pm$, and vice versa for $|\mu| \ll M_2$, for the case $|\mu| \ll M_1, M_2$ one gets $m_{\tilde{\chi}_1^0} \sim m_{\tilde{\chi}_2^0} \sim m_{\tilde{\chi}_1^\pm}$.

Now we have a look at stops, which are the main focus of this thesis. Other squarks and sleptons work similarly, but will not be discussed here. There are contributions from the superpotential, from soft SUSY breaking and from the D part of the Lagrangian. In principle there can be a mixing of the left- and right-handed squarks of all three generations, so one would have to diagonalise a 6×6 matrix in the basis $(\tilde{u}_L, \tilde{c}_L, \tilde{t}_L, \tilde{u}_R, \tilde{c}_R, \tilde{t}_R)$. Usually it is assumed that the flavour changing mixing angles are negligible (minimal flavour violation). With this simplification we get the 2×2 mass matrix in the basis $(\tilde{t}_L, \tilde{t}_R)$:

$$M_{\tilde{t}}^2 = \begin{pmatrix} m_{\tilde{t}_L}^2 + m_t^2 + D(\tilde{t}_L) & m_t(-A_t + \mu \cot\beta) \\ m_t(-A_t + \mu \cot\beta) & m_{\tilde{t}_R}^2 + m_t^2 + D(\tilde{t}_R) \end{pmatrix} \tag{2.75}$$

The eigenvalues are:

$$\begin{aligned} m_{\tilde{t}_{1,2}}^2 = & \frac{1}{2}(m_{\tilde{t}_L}^2 + m_{\tilde{t}_R}^2) + \frac{1}{4} m_Z^2 \cos 2\beta + m_t^2 \\ & \mp \left\{ \left[\frac{1}{2}(m_{\tilde{t}_L}^2 - m_{\tilde{t}_R}^2) + m_Z^2 \cos 2\beta \left(\frac{1}{4} - \frac{2}{3} \sin^2 \theta_W \right) \right]^2 + m_t^2 (\mu \cot\beta - A_t)^2 \right\}^{1/2} . \end{aligned} \tag{2.76}$$

The left-handed soft mass term for the stop is the same as for the left-handed sbottom:

$$m_{\tilde{t}_L} = m_{\tilde{b}_L} = m_{\tilde{q}_3}. \tag{2.77}$$

Since m_t appears in the last term in Eq. 2.76 we can see that the mass splitting is increased for top squarks compared to the other squarks where the appearing m_q is much smaller. This makes the lighter mass eigenstate $\tilde{t}_1$ the lightest squark in large areas of the MSSM parameter space.

2.2.4 Top-Squark phenomenology at the LHC

Top squarks can be produced in pairs via proton-proton collision at the LHC under the condition that the sum of the two masses does not exceed the centre-of-mass energy. The most contributing channels are shown in Fig. 2.3 [44]. The production cross-section at a given centre-of-mass energy depends at leading order (LO) only on the top-squark mass. At

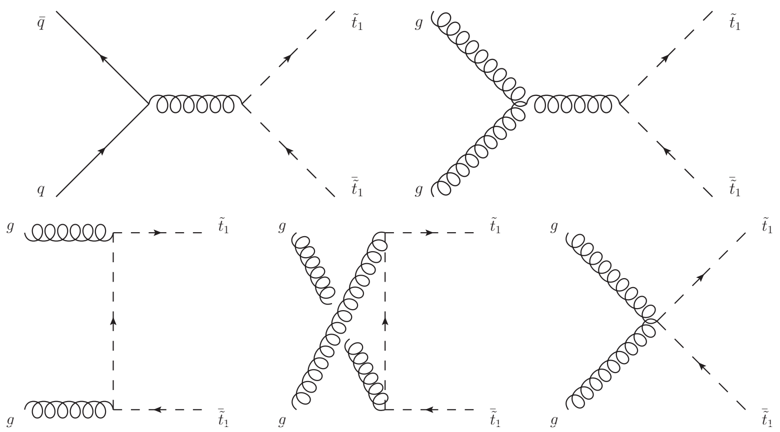
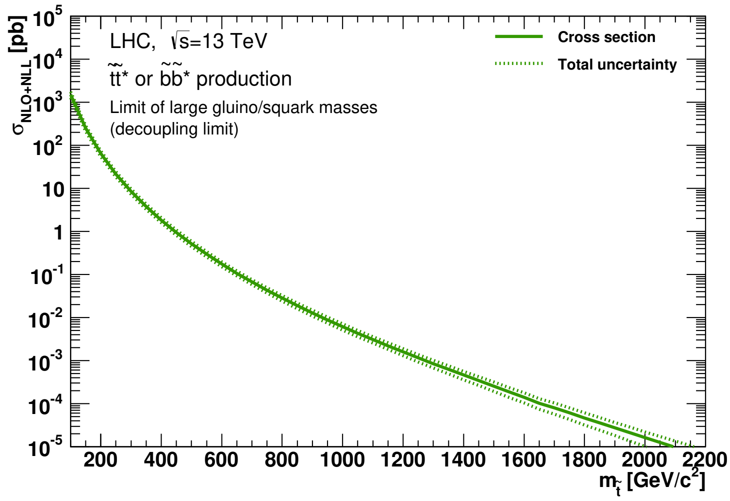
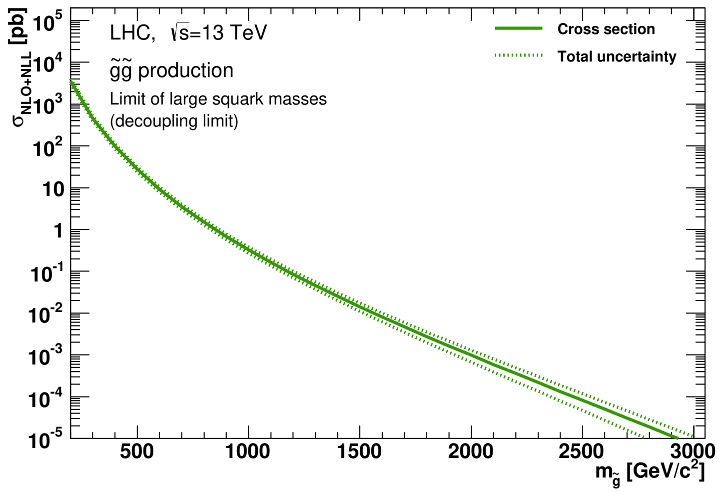


Figure 2.3: Leading-order Feynman diagrams for the production of top-squark pairs at proton-proton collisions.

next-to-leading order (NLO) other parameters can play a role but have only little numerical impact [45]. This is different from lower-generation squarks, where the $q\bar{q} \to \tilde{q}\bar{\tilde{q}}$ t-channel production with the exchange of a gluino (and thus the value of the gluino mass) plays a big role. The decay channels however depend largely on SUSY parameters, in particular on the sparticles with a lower mass. It is thus often useful to look at decoupled scenarios, where all sparticles except the $\tilde{t}_1$ and the $\tilde{\chi}_1^0$ (which would be the LSP) have a much higher mass and thus have negligible impact on the decay. The production cross-section calculated with this assumption is displayed in Fig. 2.4a (in non-decoupled scenarios it would be almost identical). In comparison, the cross section for the production of a gluino pair – depicted in Fig. 2.4b – is higher for the same sparticle mass. The dominant decay channel in this case depends on the mass difference $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$. In case of $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) \geq m(t)$ the dominant decay channel is $\tilde{t}_1 \to t\tilde{\chi}_1^0$. In the case of $m(W) < \Delta m(\tilde{t}_1, \tilde{\chi}_1^0) < m(t)$ an on-shell top cannot be produced and the 3-body decay $\tilde{t}_1 \to Wb\tilde{\chi}_1^0$ becomes the dominant decay mode. Finally in the case $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) < m(W)$ also the W is produced off-shell and the 4-body decay $\tilde{t}_1 \to bf\bar{f}\tilde{\chi}_1^0$ occurs, with f being a fermion. If further sparticles have a mass below the stop mass, they must be considered in competing decay channels. Examples are decays via an additional neutralino $\tilde{\chi}_2^0$ or a chargino $\tilde{\chi}_1^{\pm}$ which can then decay into a $\tilde{\chi}_1^0$ and either a W, Z or Higgs boson. These six decay channels are depicted in Fig. 2.5. The branching ratios of these processes can be different in each SUSY model, they depend on the $\tilde{t}_L$-$\tilde{t}_R$ mixing of the $\tilde{t}_1$



(a) Stop pair



(b) Gluino pair

Figure 2.4: Production cross-sections for pairs of stops or gluinos at $\sqrt{s} = 13\,\text{TeV}$ as a function of the sparticle mass for decoupled scenarios calculated at NLO+NLL [45]. In the decoupled scenario the calculation for stops can also be applied for sbottoms, the cross-section then depends only on the sbottom mass.

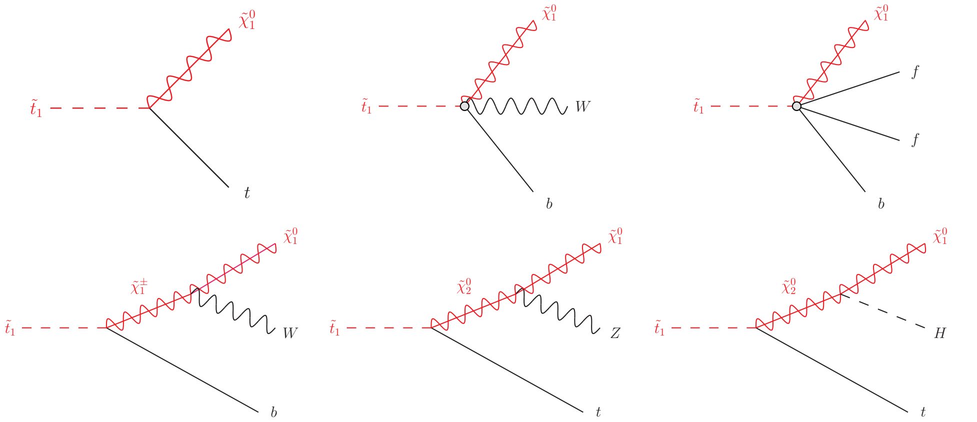


Figure 2.5: Decay modes of $\tilde{t}_1$. The decay modes in the first row are the only ones allowed if all sparticles except $\tilde{t}_1$ and $\tilde{\chi}_1^0$ are decoupled. The decay modes in the second row can occur if other sparticles are light enough. SM particles with $P_R = 1$ are drawn as black lines, sparticles with $P_R = -1$ as red lines. The circles symbolise 'effective' vertices, meaning that only virtual particles (off-shell) particles (top and or W) are created within them, all real particles are drawn as lines.

and the bino-wino-higgsino mixing of the gauginos.

W bosons decay dominantly into (first and second generation) quarks, the $W \to q\bar{q}$ branching ratio is $67.41 \pm 0.27\,\%$, the remainder goes into a lepton and its corresponding neutrino. Top quarks have a very short lifetime, hence in contrast to other quarks they do not hadronise but decay immediately, almost exclusively – due to the high value of $|V_{tb}|$ in the CKM matrix (see Eq. 2.42) – into a bottom quark and a W boson. The fully hadronic decay $t \to q\bar{q}b$ has a branching ratio of $66.5 \pm 1.4\,\%$ [14].

3 Experiment

The search for supersymmetric particles requires the conditions to produce them and a detector capable of reliably measuring their decay products. The Large Hadron Collider (LHC) is at the time of writing the most powerful particle collider in terms of beam energy and luminosity and hence opens the possibility to find new physics at energy scales that were not explored before. There are two general-purpose detectors at the LHC which are both capable of conducting measurements at comparable precisions to allow SUSY searches: ATLAS and CMS. For the work in this thesis, data taken by the ATLAS detector was used. A summary of the experimental set-up of the LHC (Sec. 3.1) and ATLAS (Sec. 3.2) is given in this chapter.

3.1 The LHC machine

The Large Hadron Collider (LHC) [46, 47] is a project that was approved by the CERN collaboration in December 1994, succeeding the proton-antiproton collider Tevatron at the Fermi National Accelerator Laboratory as the world's most powerful particle collider. The main purpose is to accelerate protons, but also heavy ions are used in special runs.

The LHC is located in the Franco-Swiss border region near Geneva, built 50 to 175 m underground in the 26.7 km long tunnel of the former Large Electron-Positron Collider (LEP). There are two beam pipes in order to accelerate particles in both directions. The main modules to lead the particles on a circular path are the 1232 dipole magnets. They operate at a temperature of 1.9K and can create a peak magnetic field of 7.74 T. More magnets (quadrupoles, sextupoles, octupoles, decapoles) are responsible for focusing the beam and smaller corrections to optimise the particles' trajectory. There are a total of 9593 magnets. The acceleration is done by 16 cavities, grouped by four in cryomodules, with two cryomodules per beam. They are operating at a temperature of 4.5 K, each cavity delivers 2 MV at 400 MHz.

Protons undergo an acceleration process in several steps: First they are accelerated at the linear accelerator Linac2 up to a kinetic energy of 50 MeV, then injected into the Proton Synchrotron Booster (PSB) which accelerates them to 1.4 GeV. The beam is then fed to the larger Proton Synchrotron (PS), where it is accelerated to 25 GeV per proton. From there the protons go into the Super Proton Synchrotron (SPS), from where they are finally transferred to the LHC after reaching an energy of 450 GeV. The accelerator complex is depicted in Fig. 3.1. In the LHC the protons are further accelerated in both directions for ~ 20 min gaining 485 keV per turn until they reach the final beam energy and circulate in the pipes for many hours during normal operating conditions. The protons in the LHC are ordered in 2808 bunches which can each be filled with up to 1.70×10^{11} protons, the nominal intensity is 1.15×10^{11} protons per bunch. The distance between bunches is 24.95 ns.

The LHC has four active interaction areas, each located in a cavern hosting one of the main detector experiments ATLAS [49], CMS [50], ALICE [51] and LHCb [52]. The ALICE detector is designed for collisions of heavy ion nuclei, while the other three detector are mainly focused on proton-proton collisions. LHCb focuses on b-physics, whereas ATLAS and CMS are designed as general-purpose detectors.

The two most important properties of a hadron collider are the centre-of-mass energy and the luminosity that can be reached. The design beam energy of the LHC for protons is 7 TeV, however in Run 2 (years 2015 to 2018), when the data for this thesis was taken, it ran with 6.5 TeV. In the case of two particles colliding head-on the centre-of-mass energy is simply the sum of both particles' energies, resulting in 13 TeV.

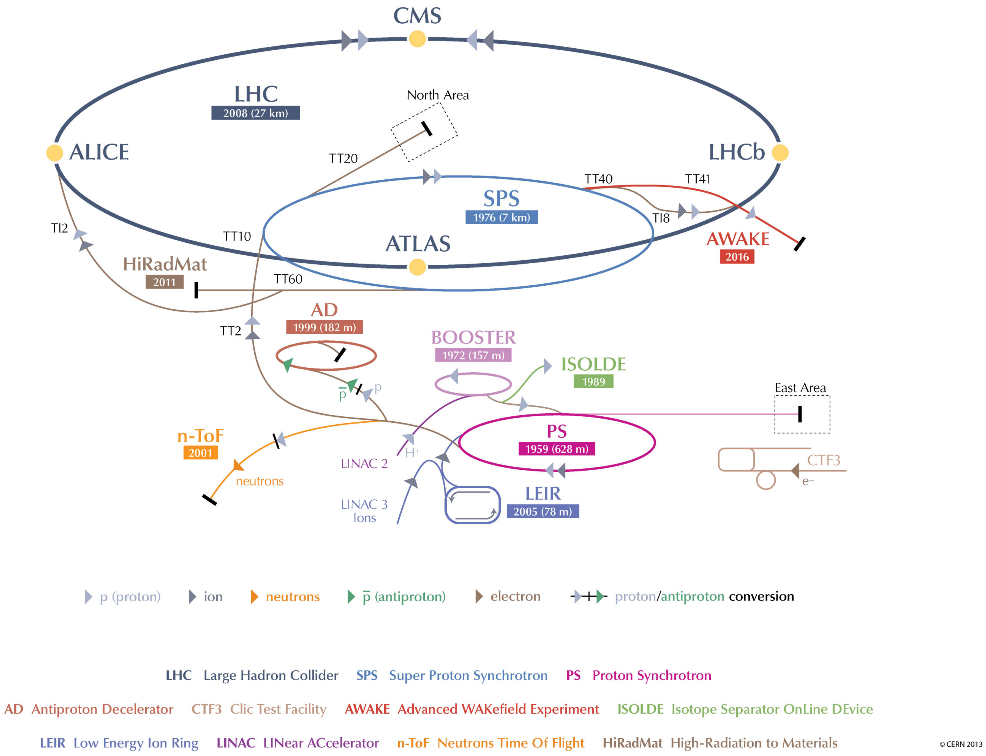


Figure 3.1: Drawing of the CERN accelerator complex, including the LHC and pre-accelerators, and experiments [48].

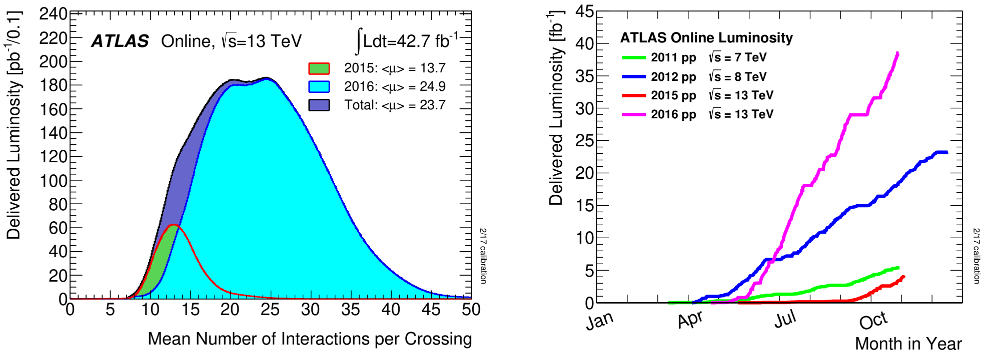
Luminosity is a measure for the number of events per time. It can be calculated from the design parameters by

$$L = \frac{f N_1 N_2}{4\pi\sigma_x\sigma_y}, \tag{3.1}$$

where N_1 and N_2 are the number of particles per bunches colliding with a frequency f. σ_x and σ_y are the the horizontal and vertical Gaussian beam width at the interaction point. LHC was designed for a nominal luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, this was exceeded in many runs, the actual luminosity was – depending on the year – often times between that and the ultimate reachable luminosity of $2.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The rate of events can be calculated by

$$N = \sigma \cdot L, \tag{3.2}$$

where σ is the cross section of the physical process. The total cross-section of proton-proton



(a) Distribution of the mean number of interactions per crossing for data taken in 2015 and 2016.

(b) Integrated luminosity versus passed time during 2015 and 2016 and in comparison during 2011 and 2012 (Run 1).

Figure 3.2: Data delivered to ATLAS by the LHC during stable beams for proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ [54].

scattering at $\sqrt{s} = 13\,\text{TeV}$ is $\sigma_{\text{tot}} = 111\,\text{mb} = 1.11 \times 10^{-25}\,\text{cm}^2$ [53], which leads at the nominal luminosity to an event rate of $1.11 \times 10^9\,\text{s}^{-1}$ or an average of 27.6 events per bunch crossing. Since the luminosity is not constant during runs the actual number of interactions per bunch crossing is a broader distribution, displayed in Fig. 3.2a. The overall data is measured in integrated luminosity $\int L\,\mathrm{d}t$; in the year 2015 the integrated luminosity delivered by the LHC was $4.2\,\text{fb}^{-1}$ and in 2016 it was $38.5\,\text{fb}^{-1}$, depicted in Fig. 3.2b. From the data in the two years, $36.1\,\text{fb}^{-1}$ is good for physics after being recorded by ATLAS.

3.2 The ATLAS detector

ATLAS[1] is a multilayer particle detector with multiple purposes, described in [49]. It is designed to measure events from proton-proton collisions at the LHC, but can also take data at heavy-ion collisions. Its main design physics goal was the search and precision measurement of the Higgs boson in multiple channels as well as sensitivity to new physics that could arise at proton-proton collisions with unprecedented centre-of-mass energy. The high luminosity of the LHC, which is needed to be sensitive to rare physics processes, poses

[1] **A** **T**oriodal **LHC** **A**paratu**S**

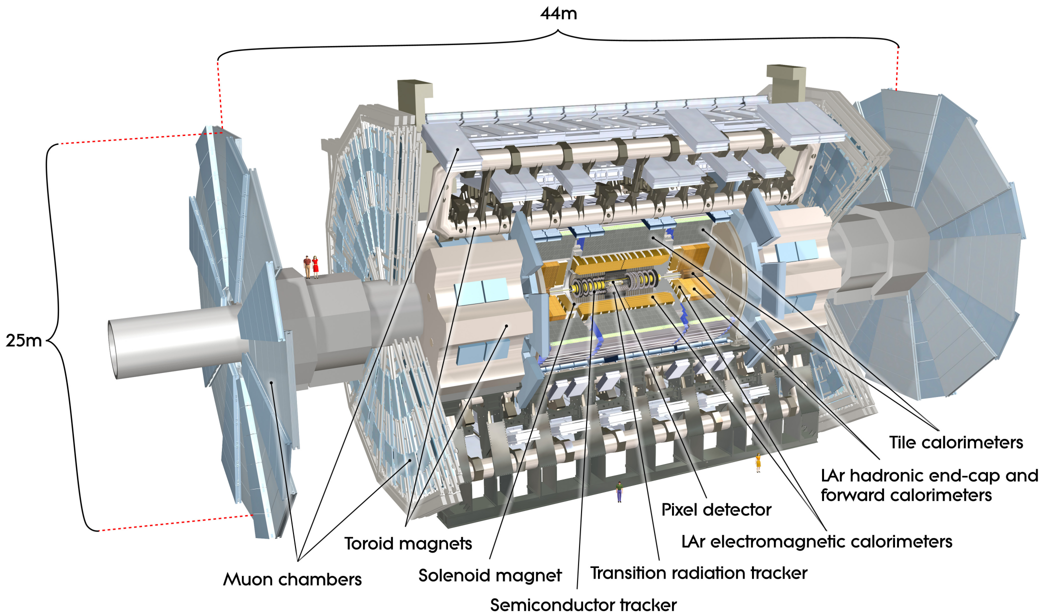


Figure 3.3: Cut-away view of the ATLAS detector [49].

challenges for the detectors: they need to have a high granularity to distinguish events from accompanying pile-up events in the same bunch-crossing; additionally the electronics and sensor elements, especially close to the interaction point, require high radiation-hardness. The detector was designed for almost full azimuthal angle coverage and with good precision for the identification of physics objects and the measurement of their momenta.

The ATLAS detector consist of several sub-detector systems, a schematic view is shown in Fig. 3.3. The right-handed coordinate system used to describe the detector (and also used in the rest of this thesis) is defined as follows: the interaction point is defined as the origin, the x- and y-axis open the plane transversal to the beam with x pointing towards the centre of the LHC and y pointing upwards; the z-axis points longitudinal to the beam. The angle ϕ is the angle in the transverse plane with $\phi = 0$ being on the x-axis. θ is the polar angle from the z-axis. Usually, instead of θ, we will use the pseudorapidity η, which is defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.3}$$

η is zero in the transverse plane and $\pm\infty$ along the beam axis. In some cases, e.g. jet reconstruction, the rapidity y is used instead:

$$y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}. \tag{3.4}$$

y has the advantage that rapidity differences between two particles are invariant under Lorentz boost along the beam axis, however it requires knowledge about the particle's energy and momentum, which is experimentally difficult. η on the other hand depends only on θ, and is thus easier to measure, and converges to y in the $m = 0$ case. For the angular distance between two objects we usually use

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.5}$$

The *inner detector*, closest to the interaction point, is a tracking system, consisting of silicon pixels and microstrip trackers (SCT[2]) and a transition radiation tracker (TRT). A solenoid coil at the outermost part immerses the inner detector in a 2 T magnetic field, which allows high-resolution momentum measurement of charged particles. The inner detector is surrounded by a high-granularity liquid-argon (LAr) *electromagnetic calorimeter* and a scintillator-tile *hadronic calorimeter*, surrounded by three large toroid magnets. The outermost part of ATLAS is the *muon spectrometer* consisting of three layers of high-precision tracking chambers. All detectors consist of a barrel-shaped part to cover the low-η region and two end-caps to cover the high-η region.
ATLAS is nominally forward-backward symmetric. Overall it is 44 m long and 25 m high and has a weight of approximately 7000 t. Its design and construction is the work of several thousand physicists, engineers, technicians and students and took place over a period of fifteen years.

[2]**S**emi**C**onductor **T**racker

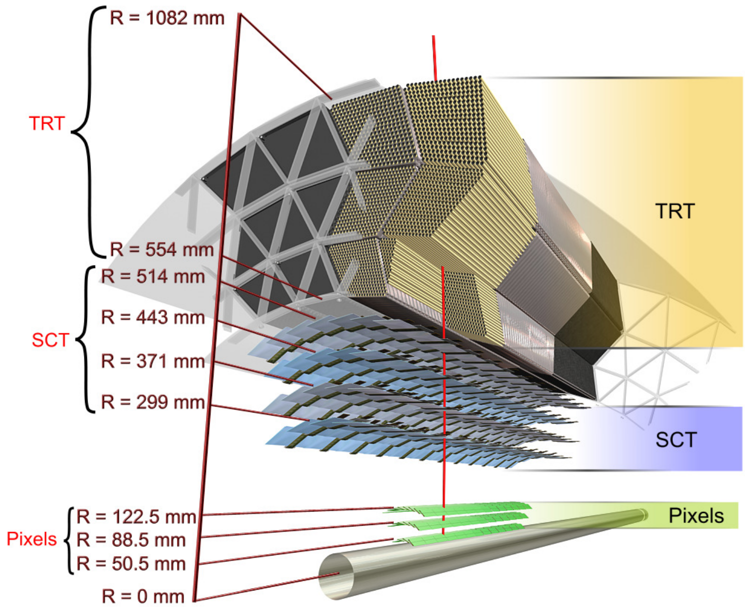


Figure 3.4: Drawing of a cut-out section of the ATLAS inner barrel detector without the IBL. The red line symbolises a charged track at $\eta = 0.3$. [49].

3.2.1 Inner detector

The inner detector is schematically displayed in Fig. 3.4 and 3.5. The viewing angle chosen in Fig. 3.4 allows to see the dimensions of the barrel region, the end-cap region is not displayed; the angle in Fig. 3.5 allows a better look on the end-cap region. The inner part of the detector is the silicon pixel detector. It consists of four cylindrical layers in the barrel region and three disc-shaped layers in the end-cap regions. It gives a very high resolution, allowing to reconstruct tracks of charged particles and vertices to a high precision. The innermost layer – 3.3 cm away from the beam axis – is the *Insertable B-layer* (IBL) [55, 56]; it was inserted between LHC Run 1 and Run 2 in order to improve the identification of secondary vertices which are used to identify decaying heavy-flavour hadrons. Most pixels are identical and have a dimension of $50 \times 250\,\mu\text{m}^2$ for the IBL and $50 \times 400\,\mu\text{m}^2$ for the other layers, which gives an intrinsic accuracy of $10\,\mu\text{m}$ in ϕ-direction and longitudinal to the particle track and $115\,\mu\text{m}$ in z direction (barrel) or radial direction R (discs). Some pixels (about 11%) are slightly longer ($600\,\mu\text{m}$) to avoid gaps. There are about 86.4 million pixels in total.

The SCT in the barrel region consists of four layers with two sets of silicon micro-strip

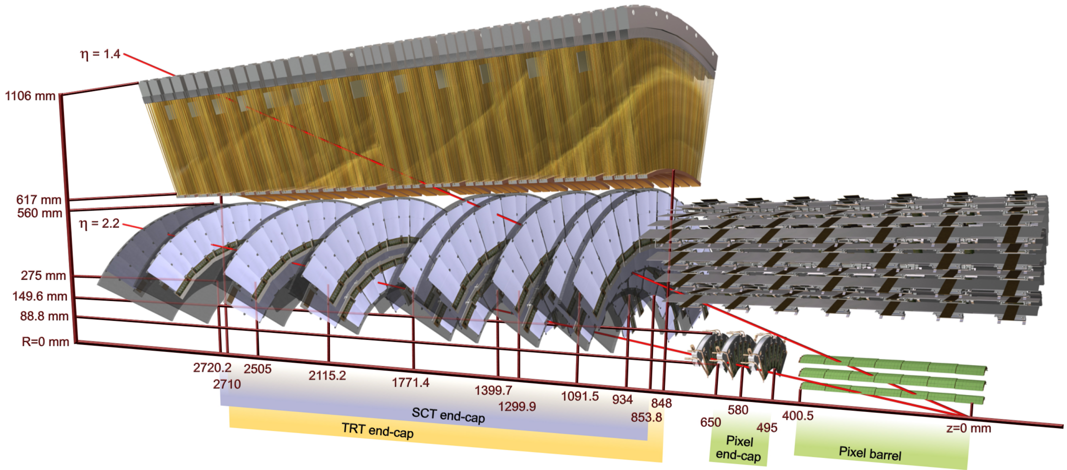


Figure 3.5: Drawing of a section of the ATLAS inner detector without the IBL. The red lines symbolise charged tracks at $\eta = 1.4$ and $\eta = 2.2$. [49].

sensors each, one set of strips is parallel to the beam direction, the other at an angle of 40 mrad, in order two obtain two-dimensional information. In the end-cap there are nine discs, consisting of a set of strips running radially and a set of strips at an angle of 40 mrad. The mean pitch of strips in both barrel and discs is $\sim 80\,\mu$m. The intrinsic accuracies are $17\,\mu$m in ϕ-direction and $580\,\mu$m in z- (barrel) or R- (end-cap) direction. The SCT has a total of approximately 6.5 million readout channels. The pixel and SCT detectors cover the region $|\eta| < 2.5$.

The TRT consists of straw tubes with a diameter of 4 mm, filled with a gas mixture of Xe (70%), CO_2 (27%) and O_2 (3%). They provide only azimuthal information with an intrinsic accuracy of $130\,\mu$m per straw (by measuring the drift-time of ionisation electrons). The barrel region consists of 73 straw planes, the tubes are arranged parallel to the beam axis, divided in halves at $\eta = 0$. In the end-cap regions there are 160 straw planes arranged radially in wheels. The TRT has approximately 351,000 readout channels. The TRT is a more economic expansion of the high-precision tracker (pixel and SCT detectors), each straw has lower precision, but this can be compensated with a higher number of hits (approximately 36 per track, fewer in the transition area between barrel and end-cap) and longer track length.

The design goal for the transverse momentum resolution in the inner detector as a whole is

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.05\% \cdot p_\mathrm{T}/\mathrm{GeV} \oplus 1\%, \tag{3.6}$$

where $\oplus$ means that the terms are added in quadrature[3]. The actual transverse momentum resolution depends on the η region and the p_T of the track, it is typically 2-6% ($p_\mathrm{T} \approx 1\,\mathrm{GeV}$) to 4-16% ($p_\mathrm{T} \approx 100\,\mathrm{GeV}$), increasing from low-$\eta$ to high-η regions.

3.2.2 Calorimeter

The calorimetry part of ATLAS consists of an electromagnetic calorimeter and a hadronic calorimeter, both divided into a barrel and end-cap region. In addition there is a forward calorimeter (FCAL) to cover the range up to $|\eta| < 4.9$. The arrangement of the sub-systems is depicted in Fig. 3.6.

The electromagnetic calorimeter (ECAL) consists of alternating layers of active liquid argon and lead plates which are passive material for absorption. The thickness of the lead plates was optimised for high energy resolution. The barrel region consists of two half-barrels, separated by a small gap at $z = 0$, it covers the region $|\eta| < 1.475$. The end-cap region is divided into an outer wheel ($1.375 < |\eta| < 2.5$) and an inner wheel ($2.5 < |\eta| < 3.2$). The accordion-like shape of the LAr detector ensures full azimuthal coverage. The design resolution of the ECAL is

$$\frac{\sigma_E}{E} = 10\% \cdot (E/\mathrm{GeV})^{-1/2} \oplus 0.7\%. \tag{3.7}$$

The ECAL has a thickness of > 22 times the radiation length X_0 in the barrel and > 24 times in the end-cap[4], making it absorb the energy of electromagnetic showers almost completely.

[3] E.g.: $a \oplus b = \sqrt{a^2 + b^2}$.

[4] The radiation length X_0 is the distance in the material after which an electron has $1/e$ of its original energy left.

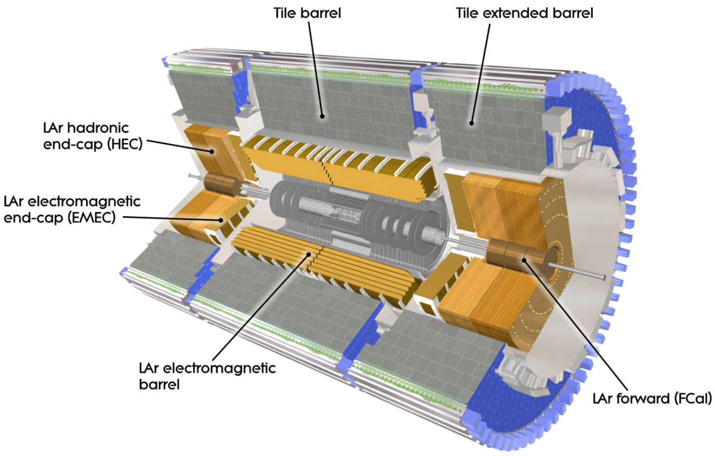


Figure 3.6: Cut-away view of the ATLAS calorimeter system. [49].

The hadronic calorimeter (HCAL) consists of a tile calorimeter in the barrel region, a LAr calorimeter in the end-cap and a LAr forward calorimeter. The tile calorimeter, which surrounds the ECAL, uses alternating layers of scintillating tile as active material and steel as absorber. At $\eta = 0$ the total detector thickness at the edge of the tile barrel is 9.7 times the interaction length[5] λ_A. In addition to the barrel region covering $|\eta| < 1.0$, there are two extended barrels covering $0.8 < |\eta| < 1.7$. The hadronic end-cap calorimeter wheels are located behind the end-cap electromagnetic calorimeters, allowing the usage of the same LAr cryostats. They cover the area $1.5 < |\eta| < 3.1$, slightly overlapping with the tile calorimeter and the FCAL. The wheels are built from copper plates, interleaved with LAr gaps as active material. The FCAL covers the region $3.1 < |\eta| < 4.9$. It is recessed by 1.2 m with respect to the ECAL and integrated into the end-cap cryostats. The FCAL has a thickness of approximately 10 interaction lengths, to reduce the background reaching the muon system in this region with high particle flux; this is achieved through high density, using copper as absorber in one module – optimised for electromagnetic calorimetry – and tungsten for two more modules to measure hadronic interactions. Each module is inlaid into a metal matrix and uses LAr in the gaps as sensitive material. The hadronic calorimetry is

[5]One interaction length is the average path length of a hadronic particle until it undergoes an elastic or inelastic interaction with the material.

designed for a resolution of

$$\frac{\sigma_E}{E} = 50\% \cdot (E/\mathrm{GeV})^{-1/2} \oplus 3\% \tag{3.8}$$

in the barrel and end-cap and

$$\frac{\sigma_E}{E} = 100\% \cdot (E/\mathrm{GeV})^{-1/2} \oplus 10\% \tag{3.9}$$

in the forward region.

3.2.3 Muon spectrometer

The outermost part of ATLAS is a set of detectors aiming to measure muons, which are the only known particles (except neutrinos, which can only interact weakly with the detector material) that traverse the whole detector without decaying or being absorbed. Their momentum is measured by tracking their trajectories in a magnetic field. In the range $|\eta| < 1.4$ the field is provided by the large barrel toroids, for $1.6 < |\eta| < 2.7$ there are two smaller end-cap magnets; in the transition area there is a combination of both magnetic fields. The bending power[6] between the innermost and outermost muon-chamber planes is 1 to 7.5 Tm, the lowest bending power is in the transition region. The muon spectrometer is a combination of different detector types, the layout is depicted in Fig. 3.7. For most of the η-range ($|\eta| < 2.7$) Monitored Drift Tubes (MDT's) are used, which provide a high precision in the measurement of the track coordinates. At $2.0 < |\eta| < 2.7$ the innermost layer is replaced by Cathode Strip Chambers (CSC's), which can withstand the high background flux. Special muon chambers are installed to provide a lower-resolution but much faster signal which can be used for triggering and measuring the second coordinate orthogonal to the precision-tracking chambers: Resistive Plate Chambers (RPC's) are used in the barrel ($|\eta| < 1.05$) and Thin Gap Chambers (TGC's) in the end-cap $1.05 < |\eta| < 2.7$. The muon spectrometer is designed

[6]The bending power is the field integral along the muon trajectory $\int B\,\mathrm{d}l$, where B is the field component normal to the muon direction.

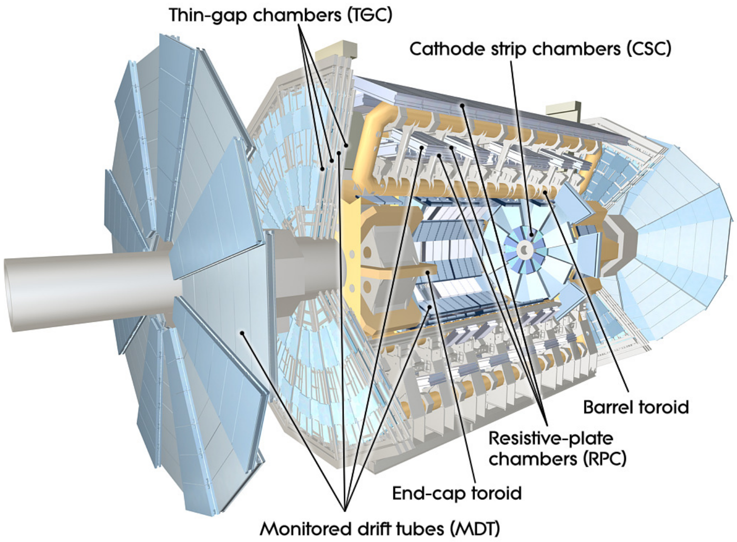


Figure 3.7: Cut-away view of the ATLAS muon system. [49].

to provide a resolution of

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 10\% \quad \text{at} \quad p_{\mathrm{T}} = 1\,\mathrm{TeV}. \tag{3.10}$$

3.2.4 Trigger system

With a bunch-crossing rate of 40 MHz it is not possible to continuously read out and store the full detector information. ATLAS uses a two-level trigger system to bring down the rate of events to be stored to approximately 1 kHz [57]. The first level (L1) is a hardware-based system, which reduces the rate to 100 kHz, the final trigger decision is made by a software based high-level trigger (HLT) system. Figure 3.8 shows the data acquisition strategy during LHC Run 2.

The L1 trigger consists of a calorimeter-based (L1Calo) and a muon-based (L1Muon) sub-system. The L1Calo system [58] looks for regions of interest based on tower clusters of different size in the electromagnetic calorimeter which exceed predefined transverse energy thresholds. The Central Trigger Processor (CTP) collects input from L1Muon and L1Calo, a topological trigger (L1Topo) performs selections based on geometry and kinematics of

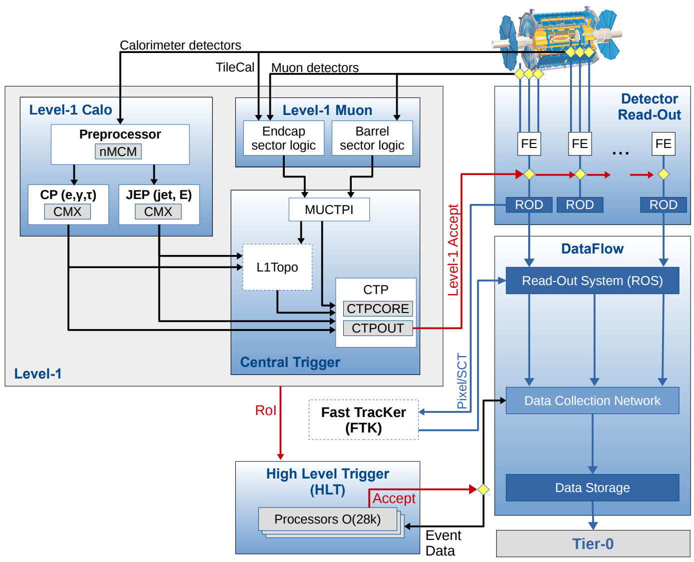


Figure 3.8: The ATLAS Trigger and Data Acquisition (TDAQ) system during LHC Run 2 [57].

trigger objects, which allows also to identify events based on missing transverse energy. If the CTP makes a positive decision, the events information is read out and buffered in the Read-Out System (ROS). The HLT then uses calorimeter information with finer granularity, precision measurements from the muon spectrometer and tracking information from the inner detector to make a final decision whether the events should be stored. There is a variety of HLT algorithms, seeded by different L1 triggers, looking for different kinds of signatures, most use a two-stage approach, first using only a fast reconstruction and rejecting the majority of events, then with a slower higher-precision approach for the remaining events. The primary triggers, which are used for physics analysis, look for signatures which are potentially interesting for the ATLAS physics programme, including electrons, photons, muons, taus, (b-)jets and missing transverse energy. Some of the HLT algorithms are discussed in Sec. 5.1.

4 Event Reconstruction

In order to use the data taken by the detector for physics analysis, the information has to be processed such that physics objects are reconstructed and identified. The methods will be described in this chapter. The first step is the use of track and vertex reconstruction with the inner detector (Sec. 4.1), which allows the identification of interaction points (*primary vertices*) and tracks of charged particles. Calorimeter cells with a measured hit that are next to each other are combined into clusters (Sec. 4.2. In a second step the information from the sub-detector systems can be used to identify muons (Sec. 4.3), electromagnetic showers (electrons or photons) (Sec. 4.4) and hadronic jets (Sec. 4.5). The signature of jets originating from a b-hadron differ from other jets due to their long lifetime, they can be identified with b-tagging algorithms (Sec. 4.6). With the information of all identified objects also the missing transverse momentum can be calculated (Sec. 4.7). Jets originating from hadronically decaying τ leptons can also be reconstructed, but since they are not used in the work presented in this thesis, this will not be further discussed. All these methods are implemented in the ATHENA software framework [59, 60]. The development of these methods and their implementation in the software framework is an effort of many people and constantly undergoes a cycle of updates to improve the physics and computational performance. This necessitates procedures to ensure that software updates only lead to the desired effects on the reconstruction results, which is routinely done within the *Physics Validation* framework (Sec. 4.8). The author contributed to the core physics validation software framework and especially to the validation of the missing transverse momentum reconstruction and thus helped to ensure stable physics results.

4.1 Track and vertex reconstruction

The track reconstruction in the inner detector uses a set of subsequent algorithms [61, 62]. A clusterisation method groups pixels and strips which measure a charge above a certain threshold that have a common edge or corner into groups, creating so-called *space-points*, which is where the charged particle has traversed the detector. In the SCT a space-point consists of clusters from both sides of a layer to measure all three coordinates, in the pixel detector one cluster suffices. The intersection point of the particle is determined from a linear approximation from the pixels contributing to a cluster. In a dense environment a cluster can be a *merged cluster*, which includes energy deposits from several particles; a neural network is trained to identify those and they can than be used for multiple tracks by the track reconstruction algorithm. After clusterisation, *seed tracks* are built from sets of three space-points, which allow an approximate momentum measurement. Depending on the sub-detectors used, these seeds tracks are classified into three categories: SCT-only, pixel-only and mixed. These classes have different purities, therefore different criteria are used to reject fake tracks; these include: momentum and impact parameters, the use of the same space-point by multiple seed tracks and the compatibility of an additional space-point with the track. Then a combinatorial Kalman filter [63] is used to build full track candidates by adding compatible space-points from other layers to the seed tracks. This can still lead to multiple possibilities. These ambiguities are resolved by assigning a *track score* to each track candidate, which takes the goodness of the track fit, cluster resolutions, holes and track momentum into account. The ambiguity solver then handles clusters shared by multiple track candidates depending on their identification as merged or single-particle clusters, favouring tracks with a higher score. This can lead to the removal of entire track candidates or the removal of clusters from a track candidate. The ambiguity solver removes tracks that do not fulfil the following criteria:

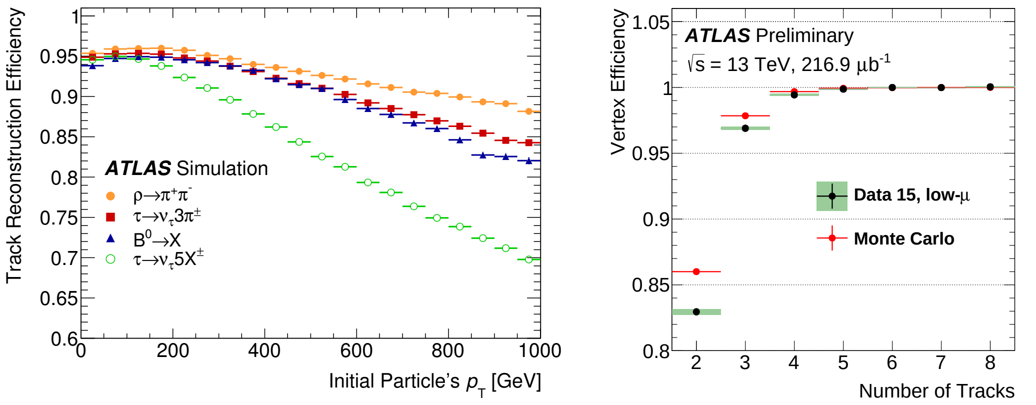
- $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ and $|\eta| < 2.5$,
- ≥ 7 pixel and SCT clusters,

- Not more than one shared pixel cluster or two shared SCT clusters,
- Not more than two holes in total and and not more than one in the pixel detector,
- $\left|d_0^{\mathrm{BL}}\right| < 2.0\,\mathrm{mm}$ and $\left|z_0^{\mathrm{BL}} \sin\theta\right| < 2.0\,\mathrm{mm}$[1].

The remaining tracks are then extended into the TRT.
An additional track-finding algorithm is then executed in backwards direction, starting in the TRT, identifying track segments with a standard Hough transform mechanism [64], using only hits that have not been used for previously reconstructed tracks. This method aims to find tracks that were not found by the former algorithms due to various reasons, e.g. tracks coming from secondary decay or material interaction vertices which might not have enough hits in the silicon detector.
The reconstruction efficiency for tracks depends mostly on the track density and therefore on the initial particle – its type and and its momentum – but also on the detector region; typically it lies in the range 80-95%. Fig. 4.1a shows the efficiency for different initial particles depending on their p_{T}.

It is important to assign reconsructed particles to the individual proton-proton interactions, hence primary vertices are reconstructed. Tracks reconstructed in the previous step are extrapolated to the beam axis and then vertex seeds are identified based on the maximum of the z-distribution of tracks. An χ^2 minimisation algorithm is then used to reconstruct the vertex from the seed [66]. This is done with an annealing algorithm, calculating weights for tracks based on their compatibility with the vertex in the previous iteration and then re-calculating the vertex with the weighted tracks. After the last iteration only tracks that lie within seven standard deviations of the vertex are associated with it, the remainder is removed and iteratively used to seed new vertices. Primary vertices are required to have at least two associated tracks. The reconstruction efficiency depends on the number of tracks;

[1] d_0^{BL} is the transverse impact parameter of a track with respect to the beam-line position, z_0^{BL} is the distance of the point used for the d_0^{BL} measurement (the ponit of closest approach on the transverse plane) from the primary vertex in z-direction.



(a) Single track reconstruction efficiency versus initial particle's p_{T} for the decay products of a primary ρ, three and five-prong τ and B^0 decaying before the IBL [62].

(b) Vertex reconstruction efficiency versus number of track in simulation low-μ data [65].

Figure 4.1: Vertex reconstruction efficiency.

as measured in a run with a low average number of ineractions per bunch-crossing in 2015, the efficiency is $\sim 83\%$ for two tracks and $97-100\%$ for three or more tracks, as depicted in Fig. 4.1b [65].

4.2 Calorimetric clusters

For the reconstruction of objects measured in the calorimeter the cell hits in the calorimeter are grouped into clusters. There are two different clustering algorithms used for object reconstruction in ATLAS: the "sliding-window" algorithm and the topological algorithm [67]. The "sliding-window" algorithm sums cells with a fixed size rectangular window. First the η-ϕ space is divided into a grid with fixed element size and the cell energy of all longitudinal layers is summed up into *CaloTowers*. Only the cells of the ECAL are used to build *EM towers*, but also *combined towers* are build with ECAL and HCAL cells, which can be used for jet reconstruction (see Sec. 4.5). Then seed pre-clusters are searched by scanning over the η-ϕ grid with a fixed window size looking for summed transverse energies of the EM towers above a threshold which are local maxima. A smaller window is used to identify the position of the pre-cluster. The actual clusters in the electromagnetic calorimeter (EM clusters) are then formed by adding up cells in a rectangular shape around the barycentre, starting with the middle layer around the barycentre of the pre-cluster seed and then continuing with the other layers using the barycentre of the middle (or other already processed) layers. Different rectangle sizes are used for different kinds of particle reconstruction and they also depend on the detector position (barrel or end-cap). The "sliding-window" algorithm is used for electromagnetic showers and jets from tau decays; ; it leads to clusters with fixed sizes, so that they can be directly compared to each other, which makes a precise calibration possible. The topological algorithm [68] starts with seed cells which have a signal-to-noise significance above a predefined seed threshold. Then cell neighbours are added iteratively to the cluster if they exceed a (lower) cell-filter threshold; if a newly added cell exceeds a growth-control threshold (lower than the seed threshold but higher than the cell-filter threshold) its neighbouring cells are then also added in a new iteration if they fulfil the criteria. This process can lead to merging of proto-clusters. A cluster splitting algorithm, which separates clusters around local maxima, prevents the formation of very large cluster structures. In contrast to the "sliding window" algorithm this method does not lead to clusters with a fixed size. The resulting clusters are used to reconstruct hadronic jets, which makes them the most frequently used calorimetric clusters in the work described in this thesis. Topological clusters are built as either electromagnetic or combined clusters.

4.3 Muons

Muons have a very clean signature in the detector: they create a track in the inner detector and hits in the muon spectrometer and deposit very little energy in the calorimeter. In the MS the signatures are reconstructed as muon tracks: Muon segments are short straight-line tracks in the MDT or CSC. A track-finding algorithm ranks segments by quality criteria and starts building tracks with high-quality segments in the middle chambers, then extending the search outwards and inwards. First the segments are matched just based on position, then a global

χ^2-fit determines if segments are compatible with a common track; if not, the segment is removed. When a track is accepted based on the fit, the associated segments cannot be used for other track candidates any more. A muon track must consist of at least two segments. With these muon tracks and the signatures in the other sub-dtetectors, there are four methods to reconstruct muons [69]:

- *combined* (CB): a combination of a MS track with an ID track; this is used for most muons in physics analysis.

- *segmented-tagged* (ST): an ID track matched with a MS segment; important mostly for low-p_{T} muons which do not necessarily reach the outer MS layers.

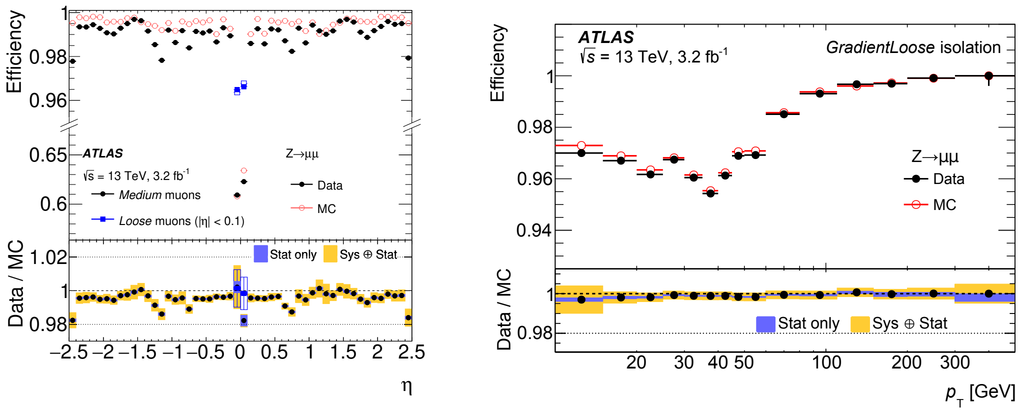
- *extrapolated* (ME): muons only reconstructed with MS data, their trajectory is extrapolated to the inner detector; this is the case mostly for the region $2.5 < |\eta| < 2.7$, which is not covered by the ID.

- *calorimeter-tagged* (CT): an ID track extrapolated through the calorimeter with energy deposits in the ECAL and HCAL compatible with a minimum-ionising particle; this method is mostly important for the region $|\eta| < 0.1$ which is not well covered by the MS.

A set of quality criteria is applied to reconstructed muon; these mostly aim to select prompt muons while rejecting secondary muons from pion or kaon decays. Four working points are defined with different criteria, depending on the muon type: *Loose*, *Medium*, *Tight* and *High-p_{T}*. The first three are inclusive, meaning that the looser selections include all muons of the tighter selections. Muons containing an ID track must have at least one hit in the pixel and five in the SCT detector and fewer than three holes in the two sub-detectors. Apart from the number of hits and wholes, criteria used for the CB identification are: the q/p significance $|{}^{q}/_{p_{\mathrm{ID}}} - {}^{q}/_{p_{\mathrm{MS}}}|/(\sigma_{\mathrm{ID}} \oplus \sigma_{\mathrm{MS}})$ which is the ratio of the charge and momentum measured in ID and MS divided by the uncertainties added in quadrature; the difference between the transverse momentum measurements in the MS and the ID divided by the p_{T}

of the combined track; the normalised χ^2 of the track fit. The *Medium* working point is the default selection for muons in physics analyses and used for control regions in the context of this thesis; it uses only CB and ME (solely for $2.5 < |\eta| < 2.7$) tracks and is designed to minimise systematic uncertainties resulting from muon reconstruction and calibration. Muons in this selection are required to have at least three hits in the MDT/CSC layers (except for $|\eta| < 0.1$) and fulfil further quality criteria to suppress the hadron background. The *Loose* selection, which is used as a baseline definition in this thesis, additionally uses CT and ST muons in the $|\eta| < 0.1$ region. The reconstruction efficiency of the two selections is displayed in Fig. 4.2a.

Often isolation is required for muons. This is done based on the variables $p_{\mathrm{T}}^{\mathrm{varcone30}}$,



(a) Identification efficiency versus pseudorapidity of the *Medium* selection for $|\eta| > 0.1$ and the*Loose* selection for $|\eta| < 0.1$ for muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$.

(b) Isolation efficiency versus p_{T} of the *GradientLoose* working point for muons satisfying the *Medium* identification criteria.

Figure 4.2: Muon efficiencies in simulation and data measured in $Z \to \mu\mu$ events [69].

which is the scalar sum of track transverse momenta $> 1\,\mathrm{GeV}$ within a cone of size $\Delta R = \min\left(10\,\mathrm{GeV}/p_{\mathrm{T}}^{\mu}, 0.3\right)$, and $E_T^{\mathrm{topocone20}}$, which is the sum of topological cluster transverse energies in a cone size $\Delta R = 0.2$ around the muon; these variables remove the expected contributions of muons originating from hadron decays. There are seven different working points with different p_{T}-dependent cuts on these variables. In this thesis the *GradientLoose* working point is used, its p_{T}-dependent efficiency is displayed in Fig. 4.2b.

Simulated muons can show slightly different results then real muons due to imperfect simulation of the detector. Their transverse momentum is therefore corrected with η dependent scale factors obtained from data-simulation comparisons in $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events. The transverse momentum is separately measured and corrected in the ID and MS and the final corrected value is a weighted average. The correction factors are $< 0.2\%$ in all detector regions. The reconstruction, identification and isolation efficiencies are also measured in data and corrections are applied to simulated events by applying weight factors.

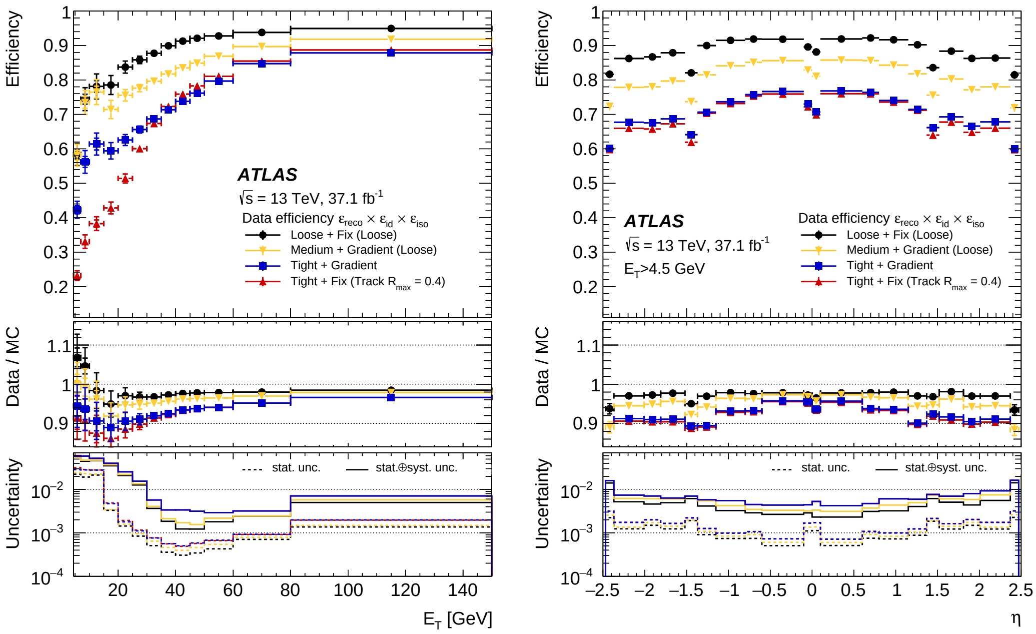
4.4 Electrons and photons

Electrons and positrons, when transversing material, lose a significant amount of energy due to bremsstrahlung. Photons can convert into an electron-positron pair, witch can then again interact with the detector material, or undergo Compton-scattering. The resulting shower of photons, electrons and positrons is usually very collimated and can normally be reconstructed as part of the same electromagnetic cluster. Dependent if a conversion happens and where the first time, zero, one or multiple tracks can be matched with a cluster originating from one primary electron or photon. EM clusters constructed with the "sliding window" algorithm are the seeds for electron and photon reconstruction, then tracks within the cluster cone are tried to be matched [70, 71]. A pattern-recognition algorithm first tries to reconstruct tracks under the pion hypothesis; if that fails, a second attempt of reconstruction is made for track seeds with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ under the hypothesis of an electron using the ATLAS Global χ^2 Track Fitter [72]. An additional fitting algorithm using an optimised Gaussian-sum filter (GSF) [73] tries to fit tracks that are only loosely matched to the EM cluster, taking into account the non-linear effects related to bremsstrahlung and increased bending in the magnetic field due to energy loss. Tracks are then matched to the EM clusters with tighter criteria, resolving ambiguities by taking into account the ΔR to the cluster barycentre in the second layer of the calorimeter and the number of hits in the silicon detectors. If a candidate track can be matched to a secondary two-track vertex, the object is considered a (converted) photon candidate if it has not hit in the innermost pixel layer. If it does not match with a photon-conversion vertex, it is considered an electron candidate. If no track is matched to the cluster, the object is a candidate for an unconverted photon. The identification of photons is mostly based on the shower shapes in the calorimeter; a *Loose* and *Tight* selection is defined [74]. In addition usually also isolation requirements are applied based on fixed cuts on fixed cones for calorimeter and tracking isolation, similar to the method used for muons. For electrons a likelihood based method is used for the identification, combining the probability density functions of 14 observables into a likelihood function and using five other observables directly as selection criteria [70]. Four working points with different efficiencies and background rejections are defined as *VeryLoose*, *Loose*, *Medium* and *Tight*. The identification criteria are optimised in bins of cluster η and E_{T}. Isolation criteria are also defined based on a variable-size cone for tracks and a fixed-size cone for clusters, as for muons. The combined reconstruction, identification and isolation efficiencies for different working points are shown in Fig. 4.3.

Electrons and positrons are calibrated by using an algorithm based on Boosted Decision Trees trained with simulated electrons and photons of known energy [75]. With this method it is taken into account that a significant part of the energy is lost outside of the calorimeter and thus cannot be measured. Correction factors are applied to account for differences between simulation and data in $Z \to ee$ events, accounting for imperfect detector simulation.

4.5 Hadronic jets

Jets coming from hadronic interactions can be reconstructed in different ways, with varying inputs and algorithms. The most common jets in ATLAS , which are widely used in this thesis, use topological clusters (see Sec. 4.2) as inputs and the anti-k_t algorithm [76] with $R = 0.4$ to construct jets [77]. The topological clusters' cell energies are measured at the EM scale, which corresponds to the deposited energy of electromagnetically interacting particles. Other possible inputs are calorimeter towers, as done by the L1 trigger system, truth particles of simulated events for *truth jets*, tracks, or jets with a smaller R-parameter (see Sec. 5.3).



(a) Efficiency versus p_T **(b)** Efficiency versus η.

Figure 4.3: Combined electron reconstruction, identification and isolation efficiencies in simulation and data measured in $J/\psi \to ee$ and $Z \to ee$ events for different working point combinations [70]; the combinations are denoted in the legend as 'Identification Working Point + Isolation Working Point'.

To cluster the inputs into jets, a distance $d_{i,j}$ between two entities (the chosen input objects) and between an entity and the beam $d_{i,B}$ are defined:

$$d_{i,j} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta_{i,j}^2}{R^2}, \tag{4.1}$$

$$d_{i,B} = k_{t,i}^{2p}, \tag{4.2}$$

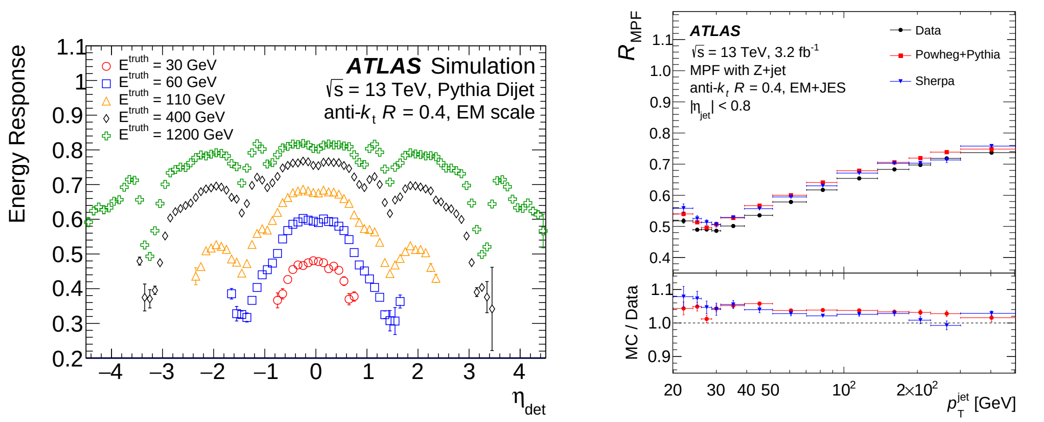
with $\Delta_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$; y is the rapidity as introduced in Eq. 3.4; k_t is the transverse momentum; R is the distance parameter which defines the characteristic size of the resulting jets; p defines the algorithm: $p = 1$ is used for the k_t algorithm, $p = -1$ for the anti-k_t algorithm, $p = 0$ is the choice for the Cambridge/Aachen algorithm [76]. The algorithm goes iteratively through the list of entities and calculates all possible $d_{i,j}$ and $d_{i,B}$. In case the smallest distance is $d_{i,j}$ the entities i and j are recombined. In case the $d_{i,B}$ is the smallest distance, the entity is removed from the list of entities and stored as a jet. This is repeated until no entities are left. The anti-k_t algorithm does not tend to cluster soft entities among themselves, but instead rather clusters soft entities to hard ones. This leads to jets with a more conical shape than other algorithms with a typical radius R. ATLAS uses the FASTJET software [78] for the execution of the jet clustering.

For $R = 0.4$ Topo jets a jet vertex tagger (JVT) [79] is used to reject background that does not come from the hard interaction by making use of information about tracks associated with the jet. Tracks are assigned to vertices if they are associated with them in the vertex reconstruction step (see Sec. 4.1) and in a second step if they are associated with the primary verex if they have $|z_0| < 3\,\mathrm{mm}$ with respect it. Tracks are assigned to jets using the ghost association method [80]. The final discriminant uses variables that depend mostly on the scalar p_T sum of tracks from the hard-scatter vertex and that sum for tracks from other vertices. The JVT is only applied for jets with $20\,\mathrm{GeV} < p_\mathrm{T} < 60\,\mathrm{GeV}$ within $|\eta| < 2.4$ (since other jets are very unlikely to originate from pile-up) and for those it has an efficiency of 92%; the false-positive rate for a pile-up jets is $< 2\%$. These rates were seen to be independent from the number of vertices.

The jet four-vectors are in a first step the sum of the four-vectors of all associated calorimeter clusters, then corrections are applied [81]. An area-based method is used to subtract the pile-up distribution dependent on the number of primary vertices N_{PV}, the mean number of interactions per bunch crossing μ and the median p_{T} density ρ:

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco}} - \rho \times A - \alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \mu, \tag{4.3}$$

where A is the area of the jet and α and β are η-dependent coefficients derived from simulation. The jet is then origin-corrected, recalulating the for-momentum such that it points to the primary vertex instead of the detector cetre. After pile-up and origin correction the absolute jet energy scale (JES) and η calibration is applied by correcting the energy and η according to simulated jets with known truth properties. The η calibration accounts for biases in the η reconstruction caused by transitions between different calorimeter techniques or granularities. In Fig. 4.4a the η dependent jet response for jets with different known energy E^{truth} is displayed. The jet response is defined as the ratio of the reconstructed and the truth energy $E^{\mathrm{reco}}/E^{\mathrm{truth}}$. A global sequential calibration is applied by using correction



(a) η-dependent jet response for jets with different $p_{\mathrm{T}}^{\mathrm{truth}}$ after pile-up and origin correction. Low jet responses are caused by gabs or transitions between the subdetectors.

(b) Comparison of the jet response in data and simulation in Z+jet events.

Figure 4.4: Jet response measurements used for the jet calibration. The η and E^{truth} dependent jet response (a) is used for the absolute jet energy scale, in-situ measurements of e.g. Z+jet events (b) are used to correct differences between data and simulation [81].

factors based on five observables that give information about the jet composition, which depends on the initiating particle[2]. Lastly a set of in-situ calibration methods is applied to account for differences between data an MC. The jet responses are obtained from data/mc comparison for different processes – dependent on the detector region – where the p_T of jets is balanced against well-measured reference objects, e.g a Z boson, as displayed in Fig. 4.4b. The final uncertainty of fully calibrated jets depends on p_T and η and is shown id Fig. 4.5.

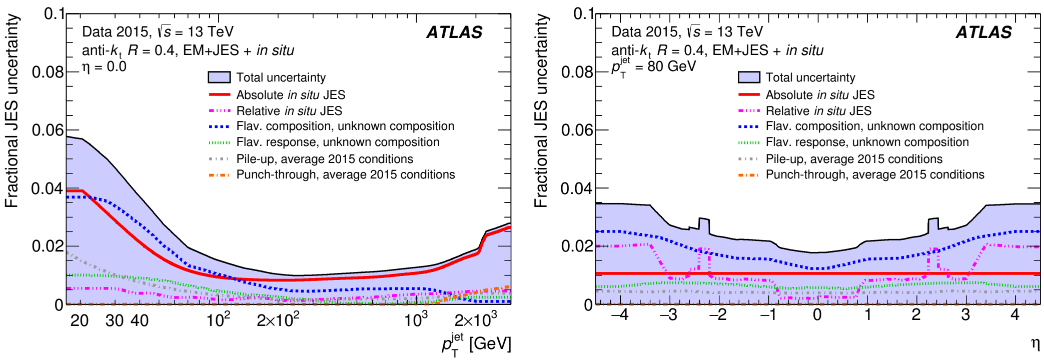


Figure 4.5: Combined uncertainty and its components in the jet energy scale for fully calibrated jets as a function of a) the jet p_T at $\eta = 0$ and b) η for a $p_T = 80\,\mathrm{GeV}$ jet [81].

4.6 b-tagging

Hadrons containing a b quark (and to a lesser degree also a c quark) have a relatively long lifetime $\mathcal{O}(10^{-12}\,\mathrm{s})$, leading to an average flight path length of several mm depending on their boost. This behaviour can be exploited to identify jets as originating from a b quark (*b-tagging*) [82,83]. One method is to use the impact parameters d_0 and $z_0 \sin\theta$ (*IP3D*), another is to search for secondary vertices (*SV*). Additionally, the *JetFitter* algorithm exploits the topological structure of the jets to reconstruct the b- or c-hadron decay chain. A multivariate method (MV2) based on Boosted Decision Trees (BDTs) combines inputs from all three methods to achieve a better discriminating power. Since c-jets behave differently from jets

[2] Quark-initiated jets tend to include higher-p_T hadrons that penetrate further into the calorimeter, gluon-initiated jets tend to contain softer but more particles.

originating from light quarks or gluons (light jets), the discriminating power of the MV2 algorithm against c- and light jets depends on the amount of c-jets used in the training procedure. Different analyses can have different needs for this, therefore several variants were trained, e.g. the variant MV2c10 (which is used throughout this thesis) was trained against a background with a 7% fraction of c-jets[3] and 93% fraction of light jets. The BDT discriminant allows the choice of different working points, defined by their b-jet efficiency. The 77% working point, which is used throughout this thesis, has an estimated c-jet rejection of 6 (meaning that 1/6 of c-jets are misidentified as b-jets), a light jet rejection of 134 and a τ rejection of 22; this performance was evaluated in simulated $t\bar{t}$ events [83].

4.7 Missing transverse momentum

Due to momentum conservation, the momenta in the transverse plane of all products of an interaction should effectively add up to zero. This would be the case if the detector could measure the momenta of all particles perfectly. However, the momentum measurement uncertainty varies for different particles, and for particles that do not deposit any energy in the detector a momentum measurement is not possible at all. Neutrinos are the only known particles that traverse the detector without any interaction, in BSM models more such 'invisible' particles can arise. The missing transverse momentum $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, which is defined as the negative vector sum of all measured transverse momenta, is hence a sign for invisible particles, in events without such it is expected to be close (because of imperfect measurement) to zero.

$\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ can be constructed with different methods [84]: One can use the p_{T} information of all reconstructed tracks in the ID that are associated with a primary vertex, one can use calorimeter clusters or one can use fully reconstructed and calibrated objects. The most common method in ATLAS is to use reconstructed objects (with the methods previously described in this chapter) and in addition use a 'soft term', which contains the information

[3]The number '10' indicates a fraction of 10%, which was originally the case. The name was retained even though the exact number was changed to optimise the rejection power [83].

of all objects (usually with low p_{T}) that have not been associated with any reconstructed physics object. A cluster-based soft term (CST) suffers from high pile-up which cannot be suppressed well, for a track-based soft term (TST) however it is possible to use only tracks associated with the primary vertex, which makes it largely independent from pile-up. Because of this the track based soft term is the preferred variant for LHC Run 2, even though it has the disadvantage that soft particles that are neutral or at $|\eta| > 2.5$ cannot be measured. With these methods the missing transverse momentum is hence defined by

$$-\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} = \sum_{\mathrm{jets}} \mathbf{p}_{\mathrm{T}} + \sum_{\mathrm{electrons}} \mathbf{p}_{\mathrm{T}} + \sum_{\mathrm{photons}} \mathbf{p}_{\mathrm{T}} + \sum_{\mathrm{taus}} \mathbf{p}_{\mathrm{T}} + \sum_{\mathrm{muons}} \mathbf{p}_{\mathrm{T}} + \sum_{\mathrm{soft,track}} \mathbf{p}_{\mathrm{T}}; \tag{4.4}$$

for the soft term tracks with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ that are associated to the primary vertex are used. An overlap removal ensures that tracks or reconstructed objects are not double-counted. The exact definition of the first five terms is specific to the object definition choices of an analysis, e.g. it is possible not to reconstruct tau jets at all, the transverse momenta of taus are then not neglected but included in the jet or soft term. The magnitude of $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ is usually referred to as missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. For $Z \to \mu\mu$ events, where there are no invisible particles, the distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the soft soft term are shown in Fig. 4.6.

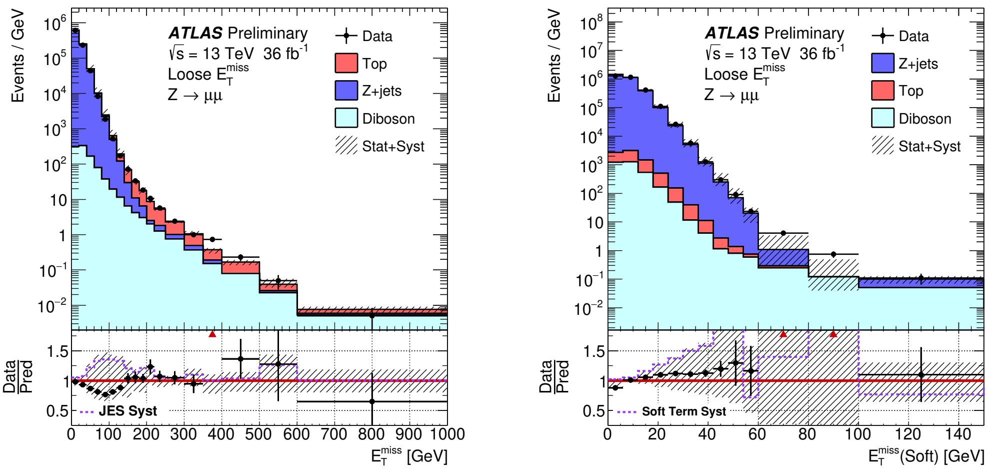


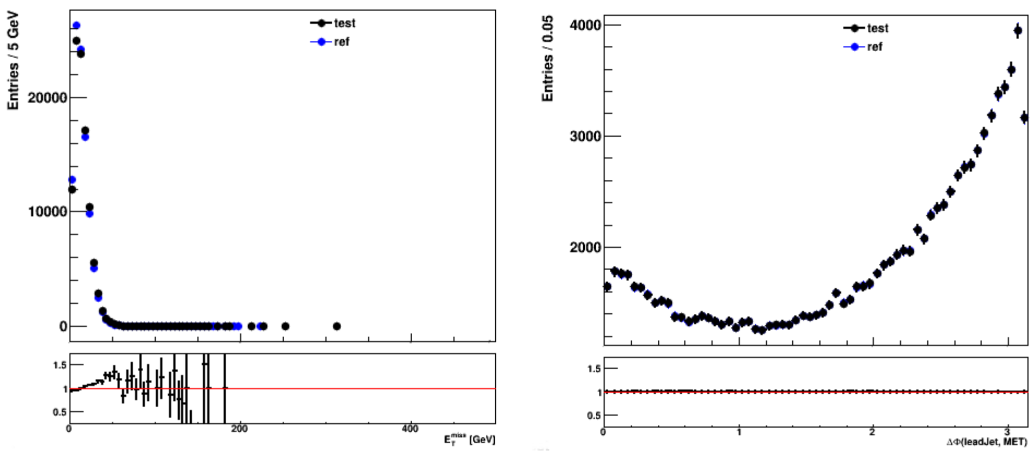
Figure 4.6: Distribution of a) $E_{\mathrm{T}}^{\mathrm{miss}}$ and b) the track-based soft term in $Z \to \mu\mu$ events in simulation and 2015+2016 data [85]. There are no invisible particles, $E_{\mathrm{T}}^{\mathrm{miss}}$ is thus expected to be close to zero.

4.8 Physics validation

Since the methods to reconstruct physics objects and the corresponding software are continuously updated and improved, the performance and technical integration of such updates has to be monitored. This is done within the *physics validation* procedure: event samples are reconstructed with an older, previously validated software with known behaviour (reference sample) and also with an updated version of the software (test sample). Distributions of the objects' properties are stored in histograms, the test and reference histograms are then compared. For $E_{\mathrm{T}}^{\mathrm{miss}}$ these distributions include: different $E_{\mathrm{T}}^{\mathrm{miss}}$ variants calculated with different jet collections and the track-based or the cluster-based soft term, as well as $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated only from tracks or clusters; for all separate terms in Eq. 4.4 the vector sum, its x- and x-component, ϕ and the scalar sum; correlations between the terms; residuals, significances and ϕ-distances to certain objects; differences between the $E_{\mathrm{T}}^{\mathrm{miss}}$ terms and naive sums of the corresponding objects. Some typical validation plots are shown in Fig. 4.7.

These tests are done regularly for updates in the object reconstruction and identification software, but can also be used for other potential changes, e.g. in simulation or treatment of pile-up events. The validation can be run on samples of simulated as well as real events, in both cases the same events are used for the test and reference samples; simulated events are usually overlaid with randomised pile-up events, which causes statistical fluctuations in the validated distributions even if the software does not cause any changes, but for data the distributions are expected to be identical. Differences between the test and reference distributions are checked by eye as well as with automated statistical tests (chi-squared and Kolmogorov-Smirnov [86, 87]). The test results are to be interpreted according to the expected changes of the software update – sometimes updates are supposed to improve the reconstruction performance, but sometimes they are not expected to change the physics results at all.

All software changes have to undergo this validation routine before they are used in physics analyses.



(a) Soft term reconstructed from tracks matched with the primary vertex in simulated events. In addition to statistical fluctuations caused by randomised pie-up there are significant shape differences. In this example the update was not expected to change the physics outcome, therefore follow-up was required before releasing the update into analysis software.

(b) Difference in ϕ between $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T}-leading jet in data. Test and reference distributions are identical, the test passed.

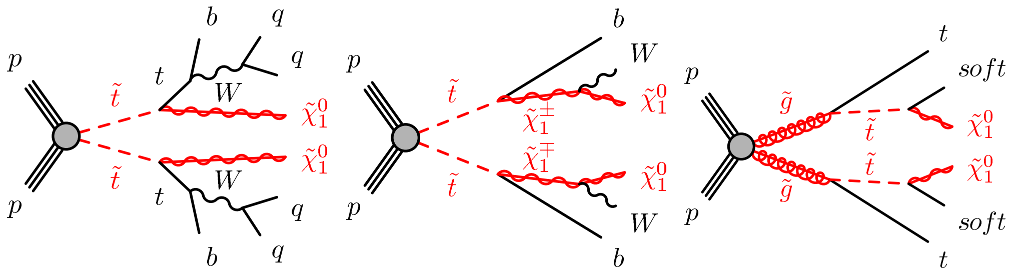
Figure 4.7: Typical tests in the $E_{\mathrm{T}}^{\mathrm{miss}}$ physics validation. Distributions of a test sample (black) are compared to a reference sample (blue), the lower panel shows the ratio of both samples.

5 Search for top squarks in the fully hadronic final state

This chapter describes the search for top squarks in fully hadronic final states at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ with an integrated luminosity of $\int L\mathrm{d}t = 36.1\,\mathrm{fb}^{-1}$ of data taken with the ATLAS detector. This search was published in JHEP [4]. The main goal is to search for evidence of stop pair production in the $t\bar{t} + E_{\mathrm{T}}^{\mathrm{miss}}$ final state, a pseudo-Feynman diagram of this process is shown in Fig. 5.1a, but also other signal models with this final state were considered. A previous search with similar methods had been conducted with $\int L\mathrm{d}t = 20.1\,\mathrm{fb}^{-1}$ of data taken with the ATLAS detector at $\sqrt{s} = 8\,\mathrm{TeV}$, which did not find evidence for BSM processes [3]. Exclusion limits on stop pair production derived from this search are shown in Fig. 5.2.

The search presented here defines several signal regions (SRs) to look for different signal models with the same final state but different kinematic properties. Each of them is optimised to achieve a high expected discovery significance for certain signal models over the SM background. SRA and SRB are both targeting the $\tilde{t}_1\tilde{t}_1 \to tt\tilde{\chi}_1^0\tilde{\chi}_1^0$ decay chain, as depicted in Fig. 5.1a, and follow similar strategies. SRA targets the regions with high stop and low neutralino mass (benchmark model: $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1000, 1)\,\mathrm{GeV}$[1]), where the top quarks are highly boosted and $E_{\mathrm{T}}^{\mathrm{miss}}$ is generally high. SRB targets regions with lower stop and higher neutralino mass (benchmark model: $m(\tilde{t}_1, \tilde{\chi}_1^0) = (600, 300)\,\mathrm{GeV}$) with lower boosted top quarks and less $E_{\mathrm{T}}^{\mathrm{miss}}$. Both SRs try two reconstruct top quarks (and in some cases W

[1]Neutralinos lighter than 1 GeV (but not zero) are theoretically allowed but would not significantly change the kinematics in the final state.



(a) $\tilde{t}_1\tilde{t}_1 \to tt\tilde{\chi}_1^0\tilde{\chi}_1^0$ **(b)** $\tilde{t}_1\tilde{t}_1 \to bb\tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\pm}$ **(c)** $\tilde{g}\tilde{g} \to \tilde{t}_1\tilde{t}_1tt$

Figure 5.1: Pseudo-Feynman diagrams showing the targeted scenarios: pair production of a top-squark pair decaying either via a top quark (left) or a chargino (middle), and pair production of gluinos decaying into top quarks and quasi-invisible top squarks. The circles symbolise a summation of all production channels. All three models have a final state with bottom and light quarks and two of the lightest neutralinos.

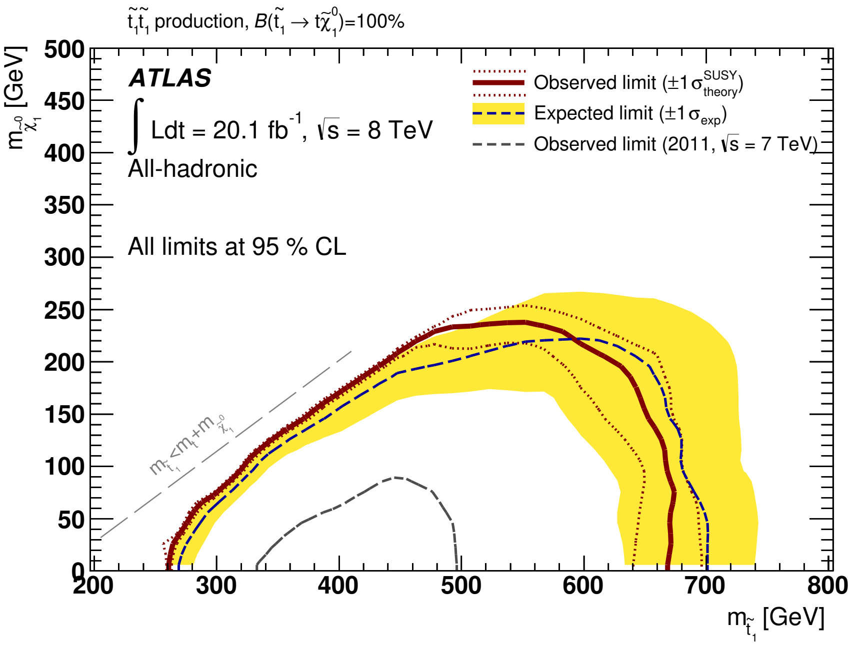


Figure 5.2: Exclusion limits prior to this thesis at 95 % confidence level on top-squark pair-production for the case when top squarks decay exclusively into a top and a neutralino [3].

bosons) and are categorised into three bins based on the reconstruction results: **TT** (two reconstructed top candidates), **TW** (one top and one W candidate), and **T0** (one top candidate only). SRC targets scenarios where $m(\tilde{t}_1) \sim m(t) + m(\tilde{\chi}_1^0)$. It selects events with a high-p_{T} jet from initial state radiation (ISR) to make them distinguishable from $t\bar{t}$. SRD targets the decay chain $\tilde{t}_1\tilde{t}_1 \to bb\tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\pm} \to bbWW\tilde{\chi}_1^0\tilde{\chi}_1^0$, as depicted in Fig. 5.1b, with the assumption $m(\tilde{\chi}_1^{\pm}) = 2m(\tilde{\chi}_1^0)$. In this scenario there are no on-shell top quarks that can be reconstructed. SRE is designed for scenarios with very highly boosted top quarks, as they can be produced in the decay of gluinos, as depicted in Fig. 5.1c, especially with large $\Delta m(\tilde{g}, \tilde{t}_1)$ and small $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ (benchmark model: $m(\tilde{g}, \tilde{t}_1, \tilde{\chi}_1^0) = (1700, 400, 395)$ GeV). The signal regions are described in more detail in Sec. 5.5.

For the signal and background processes simulated events were used, described in Sec. 5.2. The major background processes are normalised to data in dedicated control regions (CRs), which are enriched in specific backgrounds. This process is described in Sec. 5.6. Validation regions (VRs) were defined to verify that the backgrounds are well-modelled. The results in the signal regions are shown in detail in Sec. 5.9. The interpretation of these findings for different signal scenarios are described in Sec. 5.8.

The parts of this analysis with a major contribution done by the author of this thesis are trigger studies, the design of the control region for the W+jet background and the improvement of the signal region A by adding the observable M_{T2} and optimising the design of the signal region with it.

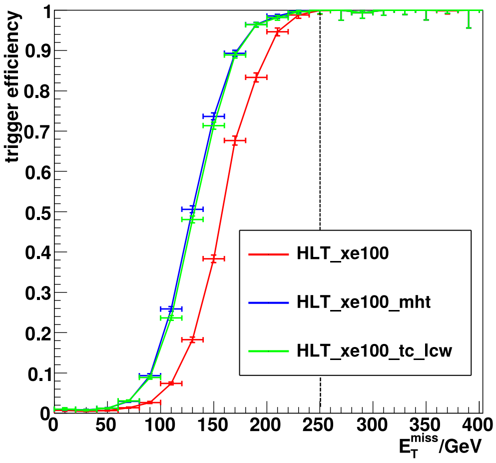
5.1 Data collection and trigger

Data for this analysis was taken from August to November 2015 and April to October 2016 at a centre-of-mass energy of 13 TeV. A *Good Runs List* (GRL) is used for the bookkeeping which of the stored data can be used for physics analysis. This results in an overall integrated luminosity of 36.1 fb^{-1}. The rate of data is much higher than can be stored offline and thus has to be brought down to $\sim$ 1 kHz. This is done by a two-level trigger system [57]:

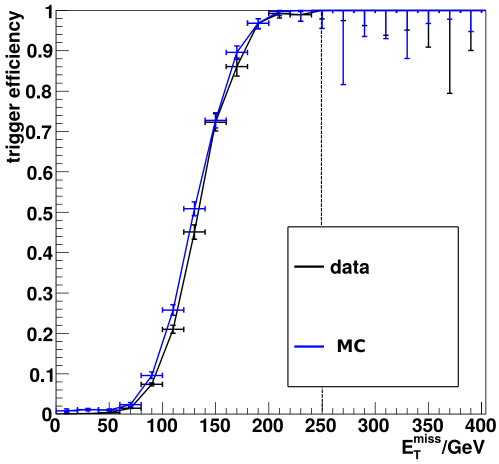
a first level (L1) hardware trigger followed by a software-based high-level trigger (HLT) (see Sec. 3.2.4). The most promising triggers for the targeted signal models in this analysis look for a significant amount of missing transverse energy. For the HLT there are different algorithms for $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction that can be used. These algorithms use simplified methods for the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation as the computation needs to run quickly during data-taking. The threshold for storing the event does therefore not correspond to a fixed threshold in offline $E_{\mathrm{T}}^{\mathrm{miss}}$ (as described in Sec. 4.7), the efficiency is described by a turn-on curve. A study was done to find out, which HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm works best in terms of efficiency and data-MC agreement and to determine the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold where the trigger is fully efficient. Since our expected signal yields are very small, only unprescaled triggers were considered[2]. In particular there were three different trigger algorithms in consideration: the cell algorithm (`xe`), which calculates $E_{\mathrm{T}}^{\mathrm{miss}}$ from LAr and Tile calorimeter cells which significantly exceed the noise level, the jet-based algorithm (`xe_mht`) which calculates $E_{\mathrm{T}}^{\mathrm{miss}}$ from trigger-level jets, and the topo-cluster algorithm (`xe_tc_lcw`), where $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated similar as in the cell algorithm but with topo-clusters built for the entire calorimeter. Other additional signatures like jet+$E_{\mathrm{T}}^{\mathrm{miss}}$ or b-tagging on trigger level were seen to not significantly improve the performance of the `xe(_[...])` triggers. This study was conducted in 2015, when due to lower rates the trigger could be run at relatively low thresholds (70 GeV for all $E_{\mathrm{T}}^{\mathrm{miss}}$ variants) but it was expected that the thresholds would increase to at least $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ and probably not all of them could be kept unprescaled. In order to study the trigger performances in the context of this analysis, a selection requiring at least four jets with $p_{\mathrm{T}}^{0-3} > (80, 80, 40, 40)\,\mathrm{GeV}$ and one jet to be b-tagged (similar to the preselection used in Sec. 5.5, but without $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss,track}}$ requirements) was applied to the background samples described in Sec. 5.2 to derive efficiency values for events with a certain $E_{\mathrm{T}}^{\mathrm{miss}}$[3]. Trigger efficiency is defined as the ratio of the number of events that pass the trigger and the number of all events with a certain $E_{\mathrm{T}}^{\mathrm{miss}}$ value. It can be seen in Fig. 5.3a that the `xe_mht` algorithm has overall the highest efficiency and reaches the plateau of full efficiency at the

[2]Prescaled triggers select only a predefined percentage of events that fulfil the criteria, whereas unprescaled triggers select all.

[3]$E_{\mathrm{T}}^{\mathrm{miss}}$ refers here and in the following to the fully reconstructed offline $E_{\mathrm{T}}^{\mathrm{miss}}$.



(a) Comparison of different $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger algorithms with simulated events.



(b) Comparison of data and simulation for the `HLT_xe100_mht` trigger.

Figure 5.3: Efficiencies of triggers with $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ threshold on trigger level as a function of the offline $E_{\mathrm{T}}^{\mathrm{miss}}$. The dashed line depicts $E_{\mathrm{T}}^{\mathrm{miss}} = 250\,\mathrm{GeV}$.

lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ value. The `xe_mht` algorithm uses methods that are closest to the methods of offline $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, especially in the fully hadronic final state, which could be the reason why it performs best. The `xe_tc_lcw` algorithm performs only slightly worse, whereas the efficiency of the `xe` algorithm is significantly lower. To determine the efficiency in data, events passing a trigger requiring 2 b jets with $p_{\mathrm{T}} > 35\,\mathrm{GeV}$ where used and the same preselection as before was applied. For simulation a set of samples was used that is expected to describe all SM processes[4]. It can be seen that there are differences between data and simulation in the turn-on. This could in principle be caused by a correlation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ an b-jet trigger efficiencies, but the same observation was made when applying a preselection that ensured full efficiency of the b-jet trigger. These differences would need to be considered if we wanted to use events with lower $E_{\mathrm{T}}^{\mathrm{miss}}$ values – but after reaching a plateau the `HLT_xe100_mht` trigger can be considered fully efficient for both data and simulated events. It can be concluded that it can be safely used above an offline $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold of 250 GeV without further considering the turn-on efficiencies. This study was considered in

[4]These samples are described in detail in Sec. 5.2.

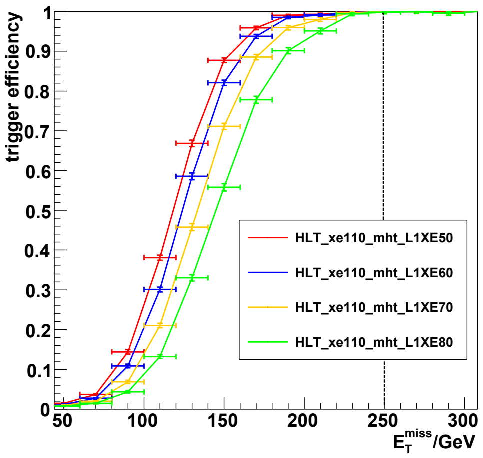
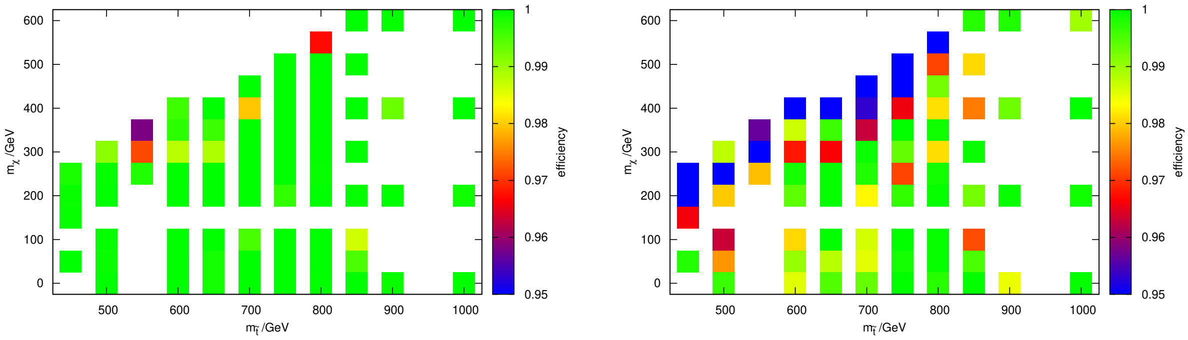


Figure 5.4: Efficiencies of `HLT_xe110_mht` triggers in simulation with different L1 seeds as a function of the offline $E_{\mathrm{T}}^{\mathrm{miss}}$. The dashed line depicts $E_{\mathrm{T}}^{\mathrm{miss}} = 250\,\mathrm{GeV}$.

the decision, which trigger to keep as the lowest unprescaled one. In 2016 `HLT_xe90_mht` was then used as lowest unprescaled in early data-taking periods, later `HLT_xe110_mht` was used in runs with higher luminosity. It was possible to keep the same L1 threshold (`L1XE50`), which also has an impact on the turn-on. The dependence on the L1 threshold is depicted in Fig. 5.4. It can be concluded that the used `HLT_xe110_mht_L1XE50` is still fully efficient at offline $E_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$.

In order to see if $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers alone are efficient enough for all signal models, the efficiencies of the triggers were determined in simulated events for all considered $\tilde{t}_1\tilde{t}_1 \to tt\tilde{\chi}_1^0\tilde{\chi}_1^0$ models. A similar preselection as the one described in Sec. 5.5 was used, but with a relaxed threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$, which was considered to be potentially useful for some of the models. The efficiencies are shown in Fig. 5.5. It can be seen that the loss due to the `HLT_xe_mht_100` trigger is below 1% in most signal models and below 5% in all, so that no other trigger needs to be taken in consideration for any signal region. The loss due to the `HLT_xe_100` trigger would be considerably larger, which is another sign that the jet-based trigger algorithm works best for the analysis under consideration.

The use of an $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is sufficient for all signal and control regions with no or one lepton. There are however two control regions (described in Sec. 5.6.2 and 5.6.3) which do



(a) `HLT_xe100_mht` trigger.

(b) `HLT_xe100` trigger.

Figure 5.5: Efficiencies of triggers for signal events with offline $E_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$.

not require high $E_{\mathrm{T}}^{\mathrm{miss}}$ and thus require a different trigger. These regions were designed such that they can use electron or muon triggers, which were found to be fully efficient at our signal definition in these regions. All used triggers are summarised in Tab. 5.1.

5.2 Simulated data

In order to estimate the contributions of different processes in all regions, a good understanding is needed of how they would look like in the detector. This means that a suitable amount of events has to be generated by Monte Carlo generators, the parton showering (PS) and hadronisation has to be simulated and then the event has to be processed in a detector simulation. The result can then be processed with the same reconstruction methods as recorded data events. For the modelling of signal and background processes different Monte Carlo generators were used.

Signal models were generated with MADGRAPH5AMC@NLO 2.2-2.4 [88], the PS and hadronisation was done with PYTHIA8 [89]. The decay of b and c hadrons was evaluated with EVTGEN 1.2.0 [90]. Cross-sections were calculated at next-to-leading order in the strong coupling constant and with next-to-leading-logarithm accuracy for resummation of

Table 5.1: Overview of the triggers used in the different data-taking periods. Data periods are defined such that the detector and trigger configuration is coherent in one period. For electrons and muons an OR combination of different triggers was used. The trigger names contain the criteria for the trigger objects. The numbers refer to the p_T of the trigger object. Electrons are identified with a likelihood-based algorithm with the three working points *lhtight, lhmedium, lhloose*. Muons are identified with a combination of ID tracks and tracks in the muon detector. The suffices *ivarloose, iloose, ivarmedium* describe isolation criteria. More information in [57].

Type	Year	HLT algorithm
E_T^{miss}	2015	`HLT_xe70_mht`
	2016	`HLT_xe90_mht` (period A-D3)
	2016	`HLT_xe100_mht` (period D4-F1)
	2016	`HLT_xe110_mht` (period F2-L)
Electron	2015	`HLT_e24_lhmedium` or `HLT_e60_lhmedium` or `HLT_e120_lhloose`
	2016	`HLT_e26_lhtight_nod0_ivarloose` or `HLT_e60_lhmedium` or `HLT_e140_lhloose_nod0`
Muon	2015	`HLT_mu20_iloose` or `HLT_mu50`
	2016	`HLT_mu26_ivarmedium` or `HLT_mu50`

the soft-gluon emissions (NLO+NLL) [44, 91, 92], the nominal cross-sections and uncertainties result from an envelope of calculations with different sets of PDF, factorisation and renormalisation scales, described in [45].

SM processes with a Z or W boson with additional jets (referred to as Z+jet and W+jet) can end up in signal regions due to hadronic activities including b- and c- jets and missing transverse energy from one or two neutrinos produced in the decay of the Z and W boson. The dominant Feynman diagrams of these processes are shown in Fig. 5.6. For the MC samples SHERPA 2.2.1 [93–96] was used with the NNPDF3.0 NNLO PDF set [97]. Generator-level filters on the presence or absence of b or c hadrons were used to increase the number of events containing heavy-flavour jets in order to improve the modelling in a selection with b-tagged jets. Samples with different filters are then separately normalised to their cross section.

Top pair production ($t\bar{t}$) can act as a background if at least one of the top quarks decays leptonically, resulting in a b quark, a charged lepton and a neutrino; the neutrino produces

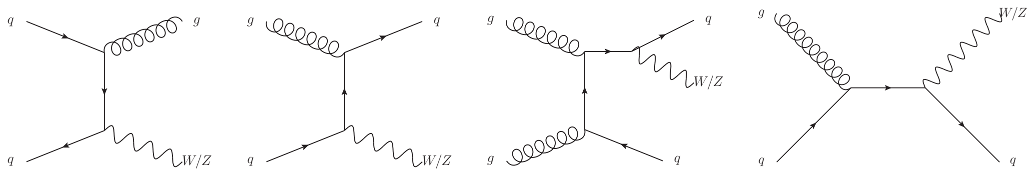


Figure 5.6: Selected Feynman diagrams for W/Z+jet production in proton-proton collisions. In case of a W the quark at the vertex changes from an up-type quark to a down-type or vice versa. In case of a Z the quark flavour does not change.

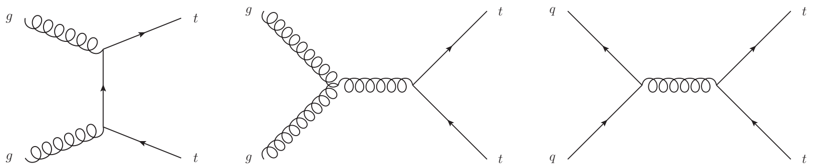
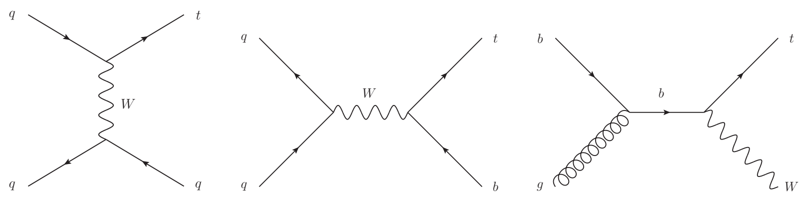


Figure 5.7: Selected Feynman diagrams for $t\bar{t}$ production in proton-proton collisions. These processes can act as backgrounds if at least one top quark decays leptonically.

missing transverse energy and the event can fulfil the selection criteria if the charged lepton is not reconstructed as such. In the case of both top quarks decaying hadronically, events can only contribute as backgrounds if there is high mismeasures $E_{\mathrm{T}}^{\mathrm{miss}}$; this happens only rarely and is thus an insignificant source of background. Feynman diagrams for $t\bar{t}$ production are shown in Fig. 5.7. The samples were generated with POWHEG-BOX 2 [98], interfaced to PYTHIA6 [99] for PS and hadronisation. A single top quark can be produced together with a second quark in the *t*- or *s*-channel or with a W boson in the *Wt*-channel, as shown in Fig. 5.8. Except for the *t*-channel, which was produced with POWHEG-BOX 1, the same configuration was used as for top pairs. In order to achieve better modelling through a higher amount of events in regions with high missing transverse energy, event-level filters were used to produce slices with different values of missing transverse momentum for the $t\bar{t}$ and *Wt* samples.

The production of a top-antitop pair in association with a Z or W boson ($t\bar{t}+V$) is a background with a low cross-section. However, if there is a pair of hadronically decaying top quarks plus a Z boson decaying via $Z \to \nu\bar{\nu}$ it results in the same final state as the signal and is therefore an *irreducible background*. $t\bar{t}+V$ samples were generated with MG5AMC@NLO [88, 100] interfaced to PYTHIA8 [89] for PS and hadronisation. The



(a) t-channel **(b)** s-channel **(c)** Wt-channel

Figure 5.8: Selected Feynman diagrams for single-top production in proton-proton collisions.These processes can act as backgrounds if at least one neutrino is produced from a top or W decay. In all three channels also an anti-top can be produced equivalently by reverting the arrows.

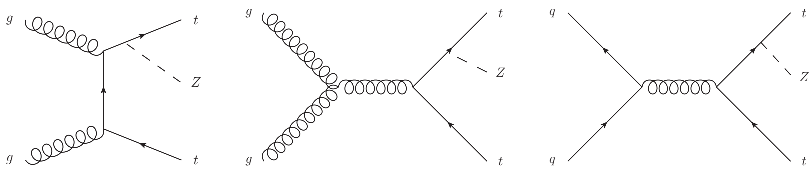


Figure 5.9: Selected Feynman diagrams for $t\bar{t}+Z$ production in proton-proton collisions. These processes can act as backgrounds if at least one top quark decays leptonically.

dominant Feynman diagrams for $t\bar{t}+Z$ are shown in Fig. 5.9.

Diboson processes are a minor background, they can contribute in signal regions due to hadronic activity and missing transverse energy from one or more neutrinos. Samples were generated with SHERPA 2.2.1 [93] with the same configuration as for Z/W+jet.

The SM process $t\bar{t}$ plus a highly energetic photon ($t\bar{t}+\gamma^{*}$) is not a background process, but is used for the estimation of the $t\bar{t}+Z$ background, as described in Sec. 5.6.3. Events for this process were generated with MG5AMC@NLO [88] interfaced to PYTHIA8 [89] for PS and hadronisation.

An overview of the configurations for each MC sample can be seen in Tab. 5.2. Further information about the background samples can be found in [104–108].

The detector simulation for all background samples was done with GEANT4 for all parts of the detector [109]. For signal samples the simulation was mostly done with a fast simulation framework which uses parametrised descriptions for showers in the electronic and hadronic calorimeters [110] and GEANT4 for the rest of the detector. This choice was made for economic reasons: while background samples can be produced centrally and then used in many analyses, each analysis has a separate demand for signal samples, often in high

Table 5.2: Overview of the configurations used for the event generation, showering, parton distribution function (PDF) and underlying-event (UE) of the MC samples.

Process	Generator	Showering	PDF set	UE tune	Precision
SUSY	MG5AMC@NLO 2.2-2.4	PYTHIA8 & EVTGEN 1.2.0	NNPDF2.3LO [97]	A14 [101]	NLO+NLL
W/Z + jets	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF3.0NNLO [97]	SHERPA default	NLO
Diboson	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF3.0NNLO [97]	SHERPA default	NLO
$t\bar{t}$	POWHEG-BOX 2	PYTHIA 6.428	CT10 (NLO) [102]	PERUGIA 2012 [103]	NNLO+NNLL
Single top:					
– t-channel	POWHEG-BOX 1	PYTHIA 6.428	CT10 F4 (NLO) [102]	PERUGIA 2012 [103]	NNLO+NNLL
– s-/Wt-channel	POWHEG-BOX 2	PYTHIA 6.428	CT10 (NLO) [102]	PERUGIA 2012 [103]	NNLO+NNLL
$t\bar{t}+V$	MG5AMC@NLO 2.2.3	PYTHIA 8.186	NNPDF3.0NNLO [97]	A14 [101]	NLO
$t\bar{t}+\gamma^{*}$	MG5AMC@NLO 2.2.3	PYTHIA 8.186	NNPDF3.0NNLO [97]	A14 [101]	NLO

numbers with different sets of model parameters; full GEANT4 simulation is computationally expensive and thus not practicable for all signal samples. For few selected signal models also samples with full simulation were produced, the differences compared to the fast simulation were seen to be small. In order to account for pp interactions in the same or nearby bunch crossings (*pile-up*) additional minimum-bias events were produced with PYTHIA8 with the A2 tune [111] and MSTW 2008 PDF set [112]. All samples were overlaid with a varying number of these events such that the number of pp interactions per bunch crossing matches the distribution in data. Events are then reconstructed with the same methods as used for data. Scale factors are used to correct for differences between data and simulation (reconstruction, identification, isolation and b-tagging efficiencies, pile-up) and to normalise the samples to the actual cross-section and integrated luminosity. These scale factors are applied as weights to the simulated events.

5.3 Object definitions

Since this search is looking only for fully hadronic final states, *jets* (including *b-jets*) and *missing transverse energy* should be the only reconstructed physics objects in the signal regions. However, also other objects are needed: *Electrons* and *muons* are vetoed against, also they are selected in 1- and 2-lepton control regions. A simple veto for hadronically decaying *taus* is also applied without using the full reconstruction. Leptonically decaying taus are not identified as such, their signature is simply a muon or electron plus $E_{\mathrm{T}}^{\mathrm{miss}}$. *Photons* are needed for the $t\bar{t}+V$ control region. An *overlap removal* is applied to ensure that a signature which satisfies several object definitions is not double counted. The details of these definitions are described in the following.

Jets

Jets are reconstructed from three-dimensional topological clusters of calorimeter cells. The reconstruction is done with the anti-k_t algorithm with a distance parameter $R = 0.4$. After a pile-up correction and calibration they have to satisfy $p_T > 20\,\mathrm{GeV}$ to be counted as a *baseline* jet. This definition is only used for the overlap removal. *Signal* jets have to fulfil additional quality criteria: they need to be within $|\eta| < 2.8$ and must pass the jet vertex tagger if they fulfil $p_T < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ to reject jets from pile-up interactions. Events with jets that are labelled as "bad" due to originating from detector noise or non-collision sources are removed [79].

b-jets

Jets labelled as b-jets must satisfy all criteria of a regular signal jet. In addition they must be within the inner detector acceptance ($|\eta| < 2.5$) and pass the 77% fixed-cut working point of the MV2c10 algorithm. Jets which are not tagged as b-jets are labelled as *light jets*.

Electrons

The baseline definition requires electrons to have $p_T > 7\,\mathrm{GeV}$ and $|\eta| < 2.47$ and satisfy the *very loose* identification requirement. This definition is used for a veto in signal regions.

Signal electrons, which are used for control region definitions, must pass the *Tight* identification criteria and satisfy $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.47$. In addition they must pass the *GradientLoose* isolation working point.

Muons

Baseline muons are required to have $p_T > 6\,\mathrm{GeV}$ and $|\eta| < 2.7$ and pass the *Loose* selection criteria. Signal muons, used for control regions, must additionally fulfil $p_T > 20\,\mathrm{GeV}$ and pass the *Medium* identification criteria and the *GradientLoose* isolation working point. Additional criteria are required to reject cosmic ($|d_0| > 0.2\,\mathrm{mm}$ and $|z_0| > 1\,\mathrm{mm}$) or poorly reconstructed (labelled as "bad") ($\sigma(q/p)/|q/p| > 0.2$) are rejected[5].

Tau veto

A simple, loose definition is used for the "τ veto": if a non-b-tagged jet within $|\eta| < 2.5$ has fewer than four tracks associated with it and the $\Delta\phi$ between the jet and the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ is less than $\pi/5$, the jet is labelled as a "tau candidate". This veto targets only the main source of τ leptons in this analysis, which is the $W \to \tau\nu$ decay with no additional source of $E_{\mathrm{T}}^{\mathrm{miss}}$, it cannot be seen as a veto against τ leptons in general.

Photons

Photons are required to be within $|\eta| < 2.37$ and satisfy the *Tight* identification criteria. In addition the *FixCutLoose* working point is used for isolation.

[5] d_0 is the transverse impact parameter of a track with respect to the primary vertex, z_0 is the distance of the point used for the d_0 measurement from the primary vertex in z-direction, $\sigma(q/p)/|q/p|$ is a measure of the momentum uncertainty for a particle with charge q.

Overlap removal

The first step of the overlap removal (OR) is to handle electrons overlapping with jets. Different methods are applied for light jets and b-jets. If the separation between a baseline electron and a light jet is $\Delta R < 0.2$, the light jet is removed and only the electron is kept. In case of a b-jet under the same conditions the jet is kept instead. This method is used so that prompt electrons are kept when they that are also identified as jets, without losing efficiency for b-jets, which would otherwise be removed..

The next step is to remove overlap of muons and jets. Non-prompt muons can be produced from in-flight decays of hadrons, in that case we want to keep the jet, whereas prompt muons should not be removed. Hadronic jets typically contain several tracks, which is used to distinguish between the two cases: if the separation between a baseline muon and a jet is $\Delta R < 0.4$, the muon is removed if the jet has at least three tracks associated with it; if it has less than three tracks, the muon is kept and the jet is removed.

The last step again handles the OR of jets and electrons, however this is done after the second step and only surviving jets are considered. If the separation between an electron and a jet is $0.2 < \Delta R < 0.4$ the electron is removed as it is likely to be a decay product of a hadron.

Missing transverse energy

The missing transverse momentum is calculated with all reconstructed objects plus a soft track-based term, as described in Sec. 4.7. The soft term is calculated from all inner-detector tracks originating from the primary vertex with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ that are not associated with any reconstructed physics object. The missing transverse momentum vector is called $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, the magnitude or 'missing transverse energy' is called $E_{\mathrm{T}}^{\mathrm{miss}}$.

Another variant of the missing transverse momentum has a minor role in this analysis: the purely track-based $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss,track}}$. It is calculated only from tracks associated with the primary vertex with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ and $|\eta| < 2.5$, without any other reconstructed objects. The reason for using this is that events without genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ (e.g. multijet and fully hadronic $t\bar{t}$ events) can enter the signal regions due to mismeasured jets or pile-up contributions. If an event has genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, it is expected to have also a significant amount of $E_{\mathrm{T}}^{\mathrm{miss,track}}$ and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ is expected to be be aligned with $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss,track}}$. This can be used to reject backgrounds without genuine $E_{\mathrm{T}}^{\mathrm{miss}}$.

5.4 Discriminating variables

Apart from the quantity and momenta of the reconstructed objects, further more complex observables are used to discriminate different processes in signal and control regions. These variables are described in the following. Transverse momenta p_{T}^{i} refer, if not indicated otherwise, to jets ordered by their p_{T}, starting with the p_{T}-leading jet with $i = 0$. Other observables are introduced in the following and some of their distributions in data and simulation are depicted in Fig. 5.10 and 5.11. More observables which are used exclusively in one signal region are introduced in Sec. 5.5.

Distances between objects

Angular distances can be calculated in the transverse plain as $\Delta\phi$, which is mainly done if the polar information is not available (when one of the objects is $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$). If the polar information is available for both objects, we can use the full information to calculate ΔR, which is defined as

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{5.1}$$

The distances between objects can give discrimination power against events with certain decay chains, because particles that are decay products of the same mother particle tend to have a smaller separation. One case where this is used is the difference between two b-jets, $\Delta R(b,b)$. This quantity is expected to be small when the b quarks are products of a gluon splitting, which is often the case for the background processes W+jet and Z+jet. If there are more than two b-jets the two with the highest b-tagging weights are used. This observable is shown in Fig. 5.11a.

Another case where ΔR is useful is to distinguish events with a leptonically decaying top from events with only a leptonically decaying W by using the distance between the b-jet and the lepton, which is expected to be small for a top but not if the b-jet originates from somewhere else. In events with two b quarks – like $t\bar{t}$ – only one b-jet is expected to be close to the lepton, thus the minimum distance of the lepton from the two b-jets with the highest b-tagging weights $\Delta R(b,l)_{\text{min}}$ is used.

Masses of reclustered large-R jets

In order to reconstruct hadronically decaying massive objects one can reconstruct jets with a larger radius. A "rule of thumb" [113] for the separation of the products of a two-body decay of an object with mass m and transverse momentum p_{T} is given by

$$\Delta R \approx \frac{2m}{p_{\text{T}}}, \tag{5.2}$$

this means that the separation of decay products becomes small if the decaying object is highly boosted and it can be reconstructed as a single jet; the R parameter of the jet reconstruction algorithm should be equal or higher than the ΔR of the decay products. The method chosen for this is to use the anti-k_t algorithm again with our regular $R = 0.4$ jets as input, but now with a larger R parameter. The main focus in this analysis is the reconstruction of top quarks, but also W bosons can be identified. As an identifier for these objects the mass of the large-R jets can be used. Two variants of collections of these jets are used in this analysis: reclustered jets with $R = 1.2$, which are mainly used to identify top quarks, and with $R = 0.8$ which are used to identify W bosons or very highly boosted top quarks. The masses of these jets labelled as $m^i_{\text{jet,R=1.2}}$ or $m^i_{\text{jet,R=0.8}}$, where i is the position of the jet in the p_{T} ordering of the corresponding jet collection. These masses are used to define the top categories in SRA and SRB: Events that satisfy $m^0_{\text{jet,R=1.2}} > 120\,\text{GeV}$ and $m^1_{\text{jet,R=1.2}} > 120\,\text{GeV}$ are categorised as **TT**, if they satisfy $m^0_{\text{jet,R=1.2}} > 120\,\text{GeV}$ and $120\,\text{GeV} < m^1_{\text{jet,R=1.2}} < 120\,\text{GeV}$ they are categorised as **TW** and if they satisfy $m^0_{\text{jet,R=1.2}} > 120\,\text{GeV}$ and $m^1_{\text{jet,R=1.2}} < 60\,\text{GeV}$ they are categorised as **T0**. **T** stands for an object that is likely to be a to, **W** for an object likely to be a W boson and **0** for an object with a comparably low mass. This notation will be used in the remainder of this thesis.

The distributions of the reclustered jet masses is shown in Fig. 5.10b, 5.10c and 5.10d. It can be seen that for signal events these reconstructed masses tend to cluster around the W-boson and the top-quark mass.

Invariant and transverse masses

If two (or more) particles are the decay products of a massive particle, we can calculate their *invariant mass* to identify the mother particle. In general the invariant mass of two particles with energies E_1, E_2 and momenta $\mathbf{p}_1$, $\mathbf{p}_2$ is given by

$$m_{12}^2 = (E_1 + E_2)^2 - |\mathbf{p}_1 + \mathbf{p}_2|^2 = m_1^2 + m_2^2 + 2(E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2). \tag{5.3}$$

In this analysis this is mainly used in the Z+jet control region to identify Z bosons decaying into an electron-positron or muon-antimuon pair. In this case the mass of the leptons is negligible and the invariant mass can be calculated as

$$m_{ll}^2 = 2(E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2) = 2 p_{\text{T}1} p_{\text{T}2} \left(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)\right), \tag{5.4}$$

where $p_{\text{T},i}$, η_i, ϕ_i ($i \in \{1,2\}$) are the properties of the two daughter particles.

In the cases where one part of the decay products is invisible one does not have the full information to calculate the invariant mass, because we have only the transverse momentum of the invisible particles. In the case of one visible and one invisible particle, which are assumed to have negligible masses, we can calculate the *transverse mass* as

$$m_{\mathrm{T}}^2 = 2(E_{\mathrm{T}1}E_{\mathrm{T}}^{\mathrm{miss}} - \mathbf{p}_{\mathrm{T}1}\cdot\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) = 2p_{\mathrm{T}1}E_{\mathrm{T}}^{\mathrm{miss}}\left(1-\cos(\phi_1-\phi_{\mathrm{miss}})\right). \tag{5.5}$$

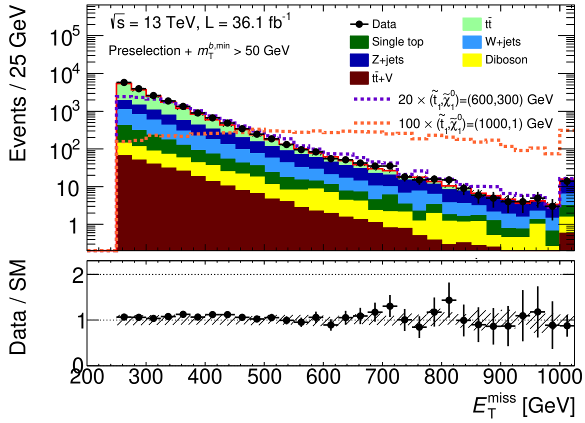
Comparing this with Eq. 5.4 one can see that the transverse mass is exactly the invariant mass if the pseudorapidity of both particles is the same, and smaller if this is not the case.

There are several important cases where this is used in the analysis. The first case is the transverse mass of the lepton and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$– denoted as $m_{\mathrm{T}}(l,\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ – in one-lepton regions. If the lepton and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ both originate from a W decay, this observable is expected to be below the W mass in contrast to events with several sources for $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, where a also higher values are expected. In addition, multijet events with a misidentified lepton and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ tend to have very low values.

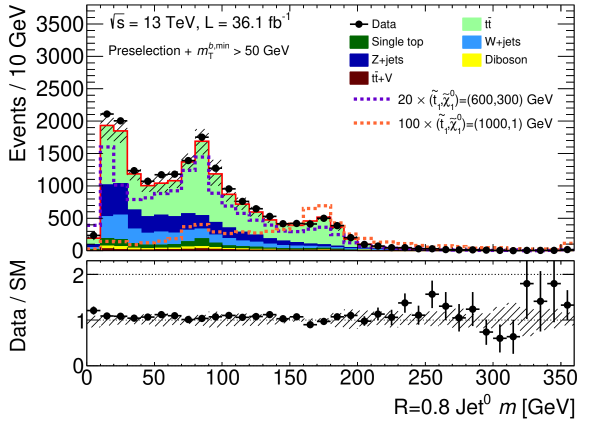
Another important case for the usage of the transverse mass is $m_T(b^{\mathrm{min}},\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ (in the following written as $m_{\mathrm{T}}^{b,\mathrm{min}}$) – the transverse mass of the b-jet which has the smallest ϕ separation with $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. In events with one leptonically decaying top this variable is expected to be below the top mass, because the transverse mass of two of the decay products (b quark and neutrino) cannot be higher than the mass of the original particle's mass, especially since the third decay particle is not taken into account at all; this can be seen in Fig. 5.10e. The variable $m_{\mathrm{T}}^{b,\mathrm{max}}$ is constructed equivalently with the b-jet with the largest $\Delta\phi(b,\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$. The distributions of $m_{\mathrm{T}}^{b,\mathrm{min}}$ and $m_{\mathrm{T}}^{b,\mathrm{max}}$ are shown in Fig. 5.10e and 5.10f.

Stransverse masses

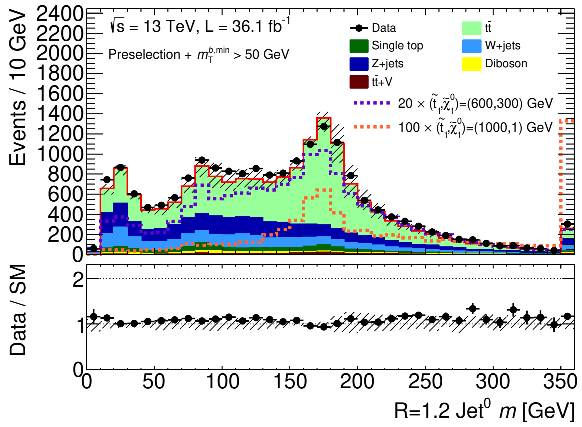
The stransverse mass $m_{\mathrm{T}2}$ is a special case of the transverse mass that aims to reconstruct the masses of a pair of particles which both decay partly invisibly [114]. This is done by splitting $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ into two parts $\mathbf{p}_1$ and $\mathbf{p}_2$ so that two transverse masses can be calculated, each constructed with of one of these parts. One scans over the possible choice for $\mathbf{p}_1$ and $\mathbf{p}_2$ and



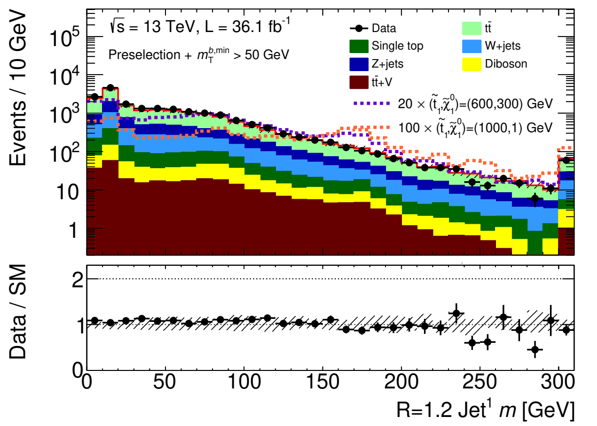
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$.



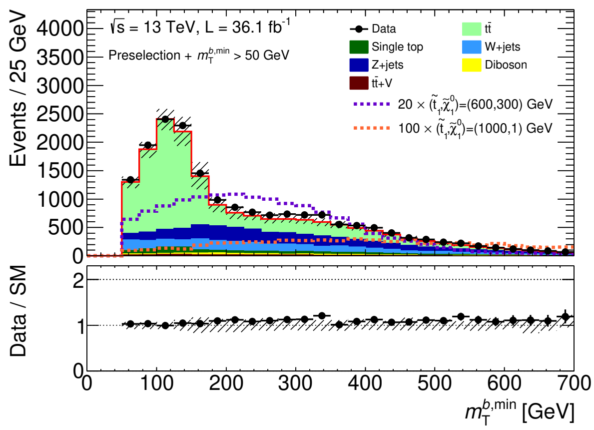
(b) Mass of p_{T}-leading $R = 0.8$ jet.



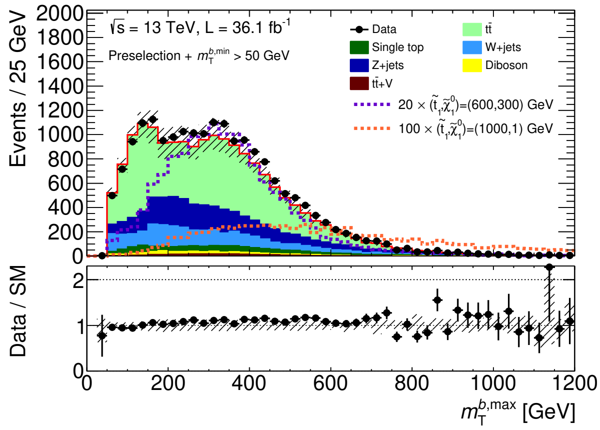
(c) Mass of p_{T}-leading $R = 1.2$ jet.



(d) Mass of p_{T}-sub-leading $R = 1.2$ jet.

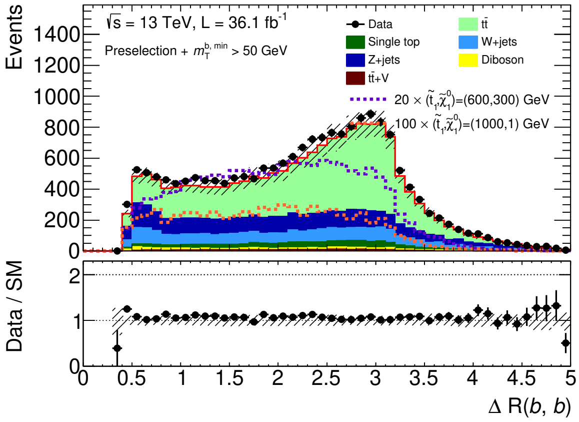


(e) $m_{\mathrm{T}}^{b,\mathrm{min}}$.

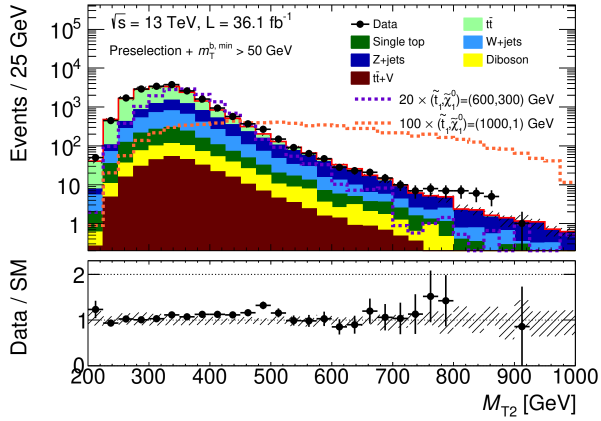


(f) $m_{\mathrm{T}}^{b,\mathrm{max}}$.

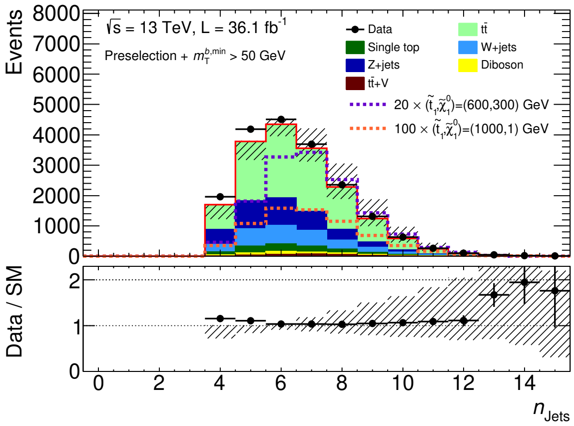
Figure 5.10: Distributions of several observables after preselection plus additional $m_{\mathrm{T}}^{b,\mathrm{min}} > 50\,\mathrm{GeV}$ requirement. Data events are shown in black, the error bars symbolise Poisson uncertainties. The SM background processes are depicted as coloured areas stacked on each other. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed lines correspond to SUSY signals with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1000, 1)\,\mathrm{GeV}$ (orange) scaled by a factor of 100 and $m(\tilde{t}_1, \tilde{\chi}_1^0) = (600, 300)\,\mathrm{GeV}$ (violet) scaled by a factor of 20, both with $\mathrm{BR}(\tilde{t}_1 \to t\tilde{\chi}_1^0) = 100\%$. The bottom panel shows the ratio of data and the sum of simulated SM events.



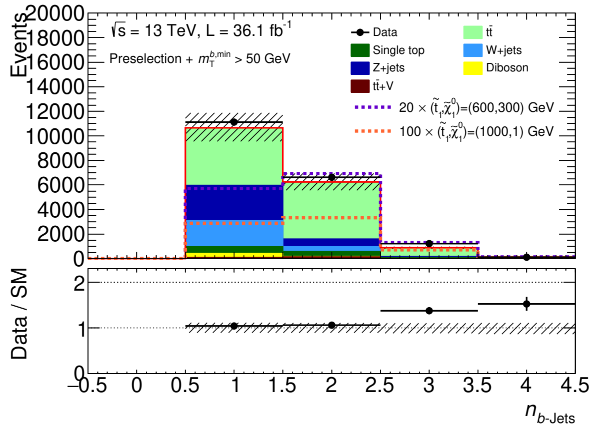
(a) $\Delta R(b,b)$



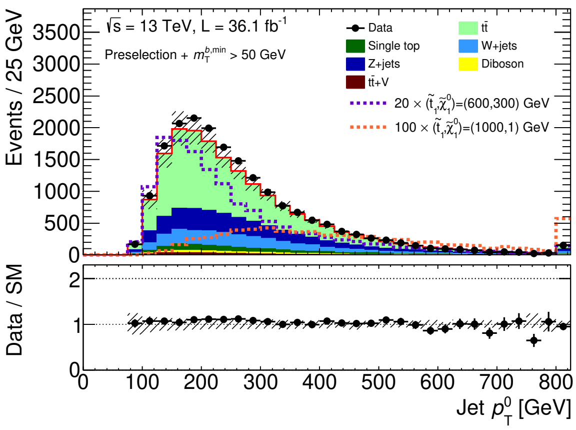
(b) m_{T2} with chi-squared method.



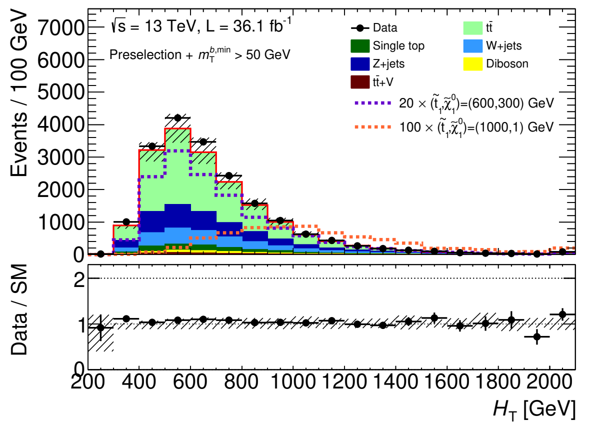
(c) Number of jets



(d) Number of b-jets.



(e) Leading jet p_T.



(f) Sum of all jet p_T's.

Figure 5.11: Distributions of several observables after preselection plus additional $m_T^{b,\min} > 50\,\text{GeV}$ requirement. Data events are shown in black, the error bars symbolise Poisson uncertainties. The SM background processes are depicted as coloured areas stacked on each other. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed lines correspond to SUSY signals with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1000, 1)\,\text{GeV}$ (orange) scaled by a factor of 100 and $m(\tilde{t}_1, \tilde{\chi}_1^0) = (600, 300)\,\text{GeV}$ (violet) scaled by a factor of 20, both with $\text{BR}(\tilde{t}_1 \to t\tilde{\chi}_1^0) = 100\%$. The bottom panel shows the ratio of data and the sum of simulated SM events.

chooses the combination that minimises the larger of the two transverse masses:

$$m_{\mathrm{T2}}^2 = \min_{\mathbf{p}_1+\mathbf{p}_2=\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}} \left[\max\left\{m_{\mathrm{T}}^2(\mathbf{p}_1,\mathbf{q}_1), m_{\mathrm{T}}^2(\mathbf{p}_2,\mathbf{q}_2)\right\}\right]. \tag{5.6}$$

This usually leads to a combination where both transverse masses are roughly equal. Here one does not necessarily need to assume massless decay products. The visible decay products can be very massive and one can also use an assumption for the mass of the invisible particle, if it is known in the scenario under consideration. This expands Eq. 5.5 to

$$m_{\mathrm{T}}^2(\mathbf{p},\mathbf{q}) = m_p^2 + m_q^2 + 2p_{\mathrm{T}}q_{\mathrm{T}}\left(1-\cos(\phi_p-\phi_q)\right). \tag{5.7}$$

In the case of the process depicted in Fig. 5.1a the visible particles $q_{1,2}$ would be the top quarks (which are fully visible in the fully hadronic channel) and the top-squark mass would be the upper limit for m_{T2}. This makes it especially useful for high-mass scenarios (SRA), because for background events m_{T2} is expected to be lower.
Using tops as the visible input is not a trivial thing to do since tops are not well defined objects on reconstruction level. A possible choice are the p_{T}-leading large-R jets $\mathrm{jet}_{\mathrm{R}=1.2}^{0,1}$ or $\mathrm{jet}_{\mathrm{R}=0.8}^{0,1}$. However since we are using also events with only one well-reconstructed top with this method in the TW and T0 category, it can be assumed that these jets will not give a good description of the original top there. Another choice could be to use a combination of light and b-jets (or just the two p_{T}-leading jets) as top candidates, based on the ΔR or invariant mass of said jets. For this there are a lot of possibilities[6], e.g. the method used in SRA in [3] or methods that find top- (and W-)candidates based on the optimisation of a chi-squared-like cost function. The latter works in principle by choosing the combination with the smallest

$$\chi^2 = \sum_i \frac{(m_{\mathrm{cand},i} - m_{\mathrm{true},i})^2}{f_i^2}, \tag{5.8}$$

where i sums over all candidates that are looked for with the mass of the searched particle $m_{\mathrm{true},i}$ and the invariant mass of the combined jets $m_{\mathrm{cand},i}$. f_i should be an estimation for

[6] A suitable method should be able to produce two top candidates for all events with ≥ 4 jets and ≥ 2 b-jets.

the uncertainty of $m_{\text{true},i}$, but can also be chosen differently if necessary. The quality of these methods can be judged with simulated events by the kinematic differences between the resulting top candidates and the top quarks on generator level, but also by the performance of m_{T2} with these input candidates in the signal region optimisation.

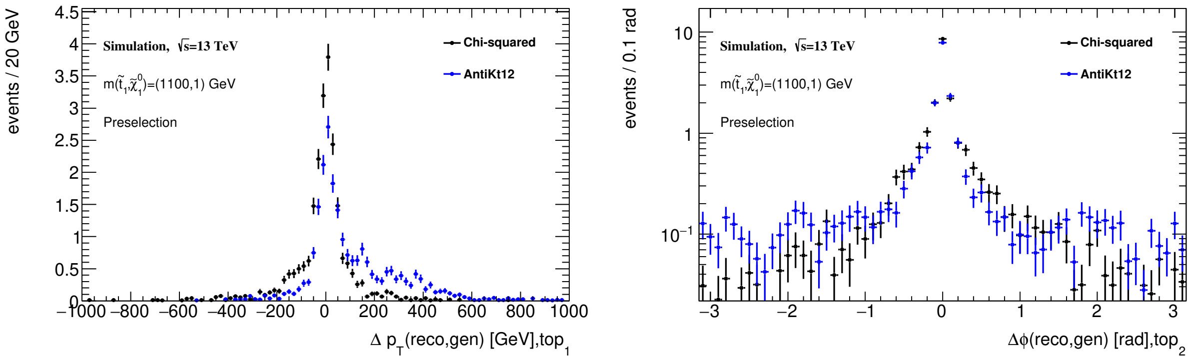
The method performing the best, which will be referred to as *chi-squared method* below, is described in the following. We are searching for two top candidates, but also two W candidates. A W candidate can consist of either one or two light jets[7], a top candidate consists of a W candidate and a b-jet. In this context only the two jets with the highest b-tagging weight are considered as b-jets, all others are considered light jets[8]. The advantage of this method is, that it constructs top candidates with equal priority while making use of the full event information and knowledge about the decay process, without being restricted to small ΔR's of the decay products. As f_i a proxy for the resolution was used: $f_i = \sqrt{m_{\text{true},i}/\text{GeV}}\,\text{GeV}$, which estimates only very roughly the mass dependency of the resolution. With that, Eq. 5.8 results in the following cost function:

$$\begin{aligned}\chi^2 =& \frac{(m^{\text{top}}_{\text{cand},1} - 173.2\,\text{GeV})^2}{173.2\,\text{GeV}^2} + \frac{(m^{\text{top}}_{\text{cand},2} - 173.2\,\text{GeV})^2}{173.2\,\text{GeV}^2} \\ &+ \frac{(m^{W}_{\text{cand},1} - 80.4\,\text{GeV})^2}{80.4\,\text{GeV}^2} + \frac{(m^{W}_{\text{cand},2} - 80.4\,\text{GeV})^2}{80.4\,\text{GeV}^2}\end{aligned} \tag{5.9}$$

This method was seen to be superior to other options which neglect W candidates or allow other combinations as W candidates. Since this choice for f_i is not highly motivated, other options for were tried, including other fixed values, but also using the uncertainty of $m_{\text{cand},i}$ calculated from propagating the uncertainties of the contained jets. It was also tried to limit the penalty for not finding W-candidates by using $(\max(m^{W}_{\text{cand},1}, m_{\text{fix}}) - 80.4\,\text{GeV})^2$ in the numerator, with different values for m_{fix}. All these methods led in practise to no or only an insignificant improvement over Eq. 5.9. The kinematic differences of the top candidates to the generator level properties in comparison to the $R = 1.2$ jets are depicted in Fig. 5.12 and 5.13. When comparing the two methods in a preselection without any top

[7] A W boson decays into two quarks, but the particle showers of those can merge into a single jet.

[8] This choice was made to ensure that the method works for all events in the signal region, because ≥ 2 light jets is not a requirement.

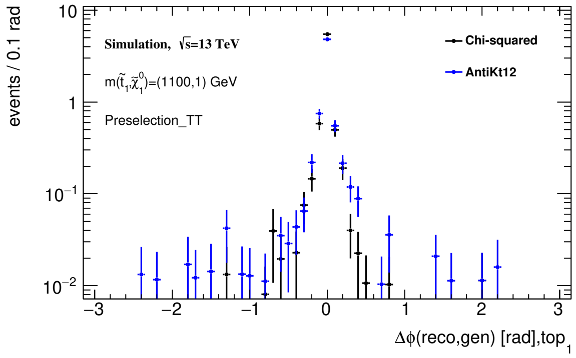


(a) Δp_{T},top$_1$ in preselection. **(b)** $\Delta\phi$,top$_2$ in preselection.

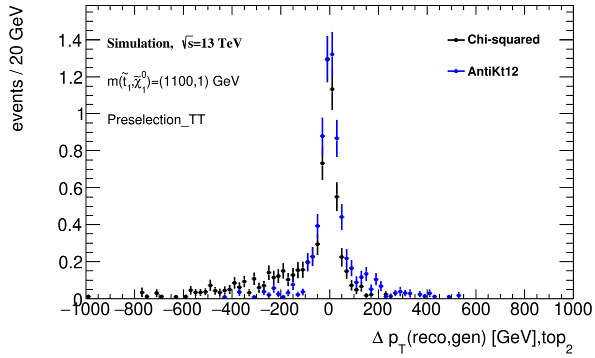
Figure 5.12: Kinematic differences of reconstructed top candidates with generator-level top quarks in simulated SUSY events with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1100, 1)$ GeV. The preselection is the same as in Sec. 5.5. Top$_1$/top$_2$ are defined as the p_{T}-leading/sub-leading $R = 1.2$ jets for the AntiKt12 method (blue) and the top-candidate with the lower/higher contribution to χ^2 for the chi-squared method (black). The differences are calculated with the generator-level top which is closer in ΔR to the top candidate. The lower panel shows the ratio of events between the chi-squared and the AntiKt12 method.

identification applied, the chi-squared method performs better in describing the p_{T} and ϕ of the original top. In a selection with the TT categorisation applied, the $R = 1.2$ jets perform unsurprisingly equally well or slightly better than the chi-squared method, but the difference is not big. However, in regions with the TW or T0 categorisation applied the $R = 1.2$ jets fail to provide a good description of both jets, whereas the chi-squared methods still performs reasonably well. It might at first be surprising, that even the "reconstructed" top jet (with $m_{\mathrm{jet,R=1.2}} > 120$ GeV) does not give a good description for the p_{T} of the top (see Fig 5.13e). The reason for this is that decay products of the other, not-reconstructed top can merge into the jet and increase its overall momentum. It is obvious from these plots that the chi-squared method is much better suited to describe the kinematic properties of the tops in SRA. It is however not suited to use the reconstructed top mass as a direct discriminator or to define signal region categories, because it tends to construct masses close to the top mass even for non-top backgrounds by design because of the method to select the 'test' combination.

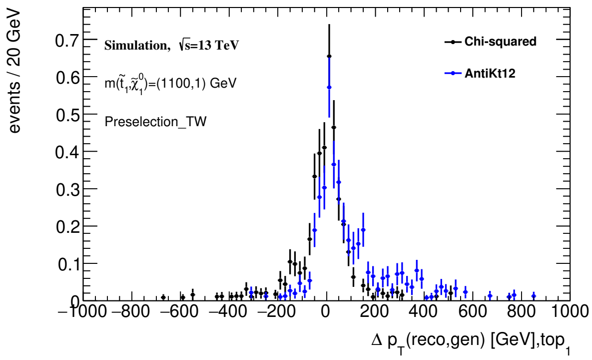
In addition to the chi-squared method, it was seen to be the most useful to use 0 GeV (as a proxy for the unknown neutralino mass) and 173.2 GeV (the top mass) as m_p and m_q in Eq. 5.7, the construction with these definitions will be used and referred to as 'm_{T2}'



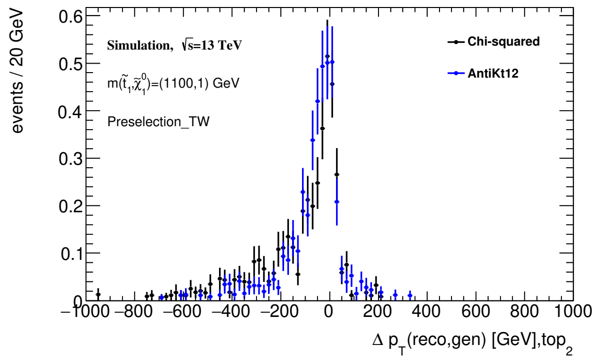
(a) $\Delta\phi$,top$_1$ in TT.



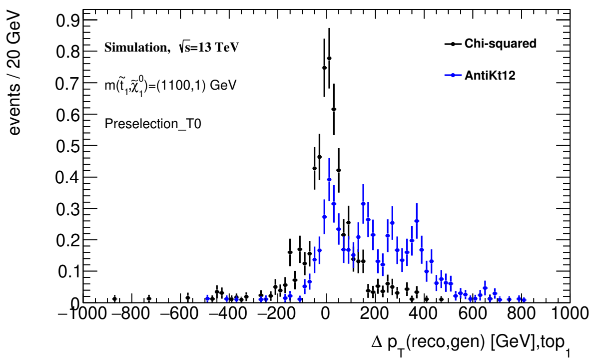
(b) Δp_{T},top$_2$ in TT.



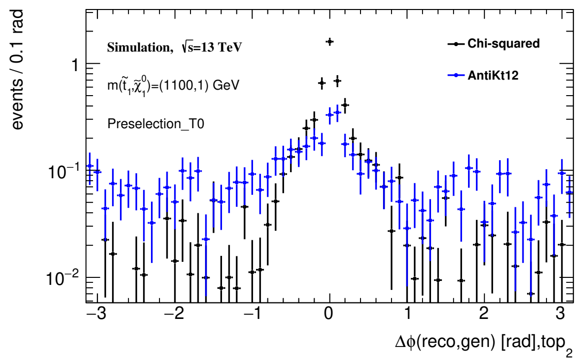
(c) Δp_{T},top$_1$ in TW.



(d) Δp_{T},top$_2$ in TW.



(e) Δp_{T},top$_1$ in T0.



(f) $\Delta\phi$,top$_2$ in T0.

Figure 5.13: Kinematic differences of reconstructed top candidates with generator-level top quarks in simulated SUSY events with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1100, 1)\,\mathrm{GeV}$. The preselection is the same as in Sec. 5.5; TT and T0 means that the corresponding top identification categorisation was applied in addition to the preselection. Top$_1$/top$_2$ are defined as the p_{T}-leading/sub-leading $R = 1.2$ jets for the AntiKt12 method (blue) and the top-candidate with the lower/higher contribution to χ^2 for the chi-squared method (black). The differences are calculated with the generator-level top which is closer in ΔR to the top candidate. The lower panel shows the ratio of events between the chi-squared and the AntiKt12 method.

in the following, its distribution is shown in Fig. 5.11b. A different mass hypothesis for neutralinos just tends to shift the resulting m_{T2} without changing the background rejection power, whereas a different 'top-mass' hypothesis (e.g. using the reconstructed mass of the candidate) tends to smear out m_{T2}, so that other choices were not seen to be of advantage for any signal model.

5.5 Signal regions

A common preselection is applied for all signal regions: a basic event cleaning is applied, which makes sure that the event has a primary vertex, belongs to a run that is white-listed by the GRL, is not vetoed by error flags of the calorimeter or SCT modules and does not have a 'bad' jet or a 'bad' or 'cosmic' muon, as defined in Sec. 5.3; the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used for all of them, together with an offline $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold of 250 GeV for full trigger efficiency; at least four jets with $p_{\mathrm{T}}^{0-3} > (80, 80, 40, 40)$ GeV are required[9]; at least one jet should be b-tagged; events with leptons fulfilling the baseline electron or muon requirement are rejected; an angular separation between the two p_{T}-leading jets and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ is required to reject events with artificial $E_{\mathrm{T}}^{\mathrm{miss}}$ from mismeasured jets: $\left|\Delta\phi\left(\mathrm{jet}^{0,1}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right| > 0.4$; further selections on $E_{\mathrm{T}}^{\mathrm{miss,track}}$ are used to reject more events with non-genuine $E_{\mathrm{T}}^{\mathrm{miss}}$: $E_{\mathrm{T}}^{\mathrm{miss,track}} > 30$ GeV and $\left|\Delta\phi\left(\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}), \mathbf{p}_{\mathrm{T}}^{\mathrm{miss,track}}\right)\right| < \pi/3$.

Several signal regions are then defined with additional selection criteria in order to enhance the signal/background ratio for different scenarios such that the expected discovery significance is optimised. All signal regions were defined only based on simulation, data was blinded during this process in order to avoid biases.

[9] This requirement was chosen such that it rejects background events while still preserving as much signal in all scenarios as possible.

Table 5.3: Common preselection for all signal regions.

Observable	**Requirement**
Event cleaning	✓
$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger	✓
N_{jets}	≥ 4
$\mathrm{Jet}^{0-3}\, p_{\mathrm{T}}$	$> (80, 80, 40, 40)\,\mathrm{GeV}$
$N_{b-\mathrm{jets}}$	≥ 1
$N_{\mathrm{baseline-leptons}}$	$= 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 250\,\mathrm{GeV}$
$\left\|\Delta\phi\left(\mathrm{jet}^{0,1}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right\|_{\mathrm{min}}$	> 0.4
$E_{\mathrm{T}}^{\mathrm{miss,track}}$	$> 30\,\mathrm{GeV}$
$\left\|\Delta\phi\left(\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss,track}}\right)\right\|$	$< \pi/3$

Signal region A

Signal region A is optimised for a high sensitivity in the scenario depicted in Fig. 5.1a with high top-squark masses, the benchmark scenario is $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1000, 1)\,\mathrm{GeV}$. At least two b-jets are required in this selection and a minimum distance of the leading three jets to $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$: $\left|\Delta\phi\left(\mathrm{jet}^{0,1,2}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right| > 0.4$. A threshold of $m_{\mathrm{T}}^{b,\mathrm{min}} > 200\,\mathrm{GeV}$ is applied to reject semi-leptonic $t\bar{t}$ events with a missed or misidentified e or μ, and the tau veto described in Sec. 5.3 is applied to reject events containig a $W \to \tau\nu$ decay. For W- and top-identification large-R jets are used: Due to the large $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ the decay products are highly boosted and with the $R = 0.8$ jet collection it is usually possible to identify at least one W boson; this is done by requiring $m > 60\,\mathrm{GeV}$ for the p_{T}-leading $R = 0.8$ jet. The $R = 1.2$ jet collection is used for the categorisation into the three signal bins **TT**, **TW** and **T0**: the p_{T}-leading jet is required to have $m > 120\,\mathrm{GeV}$, which indicates a top quark; events without such a top candidate are discarded. The sub-leading jet can then fall into three categories: $m > 120\,\mathrm{GeV}$ (TT), $m = [60, 120]\,\mathrm{GeV}$ (TW) or $m < 60\,\mathrm{GeV}$ (T0). The reason for this is that while one top is usually boosted enough to be identified as a ($R = 1.2$) jet, the other top can also be softer, so that only the decay product of the W boson or not even that are contained in one jet. The binning is used because one does not want to lose the rejection power of two identified tops, but also not lose all signal events were not both tops could be identified. For the same reason

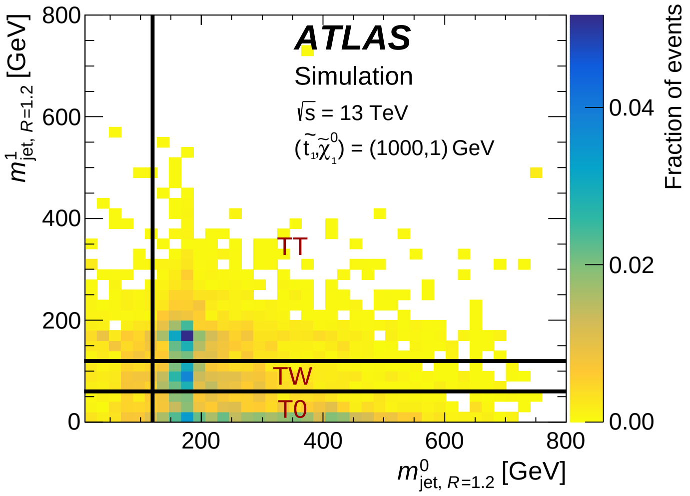


Figure 5.14: Distribution of the masses of the p_{T} leading/subleading $R = 1.2$ jets for the SRA benchmark signal ($m(\tilde{t}_1, \tilde{\chi}^0_1) = (1000, 1)$ GeV) after preselection. The areas of the TT, TW and T0 categories used for the SRA and SRB definition are sketched in. For this signal model the fractions are: 34.1% TT, 22.1% TW, 32.4% T0 and the remainder discarded due to low low mass of the leading $R = 1.2$ jets; these numbers vary for different signal models.

the usage of the sub-leading $R = 0.8$ to identify another W candidate was seen to reject too many signal events with lesser boosted objects and was thus not considered useful. Figure 5.14 shows the distribution of the two $R = 1.2$ jet masses for signal events. It can be seen that that the highest density of events is at the point where both jets have masses around the top mass, but also a significant amount of events show up around the W mass for the sub-leading jet or below.

The selections were then optimised in each category with a *simulated annealing* algorithm implementation introduced in [115]. The figure of merit for the optimisation was the median expected discovery significance Z_0 calculated with the asymptotic formula [116, 117]:

$$Z_0 = \left[2\left((s+b)\ln\left[\frac{(s+b)(b+\sigma_b^2)}{b^2+(s+b)\sigma_b^2}\right] - \frac{b^2}{\sigma_b^2}\ln\left[1+\frac{\sigma_b^2 s}{b(b+\sigma_b^2)}\right]\right)\right]^{1/2}, \tag{5.10}$$

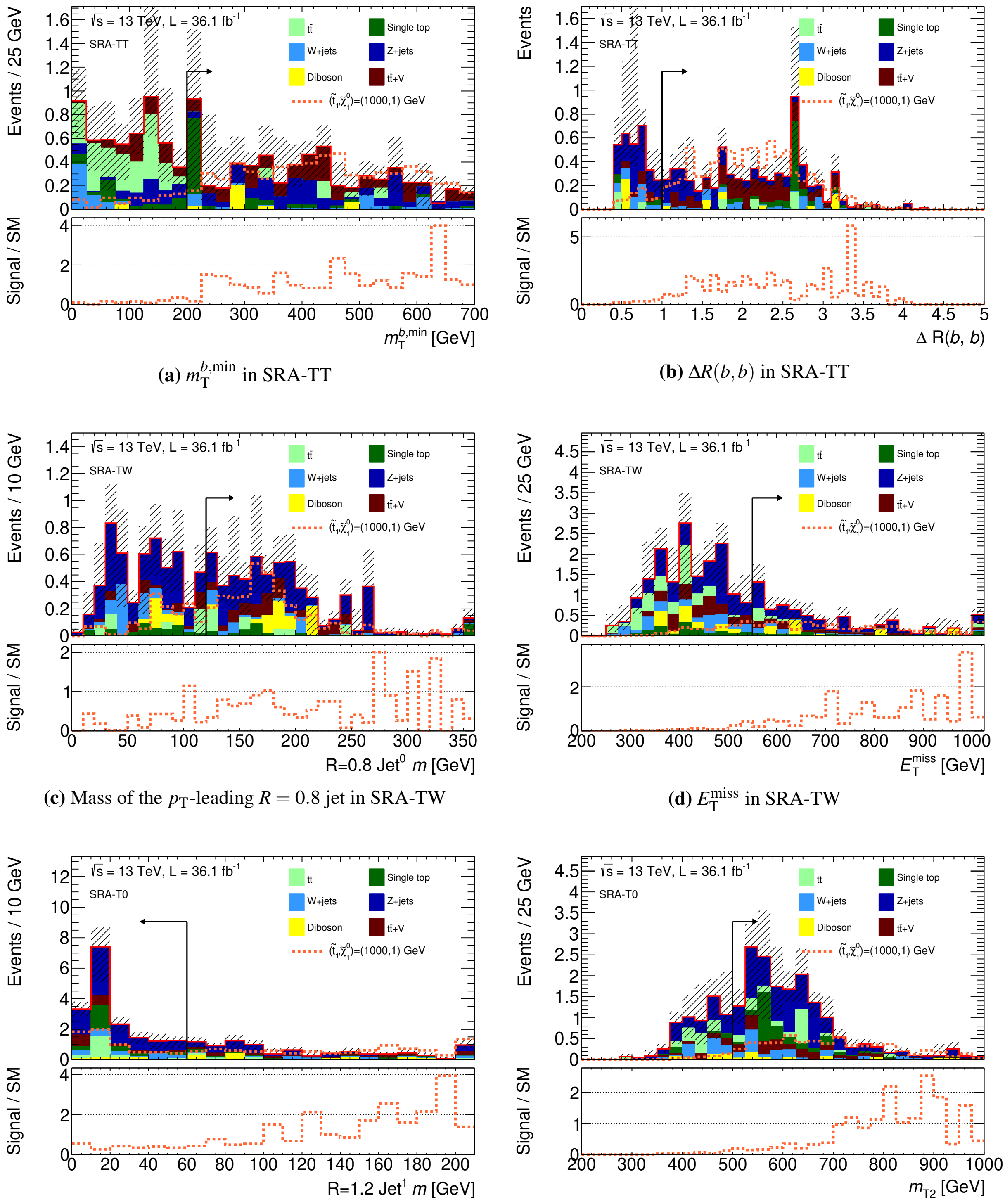
with s and b being the number of signal and background events and σ_b being the systematic uncertainty of the background yield. For σ_b a flat relative uncertainty of 30% was assumed. It was seen that the observables $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T2} were the most useful for optimising selections, whereas other observables did not significantly improve the sensitivity when changing the thresholds compared to the previously described values; only in the TT category an additional selection on $\Delta R(b,b) > 1.0$ was found to be beneficial. A summary of all criteria can be found in Tab. 5.4. Higher $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T2} thresholds could still have improved the expected significance in Eq. 5.10, but this was not done in the final SRA definition because of inaccurate background modelling in this extreme phase space. Fig. 5.15 shows the distributions of several used discriminating observables in SRA-TT , SRA-TW and SRA-T0.

Table 5.4: Selection criteria for SRA & SRB in addition to the common preselection defined in Tab. 5.3.

Signal Region	Observable	Requirement		
	Category →	TT	TW	T0
Common	$m^{0}_{\mathrm{jet},R=1.2}$	$> 120\,\mathrm{GeV}$		
	$m^{1}_{\mathrm{jet},R=1.2}$	$> 120\,\mathrm{GeV}$	$[60,120]\,\mathrm{GeV}$	$< 60\,\mathrm{GeV}$
	$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 200\,\mathrm{GeV}$		
	$N_{b-\mathrm{jets}}$	≥ 2		
	τ-veto	✓		
	$\left\|\Delta\phi\left(\mathrm{jet}^{0,1,2}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right\|_{\mathrm{min}}$	> 0.4		
A	$m^{0}_{\mathrm{jet},R=0.8}$	$> 60\,\mathrm{GeV}$		
	$\Delta R(b,b)$	> 1.0	-	
	m_{T2}	$> 400\,\mathrm{GeV}$	$> 400\,\mathrm{GeV}$	$> 500\,\mathrm{GeV}$
	$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 400\,\mathrm{GeV}$	$> 500\,\mathrm{GeV}$	$> 550\,\mathrm{GeV}$
B	$m_{\mathrm{T}}^{b,\mathrm{max}}$	$> 200\,\mathrm{GeV}$		
	$\Delta R(b,b)$	> 1.2		

Signal region B

This signal region targets scenarios with high LSP masses, the benchmark scenario is $m(\tilde{t}_1, \tilde{\chi}^0_1) = (600, 300)\,\mathrm{GeV}$. Typically all objects are less boosted than in SRA. SRB uses the same top categorisation method as SRA and a set of common selections. The usage of m_{T2} and $E_{\mathrm{T}}^{\mathrm{miss}}$ thresholds higher than preselection were not found to be useful, as well as the usage of $R = 0.8$ jets. Instead, $m_{\mathrm{T}}^{b,\mathrm{max}}$ and $\Delta R(b,b)$ were used to further suppress



(a) $m_{\mathrm{T}}^{b,\mathrm{min}}$ in SRA-TT

(b) $\Delta R(b,b)$ in SRA-TT

(c) Mass of the p_{T}-leading $R = 0.8$ jet in SRA-TW

(d) $E_{\mathrm{T}}^{\mathrm{miss}}$ in SRA-TW

(e) Mass of the p_{T}-sub-leading $R = 1.2$ jet in SRA-T0

(f) m_{T2} in SRA-T0

Figure 5.15: Distributions of several observables used in SRA. All selections of the signal region are applied except for the observable shown in the diagram. The SM background processes are depicted as coloured areas stacked on each other. Scale factors derived from control regions are not applied. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed line is the SUSY signal with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1000, 1)\,\mathrm{GeV}$. In the bottom panel the ratio of signal and SM events is shown.

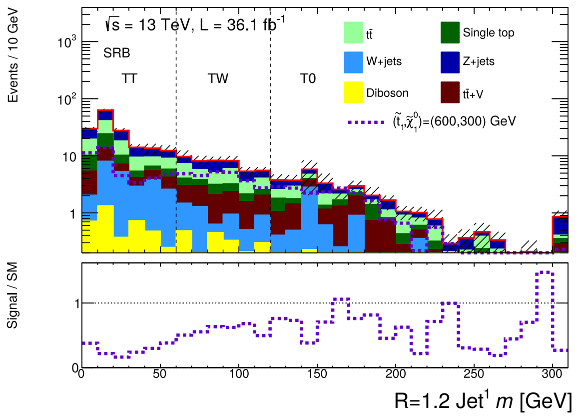
backgrounds. The summarised selection criteria can be found in Tab. 5.4. The distribution of several observables used in SRB are shown in Fig. 5.16.

Signal region C

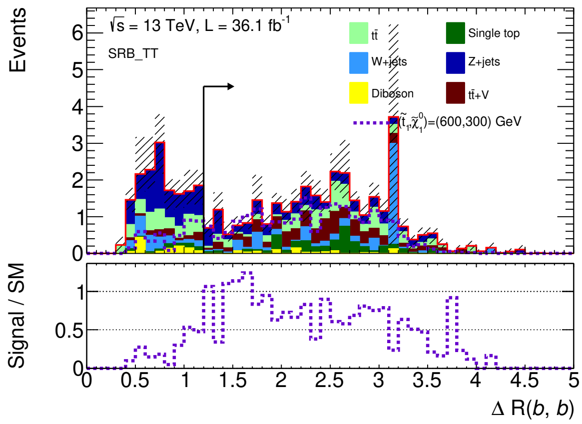
This signal region targets scenarios with $m(\tilde{t}_1) = m(t) + m(\tilde{\chi}_1^0)$, the benchmark scenarios are $m(\tilde{t}_1, \tilde{\chi}_1^0) = (300, 127)\,\text{GeV}$ and $m(\tilde{t}_1, \tilde{\chi}_1^0) = (500, 327)\,\text{GeV}$. In such a model, if the neutralinos are produced at rest in the stop-pair frame and the top squarks' momenta are back-to-back, the neutralinos' momenta are exactly opposite, which completely cancels the $E_{\mathrm{T}}^{\mathrm{miss}}$, the measurable part of such an events then looks exactly like $t\bar{t}$. In order to achieve sensitivity for such models a selection for events with initial state radiation (ISR) of a gluon is utilised. The whole stop-pair system recoils in the opposite direction of the ISR system, so that the momenta of the invisible neutralinos no longer cancel each other. In this case the ratio of the neutralino and stop mass can be deduced from $E_{\mathrm{T}}^{\mathrm{miss}}$ and the ISR jet:

$$R_{\mathrm{ISR}} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{p_{\mathrm{T}}^{\mathrm{ISR}}} \sim \frac{m_{\tilde{\chi}_1^0}}{m_{\tilde{t}_1}}. \tag{5.11}$$

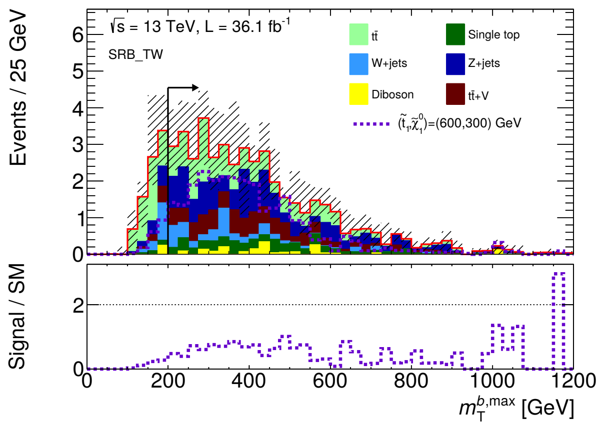
To identify the ISR system of an event, the *recursive jigsaw* [118] technique is used: All $R = 0.4$ jets are grouped into either the ISR system, which can consist of one or several jets, or the sparticle system which includes $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ and the remaining jets. Going through all possible grouping options of any number of jets in the ISR and sparticle system, the option is chosen which gives the highest momentum of the ISR system $p_{\mathrm{T}}^{\mathrm{ISR}}$ in the centre-of-mass frame (which is equivalent to the momentum of the sparticle system in the centre-of-mass frame p^{S}). With this method several new observables can be obtained that can be used to discriminate background: the ISR system must have a p_{T} of at least 400 GeV and the angular separation to $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ must be at least 3.0 (the invisible system would be at rest without ISR and is only boosted by recoiling against the ISR system); the sparticle system is required to contain at least five jets with $p_{\mathrm{T}} > 50\,\text{GeV}$ and at least one b-jet with $p_{\mathrm{T}} > 40\,\text{GeV}$; the transverse mass of the sparticle system is required to be at least 300 GeV. SRC is then



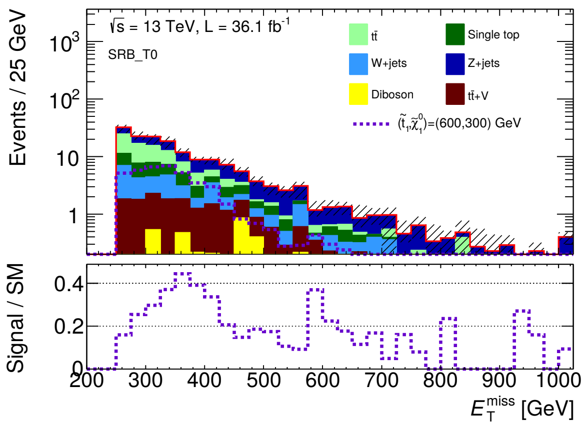
(a) Mass of the p_{T}-sub-leading $R = 1.2$ jet in SRB. The areas of the TT, TW & T0 categories are sketched in.



(b) $\Delta R(b,b)$ in SRB-TT.



(c) $m_{\mathrm{T}}^{b,\max}$ in SRB-TW.



(d) $E_{\mathrm{T}}^{\mathrm{miss}}$ in SRB-T0. Only the preselection threshold was used in SRB, cutting harder was not seen to be useful.

Figure 5.16: Distributions of several observables used in SRB. All selection cuts of the signal region are applied except for the observable shown in the diagram. The SM background processes are depicted as coloured areas stacked on each other. Scale factors derived from control regions are not applied. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed line is the SUSY signals with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (600, 300)\,\mathrm{GeV}$.

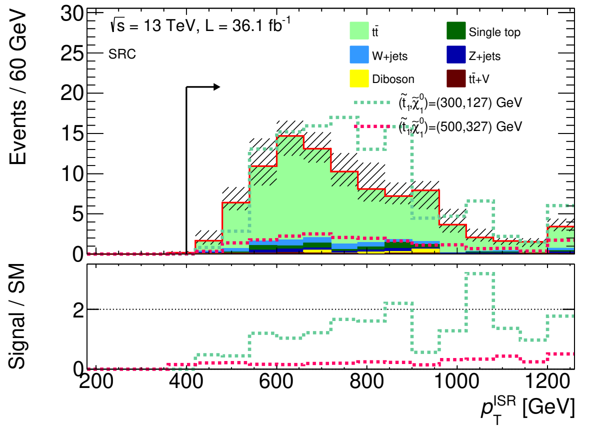
divided into five bins of R_{ISR} to achieve sensitivity for models with different $m_{\tilde{\chi}_1^0}/m_{\tilde{t}_1}$. With $E_{\mathrm{T}}^{\mathrm{miss}}$ being quite high in the used events (due to the trigger threshold), a large ratio of the sparticle system's boost often goes into the invisible system, the visible part of the sparticle system is therefore not necessarily highly boosted, which makes the usage of large-R jets not very helpful in this signal region. A summary of all selection criteria can be found in Tab. 5.5. Distributions of several recursive-jigsaw variables in this signal region are shown in Fig. 5.17. It can clearly be seen in Fig. 5.17d that the two chosen signal points have different sensitivities in the R_{ISR} bins.

Table 5.5: Selection criteria for SRC in addition to the common preselection defined in Tab. 5.3.

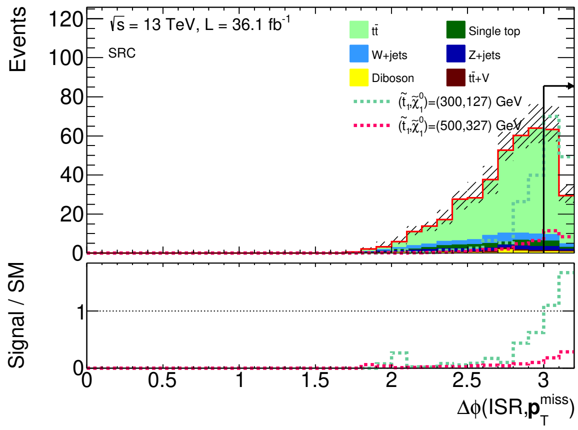
Observable	**Requirement**				
Bin $\rightarrow$	**SRC1**	**SRC2**	**SRC3**	**SRC4**	**SRC5**
$N_{b-\mathrm{jets}}^{\mathrm{S}}$	≥ 1				
$p_{\mathrm{T},b}^{0,\mathrm{S}}$	$> 40\,\mathrm{GeV}$				
$N_{\mathrm{jets}}^{\mathrm{S}}$	≥ 5				
$p_{\mathrm{T}}^{0-4,\mathrm{S}}$	$> 50\,\mathrm{GeV}$				
$m_{\mathrm{T}}^{\mathrm{S}}$	$> 300\,\mathrm{GeV}$				
$p_{\mathrm{T}}^{\mathrm{ISR}}$	$> 400\,\mathrm{GeV}$				
$\Delta\phi(\mathrm{ISR}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$	> 3.0				
R_{ISR}	$[0.3, 0.4]$	$[0.4, 0.5]$	$[0.5, 0.6]$	$[0.6, 0.7]$	$[0.7, 0.8]$

Signal region D

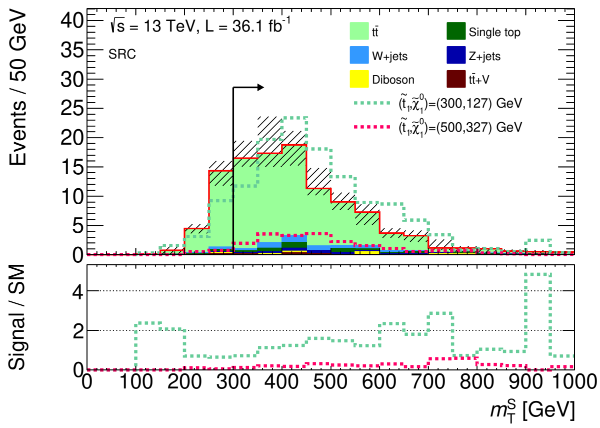
Signal region D aims to detect the process depicted in Fig. 5.1b where there are no tops in the decay chain. The top squarks decay via $\tilde{t}_1 \rightarrow b\tilde{\chi}_1^{\pm}$ with the chargino decaying via $\tilde{\chi}_1^{\pm} \rightarrow W\tilde{\chi}_1^0$ and its mass fixed at $m_{\tilde{\chi}_1^{\pm}} = 2m_{\tilde{\chi}_1^0}$. SRD is divided into two sub-regions: **SRD_low** targets the benchmark scenario $m(\tilde{t}_1, \tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = (400, 100, 50)\,\mathrm{GeV}$, **SRD_high** targets the benchmark scenario $m(\tilde{t}_1, \tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = (700, 200, 100)\,\mathrm{GeV}$. Since the branching ratio $\mathrm{BR}(\tilde{t}_1 \rightarrow b\tilde{\chi}_1^{\pm})$ can be very different in SUSY models, the signal regions were designed such that they preserve sensitivity also for the mixed decay, where one stop decays via $\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$, as a consequence no explicit veto on hadronically decaying tops was applied. Events must have at least



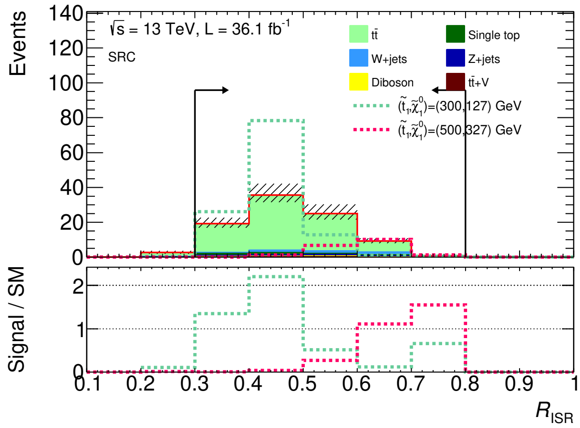
(a) p_{T} of the ISR system. Events with $p_{\mathrm{T}}^{\mathrm{ISR}} < 400\,\mathrm{GeV}$ are already negligible (but not strictly forbidden) with the rest of the selection criteria applied.



(b) Angular separation between the ISR sytem and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$.



(c) Transverse mass of the sparticle system.



(d) R_{ISR}, the bins from 0.3 to 0.8 are the SR bins SRC1 to SRC5.

Figure 5.17: Distributions of several observables used in SRC. All selection criteria of the signal region are applied except for the observable shown in the diagram and R_{ISR}. The SM background processes are depicted as coloured areas stacked on each other. Scale factors derived from control regions are not applied. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed lines correspond to the SUSY signals with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (300, 127)\,\mathrm{GeV}$ (green) and $m(\tilde{t}_1, \tilde{\chi}_1^0) = (500, 327)\,\mathrm{GeV}$ (magenta).

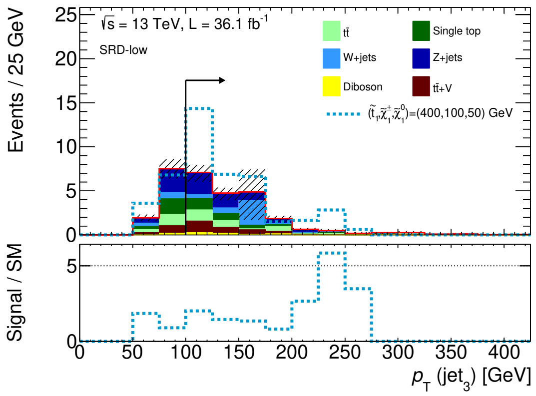
five jets with higher p_{T} thresholds than in SRA & SRB. Two b-jets are required and the observables $m_{\mathrm{T}}^{b,\mathrm{min}}$, $m_{\mathrm{T}}^{b,\mathrm{max}}$ and $\Delta R(b,b)$ were used with thresholds optimised for the signal scenarios; also the tau veto described in Sec. 5.3 and $\left|\Delta\phi\left(\mathrm{jet}^{0,1,2},\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right|_{\mathrm{min}}$ were used. Notably useful as a discriminating observable for this signal region proved the sum of the transverse momenta of the p_{T} leading and sub-leading b-jet. The exact definition of SRD can be found in Tab. 5.6. The two sub-regions are not orthogonal and can thus not be statistically combined. Distributions of some observables are shown in Fig. 5.18.

Table 5.6: Selection criteria for SRD in addition to the common preselection defined in Tab. 5.3.

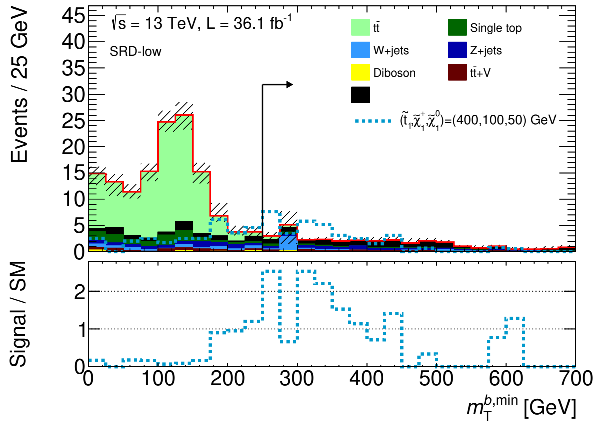
Observable	**Requirement**	
Sub-region →	**SRD-low**	**SRD-high**
$\left|\Delta\phi\left(\mathrm{jet}^{0,1,2},\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right|_{\mathrm{min}}$	> 0.4	
N_{jets}	≥ 5	
$p_{\mathrm{T}}^{0,1}$	$> 150\,\mathrm{GeV}$	
$p_{\mathrm{T}}^{2,3}$	$> 100\,\mathrm{GeV}$	$> 80\,\mathrm{GeV}$
p_{T}^{4}	$> 60\,\mathrm{GeV}$	
$N_{b-\mathrm{jets}}^{\mathrm{S}}$	≥ 2	
$\Delta R(b,b)$	> 0.8	
$p_{\mathrm{T}}^{0,b} + p_{\mathrm{T}}^{1,b}$	$> 300\,\mathrm{GeV}$	$> 400\,\mathrm{GeV}$
τ-veto	✓	
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 250\,\mathrm{GeV}$	$> 350\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{max}}$	$> 300\,\mathrm{GeV}$	$> 450\,\mathrm{GeV}$

Signal region E

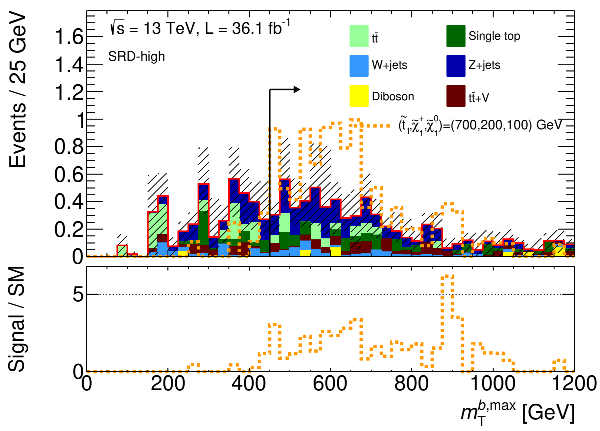
The last signal region is designed to detect very highly boosted top quarks in combination with $E_{\mathrm{T}}^{\mathrm{miss}}$. While such a topology can also arise in stop-pair production with a high stop mass, this is not the targeted scenario because the production cross-section falls quickly with increasing stop mass, which leads to scenarios that are very unlikely to be observed at the current integrated luminosity. More promising are scenarios with pair production of gluinos – due to higher cross-sections, see Fig. 2.4b – which can then decay via $\tilde{g} \to \tilde{t}_1 t$. Since we are targeting scenarios with two tops, we are looking for models with high $\Delta m(\tilde{g},\tilde{t}_1)$ but



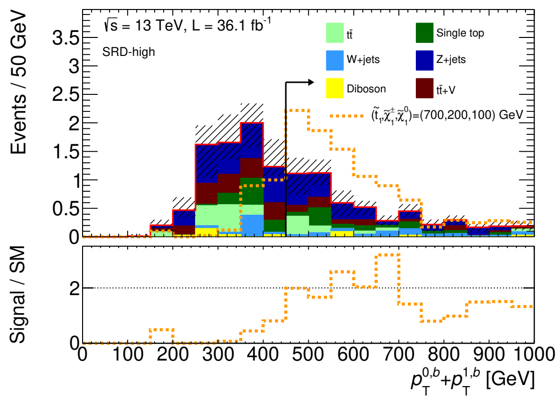
(a) Fourth-leading jet p_{T} in SRD-low. without applying a selection on the third leading jet.



(b) $m_{\mathrm{T}}^{b,\min}$ in SRD-low.



(c) $m_{\mathrm{T}}^{b,\max}$ in SRD-low.



(d) $p_{\mathrm{T}}^{0,b} + p_{\mathrm{T}}^{1,b}$ in SRD-high.

Figure 5.18: Distributions of several observables used in SRD. All selection criteria of the signal region are applied except for the observable shown in the diagram. The SM background processes are depicted as coloured areas stacked on each other. Scale factors derived from control regions are not applied. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed lines correspond to the SUSY signals with either $m(\tilde{t}_1, \tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = (400, 100, 50)$ GeV (SRD-low) or $m(\tilde{t}_1, \tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = (700, 200, 100)$ GeV (SRD-high) with $\tilde{t}_1 \to b\tilde{\chi}_1^{\pm}$.

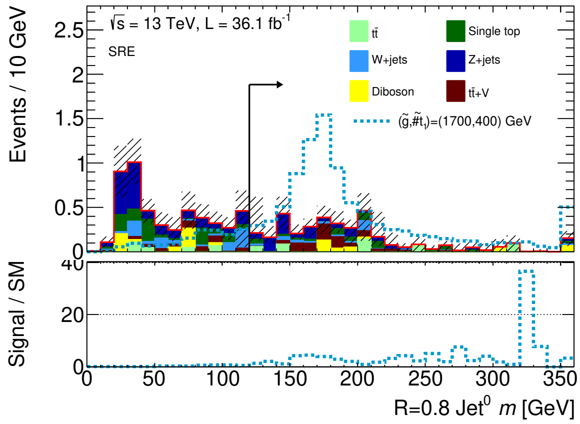
very low $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$, which makes the jets that are the visible decay products of the top squarks very soft and we can treat the top squarks as quasi-invisible. A diagram of this process is shown in Fig. 5.1c. The benchmark signal scenario for this region is $m(\tilde{g}, \tilde{t}_1, \tilde{\chi}_1^0) = (1700, 400, 395)\,\mathrm{GeV}$. We expect events to have at least 2 b-jets and the selection cuts on $m_\mathrm{T}^{b,\mathrm{min}}$ and $\left|\Delta\phi\left(\mathrm{jet}^{0,1,2}, \mathbf{p}_\mathrm{T}^\mathrm{miss}\right)\right|_\mathrm{min}$ are applied as in SRA/SRB. For the identification of the highly boosted tops $R = 0.8$ jets are used, a categorisation as in SRA/SRB was not found to be useful because the number of signal events in a hypothetical TW or T0 category would be too low. Another important variable is the sum of all jet p_T's: $H_\mathrm{T} = \sum_{i=1}^{N_\mathrm{jets}} p_\mathrm{T}^{i-1}$. This is then also used for a type of "$E_\mathrm{T}^\mathrm{miss}$ significance", which is a measure for the confidence that the reconstructed $E_\mathrm{T}^\mathrm{miss}$ is a result of invisible particles. In this case as an approximation for the uncertainty of $E_\mathrm{T}^\mathrm{miss}$, which is mostly a result of jet energy uncertainty, $\sqrt{H_\mathrm{T}/\mathrm{GeV}}\mathrm{GeV}$ is used; this results in the dimensionless observable $E_\mathrm{T}^\mathrm{miss}/(\sqrt{H_\mathrm{T}/\mathrm{GeV}}\mathrm{GeV})$. The exact criteria can be found in Tab. 5.7. Distributions of some observables are shown in Fig. 5.19.

Table 5.7: Selection criteria for SRE in addition to the common preselection defined in Tab. 5.3.

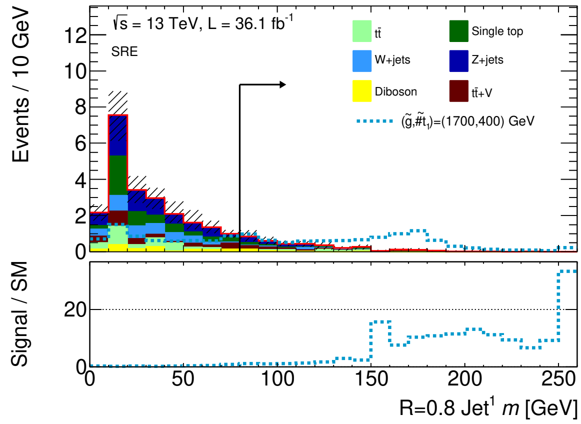
Observable	**Requirement**
$\left|\Delta\phi\left(\mathrm{jet}^{0,1,2}, \mathbf{p}_\mathrm{T}^\mathrm{miss}\right)\right|_\mathrm{min}$	> 0.4
$N_{b-\mathrm{jets}}$	≥ 2
$m^0_{\mathrm{jet},R=0.8}$	$> 120\,\mathrm{GeV}$
$m^1_{\mathrm{jet},R=0.8}$	$> 80\,\mathrm{GeV}$
$m_\mathrm{T}^{b,\mathrm{min}}$	$> 200\,\mathrm{GeV}$
$E_\mathrm{T}^\mathrm{miss}$	$> 550\,\mathrm{GeV}$
H_T	$> 800\,\mathrm{GeV}$
$E_\mathrm{T}^\mathrm{miss}/(\sqrt{H_\mathrm{T}/\mathrm{GeV}}\mathrm{GeV})$	> 18

5.6 Background estimation

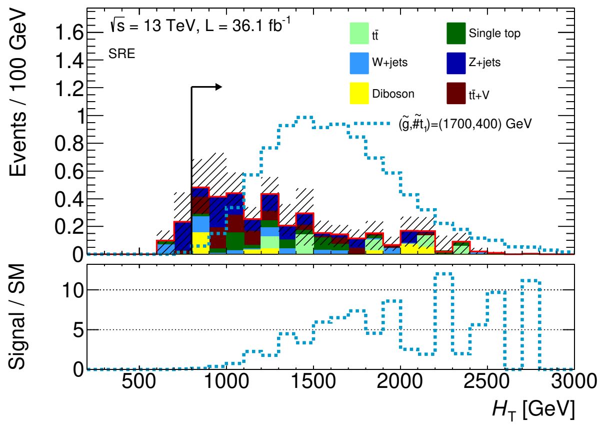
The dominant background processes in the SRs are $t\bar{t}$, Z+jet, W+jet, $t\bar{t}+Z$ and single top production. Their estimation is done by normalising each of them to data in a dedicated *control region* (CR); a scale factor for this normalisation is derived. Under the assumption



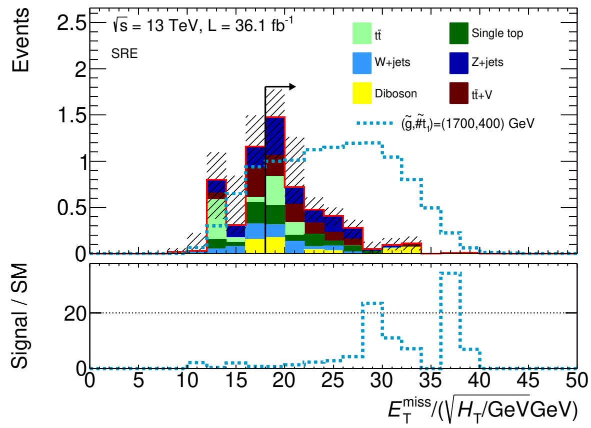
(a) Mass of the p_{T}-leading $R = 0.8$ jet.



(b) Mass of the p_{T}-sub-leading $R = 0.8$ jet.



(c) $H_{\mathrm{T}} = \sum p_{\mathrm{T}}$.



(d) $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$.

Figure 5.19: Distributions of several used observables in SRE. All selection criteria of the signal region are applied except for the observable shown in the diagram. The SM background processes are depicted as coloured areas stacked on each other. Scale factors derived from control regions are not applied. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties. The dashed line corresponds to the SUSY signal with $m(\tilde{g}, \tilde{\chi}_1^0) = (1700, 400)$ GeV. Spikes in the ratio plots are artificial, they are caused by a lack of MC events in the background description; hence only the integrated number of events in the signal region, but not in single bins, should be compared to data

that the scale factor for this process is the same in the signal region (within an uncertainty), the background process can then be exrapolated to the SR. *Validation regions* (VRs) are used to cross-check that this extrapolation works well. By using this procedure, systematic uncertainties affecting the normalisation are partially canceled out and the remaining uncertainty is smaller compared to the use of a MC-only based estimation.
The scale factors are derived in a fit, scaling all processes of the background-only model to data for the discovery test and the background-plus-signal model for the calculation of exclusion limits. They are defined such that they are enriched in events of the dedicated background process while still containing events that are kinematically close to the events that end up in SRs; all CRs are defined orthogonally. CRs use the same event-cleaning method as described in Sec. 5.5, but do not necessarily use the same preselection. The fit is done separately for each signal region, the CRs used to obtain the scale factors for each fit do not necessarily have to be the same ones. In this analysis we used several CRs for $t\bar{t}$ and Z+jet for different signal regions and categories of SRA & SRB in order to have more similar kinematics to the SRs; for the other background processes the same CR was used for all SRs. $t\bar{t}$, W+jet and single top production in the SR all have one charged lepton which is misreconstructed as a jet (this is very likely if the charged lepton is a hadronically decaying tau) or out of acceptance, thus the CRs for these processes use similar strategies.
The fit is done by maximising a likelihood function[10]. This function is a product of several likelihood functions, calculated with the number of observed data events in the control regions and their probability density function for expected events, which are the Poisson probability density functions. Systematic uncertainties are taken into account in this procedure, they are modelled as Gaussian probability density functions and treated as nuisance parameters. The background scale factors for those backgrounds with a dedicated CR are treated as free parameters in the fitting procedure and are the parameters of interest. Minor backgrounds are scaled to the integrated luminosity and the highest-precision cross-section calculation available or in case of multijet background derived from the data-driven method described later in this section. Their uncertainties are treated as nuisance parameters in the

[10]For numerical reasons the negative logarithm of the likelihood function is minimised, which is equivalent.

fit, no further scale factor is introduced for them. The signal contribution in CRs was found to be $< 8\%$ for all signal models that were not yet excluded in Run 1 [119] and was found to not significantly change the expected discovery sensitivity. For the calculation of exclusion limits this signal contribution was taken into account for each signal model (see Sec. 5.8). VRs were designed for certain backgrounds. They are not included in the fit, they have in general a lower purity than the CR in the designated process, but are closer to the SRs. They were also checked for not-excluded signal contributions; in general these are low, only some are $\sim 25\%$, namely for models with top-squark masses below 350 GeV in the different $t\bar{t}$ validation regions (VRT) and top-squark masses of 700 GeV in VRs for Z+jet (VRZD and VRZE). This needs to be taken into account in the interpretation if an excess is found in both the VR and the corresponding SR, but will normally not be a problem since the VRs are not included in the fit.

5.6.1 One-lepton backgrounds

The dominant source of background from one-lepton processes are events in which a hadronically decaying tau (originating from a $W \to \tau\nu$ decay) is misidentified as a jet. The CRs described here aim to emulate these background events. Because of lepton universality the decay into an electron or muon, which can be reconstructed more reliably, should kinematically be – apart from the negligible mass difference – the same as for taus. Therefore we define CRs for these processes by requiring exactly one signal electron or muon (in the following only these two particles are referred to as 'leptons') and then treat the lepton for the calculation of all other observables as a non-b-tagged jet. Selections on other observables are then used in order to separate these processes into different CRs and to have similar selection criteria as the SRs. A common preselection is used for all 1-lepton CRs, which is shown in Tab. 5.8. The number of baseline and signal leptons is required to be the same (referred to as N_l) and exactly one. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger and offline $E_{\mathrm{T}}^{\mathrm{miss}}$ is used in the same way as for the SRs, same for the number of jets and their p_{T} requirements; at least one b-jet is required; to reject multijet events, $\left|\Delta\phi\left(\mathrm{jet}^{0,1,}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right|_{\mathrm{min}}$ must be at least 0.4; $m_{\mathrm{T}}(l, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ is required to be below 100 GeV to make sure that the signal contribution is low.

Table 5.8: Common preselection for all one-lepton control regions.

Observable	Requirement
Event cleaning	✓
$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger	✓
N_l	1
p_{T}^{l}	$> 20\,\mathrm{GeV}$
N_{jets}	≥ 4 (including lepton)
$\mathrm{Jet}^{0-3}\, p_{\mathrm{T}}$	$> (80, 80, 40, 40)\,\mathrm{GeV}$
$N_{b-\mathrm{jets}}$	≥ 1
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 250\,\mathrm{GeV}$
$\left|\Delta\phi\left(\mathrm{jet}^{0,1}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right|_{\mathrm{min}}$	> 0.4
$m_{\mathrm{T}}(l, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$	$< 100\,\mathrm{GeV}$

W+jet control region

It is challenging to define a high-purity W+jet control region (CRW) close to the SRs, because one-lepton regions with heavy-flavour jets are usually dominated by $t\bar{t}$ events; therefore mostly $t\bar{t}$ production needs to be rejected to obtain a useful CR. This is ensured in several ways. First the selection is restricted to exactly one b-jet. This is done because $t\bar{t}$ events almost exclusively produce two b quarks, whereas in W+jet events the b-tagged jets in the SR can also come from mistagged c or light jets. As a veto against hadronic tops we can set an upper limit on the leading $R = 1.2$ jet mass; in addition to the upper threshold on $m_{\mathrm{T}}(l, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ also a lower threshold is used to reject a potential multijet contribution. Finally, for rejection against leptonic tops, the angular distance between the b-jet and the lepton was used: for leptonic tops it is expected to be small if both particles are decay products of the same tops. However, if $t\bar{t}$ events enter a 1-b region one b quark was not identified by the tagging algorithm as a b-jet; the $\Delta R(b, l)$ is only expected to be small if the identified b-jet originates from the leptonically decaying top. $\Delta R(b, l)$ could thus still give us rejection power in this case, but in the case where the identified b-jet comes from the hadronically decaying top,

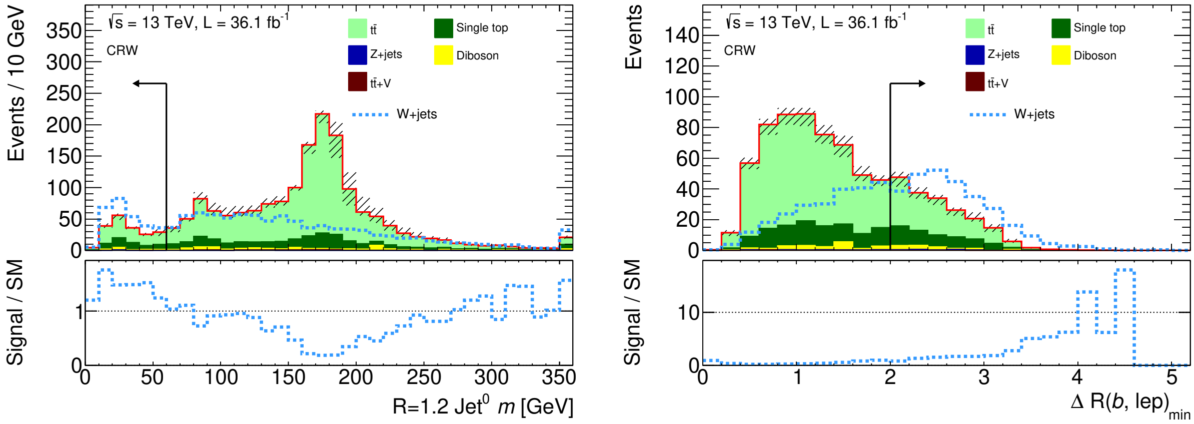
$\Delta R(b,l)$ will not be very helpful. The non-b-tagged jet with the highest b-tagging weight is likely to be the second b quark in $t\bar{t}$ events, hence a better result was seen when using $\Delta R(b^{0,1},l)_{\min}$ using the two jets with the highest b-tagging weights $b^{0,1}$. $\Delta R(b^{0,1},l)_{\min}$ was found to be useful in the single top and $t\bar{t}$ control regions as well to increase the purity. The exact requirements of the control region can be found in Tab. 5.9. The W+jet purity of this region (before applying the fit) is 60%, the majority of the reminder is $t\bar{t}$ (27%) and single top (10%).

In addition to the control region, a validation region (VRW) was designed, which is closer to the signal regions but has a smaller purity (39%) in return. It is designed similarly to the CR, but requires at least two b-jets and has an additional $m_{\mathrm{T}}^{b,\min}$ threshold. The exact requirements can also be seen in Tab. 5.9. Some observables used to increase the W+jet

Table 5.9: Selection criteria for the W+jet control and validation region in addition to the common one-lepton preselection shown in Tab.5.8.

Observable	**Requirement**	
Region	**CRW**	**VRW**
$N_{b-\mathrm{jets}}$	1	≥ 2
$m_{\mathrm{T}}(l,\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$	$<[30,100]\,\mathrm{GeV}$	
$\Delta R(b^{0,1},l)_{\min}$	>2.0	>1.8
$m^{0}_{\mathrm{jet},R=1.2}$	$<60\,\mathrm{GeV}$	$<70\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\min}$	-	$>150\,\mathrm{GeV}$

purity are shown in Fig. 5.20 to illustrate the selection choices. Additional uncertainty might come from the fact that the multiplicity of c- and b-hadrons is different in the CR and SR due to the $1b$ requirement, especially events with one c-hadron on gennerator level are a concern if occuring in significant numbers, because the heavy flavour is in this case not produced through gluon splitting but on tree level, which causes large uncertainties. This effect was studied with the help of generator-level filters. The fractions in the CR are: Events without any c or b hadrons: 19.6%, events with c hadrons but without b hadrons: 36.6%, events with b hadrons: 43.7%. The same numbers in the VR are 2.1%, 7.7% and 90%; as an example in SRB-T0 the composition is 0.1%, 11.8% and 88.1%. This shows that even though the composition in the CR is different from the SRs, the VR is designed to have



(a) Mass of the p_{T}-leading $R = 1.2$ jet.

(b) Angular distance between the lepton and b candidates.

Figure 5.20: Distributions of several observables used in CRW. All selection cuts of the control region are applied except for the observable shown in the diagram. The relevant W+jet process is depicted as a dashed line, the other SM background processes are depicted as coloured areas stacked on each other. Scale factors derived from control regions are not applied. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties.

very similar types of events as the SRs and is hence a good test for the background fit. The $W + c$ contribution in the CR is 9.73%, in the VR 5.15% and in SRD-high 1.86% (the $W + c$ contribution is in general very low in all SRs). This source of uncertainty was concluded to be of insignificant impact and is not further considered in the calculations.

Single top control region

The single-top contribution in our SRs mainly comes from the Wt channel (see Fig. 5.8c) with a hadronic top and a leptonically decaying W boson. We use a $\geq 2b$-jet region and look for one massive $R = 1.2$ jet. This makes it challenging to reduce $t\bar{t}$; for this purpose $m_{\mathrm{T}}^{b,\min}$ and $\Delta R(b^{0,1}, l)_{\min}$ are used. Furthermore, $\Delta R(b, b)$ is used to reduce events with b-jets from gluon splitting. The exact selections can be seen in Tab. 5.10. The single top purity is 45% with most of the remainder coming from $t\bar{t}$ (31%) and W+jet (20%).

Table 5.10: Selection cuts for the single-top control region in addition to the common one-lepton preselection shown in Tab.5.8.

Observable	**Requirement**
Region	**CRST**
$N_{b-\text{jets}}$	≥ 2
$m_{\mathrm{T}}(l, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$	$< [30, 100]\,\mathrm{GeV}$
$\Delta R(b^{0,1}, l)_{\min}$	> 2.0
$\Delta R(b, b)$	> 1.5
$m^{0}_{\mathrm{jet},R=1.2}$	$> 70\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\min}$	$> 200\,\mathrm{GeV}$

$\mathbf{t\bar{t}}$ control regions

For $t\bar{t}$ one can achieve a higher purity more easily because it is the dominating process in a selection with on lepton and ≥ 2 b-jets. This can be used to create CRs closer to the separate SRs. Here we created separate CRs for the three categories of SRA and SRB respectively, as well as SRC, SRD and SRE. They are all selecting low values of $\Delta R(b^{0,1}, l)_{\min}$ and are apart from that oriented at the SR cuts, with some thresholds ($E_{\mathrm{T}}^{\mathrm{miss}}$, $m_{\mathrm{T}}^{b,\min}$) used in a more relaxed way to conserve purity. The exact set of selection criteria can be found in Tab. 5.11. The $t\bar{t}$ purity is in the range 84-95%. Validation regions were defined in 0-lepton regions. These are defined equivalently to their corresponding SRs, but choosing a different region in one observable, this observable being $m_{\mathrm{T}}^{b,\min}$ for all regions except VRTC, where the cut on $\Delta\phi(\mathrm{ISR}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ was inverted instead. The selection criteria are summarised in Tabs 5.12 and 5.13.

5.6.2 Z+jet background

It is very difficult to design a selection enriched with $Z(\rightarrow \nu\nu)$+jet events, because the resonant Z mass-peak is invisible. It can, however, be assumed that events where the Z boson decays into an electron-positron or muon-antimuon pair behave very similar[11] if

[11] The lepton masses are negligible.

Table 5.11: Selection criteria for the $t\bar{t}$ control regions in addition to the common one-lepton preselection shown in Tab.5.8.

Observable	**Requirement**								
Region	**CRTA-TT**	**CRTA-TW**	**CRTA-T0**	**CRTB-TT**	**CRTB-TW**	**CRTB-T0**	**CRTC**	**CRTD**	**CRTE**
$N_{b-\text{jets}}$	≥ 2								
$\left\|\Delta\phi\left(\text{jet}^{0,1}, \mathbf{p}_{\text{T}}^{\text{miss}}\right)\right\|_{\min}$	> 0.4						-	> 0.4	
$m_{\text{T}}(l, \mathbf{p}_{\text{T}}^{\text{miss}})$	$[30, 100]$ GeV						< 100 GeV	$[30, 100]$ GeV	
$m_{\text{T}}^{b,\min}$	> 100 GeV						-	> 100 GeV	
$\Delta R(b^{0,1}, l)_{\min}$	< 1.5						< 2.5	< 1.5	
$m^{0}_{\text{jet},R=1.2}$	> 120 GeV						-		
$m^{1}_{\text{jet},R=1.2}$	> 120 GeV	$[60, 120]$ GeV	< 60 GeV	> 120 GeV	$[60, 120]$ GeV	< 60 GeV	-		
$m^{0}_{\text{jet},R=0.8}$	> 60 GeV			-					> 120 GeV
$m^{1}_{\text{jet},R=0.8}$	-								> 80 GeV
$E_{\text{T}}^{\text{miss}}$	> 250 GeV	> 300 GeV	> 350 GeV	> 250 GeV					
$\Delta R(b,b)$	> 1.0	-		> 1.2			-	> 0.8	-
$m_{\text{T}}^{b,\max}$	-			> 200 GeV			-	> 100 GeV	-
$N^{\text{S}}_{b-\text{jets}}$	-						≥ 1	-	
$N^{\text{S}}_{\text{jets}}$	-						≥ 5	-	
$p_{\text{T}}^{4,\text{S}}$	-						> 40 GeV	-	
$p_{\text{T}}^{\text{ISR}}$	-						> 400 GeV	-	
p_{T}^{1}	-							> 150 GeV	-
p_{T}^{3}	-							> 80 GeV	-
$p_{\text{T}}^{0,b} + p_{\text{T}}^{1,b}$	-							> 300 GeV	-
H_T	-								> 500 GeV

Table 5.12: Selection criteria for the $t\bar{t}$ validation regions A and B in comparison to their corresponding signal regions. The VRs have all selection criteria of their corresponding SR (shown in Tab. 5.4) except for the alternative ones given in this table, which ensure orthogonality to the SRs.

Observable	Requirement					
Region	**VRTA-TT**	**VRTA-TW**	**VRTA-T0**	**VRTB-TT**	**VRTB-TW**	**VRTB-T0**
Based on	SRA-TT	SRA-TW	SRA-T0	SRB-TT	SRB-TW	SRB-T0
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 300\,\mathrm{GeV}$	$> 400\,\mathrm{GeV}$	$> 450\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$		
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$[100, 200]\,\mathrm{GeV}$	$[140, 200]\,\mathrm{GeV}$	$[160, 200]\,\mathrm{GeV}$	$[100, 200]\,\mathrm{GeV}$	$[140, 200]\,\mathrm{GeV}$	$[160, 200]\,\mathrm{GeV}$
m_{T2}	-					
$m_{\mathrm{T}}^{b,\mathrm{max}}$	-			$> 200\,\mathrm{GeV}$		

Table 5.13: Selection cuts for the $t\bar{t}$ validation regions C, D and E in comparison to their corresponding signal regions. The VRs have all selection criteria of their corresponding SR (shown in Tab. 5.4) except for the alternative ones given in this table, which ensure orthogonality to the SRs..

Observable	Requirement		
Region	**VRTC**	**VRTD**	**VRTE**
Based on	SRC	SRD	SRE
$m_{\mathrm{T}}^{b,\mathrm{min}}$	-	$[50, 150]\,\mathrm{GeV}$	
τ-veto	-		
$p_{\mathrm{T}}^{4,\mathrm{S}}$	$> 40\,\mathrm{GeV}$	-	
$m_{\mathrm{T}}^{\mathrm{S}}$	$> 100\,\mathrm{GeV}$	-	
$m_{\mathrm{T}}^{\mathrm{V}}/m_{\mathrm{T}}^{\mathrm{S}}$	< 0.6	-	
R_{ISR}	-		
$\Delta\phi(\mathrm{ISR}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$	< 3.0	-	
$p_{\mathrm{T}}^{0,1}$	-	$> 150\,\mathrm{GeV}$	-
$p_{\mathrm{T}}^{2,3}$	-	$> 80\,\mathrm{GeV}$	-
$p_{\mathrm{T}}^{4,5}$	-		
$m_{\mathrm{T}}^{b,\mathrm{max}}$	-	$> 300\,\mathrm{GeV}$	-
$p_{\mathrm{T}}^{0,b} + p_{\mathrm{T}}^{1,b}$	-	$> 300\,\mathrm{GeV}$	-

the charged leptons are treated as if they were invisible. This allows to select events with two opposite-sign same-flavour leptons, requiring the invariant mass of the lepton pair to be close to the Z mass, which results in a high Z+jet purity. Such events should not have genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, which makes it necessary to use a different trigger. Muon & electron triggers are used here and only leptons with $p_{\mathrm{T}} > 28\,\mathrm{GeV}$ are selected so that the triggers are fully efficient. One can then calculate a new variant of 'missing transverse momentum' with the lepton-antilepton pair treated as invisible:

$$\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}'} = \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} + \mathbf{p}_{\mathrm{T}}^{\mathrm{lepton}} + \mathbf{p}_{\mathrm{T}}^{\mathrm{antilepton}}. \tag{5.12}$$

All observables that use $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ in the SRs use $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}'}$ instead in the Z+jet control region (CRZ) and are also marked with a prime. The original $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be low. Other criteria are chosen such that they are similar to the SRs. There are four variants of CRZ: One for the TT and TW category and one for the T0 category of SRA and SRB, one for SRD and one for SRE. The Z+jet background in SRC is negligible so that a CRZ for this region is not needed. The exact selection criteria can be found in Tab. 5.14. The Z+jet purity in these regions is 73-83%. As validation regions we used 0-lepton regions with similar cuts as the SRs, but with the $\Delta R(b,b)$ cut inverted. One region for SRA and SRB was defined, one for SRD and one for SRE. The exact cuts can be found in Tab. 5.15. The Z+jet purity in these regions is significantly smaller than in the CRs (41-48%).

5.6.3 $\mathbf{t\bar{t}+Z}$ background

The $t\bar{t}+Z$ production has a relatively small cross section, but is nevertheless an important background process, because it results in the same final state as our signals if the Z boson is decaying into neutrinos, whereas all other background processes are the result of missed or fake objects. Defining a CR similarly as for Z+jet for the $Z \to l\bar{l}$ case leads to large contributions from other backgrounds (mainly $t\bar{t}$ and Z+jet); the option of using leptonic decay modes of the top by defining a 3-lepton or 4-lepton CR was investigated but was found

Table 5.14: Selection cuts for the different Z+jet control regions.

<table>
<tr><th>Observable</th><th colspan="4">Requirement</th></tr>
<tr><th>Region</th><th>CRZAB-TT-TW</th><th>CRZAB-T0</th><th>CRZD</th><th>CRZE</th></tr>
<tr><td>Event cleaning</td><td colspan="4">✓</td></tr>
<tr><td>Trigger</td><td colspan="4">electron or muon</td></tr>
<tr><td>N_l</td><td colspan="4">2, opposite charge, same flavour</td></tr>
<tr><td>p_{T}^{l}</td><td colspan="4">$> 28\,\mathrm{GeV}$</td></tr>
<tr><td>m_{ll}</td><td colspan="4">$[86, 96]\,\mathrm{GeV}$</td></tr>
<tr><td>N_{jets}</td><td colspan="4">≥ 4</td></tr>
<tr><td>$\mathrm{Jet}^{0-3}\,p_{\mathrm{T}}$</td><td colspan="4">$> (80, 80, 40, 40)\,\mathrm{GeV}$</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}$</td><td colspan="4">$< 50\,\mathrm{GeV}$</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}\prime}$</td><td colspan="4">$> 100\,\mathrm{GeV}$</td></tr>
<tr><td>$N_{b-\mathrm{jets}}$</td><td colspan="4">≥ 2</td></tr>
<tr><td>$m_{\mathrm{jet},R=1.2}^{0}$</td><td colspan="2">$> 120\,\mathrm{GeV}$</td><td colspan="2">-</td></tr>
<tr><td>$m_{\mathrm{jet},R=1.2}^{1}$</td><td>$> 60\,\mathrm{GeV}$</td><td>$< 60\,\mathrm{GeV}$</td><td colspan="2">-</td></tr>
<tr><td>$m_{\mathrm{T}}^{b,\mathrm{min}\prime}$</td><td colspan="2">-</td><td colspan="2">$> 200\,\mathrm{GeV}$</td></tr>
<tr><td>$m_{\mathrm{T}}^{b,\mathrm{max}\prime}$</td><td colspan="2">-</td><td>$> 200\,\mathrm{GeV}$</td><td>-</td></tr>
<tr><td>H_{T}</td><td colspan="3">-</td><td>$> 500\,\mathrm{GeV}$</td></tr>
</table>

Table 5.15: Selection cuts for the different Z+jet validation regions in addition to the common preselection defined in Tab. 5.3.

<table>
<tr><th>Observable</th><th colspan="3">Requirement</th></tr>
<tr><th>Region</th><th>VRZAB</th><th>VRZD</th><th>VRZE</th></tr>
<tr><td>N_{jets}</td><td>≥ 4</td><td>≥ 5</td><td>≥ 4</td></tr>
<tr><td>$N_{b-\mathrm{jets}}$</td><td colspan="3">≥ 2</td></tr>
<tr><td>τ veto</td><td colspan="2">✓</td><td>-</td></tr>
<tr><td>$m_{\mathrm{T}}^{b,\mathrm{min}}$</td><td colspan="3">$200\,\mathrm{GeV}$</td></tr>
<tr><td>$m_{\mathrm{jet},R=1.2}^{0}$</td><td>$> 120\,\mathrm{GeV}$</td><td colspan="2">-</td></tr>
<tr><td>$\Delta R(b,b)$</td><td>< 1.0</td><td>< 0.8</td><td>< 1.0</td></tr>
<tr><td>$m_{\mathrm{T}}^{b,\mathrm{max}}$</td><td>-</td><td>$> 200\,\mathrm{GeV}$</td><td>-</td></tr>
<tr><td>H_{T}</td><td colspan="2">-</td><td>$> 500\,\mathrm{GeV}$</td></tr>
<tr><td>$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$</td><td colspan="2">-</td><td>$> 14\sqrt{\mathrm{GeV}}$</td></tr>
<tr><td>$m_{\mathrm{jet},R=0.8}^{0}$</td><td colspan="2">-</td><td>$> 120\,\mathrm{GeV}$</td></tr>
</table>

to be not very practical due to low number of events. Instead the option was chosen to design a $t\bar{t}+\gamma$ CR. This method was introduced in [120] and can be done because the processes are very similar in terms of Feynman diagrams[12]; the MC samples were also generated with the same configuration. If the transverse momentum of the photon is high enough, it can be assumed to be kinematically very similar to the Z boson because the mass difference becomes negligible; the photon $\mathbf{p}_{\mathrm{T}}$ can than be used to describe the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ caused by $Z \to \nu\nu$ and replace it in the calculation of observables similar to $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}'}$ in the Z+jet CR. A 1-lepton option was chosen, meaning that the $t\bar{t}$ part of the event decays semi-leptonically; lepton triggers were used and the lepton p_{T} threshold chosen such that the triggers are fully efficient. The exact set of cuts can be seen in Tab. 5.16.

Table 5.16: Selection cuts for the $t\bar{t}+\gamma$ control region.

Observable	**Requirement**
Event cleaning	✓
Trigger	electron or muon
N_l	1
p_{T}^{l}	$> 28\,\mathrm{GeV}$
N_γ	1
p_{T}^{γ}	$> 150\,\mathrm{GeV}$
N_{jets}	≥ 4
Jet$^{0-3}$ p_{T}	$> (80, 80, 40, 40)\,\mathrm{GeV}$

5.6.4 Multijet background

The QCD multijet background has a very high cross-section, but a very low probability to end up in our signal regions because multijet events do not produce genuine $E_{\mathrm{T}}^{\mathrm{miss}}$. A reliable modelling with simulated events cannot be achieved due to the impractically large amount of events that would be needed and because effects that can lead to fake $E_{\mathrm{T}}^{\mathrm{miss}}$ (mostly due to mismeasurement of hadronic jets) are not expected to be reliably described by the detector simulation. Therefore the multijet background is estimated with the data-driven *jet-smearing*

[12]The dominant diagrams for $t\bar{t}+\gamma$ production are the same as in Fig. 5.9, replacing the Z boson with a photon.

method introduced in [121] and described in detail for this analysis in [115]. It is based on the assumption that fake $E_{\mathrm{T}}^{\mathrm{miss}}$ is produced by an unfortunate combination of mismeasured jets. The jet-smearing methods uses *seed-events* ('well-measured' data events) and smears the jet p_{T}'s randomly according to a response function. By doing this repeatedly, we obtain a large amount of pseudo-data, among these are a significant amount of high-$E_{\mathrm{T}}^{\mathrm{miss}}$ events. The seed events are events obtained from jet-triggers with ≥ 4 jets and $\geq 1b$-jets with a low $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, which is defined here as

$$E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} = \frac{E_{\mathrm{T}}^{\mathrm{miss}} - 8\,\mathrm{GeV}}{\sqrt{H_{\mathrm{T}}}}. \tag{5.13}$$

Since $E_{\mathrm{T}}^{\mathrm{miss}}$ is not only calculated from jets, but with an additional soft term, whereas H_{T} in the denominator is only calculated from reconstructed jets, the correction term '$-8\,\mathrm{GeV}$' was added in the numerator to remove this bias. The exact threshold is dependent on the number of b-jets because those can produce neutrinos in b-hadron decays:

$$E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} < 0.3 + 0.1 \cdot N_{b-\mathrm{jets}}. \tag{5.14}$$

The smearing is done with ATLAS internal software, using a jet response function derived from simulation by comparing generator-level jet properties with their corresponding reconstructed properties; this is done 5000 times for each seed event. The arbitrary amount of pseudo-data is then normalised to data in control regions, which are close to the signal regions, but with $\left|\Delta\phi\left(\mathrm{jet}^{0,1}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right|_{\mathrm{min}} < 0.1$ and some other criteria like $E_{\mathrm{T}}^{\mathrm{miss}}$ relaxed to increase the number of events to reduce statistical uncertainty. The exact selection criteria for the CRs can be found in Tab. 5.17. The amount of multijet events in the SRs can then be estimated by applying the SR criteria to the pseudo-data scaled with the factor derived from the normalisation in the CR.

The uncertainty of this method is obtained by varying the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. threshold for the seed events up ($E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} < 0.6 + 0.2 \cdot N_{b-\mathrm{jets}}$) and down ($E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} < 0.2 + 0.05 \cdot N_{b-\mathrm{jets}}$) in addition to a 30% uncertainty added quadratically to account for low-side-tail modifications of the response function, seen in Run 1.

Table 5.17: Selection cuts for QCD control regions in addition to the preselection defined in Tab. 5.3 or replacing them in case of the same observables in both tables.

<table>
<tr><th>Observable</th><th colspan="5">Requirement</th></tr>
<tr><th>Region</th><th>CRQA</th><th>CRQB</th><th>CRQC</th><th>CRQD</th><th>CRQE</th></tr>
<tr><td>$\left|\Delta\phi\left(\text{jet}^{0,1}, \mathbf{p}_{\mathrm{T}}^{\text{miss}}\right)\right|_{\min}$</td><td colspan="2">$< 0.1$</td><td>$[0.05, 0.1]$</td><td colspan="2">$< 0.1$</td></tr>
<tr><td>N_{jets}</td><td colspan="3">≥ 4</td><td>≥ 5</td><td>≥ 4</td></tr>
<tr><td>$N_{b-\text{jets}}$</td><td colspan="2">≥ 2</td><td>≥ 1</td><td colspan="2">≥ 2</td></tr>
<tr><td>τ veto</td><td colspan="2">✓</td><td>-</td><td>✓</td><td>-</td></tr>
<tr><td>$m_{\mathrm{T}}^{b,\min}$</td><td colspan="2">$> 100\,\text{GeV}$</td><td>-</td><td colspan="2">$> 100\,\text{GeV}$</td></tr>
<tr><td>$E_{\mathrm{T}}^{\text{miss}}$</td><td>$> 300\,\text{GeV}$</td><td colspan="4">$> 250\,\text{GeV}$</td></tr>
<tr><td>$m_{\text{jet},R=1.2}^{0}$</td><td colspan="2">$> 120\,\text{GeV}$</td><td colspan="3">-</td></tr>
<tr><td>$m_{\text{jet},R=0.8}^{0}$</td><td>$> 60\,\text{GeV}$</td><td colspan="3">-</td><td>$> 120\,\text{GeV}$</td></tr>
<tr><td>$m_{\text{jet},R=0.8}^{0}$</td><td colspan="4">-</td><td>$> 80\,\text{GeV}$</td></tr>
<tr><td>$m_{\mathrm{T}}^{b,\max}$</td><td>-</td><td>$> 200\,\text{GeV}$</td><td colspan="3">-</td></tr>
<tr><td>$\Delta R(b,b)$</td><td>-</td><td>> 1.2</td><td>-</td><td>> 0.8</td><td>-</td></tr>
<tr><td>$N_{b-\text{jets}}^{\mathrm{S}}$</td><td colspan="2">-</td><td>≥ 1</td><td colspan="2">-</td></tr>
<tr><td>$N_{\text{jets}}^{\mathrm{S}}$</td><td colspan="2">-</td><td>≥ 5</td><td colspan="2">-</td></tr>
<tr><td>$p_{\mathrm{T}}^{4,\mathrm{S}}$</td><td colspan="2">-</td><td>$> 40\,\text{GeV}$</td><td colspan="2">-</td></tr>
<tr><td>$p_{\mathrm{T}}^{\text{ISR}}$</td><td colspan="2">-</td><td>$> 150\,\text{GeV}$</td><td colspan="2">-</td></tr>
<tr><td>$\Delta\phi(\text{ISR}, \mathbf{p}_{\mathrm{T}}^{\text{miss}})$</td><td colspan="2">-</td><td>> 2.0</td><td colspan="2">-</td></tr>
<tr><td>$E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$</td><td colspan="5">-</td></tr>
<tr><td>H_{T}</td><td colspan="4">-</td><td>$> 800\,\text{GeV}$</td></tr>
</table>

5.7 Systematic Uncertainties

For the correct interpretation of results, we need a good estimation for the uncertainties of SM and signal processes. In most SRs systematic uncertainties are not the main limiting factor, because statistical uncertainties dominate in regions with a low number of events; however systematic uncertainties are not negligible and there are also regions where they are the dominant source of uncertainty, e.g. in SRB-T0. Systematic uncertainties consist of both experimental and theoretical uncertainties.

5.7.1 Experimental uncertainties

Experimental uncertainties can come from imperfect knowledge of the generator modelling or imperfect detector simulation due to effects that are difficult to quantify. Such uncertainties can affect the probability of events ending up in certain selections, e.g. because of different object reconstruction efficiencies in data and MC. Such deviations are usually handled by applying scale factors on events derived from comparison of data and simulation, e.g. by choosing a well-understood selection with electrons and then comparing the efficiency of electrons in data and MC. These uncertainties are estimated by varying these scale factors by one standard deviation of their uncertainty up and down, which then results in yields differing from the nominal case. Other experimental uncertainties can affect the kinematics of objects and can hence affect the distribution of observables. The effect of such uncertainties are estimated by varying the calibration of these objects up or down by one standard deviation, then calculate all observables with the re-calibrated objects and apply selection cuts on these. There are different methods to estimate the size of these uncertainties for different systematics.

Jet energy scale & resolution

The jet energy calibration is performed by comparing jet energies at generator level with reconstructed jets, the p_{T} and η dependent uncertainties are estimated using different in-situ methods, the majority using events where a jet recoils against a photon or Z boson; also uncertainties from flavour composition and pile-up effects are taken into account [122]. In principal the jet calibration includes 80 nuisance parameters, but in practise the description of the jet energy scale (JES) uncertainty can be modelled with a precision sufficient for this analysis with a compressed set of four nuisance parameters, which are combinations of the 80. In addition to the scale, the jet energy resolution (JER) effects are estimated by smearing the jet p_{T}'s and η's according to resolutions measured in a dijet balance method [123]. The impact of the JES uncertainty on the event yield is $< 10\%$ in all SRs except SRC5 (17%). JES usually has a smaller impact, howerer in signal regions where jet reclustering is used it can influence which sets of $R = 0.4$ jets are combined to large-R jets, thus it reaches 10-12% in some SRs, mostly it is lower.

b-tagging efficiencies

b-tagging efficiencies are derived separately for b-, c- and light jets. Scale factors are applied to correct for differences in efficiencies in data and simulation. These factors and their uncertainties have been measured with a set of complementary methods and combined, taking correlations into account [124]. The impact of the b-tagging uncertainty on the yields is mostly small in our SRs, only in SRD they have a bigger impact (9 and 7%) due to the explicit cut on b-jet p_{T}'s.

Pile-up uncertainties

Pile-up related correction factors are applied to simulated events to account for the different distributions of average number of interactions per bunch-crossing, $\langle\mu\rangle$, in simulation and data. Uncertainties arise from this difference (which is treated as a two-sided uncertainty) as well as from the measurement of $\langle\mu\rangle$ [125]. The impact of this uncertainty on the yields ranges from 0.5% to 14% in the different SRs. The large uncertainty in some SRs is mainly an artifact of the method used for obtaining the pile-up correction factors which can lead to events with large weights.

$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term

Since $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from all measured objects, most of its uncertainty is already handled by treating the uncertainties of those objects, only the soft term remains. The uncertainty of the soft term is measured with in-situ methods using $Z \to \mu\mu$ events without additional jets and comparing simulation to data [126]. The imapct of this uncertainty is below 5% for all SRs except SRC5 (15%).

Lepton uncertainties

Lepton uncertainties have to be taken into account because of the use of leptons in the CRs but also because of the efficiency of the lepton veto in the SRs. The energy scale of electrons and muons is obtained by selecting events with two same-flavour opposite-sign leptons and fitting the di-lepton invariant mass to the known mass peaks of $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ and $J/\Psi \to \mu^+\mu^-$. Uncertainties arise from this fitting procedure [127, 128]. Additional uncertainties coming from trigger [57], reconstruction, identification and isolation efficiencies are also taken into account. These uncertainties are estimated for with 'tag-and-probe' methods utilising $Z \to e^+e^-(\mu^+\mu^-)$ and $J/\Psi \to e^+e^-(\mu^+\mu^-)$ events [128, 129]. The impact of lepton uncertainties is rather small in this analysis, they do not exceed 1% in any SR.

Luminosity

The overall uncertainty of the luminosity for the combined 2015+2016 data is 3.2%, it is measured with a x-y beam-separation method, as described in [130]. With normalisation mostly calculated from CRs, the impact of the luminosity uncertainty is small compared to other uncertainties for this analysis; it only has an impact on the signal samples as well as the diboson background, which iss not normalised in a dedicated CR.

5.7.2 Theory uncertainties

Theory uncertainties of background and signal samples can affect both the normalisation and the kinematic distributions. In case the background is normalised in a control region, the theory uncertainty on the normalisation has no impact. The theory uncertainties are thus calculated as uncertainties on the transfer factors T, which are defined as the ratio of the yields of a process X in a SR Y and their corresponding CR:

$$T_{X,Y} = \frac{N_{X,\mathrm{SR}Y}}{N_{X,\mathrm{CR}Y}}. \tag{5.15}$$

Relative uncertainties are then calculated by comparing the nominal yield to the yield with the theory uncertainty Z applied:

$$\sigma(X,Y,Z) = \frac{T_{X,Y}^{\mathrm{nominal}} - T_{X,Y}^{Z}}{T_{X,Y}^{\mathrm{nominal}}}. \tag{5.16}$$

Upwards and downwards uncertainties are treated separately; if several theoretical uncertainties are considered, they are added quadratically.
The uncertainty on the Z+jet and W+jet backgrounds was calculated by applying weights provided by the SHERPA generator. These weights correspond to varying the renormalisation, factorisation, merging and resummation scales each by a factor of two up and down. This method is cleaner than comparing it to a different generator, where one can usually only see the impact of several uncertainties, however a comparison with a different generator can

be found in [106]. The impact of the Z+jet and W+jet uncertainties on the event yield was found to be 3% or less. No additional uncertainty was applied for the $W+c$ contribution (discussed in Sec. 5.6.1) because the validation region was seen to match the data well, even though it has a different flavour composition than the control region, and it is thus assumed that this has a negligible impact.

The $t\bar{t}$ uncertainties were evaluated by comparing the nominal sample to samples from different generators. For the hard scattering generation uncertainty this additional sample for comparing was created with SHERPA 2.2.1 and for the parton shower uncertainty with POWHEGHERWIG++ [131] with the UE5C6L1 CT10 tuning. For the ISR/FSR radiation, additional samples were generated with POWHEGPYTHIA with decreased or increased amount of radiation (each by a factor of two). The impact of the $t\bar{t}$ uncertainty is typically 12% or below in all SRs except in the $t\bar{t}$ dominated SRC, where it reaches up to 71% in one bin. These large uncertainties are causes mostly by a low amount of MC events in the bins.

The uncertainties on the $t\bar{t}+W/Z$ yield were calculated by varying the factorisation and renormalisation scales and in addition comparing the nominal yields to a sample created with OPENLOOPS+SHERPA. The same variations were also applied to the $t\bar{t}+\gamma$ sample. That way the use of the control region for this background will only decrease the uncertainty if the assumption that both processes behave similarly holds true, this seems to be the case since the impact of the $t\bar{t}+V/\gamma$ uncertainties on the yield in signal regions is only 2% or less.

The single top uncertainty was evaluated by comparing the nominal sample to a sample generated with POWHEGHERWIG++. Also the ISR/FSR uncertainties were evaluated similarly as for $t\bar{t}$. An additional uncertainty comes from the fact that there is interference of $t\bar{t}$ with the dominant Wt channel if an additional b is produced (in beyond-LO processes) [132]. This uncertainty was estimated by comparing the nominal $t\bar{t}$ plus single-top+b yields with a $WWbb$ sample including resonant and non-resonant production. A relative uncertainty of 30% was found to be the upper limit in SRs and was used as a conservative estimate. The impact of the single-top theory uncertainty on the event yield is 12% in SRE, but in all other SRs only 6% and below.

For diboson processes an uncertainty of 50% is assumed as a conservative estimation. Due to the very small contribution to SRs this is still negligible.

Signal uncertainties were evaluated by varying parameters in the simulation, including the QCD coupling constant, renormalisation and factorisation scales, CKKW matching scale and parton shower tune variations (each varied up and down by a factor of two). The impact of these uncertainties on the signal yield range between 10% and 25% for the $\tilde{t}_1 \to t\tilde{\chi}_1^0$ models, but can become higher in other signal models. Due to the high number of signal models, only for those the variations were performed which are at the edge of exclusion at 95% confidence level[13]. The uncertainty of the production cross-section ($\sim 15-20\%$ for direct top-squark production [45]) was also taken into account.

Tables 5.18 and 5.19 show a summary of the dominating experimental and theory systematic uncertainties.

Table 5.18: All systematic uncertainties are given in percent relative to the total background yield if their impact is greater than 1% in at least one SR for SRA and SRB. μ_X are the uncertainties on the normalisation factors of the process X that come from statistical uncertainties in the CRs. MC statistical uncertainties were found to be negligible.

Region $\rightarrow$	**SRA-TT**	**SRA-TW**	**SRA-T0**	**SRB-TT**	**SRB-TW**	**SRB-T0**
Total syst. unc.	24	23	15	19	14	15
$t\bar{t}$ theory	10	6	3	10	11	12
$t\bar{t}+V/\gamma$ theory	2	< 1	< 1	1	< 1	< 1
Z theory	1	3	2	< 1	1	< 1
Single top theory	6	3	5	3	4	5
Diboson theory	< 1	2	< 1	< 1	< 1	< 1
$\mu_{t\bar{t}}$	< 1	< 1	< 1	2	2	1
$\mu_{t\bar{t}+Z}$	6	3	2	4	3	2
μ_Z	6	10	7	5	6	4
μ_W	1	1	1	2	1	2
$\mu_{\text{singletop}}$	5	3	5	4	4	5
JER	10	12	4	3	4	3
JES	4	7	1	7	4	<1
b-tagging	1	3	2	5	4	4
$E_{\text{T}}^{\text{miss}}$ soft term	2	2	< 1	1	< 1	< 1
Multijet estimate	1	< 1	< 1	2	2	< 1
Pileup	10	5	5	8	1	3

[13]Typical signal models with a CL_s value close to 0.05 – see Sec. 5.8 – where chosen by hand for this.

Table 5.19: All systematic uncertainties are given in percent relative to the total background yield if their impact is greater than 1% in at least one SR for SRC, SRD and SRE. μ_X are the uncertainties on the normalisation factors of the process X that come from statistical uncertainties in the CRs. MC statistical uncertainties were found to be negligible. Uncertainties in several SRC bins are in general quite large, which is mostly due to the limited sze of the MC samples that were used to estimate these uncertainties.

Region $\rightarrow$	**SRC1**	**SRC2**	**SRC3**	**SRC4**	**SRC5**	**SRD-low**	**SRD-high**	**SRE**
Total syst. unc.	31	18	18	16	80	25	18	22
$t\bar{t}$ theory	27	11	14	11	71	12	10	11
$t\bar{t}+V$ theory	< 1	< 1	< 1	< 1	< 1	< 1	1	
Z theory	< 1	< 1	< 1	< 1	< 1	< 1	< 1	2
W theory	< 1	< 1	1	3	2	< 1	< 1	1
Single top theory	3	2	2	3	< 1	5	6	12
$\mu_{t\bar{t}}$	4	6	6	5	5	1	1	< 1
$\mu_{t\bar{t}+Z}$	< 1	< 1	< 1	< 1	< 1	2	2	4
μ_Z	< 1	< 1	< 1	< 1	< 1	4	5	5
μ_W	< 1	< 1	1	3	3	3	1	2
$\mu_{\text{singletop}}$	3	2	2	3	< 1	5	6	6
JER	4	10	6	5	10	3	6	4
JES	4	5	2	2	17	8	4	5
b-tagging	2	2	< 1	2	4	9	7	< 1
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term	1	3	2	3	15	4	3	2
Multijet estimate	12	3	< 1	< 1	< 1	2	2	< 1
Pileup	< 1	1	< 1	2	14	9	< 1	2

5.8 Statistical interpretation

The goal of the statistical interpretation is to test the hypothesis that there are only SM processes, called H_0, against a hypothesis H_1 where there is additionally a signal. The methods for this test are described in the following. For a discovery significance we only need the p-value of H_0, which is defined as

$$p(x \geq x^{\text{observed}}|H_0) = 1 - \text{CL}_b = 1 - p(x < x^{\text{observed}}|H_0), \tag{5.17}$$

which is the likelihood of a test statistics x to be equal or above the observed value x^{observed} under the assumption that H_0 is true and that a potential signal would lead to higher values of x. The test statistics x can be chosen according to the needs (but should be chosen prior to unblinding to avoid a bias), but its probability density distribution should be known; in the simple case of a cut-and-count analysis with just one bin one can use the number of events in the signal region (x in H_0 corresponds then to the number of expected background events). The popularly quoted Z-value, which is the number of standard deviations in a normal distribution, is connected with the p-value by the equation

$$\int_Z^{+\infty} \mathscr{G}(x)\,\mathrm{d}x = p, \tag{5.18}$$

where $\mathscr{G}(x)$ is a standard Gaussian distribution. An observation of a signal can only be claimed if the Z-value exceeds 5, which corresponds to a p-value of $2.87 \cdot 10^{-7}$.
If no signal is observed that way, we want to evaluate if certain signal models can be excluded. For this we need the CL_{s+b} value, which is defined as

$$\text{CL}_{s+b} = p(x \leq x^{\text{observed}}|H_1), \tag{5.19}$$

which is the probability of x being below the observed value under the assumption that H_1 (which is the signal plus background hypothesis) is true. The CL_{s+b} value is not commonly used by ATLAS for exclusion, because it could happen (e.g. in an under-fluctuation in a

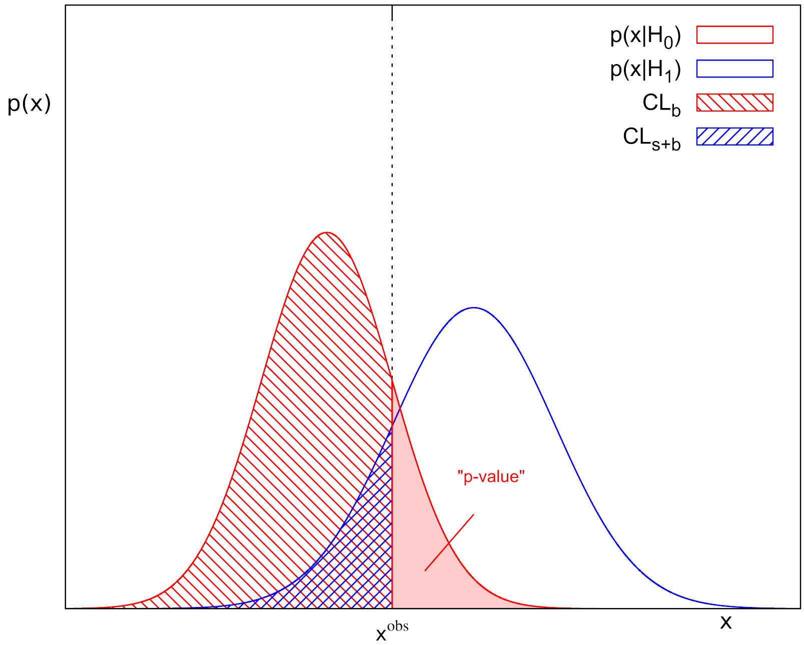


Figure 5.21: Illustration of the CL_{s+b} and CL_b values. The lines represent the probability density function for a certain value of the test statistic x assuming either the background-only hypothesis (H_0, red) or the signal+background hypothesis (H_1, blue). The hatched areas represent the integrated probabilities that define CL_{s+b} and CL_b. The integrated probability over all values of x is normalised to unity for both hypotheses.

cut-and-count signal region) that both the CL_b and CL_{s+b} values are low and we would exclude a signal even though we cannot confirm the background-only hypothesis and would not expect sensitivity to the signal model. To account for that, instead the CL_s value [133] is used, which is defined as

$$\mathrm{CL}_s = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b}. \tag{5.20}$$

The meaning of the CL_{s+b} and CL_b values are illustrated in Fig. 5.21. If $\mathrm{CL}_\mathrm{s} < 0.05$ the signal model is considered excluded at 95% confidence level.

The most powerful test statistic x to reject a hypothesis H_0 in favour of H_1 is according to the Neyman-Pearson lemma the likelihood ratio of the two hypotheses [134]. For numerical reasons usually the negative logarithm of this ratio q is used:

$$q_\mathrm{NP} = -2\ln\frac{L(\mathrm{observation}|H_0)}{L(\mathrm{observation}|H_1)}. \tag{5.21}$$

We use instead the Profile Likelihood test statistic, which is in general the choice for LHC searches and works well in the presence of nuisance parameters:

$$q_\mu = -2\ln\frac{L(\mu s+b,\hat{\hat{\theta}}_\mu)}{L(\hat{\mu}s+b,\hat{\theta})}. \tag{5.22}$$

The signal strength parameter μ is set to 0 for discovery (one wants to reject the background-only hypothesis b) and to 1 for exclusion (to reject the background + signal hypothesis s+b). $\hat{\hat{\theta}}_\mu$ is a set of nuisance parameters which is optimised for the fixed μ to maximise the likelihood. In the denominator $\hat{\mu}$ and the nuisance parameters $\hat{\theta}$ are both optimised for the maximum likelihood. A big advantage of the Profile Likelihood is that according to Wilks' theorem for a sufficiently high number of events q_μ follows a χ^2 distribution and therefore allows the use of asymptotic formulas [135, 136]. The likelihood function L is a product of the Poisson probability density functions $\mathscr{P}$ in each bin for the number of observed events in this bin for a given set of nuisance parameters and signal strength μ and the probability density function of the nuisance parameters NP, usually treated as standard Gaussian functions $\mathscr{G}$:

$$L(\text{obs}|\mu,\theta) = \prod_i^{\text{bins}} \mathscr{P}(N_i^{\text{obs}}|\mu\cdot s_i(\theta)+b_i(\theta))\cdot\Gamma(\theta_i^{\text{stat}}|\beta_i)\cdot\prod_j^{\text{NP}}\mathscr{G}(\theta_{ij}^{\text{syst}}|\text{mean}=0,\sigma=1), \tag{5.23}$$

where 'obs' is the entire observation in the signal region with N_i^{obs} being the observed number of events in each bin. The gamma functions $\Gamma(\theta_i^{\text{stat}}|\beta_i)$ describe the MC statistical uncertainty. Several bins can only be combined in this way into one test statistic if they are orthogonal. This means that the three bins of SRA can be combined with each other, as well as the three bins of SRB and the five bins of SRC, but not the two sub-regions of SRD or the different SRs with each other.

These methods are implemented in the software framework HistFitter [137] which is used for the statistical interpretation. It also processes the normalisation of the backgrounds in the CRs deriving scale factors f_α for each process α in its dedicated control region CRα.

This is done by the formula

$$f_\alpha = \frac{n_{\text{data,CR}\alpha} - n^{\text{others}}_{\text{sim,CR}\alpha}}{n^{\alpha}_{\text{sim,CR}\alpha}}, \tag{5.24}$$

'others' are in the case of the discovery fit all other background processes (background-only fit). n^{α}_{sim} corresponds to a naive normalisation to the best known cross section and luminosity. The other background processes in this formula are also scaled by the scale factor derived in their corresponding control region, which in return again depend on the other scale factors. This means that the fit has to be done simultaneously in all control regions. Since there are multiple CRs for $t\bar{t}$ and Z+jet, a choice has to be made, which one to use for the calculation of the scale factors of other processes. For the W+jet scale factor it was decided to use CRTD and CRZD. This was chosen because all other CRs explicitly select top quarks, which CRW vetoes. For CRST the regions CRTB-T0 and CRZAB-T0 were chosen because it was seen that most single-top events fall into the T0 category and the SRB-like regions are closer to CRST in terms of $E_{\text{T}}^{\text{miss}}$ requirements. The same regions were also used for the $t\bar{t}+\gamma$ CR, where most of the $t\bar{t}$ and Z+jet events fall into the T0 category.
For the calculation of discovery significances the background processes are fitted in the CRs and SRs, which is more conservative than fitting only in the CRs and then using the scale factors in the SRs. For the calculation of upper limits (exclusion) the fit is also done in the signal region and the signal strength μ is used as an additional parameter to be fitted in the CRs and SR.

5.9 Results

Table 5.20 shows the calculated scale factors derived in the a background-only fit, where the background scale factors are derived only in the CRs without assuming a signal. All scale factors are reasonably close to 1, which means that the amount of backgrounds in the measured data does not deviate too far from the nominal estimation based on simulation. The scale factors in the exclusion fit, which takes the signal contribution in the CRs into

Table 5.20: Scale factors derived from CRs in the background-only fit [4].

Sample	**Fitted scale factor**
W+jet (all SRs)	1.267 ± 0.146
$t\bar{t}$ (SRA-TT)	1.173 ± 0.146
$t\bar{t}$ (SRA-TW)	1.138 ± 0.112
$t\bar{t}$ (SRA-T0)	0.898 ± 0.121
$t\bar{t}$ (SRB-TT)	1.202 ± 0.156
$t\bar{t}$ (SRB-TW)	0.969 ± 0.0681
$t\bar{t}$ (SRB-T0)	0.924 ± 0.0525
$t\bar{t}$ (SRC)	0.707 ± 0.0498
$t\bar{t}$ (SRD)	0.945 ± 0.103
$t\bar{t}$ (SRE)	1.012 ± 0.180
Single top (all SRs)	1.166 ± 0.390
Z+jet (SRA,B TT and TW)	1.170 ± 0.238
Z+jet (SRA,B T0)	1.131 ± 0.144
Z+jet (SRD)	1.035 ± 0.146
Z+jet (SRE)	1.185 ± 0.152
$t\bar{t}+V$	1.290 ± 0.204

account, vary for every signal model and will therefore not explicitly be given here.
After applying the scale factors to the simulated yields, the predicted yields in the VRs and SRs can be compared to data. Figure 5.22 shows a summary of the yields in all VRs compared to the post-fit expectation. The differences fluctuate only within the expected uncertainty, giving us confidence that our methods works well.

The expected and measured yields in the SRs are shown in Tabs 5.21, 5.22 and 5.23. A graphical summary is shown in 5.23.

It can be seen that we do not observe any significant excess over the background expectation. The largest upwards deviation is seen in SRB-T0 with 179 ± 26 events expected but 206 observed, which corresponds to a p-value of 0.13 or a significance of 1.15 [4]. SRC4 shows a larger deviation with 7.7 events expected but only 1 event observed, which corresponds to a CL_b value of 0.01, however this does not corespond to a discovery significance , since only positive deviations are interpreted as evidence for a signal. Overall there is no significant excess and thus no finding of new physics can be reported. The distributions of some

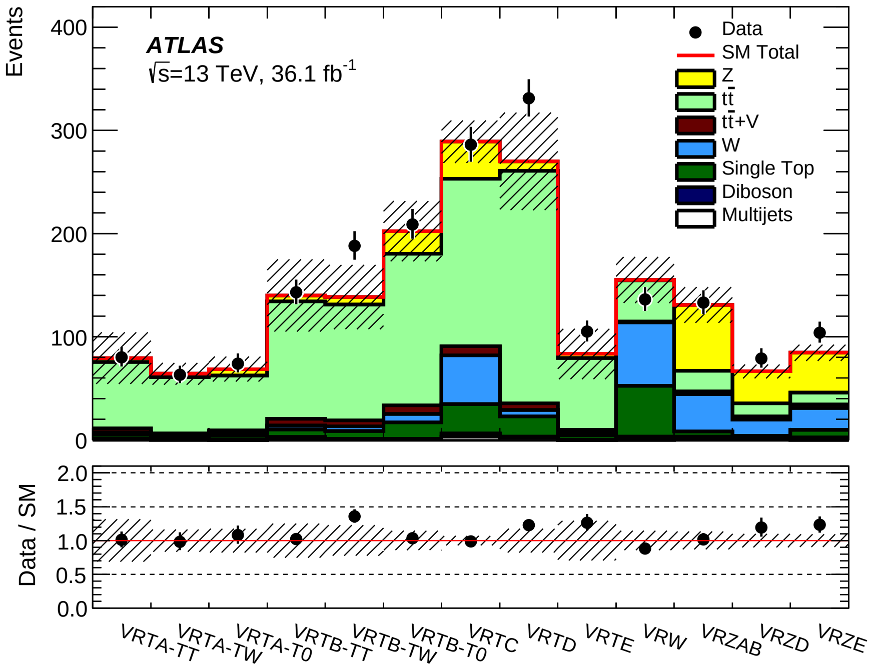


Figure 5.22: Summary of the yields in all validation regions after the likelihood fit [4]. SM background processes are depicted as coloured areas stacked on each other. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties, MC statistical uncertainties, and theoretical uncertainties in the extrapolation from the control region.

Table 5.21: Observed and expected yields in SRA and SRB after applying the scale factors from the discovery fit. The uncertainties include MC statistical uncertainties, detector-related systematic uncertainties, and theoretical uncertainties in the extrapolation from CR to SR [4].

Region →	**SRA-TT**	**SRA-TW**	**SRA-T0**	**SRB-TT**	**SRB-TW**	**SRB-T0**
Observed	11	9	18	38	53	206
Total SM	8.6 ± 2.1	9.3 ± 2.2	19.7 ± 2.7	39.3 ± 7.6	52.4 ± 7.4	179 ± 26
$t\bar{t}$	$0.71^{+0.91}_{-0.71}$	$0.51^{+0.55}_{-0.51}$	1.31 ± 0.64	7.3 ± 4.3	12.4 ± 5.9	43 ± 22
W+jet	0.82 ± 0.15	0.89 ± 0.56	2.00 ± 0.83	7.8 ± 2.8	4.8 ± 1.2	25.8 ± 8.8
Z+jet	2.5 ± 1.3	4.9 ± 1.9	9.8 ± 1.6	9.0 ± 2.8	16.8 ± 4.1	60.7 ± 9.6
$t\bar{t}+V$	3.16 ± 0.66	1.84 ± 0.39	2.60 ± 0.53	9.3 ± 1.7	10.8 ± 1.6	20.5 ± 3.2
Single top	1.20 ± 0.81	0.70 ± 0.42	2.9 ± 1.5	4.2 ± 2.2	5.9 ± 2.8	26 ± 13
Diboson	-	0.35 ± 0.26	-	0.13 ± 0.07	0.60 ± 0.43	1.04 ± 0.73
Multijet	0.21 ± 0.10	0.14 ± 0.09	0.12 ± 0.07	1.54 ± 0.64	1.01 ± 0.88	1.8 ± 1.5

Table 5.22: Observed and expected yields in SRC after applying the scale factors from the discovery fit. The uncertainties include MC statistical uncertainties, detector-related systematic uncertainties, and theoretical uncertainties in the extrapolation from CR to SR [4].

Region $\rightarrow$	**SRC1**	**SRC2**	**SRC3**	**SRC4**	**SRC5**
Observed	20	22	22	1	0
Total SM	20.6 ± 6.5	27.6 ± 4.9	18.9 ± 3.4	7.7 ± 1.2	0.91 ± 0.73
$t\bar{t}$	12.9 ± 5.9	22.1 ± 4.3	14.6 ± 3.2	4.91 ± 0.97	$0.63^{+0.70}_{-0.63}$
W+jet	0.80 ± 0.37	1.93 ± 0.49	1.91 ± 0.62	1.93 ± 0.46	0.21 ± 0.12
Z+jet	-	-	-	-	-
$t\bar{t}+V$	0.29 ± 0.16	0.59 ± 0.38	0.56 ± 0.31	0.08 ± 0.08	0.06 ± 0.02
Single top	1.7 ± 1.3	$1.2^{+1.4}_{-1.2}$	1.22 ± 0.69	0.72 ± 0.37	-
Diboson	0.39 ± 0.33	$0.21^{+0.23}_{-0.21}$	0.28 ± 0.18	-	-
Multijet	4.6 ± 2.4	1.58 ± 0.77	0.32 ± 0.17	0.04 ± 0.02	-

Table 5.23: Observed and expected yields in SRD and SRE after applying the scale factors from the discovery fit. The uncertainties include MC statistical uncertainties, detector-related systematic uncertainties, and theoretical uncertainties in the extrapolation from CR to SR [4].

Region $\rightarrow$	**SRD-low**	**SRD-high**	**SRCE**
Observed	27	11	3
Total SM	25.1 ± 6.2	8.5 ± 1.5	3.64 ± 0.79
$t\bar{t}$	3.3 ± 3.3	0.98 ± 0.88	$0.21^{+0.39}_{-0.21}$
W+jet	6.1 ± 2.9	1.06 ± 0.34	0.52 ± 0.27
Z+jet	6.9 ± 1.5	3.21 ± 0.62	1.36 ± 0.25
$t\bar{t}+V$	3.94 ± 0.85	1.37 ± 0.32	0.89 ± 0.19
Single top	3.8 ± 2.1	1.51 ± 0.74	0.66 ± 0.49
Diboson	-	-	-
Multijet	1.12 ± 0.37	0.40 ± 0.15	-

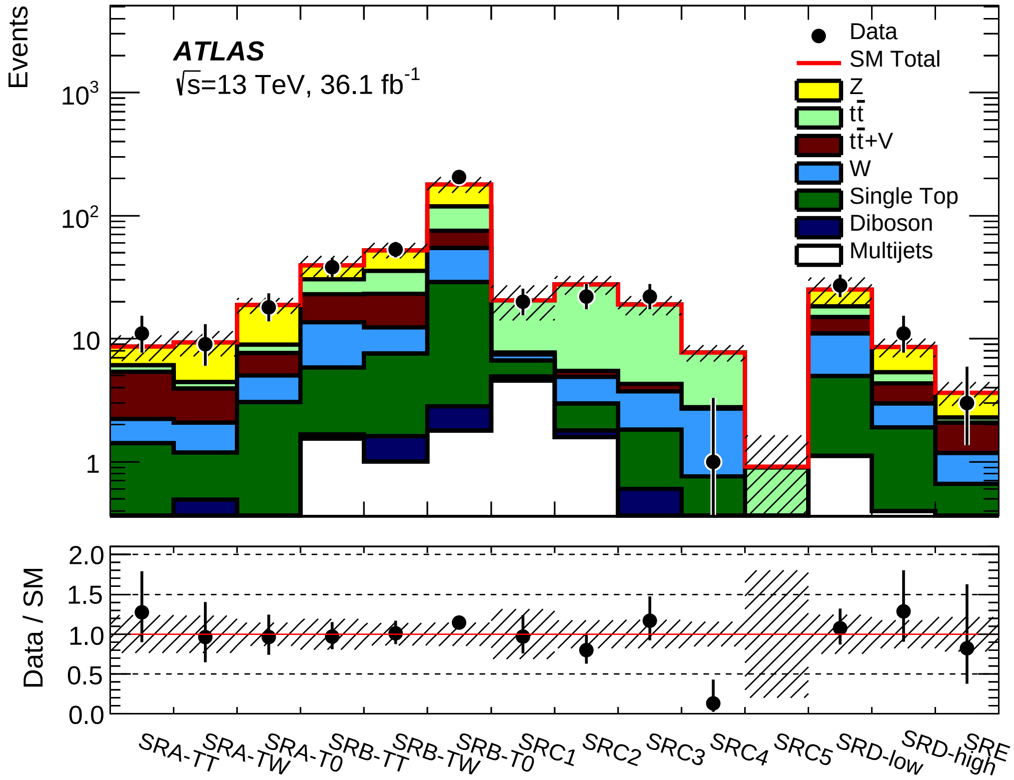
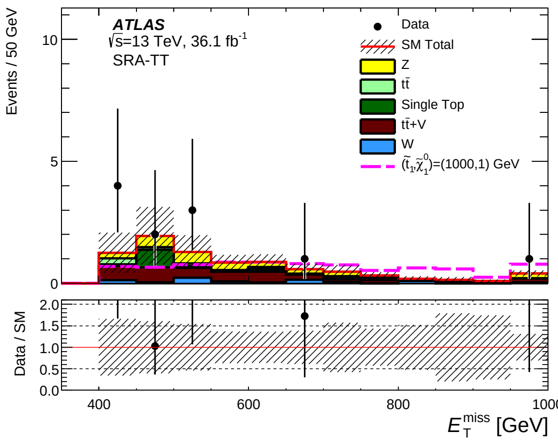


Figure 5.23: Summary of the yields in all signal regions after the likelihood fit [4]. SM background processes are depicted as coloured areas stacked on each other. The hatched band shows the MC uncertainty, consisting of statistical and detector-related systematic uncertainties, MC statistical uncertainties, and theoretical uncertainties in the extrapolation from the control region.

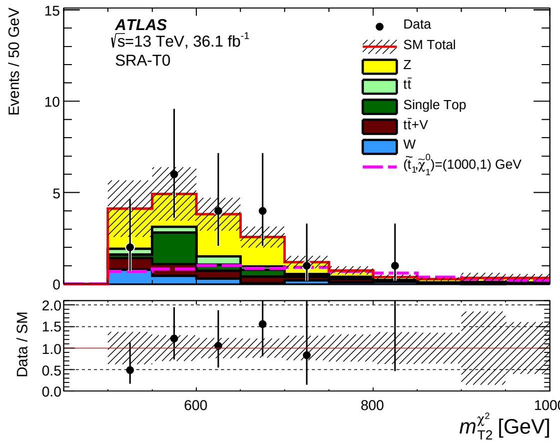
observables of data compared to the expected yields in chosen SRs are shown in fig. 5.24. For the interpretation into exclusion limits, for each signal model the SR with the lowest expected CL_s value (calculated as if we observed exactly the background expectation) was chosen and the observed CL_s was calculated for this SR to evaluate if the signal point is excluded. This results in exclusion contours, which are usually displayed in a two-dimensional plane, reducing the signal models to a grid of only two parameters. Regions between two signal points were interpolated based on the known evolution of the signal cross-section; the neighbouring signal points usually differ by 50 GeV in the dispayed particle masses, along the $m(\tilde{t}_1) = m(t) + m(\tilde{\chi}_1^0)$ diagonal a higher granularity of signal models was used. All signal models within the drawn contour are excluded at 95% confidence level. For models outside of the contour no statement can be made.

The first interpretation is done for the simplified model where only $\tilde{t}_1$ and $\tilde{\chi}_1^0$ are considered and the top squark decays purely via $\tilde{t}_1 \to t\tilde{\chi}_1^0$ (process depicted in Fig. 5.1a). The exclusion contour is shown as a function of $m(\tilde{t}_1)$ and $m(\tilde{\chi}_1^0)$ in Fig. 5.25. The different areas here were targeted by SRA (high $m(\tilde{t}_1)$), SRB (bulk region) and SRC (close to the $m(\tilde{t}_1) = m(t) + m(\tilde{\chi}_1^0)$ diagonal). Fig. 5.26 shows which signal region was the most sensitive for each signal point.

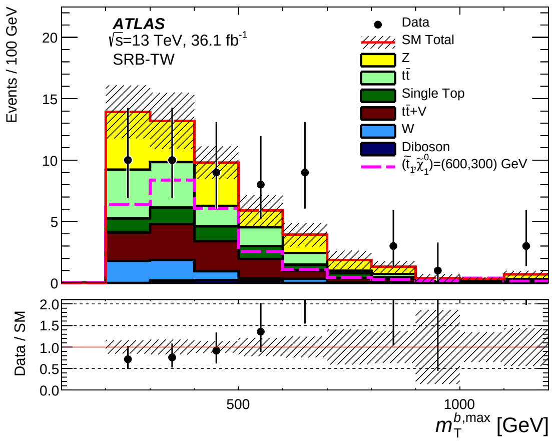
Then we can interpret the results for the case where $r := \mathrm{BR}(\tilde{t}_1 \to t\tilde{\chi}_1^0) < 100\%$ and the remainder decays via $\tilde{t}_1 \to b\tilde{\chi}_1^{\pm}$. This means that in r^2 of the events both top squarks decay via $\tilde{t}_1 \to t\tilde{\chi}_1^0$, in $(1-r)^2$ of the events both via $\tilde{t}_1 \to b\tilde{\chi}_1^{\pm}$ (process depicted in Fig. 5.1b) and for the remainder it is a mixed decay. This is depicted in Fig. 5.27 for different branching ratios, the chargino mass is set here to $m(\tilde{\chi}_1^{\pm}) = m(\tilde{\chi}_1^0) + 1\,\mathrm{GeV}$. This is motivated by the *naturalness* argument (see Sec. 2.2.1), reducing the necessary fine-tuning of the model, which leads to the lightest neutralino and chargino being quasi-degenerate [138]. It can be seen that the exclusion contour shrinks for lower r, because the signal regions are not as sensitive to events without top quarks. Especially SRA, which has low expected event numbers, suffers from lower expected signal yields, which decreases the exclusion from $m(\tilde{t}_1) < 1000\,\mathrm{GeV}$ at $r = 100\%$ to $m(\tilde{t}_1) < 820\,\mathrm{GeV}$ at $r = 0\%$ at $m(\tilde{\chi}_1^0) < 150\,\mathrm{GeV}$. The sensitivity in the region covored by SRB, which has higher expected event numbers, also decreases, but slightly less rapidly.



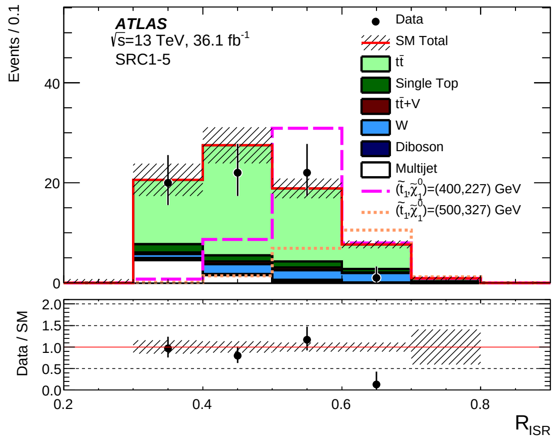
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ in SRA-TT.



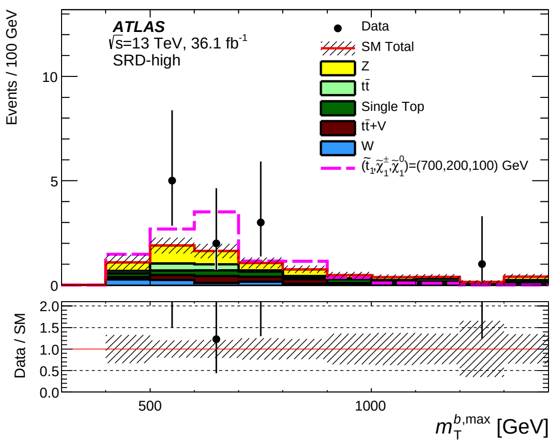
(b) m_{T2} in SRA-T0.



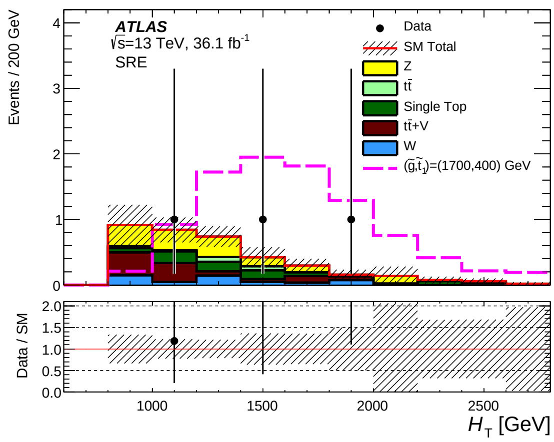
(c) $m_{\mathrm{T}}^{b,\mathrm{max}}$ in SRB-TW.



(d) R_{ISR} in SRC1–5.



(e) $m_{\mathrm{T}}^{b,\mathrm{max}}$ in SRD-high.



(f) H_{T} in SRE.

Figure 5.24: Distributions of several observables used in different signal regions in data and background after the likelihood fit [4]. The SM background processes are depicted as coloured areas stacked on each other. The hatched band shows the MC statistical and detector-related systematic uncertainties. The dashed lines correspond to the different SUSY benchmark signals explained in Sec. 5.5.

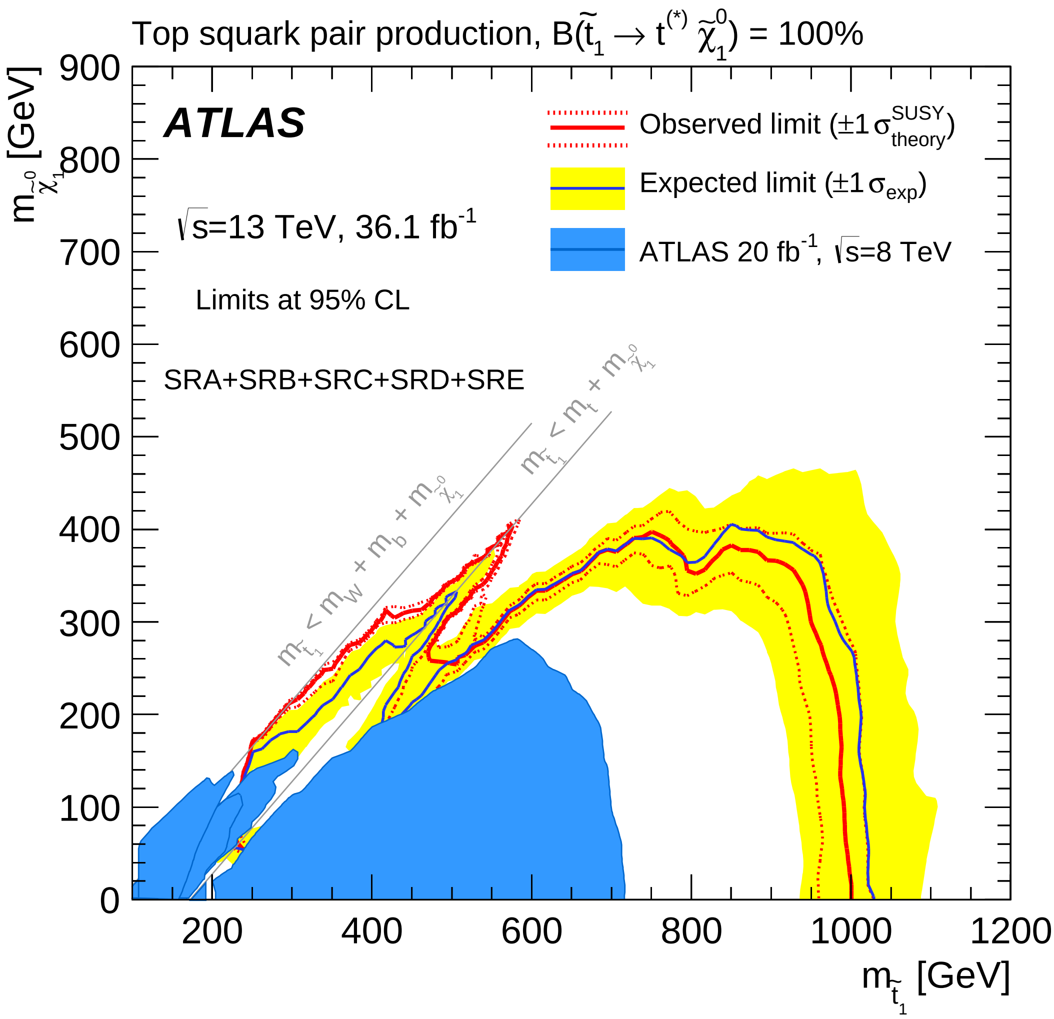


Figure 5.25: Observed (red line) and expected (blue line) exclusion contours for stop pair production in the case where $\mathrm{BR}(\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0) = 100\%$ in the $\tilde{t}_1$-$\tilde{\chi}_1^0$ mass plane [4]. The red dashed lines correspond to $\pm 1\sigma$ of the theoretical signal uncertainties. The yellow band is the $\pm 1\sigma$ variation of the limit expectation based on systematic and statistical uncertainty of the expected background yield. The area that was already excluded by different searches at $\sqrt{s} = 8\,\mathrm{TeV}$ is overlaid in blue [119].

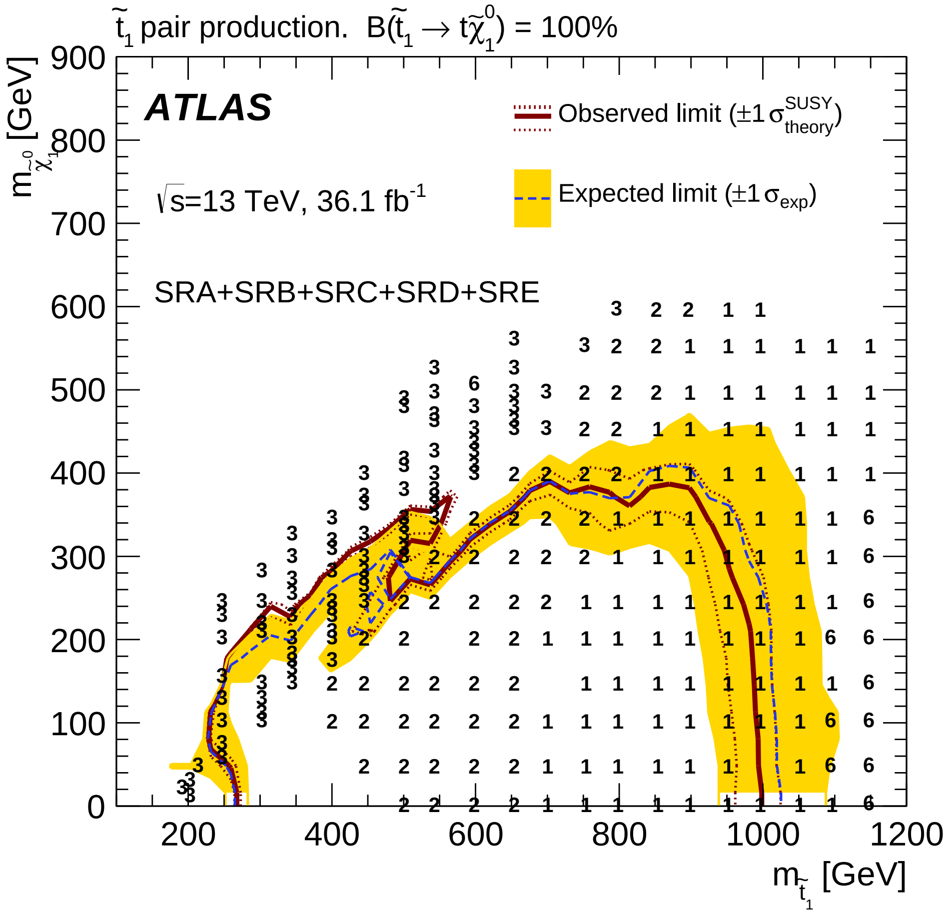


Figure 5.26: The same expected and observed exclusion contours as in Fig. 5.25 are shown, additionally for each signal point a number indicates which signal region gave the best expected CL_s value [4]. The numbers 1, 2, 3, 4, 5, 6 correspond to SRA, SRB, SRC, SRD-low, SRD-high, SRE, respectively. Itcan can be seen that SRA, SRB and SCR cover the regions they were designed for, SRE would be sensitive for models with higher $\tilde{t}_1$ masses, which cannot be included at this point. The SRD regions are unsurprisingly not sensitive for these signal models.

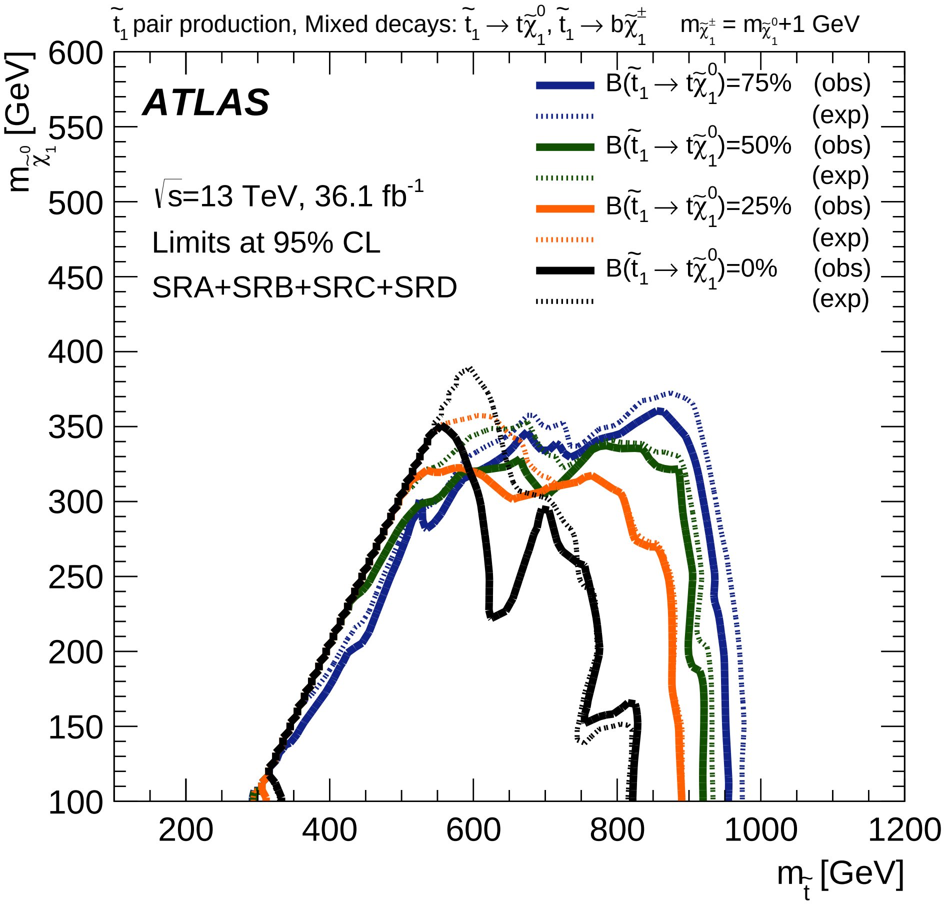


Figure 5.27: Exclusion contours in the $\tilde{t}_1$-$\tilde{\chi}_1^0$ mass plane for stop pair production different scenarios with different values for BR$(\tilde{t}_1 \to t\tilde{\chi}_1^0) = 100\%$ and the remainder decaying via $\tilde{t}_1 \to b\tilde{\chi}_1^{\pm}$, where $m(\tilde{\chi}_1^{\pm}) = m(\tilde{\chi}_1^0 + 1\,\text{GeV}$ [4]. The solid lines are the observed limits, the dashed lines the expectation. For the sake of clarity uncertainty bands are not drawn here, but can seen in dedicated exclusion graphics for each of the four scenarios in [4].

The last targeted simplified model is the gluino pair production, where gluinos decay via $\tilde{g} \to t\tilde{t}_1 \to t\tilde{\chi}_1^0 + \text{soft}$ (process depicted in Fig. 5.1c). The exclusion contour is shown in Fig. 5.28 as a function of $m(\tilde{g})$ and $m(\tilde{t}_1)$. The neutralino mass is set to $m(\tilde{\chi}_1^0) = m(\tilde{t}_1) - 5\,\text{GeV}$, because only a small mass gap leads to the targeted scenario with two tops and quasi-invisible stops (see Fig. 5.1c). For $m(\tilde{t}_1) < 600\,\text{GeV}$ signal models with $m(\tilde{g}) < 1800\,\text{GeV}$ can be excluded in this scenario.

The results can also be interpreted in full, more realistic models. One approach is the so-called *phenomenological MSSM* (pMSSM) [142, 143]. This SUSY framework reduces the

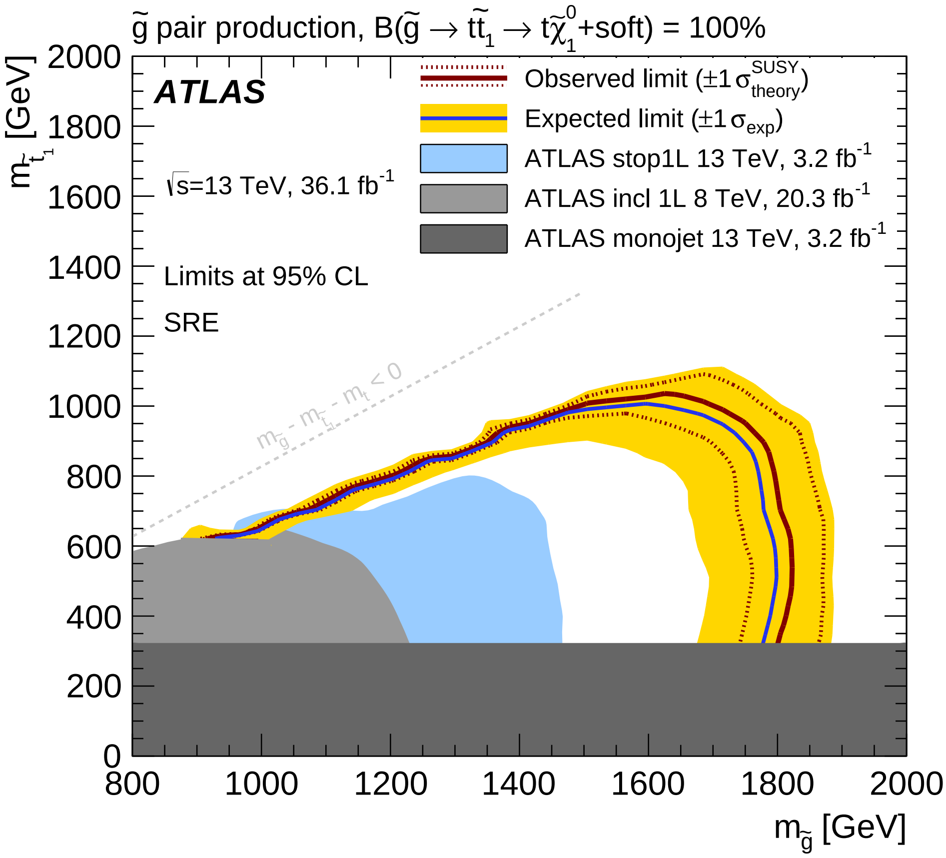


Figure 5.28: Observed (red line) and expected (blue line) exclusion contour in the $\tilde{g}$-$\tilde{t}_1$ mass plane for gluino pair production where the gluinos decay via $\tilde{g} \rightarrow t\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0 + \text{soft}$ [4]. The red dashed lines correspond to $\pm 1\sigma$ of the theoretical signal uncertainties. The yellow band is the $\pm 1\sigma$ variation of the limit expectation. Limits from other searches are overlaid [139–141].

number of free parameters by making some assumptions. These assumptions are: no R-parity violation; no new sources of CP and flavour violation; the first two generations are mass degenerate and their trilinear couplings are negligible. 19 free parameters remain, which are usually defined at the mean stop-mass scale: the gaugino mass parameters M_1 (bino), M_2 (wino) and M_3 (gluino); the ratio of the Higgs vacuum expectation values $\tan\beta$; the higgsino mass parameter μ and the mass of the pseudoscalar Higgs boson m_A; the sfermion mass parameters $m_{\tilde{f}}$, with $\tilde{f} = \tilde{Q}_k, \tilde{u}_k, \tilde{d}_k, \tilde{L}_k, \tilde{e}_k$ ($k = 1,2,3$, but the parameters are identical for $k = 1$ and $k = 2$); the trilinear couplings of the third generation A_t, A_b, A_τ. These parameters can then be set to build well-motivated, more realistic models, which usually have more than one decay channel for top squarks, and we can check if these models can be excluded with our results. One example that will be presented here is the so-called 'well-tempered neutralino' model [144]. Other interpretations can be found in [4]. The well-tempered neutralino model aims to obtain a good CDM candidate (see Sec. 2.2.1). The LSP is not necessarily a good CDM candidate in all SUSY models, because the resulting relic density deviates too far from the measurement of the Planck experiment $\Omega_{CDM}h^2 = 0.1188 \pm 0.0010$ [31]. In the considered models a relic density of $0.10 < \Omega_{CDM}h^2 < 0.12$ was achieved by setting $M_1 \approx -\mu$, which makes the LSP a bino-higgsino mixed state[14]. Signal points were created in the μ vs. $m_{\tilde{u}_3}$ and μ vs. $m_{\tilde{Q}_3}$ planes with reasonable choices for the other parameter, i.e. resulting in a Higgs boson mass close to 125 GeV, and the production of top and bottom squarks was considered. In these models we typically get $\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^0$ and $\tilde{\chi}_3^0$ masses which are not more than 70 GeV higher than the $\tilde{\chi}_1^0$ mass, which leads to several competing decay channels. The interpretation is done on the $\tilde{t}_1$-$\tilde{\chi}_1^0$ mass plane for the two different cases $\tilde{t}_1 = \tilde{t}_L$ and $\tilde{t}_1 = \tilde{t}_R$. In the former case top squarks with $m(\tilde{t}_1) < 800$ GeV can be excluded for $m(\tilde{\chi}_1^0) < 200$ GeV. In the latter case $m_{\tilde{Q}_3}$ is too high to give any contributions from sbottom production, which significantly decreases the sensitivity. The exclusions are shown in Fig. 5.29.

[14] This is not the only possibility to achieve a feasible relic density.

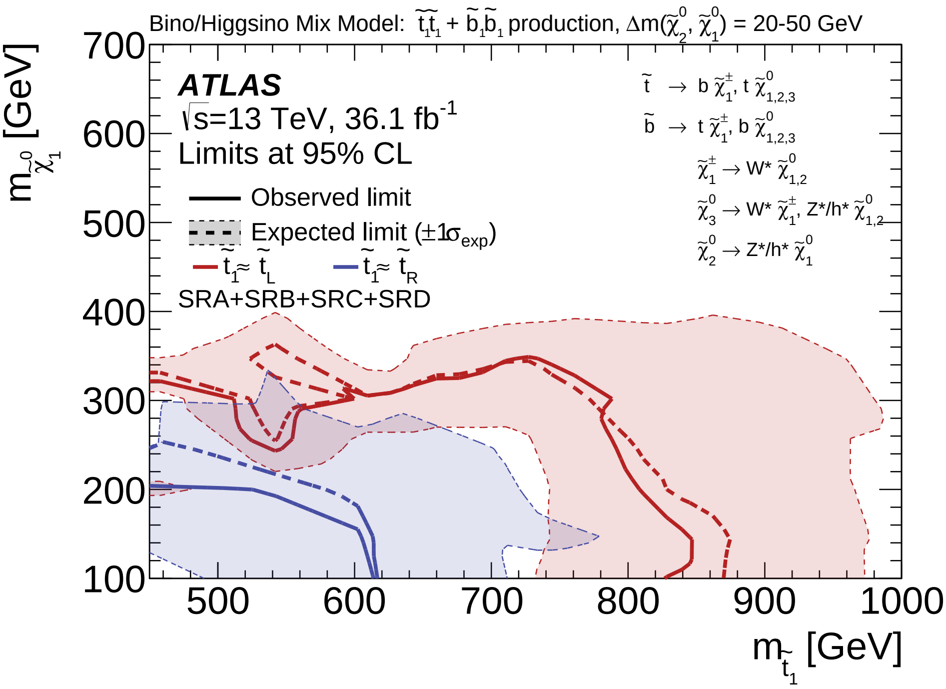


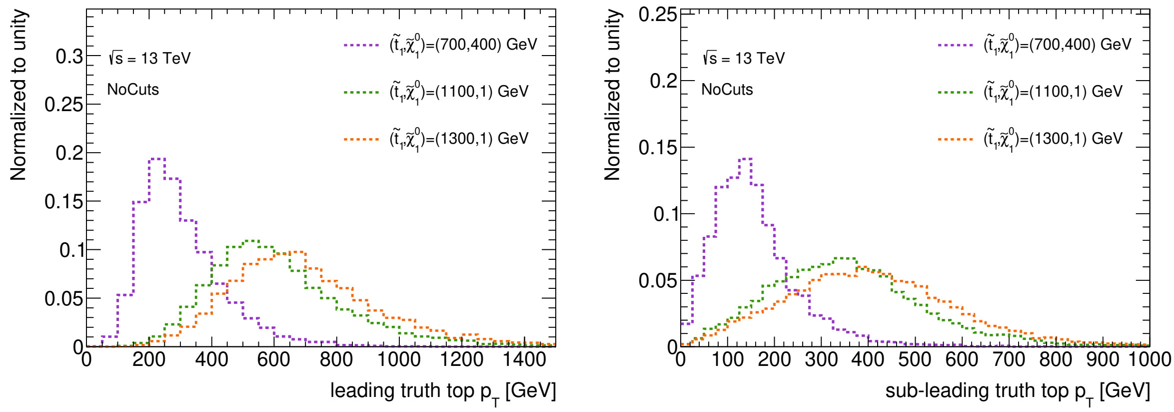
Figure 5.29: Observed (solid) and expected (dashed) exclusion contours for 'well-tempered neutralino' models in the $\tilde{t}_1$-$\tilde{\chi}^0_1$ mass plane [4]. The case $\tilde{t}_1 = \tilde{t}_L$ is drawn in red, the case $\tilde{t}_1 = \tilde{t}_R$ in blue. The coloured band corresponds to the $\pm 1\sigma$ variation of the limit expectation.

6 Studies with new methods

The following studies were conducted by the author in order to find methods to improve the results in Chpt. 5 in a follow-up analysis with a larger dataset. They were done with a set of MC samples that are mostly equivalent to the ones described in Sec. 5.2, but with a larger number of events and simulated conditions that represent the full data taken in 2015-2018. Real data was not used for these studies (including data-driven methods for background estimation) and the used object definitions (calibration and selection) do not exactly correspond to what will be the final recommendations. This should however only slightly change the results of these studies and is not expected to negate any of the conclusions. As a preselection we use similar criteria to Tab. 5.3, but use $\left|\Delta\phi\left(\text{jet}^{0-3}, \mathbf{p}_{\mathrm{T}}^{\text{miss}}\right)\right|_{\text{min}} > 0.4$ and avoid using $\mathbf{p}_{\mathrm{T}}^{\text{miss,track}}$, which might not be needed any more. Additionally $N_{b-\text{jets}} \geq 2$ is required, since the goal of this study is to improve the previous SRA and SRB in Chpt. 5.

6.1 Top identification

After the results in Chpt. 5 the focus is now on targeting models that are not yet excluded. This means that the new SRA should target scenarios with $m(\tilde{t}_1) > 1\,\text{TeV}$ and the new SRB should cover the gap between the bulk region and the $m(\tilde{t}_1) = m(t) + m(\tilde{\chi}_1^0)$ diagonal in Fig. 5.25. For SRA it means that the top quarks are in general more boosted and since new top identification methods for boosted tops have been developed, they should be tested for their potential to improve the analysis compared to our current method with $R = 1.2$ reclustered jets, described in Sec. 5.4. SRB has in general less boosted objects and one can try if other methods can improve the sensitivity here. The transverse momenta of the tops on generator level for different benchmark models are shown in Fig. 6.1. The goal is to optimise for discovery, therefore it was tried to improve the expected discovery significance (see Eq. 5.10).



(a) Leading top p_{T}.

(b) Sub-leading top p_{T}.

Figure 6.1: Transverse momentum of the leading and sub-leading top on generator level for different signal models.

6.1.1 Large-R-jet tagging

Jet-substructure based top-tagging methods have been developed specifically to label large-R jets as being likely to originate from a top quark and in similar ways also for other heavy objects like a W boson [145]. They are designed for a specific jet collection which reconstructs jets from individually calibrated noise-suppressed topological clusters of calorimeter cells with the anti-k_t algorithm using a radius parameter of $R = 1.0$. A grooming technique is applied to remove effects of pile-up or the underlying event. The used method (*trimming* [146]) reclusters the used jet constituents into smaller subjets (here with the k_t algorithm using a radius parameter of $R = 0.2$) and then discards all subjets with p_{T} smaller than a specific fraction of the total p_{T} of the original jet. The jets are then calibrated in two steps, correcting the jet energy scale and the jet mass scale [122, 147].

Jet moments are observables that quantify the jet substructure, calculated from jet constituents. Jet tagging algorithms make use of the different distributions in some of these moments of jets originating from light quarks or gluons and tops or other heavy objects. Some important jet moments are defined in the following:

- The *N-subjettiness* is calculated with the formula

$$\tau_N = \frac{1}{d_0} \sum_k p_{\mathrm{T},k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, ..., \Delta R_{N,k}\}, \tag{6.1}$$

  d_0 is the normalisation factor $d_0 = \sum_k p_{\mathrm{T},k} R_0$ ($R_0 = 1.0$ is the characteristic jet radius of the original jet); k runs over N jet constituents, which are obtained by using the exclusive-k_t algorithm, forcing the original jet into exactly N subjets [148]. τ_N of a jet quantifies to what degree it can be regarded as a composition of N subjets. Often a ratio is used:

$$\tau_{ij} = \frac{\tau_i}{\tau_j} \tag{6.2}$$

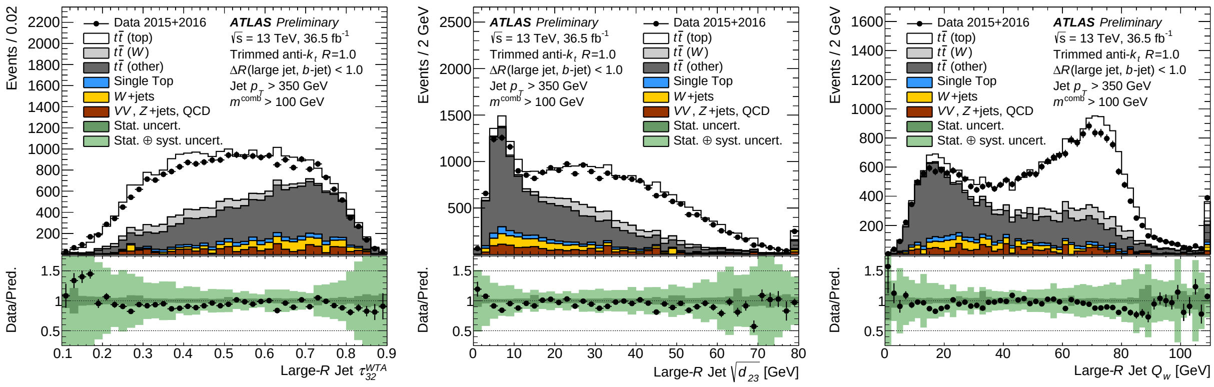
  The jet moment τ_{32}, which is important for top tagging, is depicted in Fig. 6.2a.

- The *splitting scale* $\sqrt{d_{ij}}$ is determined during reclustering the jet constituents with the k_t algorithm, which tends to cluster high-p_{T} constituents with large distances last. It is defined as

$$\sqrt{d_{ij}} = \min(p_{\mathrm{T},i}, p_{\mathrm{T},j}) \times \frac{\Delta R_{ij}}{R_0}, \tag{6.3}$$

  where $R_0 = 1.0$ is the distance parameter of the jet [149]. The splitting scale $\sqrt{d_{12}}$ is determined in the last clustering step; in case of a top i and j are expected to represent the b quark and the W boson. $\sqrt{d_{23}}$ is determined in the second-to-last clustering step and the subjets in case of a top are expected to represent the decay products of the W boson, the distribution is depicted in Fig. 6.2b.

- The variable Q_W is determined in the second-to-last reclustering step with the k_t algorithm. It is calculated as the minimum pair-wise invariant mass of the three subjets, which is expected to be at the W mass in case of a top [150]. This can be seen in Fig. 6.2c.



(a) τ_{ij}. **(b)** $\sqrt{d_{23}}$. **(c)** Q_W.

Figure 6.2: Several jet moments in data and MC [151]. $t\bar{t}$ events where a generator-level top can be matched with the jet are depicted in white. A preselection was applied to increase the number of top-matched events. The bottom panel shows the data/MC ratio and systematic uncertainties of experiment and theory.

- *Energy correlation functions* are subjet-independent methods to identify the N-pronged substructure by studying the angular separation and transverse momentum of combinations of the constituents [152]. One defines

$$e_2 = \frac{1}{(p_\mathrm{T}^\mathrm{jet})^2} \sum_{i<j\in J} p_{\mathrm{T}i} p_{\mathrm{T}j} R_{ij} \quad \& \quad e_3 = \frac{1}{(p_\mathrm{T}^\mathrm{jet})^3} \sum_{i<j<k\in J} p_{\mathrm{T}i} p_{\mathrm{T}j} p_{\mathrm{T}k} R_{ij} R_{jk} R_{ik}. \quad (6.4)$$

The indices i, j, k run over the jet constituents J. e_{N+1} is expected to be much smaller than e_N if a jet has N constituents. From these quantities one derives the variables

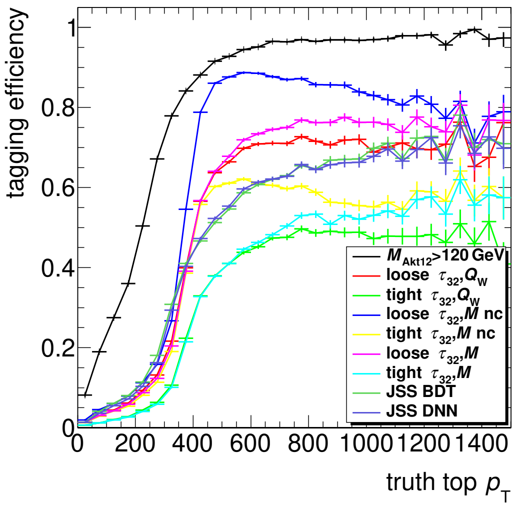
$$D_2 = \frac{e_3}{(e_2)^3} \quad \& \quad C_2 = \frac{e_3}{(e_2)^2}. \quad (6.5)$$

The e_i as well as D_2 and C_2 are dimensionless.

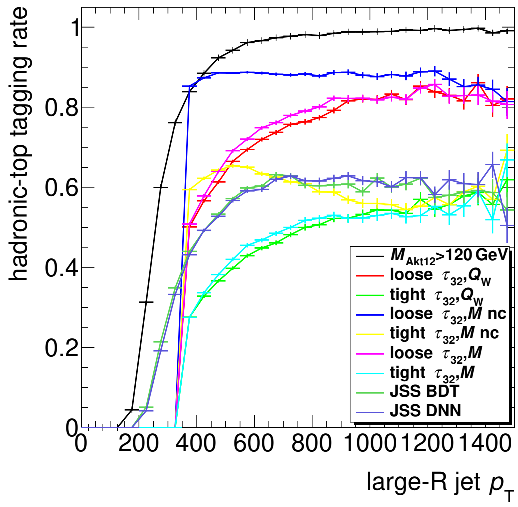
- The *combined mass* m_c is the linear combination of the jet masses calculated with a calorimeter-based method and a track-assisted method [147]. It is not considered a jet moment, but is also an important quantity for the identification of heavy objects.

Several simple two-variable tagging algorithms are available. For top-tagging there are three variants of variable combination, each one using an upper threshold on τ_{32} plus one criterion on another jet moment; the options are a lower threshold on the combined mass m_c, on Q_W or on $\sqrt{d_{23}}$. For W-tagging there is only one option using a lower and upper threshold on the combined mass and an upper threshold on D_2. The exact values are p_T-dependent and are optimised such that they result in a constant efficiency over the full p_T range for jets that fully contain all top decay products. This efficiency is tuned to either 50% (*tight* working point) or 80% (*loose* working point). Since especially decaying objects with low p_T do not necessarily result in all decay products being contained in one jet, an attempt was taken to optimise the tagger configuration also to increase the efficiency for jets where not all decay products are fully contained ('nc'='not-contained'). These configurations exist for the W-tagger and for the top-tagger that uses τ_{32} and m_c. More complex taggers use machine-learning techniques to analyse a large set of substructure variables. These work with either *Boosted Decision Trees* (BDT) or *Deep Neural Networks* (DNN) and are each configured as 80% efficiency tagger for tops or 50% efficiency taggers for W bosons.

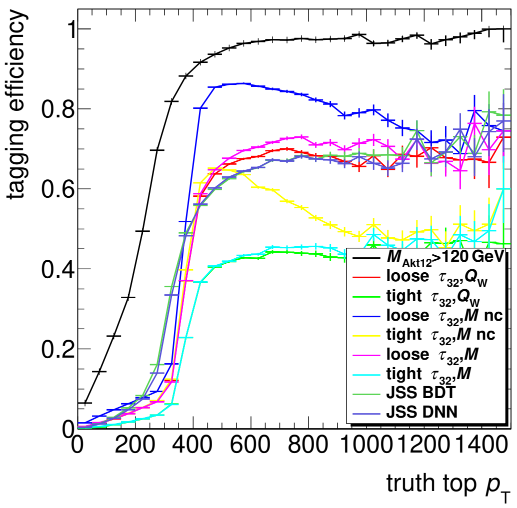
In our analysis the tops' transverse momenta are often not high enough (see Fig 6.1) that all decay products are 'fully-contained' in the jets. The efficiency for all kind of tops in the context of this analysis should be known, the design efficiency can only be expected at the high-p_T end of the spectrum. In addition we want to know how well the different taggers are performing in terms of background rejection in the context of this analysis. For this study we defined top-jets as such if the generator-level top quark is within $\Delta R < 1.0$ of the jet, without the need of its decay products being within the jet, but with the condition that the top quark decays hadronically. For the test with signal events a set of $\tilde{t}_1\tilde{t}_1 \rightarrow tt\tilde{\chi}_1^0\tilde{\chi}_1^0$ signal samples with different $\tilde{t}_1$ and $\tilde{\chi}_1^0$ masses was used to cover tops with different boosts. The efficiency can be depicted either as a function of the generator-level top p_T or the large-R jet p_T. Both versions are shown in Fig 6.3. In the first case we define the efficiency as the ratio of the number of generator-level tops matched with top-tagged jets over all generator-level tops, in the second case we define the hadronic top tagging rate as the number of tagged jets matched with a generator-level top over all jets matched with a generator-level top. As an



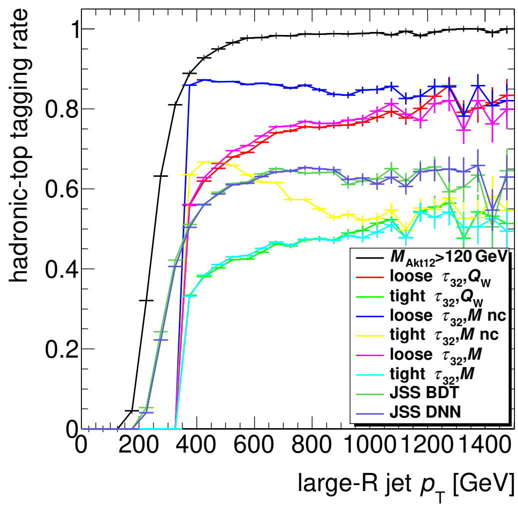
(a) Efficiencies in signal samples dependent on top p_T.



(b) Hadronic top tagging rate in signal samples dependent on jet p_T.



(c) Efficiencies in $t\bar{t}$ dependent on top p_T.



(d) Hadronic top tagging rate in $t\bar{t}$ dependent on jet p_T.

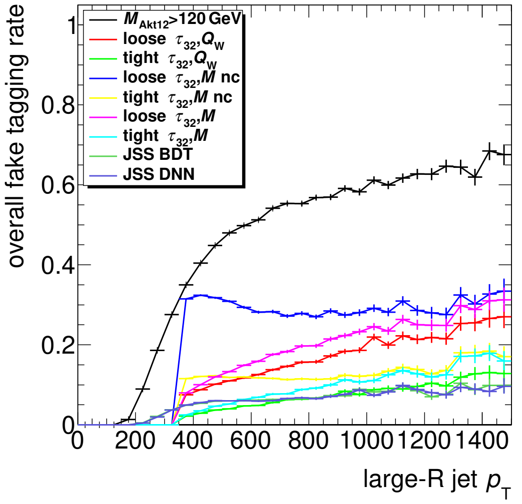
Figure 6.3: Efficiencies of different tagging algorithms. 'nc' = not-contained, 'loose' = 80% working point, 'tight' = 50% working point. A jet and a top are counted as matched, if $\Delta R(\text{top},\text{jet}) < 1.0$ for $R = 1.0$ jets and $\Delta R(\text{top},\text{jet}) < 1.2$ for $R = 1.2$ reclustered jets.

alternative to the signal samples also a $t\bar{t}$ sample was tested. The tested events have passed the preselection, as defined in Sec. 5.5. As a comparison to the new tagging methods also the method described in Sec. 5.4 is used, which uses reclustered $R = 1.2$ jets and counts jets as 'top-tagged' if $m_{\mathrm{jet},R=1.2} > 120\,\mathrm{GeV}$. It must be considered that the minimum jet p_{T} for the two-variable taggers is $350\,\mathrm{GeV}$ by design, $200\,\mathrm{GeV}$ for the BDT and DNN taggers and the reclustering method does not require a minimum p_{T}.

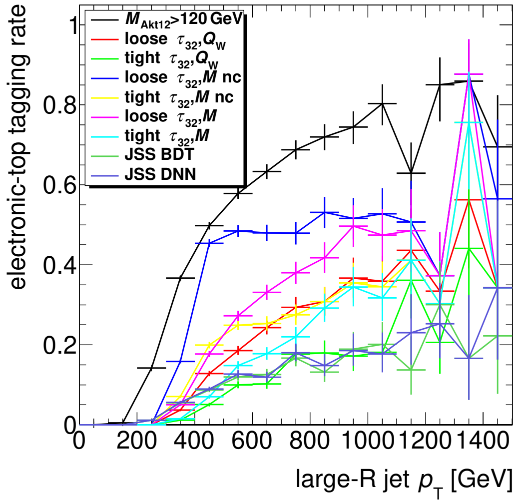
We are not displaying the efficiencies for the top-tagger based on τ_{32} and $\sqrt{d_{23}}$ or discussing it any further, because it is known to be more suited for highly boosted tops and was seen early on to perform worse than the others in the context of the analysis.

Since efficiencies do not show the whole picture about how useful the taggers are, we are also looking at the fake rates, meaning the tagging efficiencies of jets that do are not matched with a hadronically decaying top on generator level. Fake rates can look different depending on the process. In fig. 6.4 the fake rates for $t\bar{t}$ are shown; this process has hadronic tops, but only jets are considered that are not matched with a generator-level top. Separate plots are depicted for jets that are matched with leptonically decaying top (divided into decays with electrons, muons and taus) and jets without a top. Also the fake rate of Z+jets is shown, where there are no tops in any events.

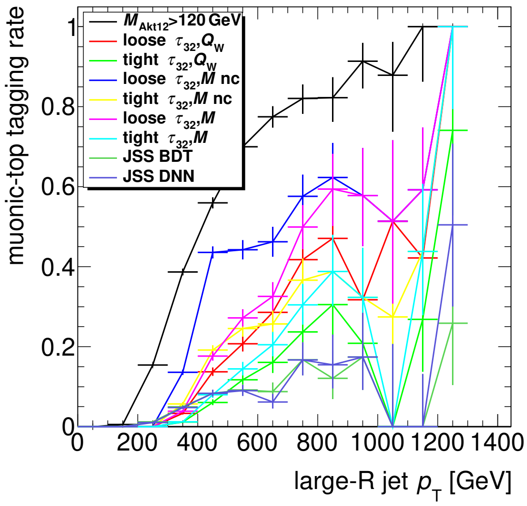
It can be seen that the reclustering method (black) ensures the highest efficiency, but also has very high fake rates, thus having no particularly good background rejection power. The MVA based taggers (DNN (dark blue) and BDT (dark green)) on the other hand have very low fake rates (only at low jet p_{T} the tight taggers perform better) while still obtaining a reasonably high efficiency. The fake rate is especially high if the large-R jet contains a top decaying via a tau, but is in general higher for leptonically decaying tops than for non-top jets.



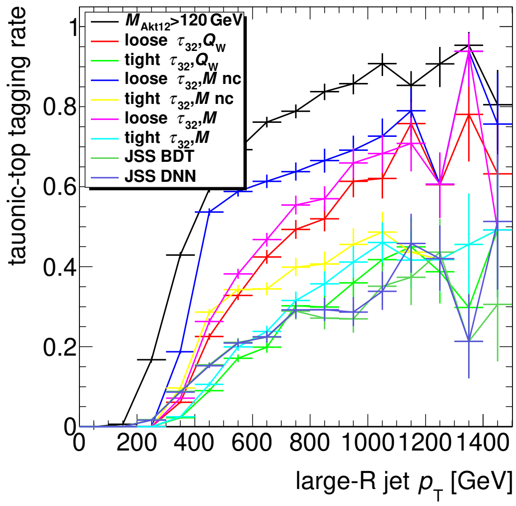
(a) all except $t \to bqq$



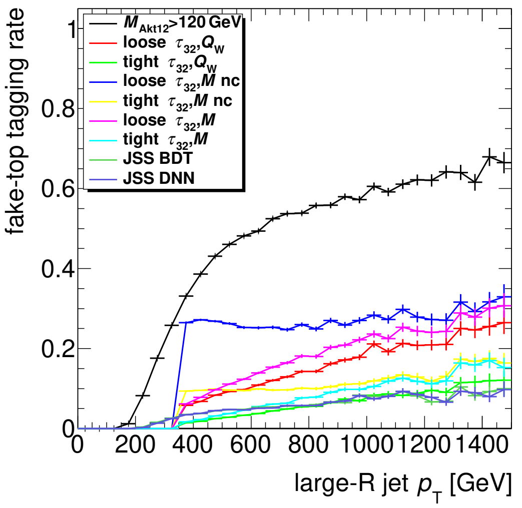
(b) $t \to be\nu$



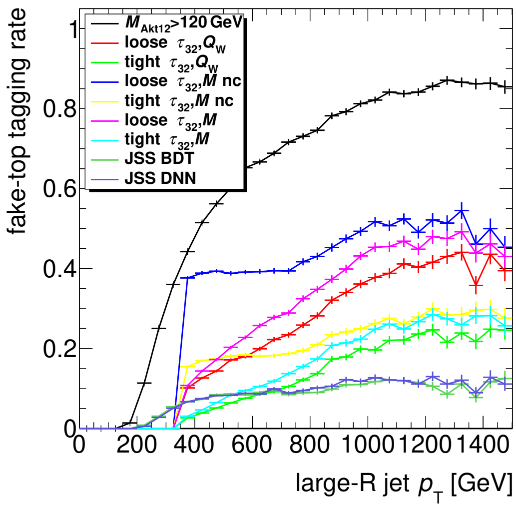
(c) $t \to b\mu\nu$



(d) $t \to b\tau\nu$



(e) no t

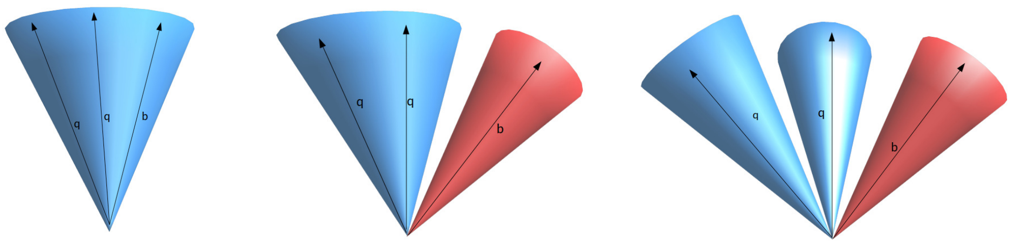


(f) no t (Z+jets sample)

Figure 6.4: Fake rates of different taggers ('nc' = not-contained) dependent on jet p_{T}. A jet and a top are counted as matched, if $\Delta R(\mathrm{top},\mathrm{jet}) < 1.0$ for $R = 1.0$ jets and $\Delta R(\mathrm{top},\mathrm{jet}) < 1.2$ for $R = 1.2$ reclustered jets. Rates were evaluated in $t\bar{t}$ samples in a)–e) for jets matched with decaying tops or no top, in Z+jets samples for f).

6.2 Top identification in semi-boosted scenarios

In regions with semi-boosted tops, as in our current SRB, top-tagging has a low efficiency and is thus not the best choice when used as the only top-identification method, but it is still possible to find a better method than the mass of $R = 1.2$ jets. The method presented here extends the top reconstruction by allowing several definitions based on the decay kinematics, which are symbolically depicted in Fig. 6.5. The categories are:



(a) Monojet scenario. **(b)** Dijet scenario. **(c)** Trijet scenario.

Figure 6.5: Symbolic representation of the three different decay scenarios of a semi-boosted top quark.

- **Monojet tops:** All decay products have merged into one single jet. We identify it with top tagging, using the 50% working point of the τ_{32} and Q_W based 'contained' tagger. The choice for this tagging algorithm was made because the focus is here on high-p_T tops with this method, whereas low-p_T tops should be covered by the other two methods.

- **Dijet tops:** The decay products of the W boson have merged into one jet, the b-jet is separate[1]. In this case we identify the W-jet with the 50% 'contained' W-tagging algorithm and the b-jet with the usual MV2c10 algorithm. Since the two jets are not part of the same jet collection and an overlap removal between the two collections was not enforced, we must make sure that the two jets do not overlap. We can also

[1]The case where a b-jet and a light jet merge into one large-R jet is not considered here since such a jet is difficult to identify.

assume that the angular separation between the jets is not too large and we can use the invariant mass as identification for tops. Thus we demand:

$$1.0 < \Delta R(b,W) < 1.4 \quad \& \quad 100\,\text{GeV} < m(b,W) < 250\,\text{GeV}. \tag{6.6}$$

- **Trijet tops:** The b-jet and the two light jets are all separate, but are still expected to be close to each other. We identify a top by checking every possible combination of three $R = 0.4$ jets i, j, k for the following criteria: at least one of the jets is b-tagged; the angular distance between the jets is not too large.

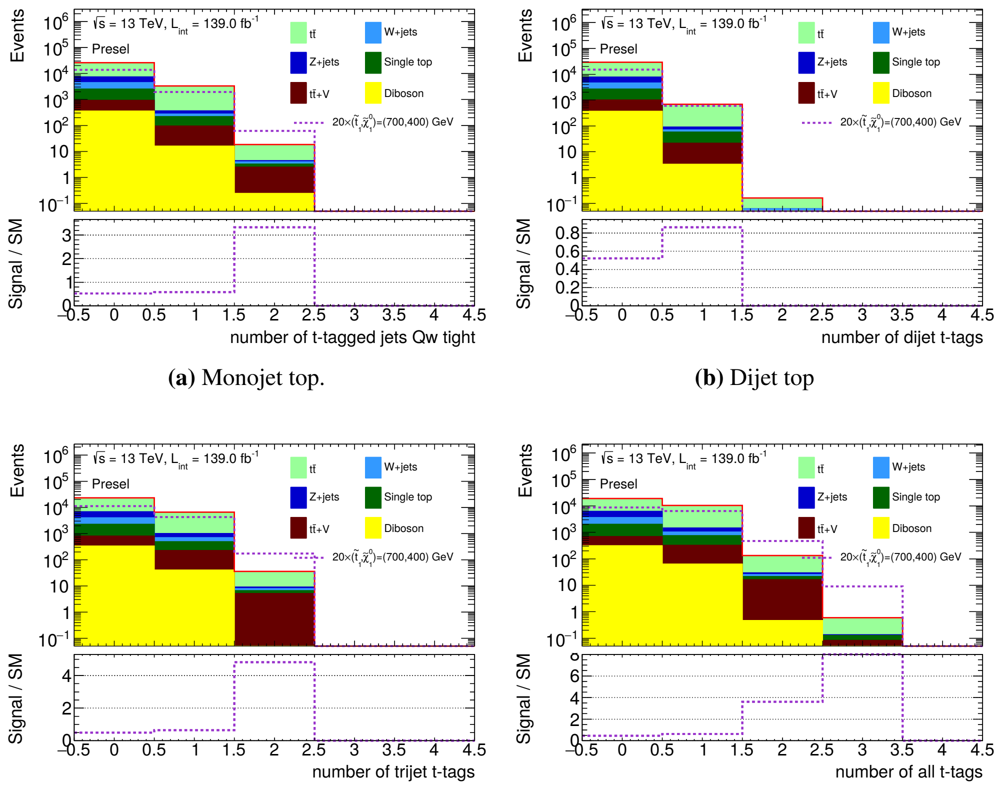
$$\max\{\Delta R(i,j), \Delta R(j,k), \Delta R(i,k)\} < 1.4. \tag{6.7}$$

  The invariant mass is required to be close to the top mass, the ratio of the invariant mass of the W candidate (defined as the combination of two of the jets with an invariant mass closest to the W mass) and the top candidate (all three jets) must be close to the ratio of the known particle masses.

$$120\,\text{GeV} < m(t_\text{cand}) < 230\,\text{GeV} \quad \& \quad 0.85 < \frac{m(t_\text{cand})/m(t)}{m(W_\text{cand})/m(W)} < 1.25. \tag{6.8}$$

  This is still a rather loose definition, but tightening the used criteria was not found to be favourable. New methods to further filter top candidates could be beneficial for this method.

It was taken care of that no object is double-counted, meaning that the large R-jets used successfully for the monojet identification and the $R = 0.4$ jets that are within $\Delta R < 1.0$ of them are not reused for the dijet and trijet identification, and in the same way objects used for the dijet identification were not re-used for the trijet case. The exact values of these criteria were tuned so that the background contribution is still acceptable (which means that the mass requirements in the trijet case need to be stricter). Counting the number of top tags with these methods one gets the distributions shown in Fig. 6.6. Most tags for the



(a) Monojet top. **(b)** Dijet top

(c) Trijet top. **(d)** All combined

Figure 6.6: Number of top tags with the different methods after preselection. The SM background processes are depicted as coloured areas stacked on each other. The dashed line is a SUSY signal with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (700, 400)$ GeV scaled by a factor of 20. In d) all tags with all three methods are counted.

$m(\tilde{t}_1, \tilde{\chi}_1^0) = (700, 400)$ GeV signal come from the trijet method, but in a significant number of events at least one top is identified with the mono- or dijet method. Adding up all tags we get a good signal/background ratio for events with two tags while maintaining high statistics. Due to fakes, signal events with three tags can also occur in some cases, so signal regions should not exclude these events.

6.3 Conclusions of top identification study

It was seen that the MVA taggers are the most interesting ones as stand-alone to-identification method in the context of this analysis for regions with high $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ (SRA in Sec. 5.5), both perform very similarly, the DNN version slightly better than the BDT version. The not-contained taggers are also interesting, offering a good efficiency at low p_{T}, but especially the 'loose' version allowing significantly more fakes than other tagging algorithms. The 'contained' taggers perform especially well in the high-p_{T} region, which is in the context of this analysis only interesting if combined with other methods that target the low-p_{T} region. Such methods were introduced in Sec. 6.2 as the dijet and trijet methods. The combination of these methods is potentially interesting for regions with low $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ (SRB in Sec. 5.5). All tagging algorithms' performances differs vastly from the reclustering method, which suggests that the top identification method cannot be simply replaced in the signal regions discussed in Sec. 5.5, but a re-optimisation of the signal region has to be done, which should be done anyway for a higher integrated luminosity.
Later in this chapter optimisation attempts will be presented using the DNN tagger for SRA in Sec. 6.5.1 and the combined semi-boosted method for SRB in Sec. 6.5.2. Optimisation with other tagging algorithms was attempted as well, but did not lead to a better result and will hence not be presented here.

6.4 Object-based $\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}$ significance

A type of "$E_{\mathrm{T}}^{\mathrm{miss}}$ significance" was already used in Sec. 5.5. However, a newer method can be used, which calculates for each event the p-value of the null-hypothesis of having zero real $E_{\mathrm{T}}^{\mathrm{miss}}$ by taking into account the transverse momentum resolutions of all objects instead of using the estimator $\sqrt{H_{\mathrm{T}}}$ for the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution [153]. The $\mathbf{p}_{\mathrm{T}}$ resolution consists of two parts: the resolution parallel to the jet direction and the one perpendicular to it (related to the angular resolution). In our signal regions jets are the only objects, their resolution depends on their momentum and the detector region. A central jet with $p_{\mathrm{T}} = 20\,\mathrm{GeV}$ has a resolution in parallel direction of $\sim 22\%$ and in perpendicular directions of 5-7%, for a $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ jet it is only 7% in parallel and 1.1-1.6% in perpendicular directions [122]. An additional uncertainty comes from the probability that a jet is produced from pile-up, this probability can be estimated as a function of the JVT discriminant. For the soft term resolution, 10 GeV was used as a conservative estimator, obtained from $Z \to \mu\mu$ events with no additional jets [126].

The measurement of each object is considered independent and the measurement of the transverse momentum p_{T} and the azimuthal angle ϕ are considered uncorrelated. All the resolutions add up to a total variance component longitudinal (σ_{L}) and transverse (σ_{T}) to $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, and a correlation ρ_{LT}. The correlation is only introduced by rotating the coordinate system into the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ system, in the system of the jet the longitudinal and transverse variance are uncorrelated. The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance can be written in the L-T basis as

$$\mathscr{S}^2 = \frac{\left|E_{\mathrm{T}}^{\mathrm{miss}}\right|^2}{\sigma_{\mathrm{L}}^2(1-\rho_{\mathrm{LT}}^2)}. \tag{6.9}$$

The distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance after preselection, plus additionally the τ-veto and $m_{\mathrm{T}}^{b,\mathrm{min}} > 175\,\mathrm{GeV}$ applied to be closer to SRA and SRB, is displayed in Fig. 6.7. It can be seen that $E_{\mathrm{T}}^{\mathrm{miss}}$ significance could potentially be useful in both SRA and SRB. With $E_{\mathrm{T}}^{\mathrm{miss}}$ being in the numerator of Eq. 6.9 and the denominator not necessarily increasing for high genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, one can assume that there is some correlation of $\mathscr{S}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$; thus it has to be seen in the optimisation if using requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $E_{\mathrm{T}}^{\mathrm{miss}}$ or a combination of the two performs better in the different SRs.

6.5 Signal region optimisation

It was attempted to optimise the SRA and SRB described in Sec. 5.5 by applying these new methods. We assume an integrated luminosity of $139.0\,\mathrm{fb}^{-1}$ and a flat systematic uncertainty,

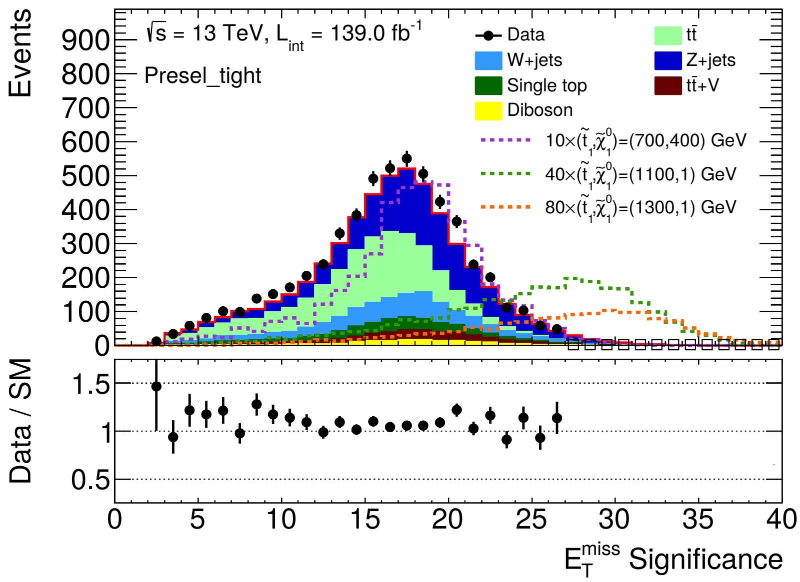


Figure 6.7: Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance after a preselection. Data events are shown in black, the error bars symbolise Poisson uncertainties. The SM background processes are depicted as coloured areas stacked on each other. Systematic errors were not evaluated in this preliminary phase and no scale factors from CRs were applied. The dashed lines correspond to SUSY signals with $m(\tilde{t}_1, \tilde{\chi}_1^0) = (700, 400)\,\mathrm{GeV}$ scaled by a factor of 10, $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1100, 1)\,\mathrm{GeV}$ scaled by a factor of 40 and $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1300, 1)\,\mathrm{GeV}$ scaled by a factor of 80. Data was blinded (indicated by boxes) in bins where a potential signal contribution would exceed 15%.

based on Tab. 5.18 and an estimation about future improvements[2]. All selection criteria are in addition to the preselection described in the beginning of this chapter.

6.5.1 SRA optimisation

An optimisation of SRA was performed, assuming a flat 20% systematic uncertainty on the background yields in addition to the MC statistical uncertainty. The benchmark signal $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1300, 1)$ GeV was used. Signal regions with only very few MC events would be problematic even though the calculated expected significance would suggest otherwise, because some SM processes would be considered as zero without any uncertainty in the calculation although this could happen just because of a lack of MC statistics. In order to prevent this problem, selections with too few events were avoided. We use the DNN taggers which were seen to work best. The number of tags is used to sort events into equivalent top identification categories (TT, TW, T0). All other variables are based on the strategy of the $36.1\,\mathrm{fb}^{-1}$ search with the exception of the newly added $E_{\mathrm{T}}^{\mathrm{miss}}$ significance. The first optimisation shows the optimal case, if DNN top- and W-tagging are fully usable[3]. A summary of the selection criteria is shown in Tab. 6.1. All selections are always on top of the preselection.

The second optimisation was done without W-tagging and with top-tagging only applied for jets with $p_{\mathrm{T}} > 300$ GeV. The reason for this optimisation is that top-tagging for jets with $p_{\mathrm{T}} < 300$ GeV and W-tagging was seen to differ significantly in fast calorimeter simulation vs. full simulation. Here, there is still a TT category, but instead of TW and T0, T-high ($E_{\mathrm{T}}^{\mathrm{miss}}$) and T-low ($E_{\mathrm{T}}^{\mathrm{miss}}$) is used. A summary of the selection criteria is shown in Tab. 6.2.

The third optimisation does the same, but not using thresholds on $E_{\mathrm{T}}^{\mathrm{miss}}$ or m_{T2} higher than 700 GeV in order to avoid cutting too much into the tails, which can be modelled unreliably. A summary of the selection criteria is shown in Tab. 6.3.

[2] In most regions the exact numbers for the systematic uncertainties would not have a big impact on the optimisation as the statistical uncertainties are dominating.

[3] Jet substructure variables are not well described by the fast calorimeter simulation, which was used for signal samples in this test. Especially all W-tagging algorithms were seen to have vastly different efficiencies when using full detector simulation. For the DNN top-tagger efficiencies were seen to be compatible.

Table 6.1: Summary of optimised SRA selection criteria obtained with top- and W-tagging, in addition to the preselection described in the text.

Variable	SRA-TT	SRA-TW	SRA-T0
τ-veto		✓	
$N_{\mathrm{top,DNN}}$	≥ 2	$= 1$	
$N_{\mathrm{W,DNN}}$		≥ 1	$= 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 700\,\mathrm{GeV}$	$> 550\,\mathrm{GeV}$	$> 900\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 100\,\mathrm{GeV}$	$> 100\,\mathrm{GeV}$	$> 100\,\mathrm{GeV}$
m_{T2}	$> 500\,\mathrm{GeV}$	$> 550\,\mathrm{GeV}$	$> 800\,\mathrm{GeV}$
$E_{\mathrm{T}}^{\mathrm{miss}}$ Sig.	> 5	> 5	> 5

Table 6.2: Summary of optimised SRA selection criteria obtained with top-tagging for jets with $p_{\mathrm{T}} > 300\,\mathrm{GeV}$, in addition to the preselection described in the text.

Variable	SRA-TT	SRA-T-high	SRA-T-low
τ-veto		✓	
$N_{\mathrm{top,DNN}}$	≥ 2	$= 1$	
$p_{\mathrm{T top_1,DNN}}$		$> 300\,\mathrm{GeV}$	
$p_{\mathrm{T top_2,DNN}}$	$> 300\,\mathrm{GeV}$	-	-
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 700\,\mathrm{GeV}$	$> 900\,\mathrm{GeV}$	$[600, 900]\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 75\,\mathrm{GeV}$	$> 100\,\mathrm{GeV}$	$> 150\,\mathrm{GeV}$
m_{T2}	$> 450\,\mathrm{GeV}$	$> 800\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$
$\Delta R(b,b)$	-	-	> 1.2
$E_{\mathrm{T}}^{\mathrm{miss}}$ Sig.	> 5	> 14	> 12

Table 6.3: Summary of optimised SRA selection criteria obtained with top-tagging for jets with $p_{\mathrm{T}} > 300\,\mathrm{GeV}$ and limits on $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T2} thresholds, in addition to the preselection described in the text.

Variable	SRA-TT	SRA-T-high	SRA-T-low
τ-veto		✓	
$N_{\mathrm{top,DNN}}$	≥ 2	$= 1$	
$p_{\mathrm{T}\,\mathrm{top}_1,\mathrm{DNN}}$		$> 300\,\mathrm{GeV}$	
$p_{\mathrm{T}\,\mathrm{top}_2,\mathrm{DNN}}$	$> 300\,\mathrm{GeV}$	-	-
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 700\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$	$[600, 700]\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 75\,\mathrm{GeV}$	$> 175\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$
m_{T2}	$> 450\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$	$> 550\,\mathrm{GeV}$
$\Delta R(b,b)$	-	> 0.8	> 1.2
$E_{\mathrm{T}}^{\mathrm{miss}}$ Sig.	> 5	> 14	> 5

At last an optimisation using the old $R = 1.2$ reclustering method is performed as a comparison, as before not using thresholds on $E_{\mathrm{T}}^{\mathrm{miss}}$ or m_{T2} higher than $700\,\mathrm{GeV}$. These selection criteria are shown in Tab. 6.4.

Table 6.4: Summary of optimised SRA selection criteria obtained with $R = 1.2$ reclustering in addition to the preselection described in the text.

Variable	SRA-TT	SRA-TW	SRA-T0
τ-veto		✓	
$m_{\mathrm{jet},R=0.8}^{0}$		$> 60\,\mathrm{GeV}$	
$m_{\mathrm{jet},R=1.2}^{0}$		$> 120\,\mathrm{GeV}$	
$m_{\mathrm{jet},R=1.2}^{0}$	$> 120\,\mathrm{GeV}$	$[60, 120]\,\mathrm{GeV}$	$< 60\,\mathrm{GeV}$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 700\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 175\,\mathrm{GeV}$	$> 225\,\mathrm{GeV}$	$> 300\,\mathrm{GeV}$
m_{T2}	$> 700\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$	$> 700\,\mathrm{GeV}$
$\Delta R(b,b)$	> 1.2	> 1.0	> 0.8
$E_{\mathrm{T}}^{\mathrm{miss}}$ Sig.	> 6	> 8	> 15

A comparison of the expected significances can be found in Tab. 6.5. The significances are calculated with the asymptotic formula (see Eq. 5.10). The combined significance is calculated by adding the significances in the single bins in quadrature.

It can be seen that “low”-$E_{\mathrm{T}}^{\mathrm{miss}}$ regions do not seem to be very useful. $E_{\mathrm{T}}^{\mathrm{miss}}$ significance does

Table 6.5: Summary of expected significances for $m(\tilde{t}_1, \tilde{\chi}_1^0) = (1300, 1)\,\mathrm{GeV}$ with different set of selection criteria in $139.0\,\mathrm{fb}^{-1}$. The T1 and T2 categories cannot directly be compared with each other, because their definition is very different. SRA-T1 means SRA-TW in the first and fourth case and SRA-T(high $E_{\mathrm{T}}^{\mathrm{miss}}$) in the second an third; SRA-T2 is SRA-TW in the first and fourth case and SRB-T(low $E_{\mathrm{T}}^{\mathrm{miss}}$) in the second an third. Events that are in the T1 category for one definition can be in the T2 category of another and vice versa.

	Significance ($\sigma = 20$ %) at $139.0\,\mathrm{fb}^{-1}$			
	SRA-TT	SRA-T1	SRA-T2	combined
Top- and W-tagging	1.66	0.78	1.62	2.44
Only top-tagging with $p_{\mathrm{T}}^{\mathrm{jet}} > 300\,\mathrm{GeV}$	1.70	1.68	0.83	2.53
+ $E_{\mathrm{T}}^{\mathrm{miss}}$&$m_{\mathrm{T2}} < 700\,\mathrm{GeV}$	1.70	1.51	0.34	2.30
$R = 1.2$ reclustering	1.72	0.69	0.70	1.98

not significantly improve the sensitivity here, just using a baseline threshold of $E_{\mathrm{T}}^{\mathrm{miss}}$ Sig. > 5 to reject multijet background instead of the presented thresholds here would not change much. The usage of the DNN top tagger however improves the sensitivity significantly. It can be seen, that especially regions with only one identified top contribute much more than with the reclustering method. In regions with two identified tops the reclustering method has a better sensitivity despite the higher background contribution. This is because with the current amount of data taken, events with two top tags are still very rare and the discovery sensitivity is statistically limited.

The expected yields of the third variant (only top-tagging for $p_{\mathrm{T}} > 300\,\mathrm{GeV}$ and limited values for $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T2}) can be found in Tab. 6.6. The background predictions of this of this variant are more reliable because phace space with potentially problematic MC modelling is excluded, however the relative background composition is similar to the second variant. The signal and background yields are both very small, which is mainly caused by the relatively low efficiency and high background rejection of the DNN tagger. This could lead to challenges for a reliable background modelling. The gain in expected discovery significance compared to the reclustering method however motivates to consider top-tagging in SRA.

Table 6.6: Summary of yields obtained with moderately optimised SRA with the selections shown in Tab. 6.3 assuming an integrated luminosity of $139.0\,\mathrm{fb}^{-1}$.

Process	SRA-TT	SRA-T-high	SRA-T-low
$t\bar{t}$	0.04 ± 0.01	0.11 ± 0.03	0.22 ± 0.05
$t\bar{t}+V$	0.30 ± 0.22	0.94 ± 0.15	0.97 ± 0.17
Single Top	-	0.28 ± 0.10	0.42 ± 0.24
Z+jets	0.09 ± 0.03	1.26 ± 0.10	1.12 ± 0.12
W+jets	0.02 ± 0.02	0.16 ± 0.05	0.02 ± 0.01
diboson	-	0.14 ± 0.10	0.13 ± 0.10
SM total	0.45 ± 0.10	2.91 ± 0.24	2.88 ± 0.44
$m(\tilde{t}_1, \tilde{\chi}_1^0) = (1100, 1)\,\mathrm{GeV}$	3.28 ± 0.23	8.07 ± 0.36	1.92 ± 0.17
$m(\tilde{t}_1, \tilde{\chi}_1^0) = (1300, 1)\,\mathrm{GeV}$	1.63 ± 0.08	3.22 ± 0.12	0.65 ± 0.05

Table 6.7: Summary of optimised SRB selection criteria obtained without new methods, in addition to the preselection described in the text.

Variable	SRB-TT	SRB-TW	SRB-T0
τ-veto		✓	
$m^0_{\mathrm{jet},R=1.2}$		$> 120\,\mathrm{GeV}$	
$m^1_{\mathrm{jet},R=1.2}$	$> 120\,\mathrm{GeV}$	$[60, 120]\,\mathrm{GeV}$	$< 60\,\mathrm{GeV}$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 300\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$	$> 300\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 200\,\mathrm{GeV}$	$> 225\,\mathrm{GeV}$	$> 200\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{max}}$	$> 250\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$	$> 200\,\mathrm{GeV}$
$\Delta R(b,b)$	> 1.2	> 1.2	> 1.2

6.5.2 SRB optimisation

An optimisation of SRB was performed, assuming a flat 15% systematic uncertainty[4] in addition to the MC statistical uncertainty. It was attempted to improve the expected discovery significance by using the top identification method introduced in Sec. 6.2. The benchmark signal $m(\tilde{t}_1, \tilde{\chi}_1^0) = (700, 400)\,\mathrm{GeV}$ was used.

The first optimisation is as a reference, only optimising the selections already used in Sec. 5.5. The resulting selection criteria can be found in Tab. 6.7.

The second optimisation uses the semi-boosted top identification method by adding up all reconstructed tops with the three methods introduced in Sec. 6.2 and depicted in Fig. 6.6.

[4] It was seen that the sum of systematic uncertainties are smaller in SRB than in SRA, see Tab. 5.18.

Table 6.8: Summary of optimised SRB selection criteria obtained with top-tagging for jets with $p_{\mathrm{T}} > 300\,\mathrm{GeV}$ and lower thresholds on $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T2}, in addition to the preselection described in the text.

Variable	SRB-TT	SRB-TW	SRB-T0
τ-veto		✓	
$N_{\mathrm{top,all}}$	≥ 2	$= 1$	
$N_{\mathrm{W,tight}}$	-	≥ 1	$= 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 300\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$	$> 300\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 200\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$	$> 200\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{max}}$	$> 300\,\mathrm{GeV}$	$> 200\,\mathrm{GeV}$	$> 300\,\mathrm{GeV}$
$\Delta R(b,b)$	> 1.0	> 1.0	> 1.4

For the 1-top regions a splitting into the categories TW and T0 was done by looking for additional W-tagged jets (by using the tight contained W-tagger on jets that were not yet used for the top identification). The resulting selection criteria can be found in Tab. 6.8.

The third optimisation in addition uses a lower threshold on $E_{\mathrm{T}}^{\mathrm{miss}}$ significance. Looking at ROC ('receiver operating characteristics') curves for lower thresholds in all of these variables, shown in Fig. 6.8, it can be seen that $E_{\mathrm{T}}^{\mathrm{miss}}$ significance has the best background rejection power while still maintaining a reasonably high signal acceptance. It should be noted that $m_{\mathrm{T}}^{b,\mathrm{min}}$ is still a very powerful discriminating variable; a threshold of $m_{\mathrm{T}}^{b,\mathrm{min}} > 175\,\mathrm{GeV}$, which removes a big portion of the $t\bar{t}$ background, was already applied for these curves, it is just not particularly useful to use a much higher threshold. $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is expected to have good discrimination power against processes with fake $E_{\mathrm{T}}^{\mathrm{miss}}$; since we are only using MC samples with genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, this effect cannot be seen here. The resulting selection criteria can be found in Tab. 6.9. In principle an upper threshold on $E_{\mathrm{T}}^{\mathrm{miss}}$ significance could also be used, which would further increase the discovery significance for the benchmark signal, since the background has a broader distribution. However, this could potentially damage the sensitivity for other signal points than the benchmark point, which can have different $E_{\mathrm{T}}^{\mathrm{miss}}$ significance distributions. This motivates an analysis with multiple bins in $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, which could potentially increase the sensitivity without any loss for other signal points. This possibility is not further explored here.

Another possibility to improve SRB is the usage of m_{T2}. Since signal models targeted by

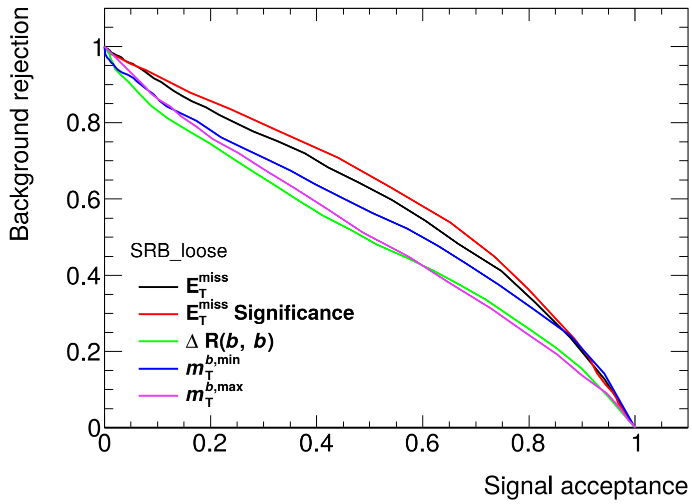


Figure 6.8: Background rejection vs. signal acceptance for all possible lower thresholds in the used variables in SRB. A good working point should be close to the upper right corner. For these curves a loose SRB-like selection was applied, meaning preselection, $m^0_{\text{jet},R=1.2} > 120\,\text{GeV}$, $m_T^{b,\text{min}} > 175\,\text{GeV}$ and $\Delta R(b,b) > 1.0$.

Table 6.9: Summary of optimised SRB selection criteria obtained with semi-boosted top-tagging and lower thresholds on E_T^{miss} significance in addition to the preselection described in the text.

Variable	SRB-TT	SRB-TW	SRB-T0
τ-veto		✓	
$N_{\text{top,all}}$	≥ 2	$= 1$	
$N_{\text{W,tight}}$	-	≥ 1	$= 0$
E_T^{miss}	$> 250\,\text{GeV}$	$> 250\,\text{GeV}$	$> 300\,\text{GeV}$
$m_T^{b,\text{min}}$	$> 200\,\text{GeV}$	$> 250\,\text{GeV}$	$> 200\,\text{GeV}$
$m_T^{b,\text{max}}$	$> 300\,\text{GeV}$	$> 200\,\text{GeV}$	$> 200\,\text{GeV}$
$\Delta R(b,b)$	> 1.0	> 1.0	> 1.4
E_T^{miss} Sig.	> 16	> 13	> 17

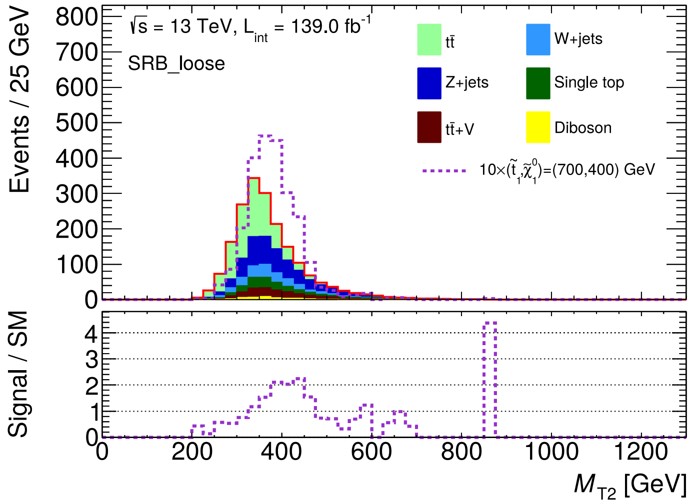


Figure 6.9: m_{T2} in a loose SRB-like selection, meaning preselection, $m^0_{\mathrm{jet},R=1.2} > 120\,\mathrm{GeV}$, $m_{\mathrm{T}}^{b,\mathrm{min}} > 175\,\mathrm{GeV}$ and $\Delta R(b,b) > 1.0$. The SM background processes are depicted as coloured areas stacked on each other. The dashed line corresponds to a SUSY signals with $m(\tilde{t}_1, \tilde{\chi}^0_1) = 700, 400\,\mathrm{GeV}$ scaled by a factor of 10.

SRB are expected to have slightly higher values and also a harder cut-off than backgrounds (see Sec. 5.4). This can be seen in Fig 6.9.

One can in principle use a lower and upper threshold. However it is better to avoid a lower thresholds since m_{T2} usually becomes lower for higher $m(\tilde{\chi}^0_1)$ and to restrict the sensitivity in this direction is undesirable. An advantage of an upper threshold is, that it would make SRB orthogonal to SRA, which opens the possibility to combine them statistically. This could smoothen exclusion limits in regions where both SRs are sensitive; since we are only looking at the expected significance of the benchmark point, this effect cannot be seen here. A set of selection criteria using a fixed upper threshold of $m_{\mathrm{T2}} < 450\,\mathrm{GeV}$ can be found in Tab. 6.10.

A comparison of the expected significances can be found in Tab. 6.11. The significances are calculated with the asymptotic formula (see Eq. 5.10). The combined significance is calculated by adding the significances in the single bins in quadrature.

Each of the tested methods can significantly improve the expected discovery significance compared to a SR that was just re-optimised with the old methods and should hence be considered for the final definition of SRB with the full Run 2 dataset. It can be seen that the

Table 6.10: Summary of optimised SRB selection criteria obtained with semi-boosted top-tagging, lower thresholds on $E_{\mathrm{T}}^{\mathrm{miss}}$ significance and a fixed upper threshold on m_{T2} in addition to the preselection described in the text.

Variable	SRB-TT	SRB-TW	SRB-T0
τ-veto		✓	
$N_{\mathrm{top,all}}$	≥ 2	$= 1$	
$N_{\mathrm{W,tight}}$	-	≥ 1	$= 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 300\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$	$> 300\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{min}}$	$> 200\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$	$> 250\,\mathrm{GeV}$
$m_{\mathrm{T}}^{b,\mathrm{max}}$	$> 200\,\mathrm{GeV}$	$> 200\,\mathrm{GeV}$	$> 200\,\mathrm{GeV}$
$\Delta R(b,b)$	> 1.0	> 1.0	> 1.4
m_{T2}		$< 450\,\mathrm{GeV}$	
$E_{\mathrm{T}}^{\mathrm{miss}}$ Sig.	> 15	> 13	> 18

Table 6.11: Summary expected significance for $m(\tilde{t}_1, \tilde{\chi}_1^0) = (700, 400)\,\mathrm{GeV}$ with different set of selection criteria in $139.0\,\mathrm{fb}^{-1}$.

	Significance ($\sigma = 15$ %) at $139.0\,\mathrm{fb}^{-1}$			
	SRB-TT	SRB-TW	SRB-T0	combined
Old	1.69	1.40	0.88	2.36
Semi-boosted top identification	2.43	1.45	1.55	3.22
+lower threshold on $E_{\mathrm{T}}^{\mathrm{miss}}$ Sig.	2.63	1.52	1.75	3.51
+ $m_{\mathrm{T2}} < 450\,\mathrm{GeV}$	3.16	1.68	2.24	4.22

Table 6.12: Expected yields in the optimised SRB obtained with the selections shown in Tab. 6.10 assuming an integrated luminosity of $139.0\,\text{fb}^{-1}$.

Process	SRB-TT	SRB-TW	SRB-T0
$t\bar{t}$	0.51 ± 0.06	0.61 ± 0.11	11.65 ± 0.67
$t\bar{t}+V$	1.98 ± 0.24	1.20 ± 0.17	13.36 ± 0.56
Single Top	0.13 ± 0.07	0.30 ± 0.23	11.94 ± 1.44
Z+jets	0.18 ± 0.04	0.30 ± 0.11	17.34 ± 1.24
W+jets	0.07 ± 0.04	0.16 ± 0.12	5.51 ± 0.95
diboson	-	-	1.13 ± 0.34
SM total	2.93 ± 0.27	2.67 ± 0.35	61.47 ± 2.33
$m(\tilde{t}_1, \tilde{\chi}_1^0) = (700, 400)\,\text{GeV}$	7.49 ± 1.22	3.44 ± 0.74	28.77 ± 2.26

TW category often does not contribute much, a splitting of the 1-top-tag region into TW and T0 might not be necessary.

The expected yields of the last variant (using semi-boosted top-tagging, a lower threshold on $E_{\text{T}}^{\text{miss}}$ significance and an upper threshold on m_{T2}) can be found in Tab. 6.12.

7 Conclusion

The Standard Model of Particle Physics is a very successful theory, bringing together the knowledge about particle properties and interactions collected in many experiments. Its predictions have been confirmed whenever an experiment was able to measure previously unknown phase space and it has withstood all attempts to directly disproof it. Despite its success there is evidence that it is not sufficient to fully describe the universe. Supersymmetry is a promising theory that can address several of the unsolved problems of the Standard Model, but none of its predicted additional particles has been observed so far.

The Large Hadron Collider and the ATLAS detector were built with the goals to find the last missing piece of the Standard Model – the Higgs Boson – and to search for evidence of physics beyond it. While Run 1 of the LHC was a big success in that the Higgs Boson was discovered, no discovery was made in the search for physics beyond the Standard Model. In Run 2 the centre-of-mass energy was increased from 8 to 13 TeV, which could allow to observe processes that were not seen before. Additionally the luminosity was increased, resulting in a total of $L_{\text{int}} = 139\,\text{fb}^{-1}$ of proton-proton collisions usable for physics recorded in the years 2015 - 2018.

A search for the top squark was conducted, which is a hypothetical scalar partner of the top quark predicted by supersymmetric models. A large set of R-parity conserving SUSY models with different parameters was considered. The most promising channel, resulting in the best discovery significance, is the production of a top-squark pair, each decaying into a top quark and a neutralino ($\tilde{t}_1\tilde{t}_1 \to tt\tilde{\chi}_1^0\tilde{\chi}_1^0$); the neutralino is considered the lightest SUSY particle and stable. This decay is kinematically possible if the top-squark mass exceeds the sum of the top and neutralino masses. Only hadronic top decays were considered for this search. Thus, final states with jets (including b-jets) and missing transverse momentum were considered. Signal regions were designed to optimise the expected discovery sensitivity for different models. Major Standard Model backgrounds were modelled by using simulated events, normalised to data in dedicated control regions. Data taken in 2015 and 2016 was used for this analysis, resulting in $L_{\text{int}} = 36.1\,\text{fb}^{-1}$. No evidence for a signal exceeding the Standard Model expectation was found. The results were interpreted as excluded areas of parameter space for different models, greatly superseding the limits obtained from Run 1. In the simplified model with $\text{BR}(\tilde{t}_1 \to t\tilde{\chi}_1^0) = 100\%$, top squarks with a mass below 950 GeV can be excluded for neutralinos below a mass of 350 GeV, in other models the exclusion limits are less stringent.

The full Run 2 data has not been analysed yet and the strategy could change with respect to this iteration. Several studies of new methods were conducted and their potential to improve the sensitivity was estimated by designing signal region suggestions. While models with high top-squark masses could profit from machine-learning technique based top tagging methods, for models with high neutralino masses the sensitivity could be improved by combining several top-reconstructing techniques that consider top signatures consisting of one, two or three jets, and by making use of $E_{\text{T}}^{\text{miss}}$ significance, a new observable which is a measure for the likelihood of an event to contain invisible particles. The conclusions of these studies will be considered in a subsequent analysis.

Thanks to the larger dataset, but also improved methods, the sensitivity can be expected to be widely expanded with the full Run-2 data, which will either result in an even larger exclusion of SUSY models – or a discovery.

8 Acknowledgements

First and foremost, I would like to thank Prof. Karl Jakobs, for giving me the opportunity to do the work for this thesis in his group in Freiburg and creating a nice atmosphere to work in. I would also like to thank Christian Weiser for his supervision and consultation during the second half of this time. Thanks go out to the ttbarMET0l analysis team around Walter Hopkins, Vasiliki Kouskoura and Matthias Saimpert, it was a great pleasure to be an active part of the team.
Special thanks go to Frederik Rühr, who was always a great help for me, thank you for guiding and mentoring me during these years. I also want to thank the other people that I worked with the closest in the Freiburg SUSY group: Kilian Rosbach, Julian Wollrath, Krzysztof Bożek and especially Christian Lüdtke, who has been my companion for a long time; thank you all for many fruitful discussions, being helpful in many ways and also a lot of mutual activities during and after work. I want to thank my lovely colleagues for an unforgettable time in Freiburg: Giulia Gonella, Théo Megy, Cedric Hönig, Frank Sauerburger, Ralf Gugel, Sven Mägdefessel, Arturo Rodriguez Rodriguez, Simona Gargiulo, Leena Diehl, Hannah Arnold, Felix Bührer, Philip Sommer, Daniel Büscher, Vakhtang Tsiskaridze, Carsten Burgard, Phuong Nguyen Dang, Claudia Giuliani, Liv Wiik-Fuchs, Stephen Jiggins, Shigeki Hirose, Lei Zhang, Karsten Köneke, Ulrich Parzefall and Ines Veiga; especially mentioned should be Christina Skorek, who has always had an open ear for my problems and was of great help with all kinds of administrative things. I also want to thank Prof. Wim de Boer, who convinced my to engage with high-energy physics in the first place.

Mein Dank geht natürlich auch an meine Familie, vor allem an meine tollen Eltern Elisabeth und Erwin Mogg. Vielen Dank für all den Rückhalt und die Unterstützung.

Zusammenfassung

In dieser Thesis wird die Suche nach einem supersymmetrischen Partner des Top Quarks mit dem ATLAS-Detektor bei $\sqrt{s} = 13\,\text{TeV}$ diskutiert. Das Standard Modell der Elementarteilchenphysik kann, obwohl es eine sehr erfolgreiche Theorie ist, keine alles beschreibende Theorie sein. Supersymmetrie ist eine mögliche Erweiterung, die als Lösungsansatz für mehrere offene Probleme des Standardmodells eingeführt werden kann. Eine ihrer Vorhersagen ist die Existenz von hypothetischen «Top-Squarks», die bei Proton-Proton-Kollisionen am Large Hadron Collider als Paar erzeugt werden könnten; das kann zu einem Topquark-Paar und fehlendem Transversalimpuls im Detektor führen. In dieser Thesis wird die Suche nach diesem Prozess im vollhadronischen Endzustand beschrieben; dies beinhaltet die Suchstrategie, die Untergrundabschätzung und die Untersuchung neuer Methoden, um die Sensitivität in der durchgeführten und einer nachfolgenden Suche zu verbessern. Mit $36\,\text{fb}^{-1}$ Proton-Proton-Kollisionsdaten, durchgeführt am LHC und aufgnommen von ATLAS in den Jahren 2015 und 2016, wurden in dieser Analyse keine Hinweise auf neue Teilchen gefunden; im vereinfachten Modell, das von einem Zerfall von Top-Squarks in je ein Topquark und ein Neutralino ausgeht, können Top-Squarks mit einer Masse unter $950\,\text{GeV}$ für Neutralino-Massen unter $350\,\text{GeV}$ ausgeschlossen werden. In dieser Thesis werden darüber hinaus Studien von neuen Methoden vorgestellt, die zur besseren Identifizierung von Top Quarks und von unsichtbaren Teilchen führen und die Sensitivität einer nachfolgenden Analyse verbessern können.

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.*, vol. B716, pp. 1–29, 2012, arXiv:1207.7214.

[2] CMS collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," *Phys. Lett.*, vol. B716, pp. 30–61, 2012, arXiv:1207.7235.

[3] ATLAS Collaboration, "Search for direct pair production of the top squark in all-hadronic final states in proton–proton collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector," *JHEP*, vol. 09, p. 015, 2014, arXiv:1406.1122.

[4] ATLAS Collaboration, "Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector," *JHEP*, vol. 12, p. 085, 2017, arXiv:1709.04183.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995.

[6] G. Altarelli, "Collider Physics within the Standard Model: a Primer," 2013, arXiv:1303.2842.

[7] D. J. Griffiths, *Introduction to Elementary Particles*. New York, USA: Wiley, 1987.

[8] E. Noether, “Invariante Variationsprobleme,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1918, pp. 235–257, 1918.

[9] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, “Advantages of the color octet gluon picture,” *Physics Letters B*, vol. 47, no. 4, pp. 365 – 368, 1973.

[10] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[11] A. Salam, “Weak and Electromagnetic Interactions,” *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.

[12] T. Nakano and K. Nishijima, “Charge Independence for V-particles,” *Progress of Theoretical Physics*, vol. 10, pp. 581–582, 11 1953.

[13] J. J. Thomson, “Cathode rays,” *Phil. Mag. Ser.5*, vol. 44, pp. 293–316, 1897.

[14] M. Tanabashi *et al.*, “Review of Particle Physics,” *Phys. Rev.*, vol. D98, no. 3, p. 030001, 2018.

[15] C. Patrignani *et al.*, “Review of Particle Physics,” *Chin. Phys.*, vol. C40, no. 10, p. 100001, 2016.

[16] S. T. Petcov, “The Nature of Massive Neutrinos,” *Adv. High Energy Phys.*, vol. 2013, p. 852987, 2013, arXiv:1303.5819.

[17] P. A. M. Dirac, “The quantum theory of the electron,” *Proc. Roy. Soc. Lond.*, vol. A117, pp. 610–624, 1928.

[18] M. Gell-Mann, “Symmetries of baryons and mesons,” *Phys. Rev.*, vol. 125, pp. 1067–1084, 1962.

[19] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[20] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[21] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

[22] S. P. Martin, "A Supersymmetry primer," 1997, arXiv:hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[23] M. M. Mühlleitner, "Supersymmetrie an Collidern," Lecture notes, 2012.

[24] H. Kalka and G. Soff, *Supersymmetry. (In German)*. Stuttgart, Germany: Teubner, 1997.

[25] H. Baer and X. Tata, *Weak scale supersymmetry: From superfields to scattering events*. Cambridge University Press, 2006.

[26] R. Barbieri and G. F. Giudice, "Upper Bounds on Supersymmetric Particle Masses," *Nucl. Phys. B*, vol. 306, p. 63, 1988.

[27] B. de Carlos and J. A. Casas, "One loop analysis of the electroweak breaking in supersymmetric models and the fine tuning problem," *Phys. Lett. B*, vol. 309, pp. 320–328, 1993, arXiv:hep-ph/9303291.

[28] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helv. Phys. Acta*, vol. 6, pp. 110–127, 1933. [Gen. Rel. Grav.41,207(2009)].

[29] E. Corbelli and P. Salucci, "The extended rotation curve and the dark matter halo of M33," *Monthly Notices of the Royal Astronomical Society*, vol. 311, pp. 441–447, 01 2000.

[30] D. Clowe, A. Gonzalez, and M. Markevitch, "Weak-lensing mass reconstruction of the interacting cluster 1e 0657-558: Direct evidence for the existence of dark matter," *The Astrophysical Journal*, vol. 604, pp. 596–603, 04 2004.

[31] P. A. R. Ade *et al.*, “Planck 2015 results. XIII. Cosmological parameters,” *Astron. Astrophys.*, vol. 594, p. A13, 2016, arXiv:1502.01589.

[32] U. Amaldi, W. de Boer, and H. Furstenau, “Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP,” *Phys. Lett.*, vol. B260, pp. 447–455, 1991.

[33] D. I. Kazakov, “Beyond the standard model: In search of supersymmetry,” in *2000 European School of high-energy physics, Caramulo, Portugal, 20 Aug-2 Sep 2000: Proceedings*, pp. 125–199, 2000, arXiv:hep-ph/0012288.

[34] S. R. Coleman and J. Mandula, “All Possible Symmetries of the S Matrix,” *Phys. Rev.*, vol. 159, pp. 1251–1256, 1967.

[35] R. Haag, J. T. Łopuszański, and M. Sohnius, “All possible generators of supersymmetries of the s-matrix,” *Nuclear Physics B*, vol. 88, no. 2, pp. 257 – 274, 1975.

[36] L. Girardello and M. T. Grisaru, “Soft Breaking of Supersymmetry,” *Nucl. Phys. B*, vol. 194, p. 65, 1982.

[37] P. Fayet, “Supersymmetry and Weak, Electromagnetic and Strong Interactions,” *Phys. Lett. B*, vol. 64, p. 159, 1976.

[38] P. Fayet, “Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions,” *Phys. Lett. B*, vol. 69, p. 489, 1977.

[39] H. Nishino *et al.*, “Search for Proton Decay via p —> e+ pi0 and p —> mu+ pi0 in a Large Water Cherenkov Detector,” *Phys. Rev. Lett.*, vol. 102, p. 141801, 2009, arXiv:0903.0676.

[40] B. Bajc, J. Hisano, T. Kuwahara, and Y. Omura, “Threshold corrections to dimension-six proton decay operators in non-minimal SUSY SU (5) GUTs,” *Nucl. Phys.*, vol. B910, pp. 1–22, 2016, arXiv:1603.03568.

[41] G. R. Farrar and P. Fayet, “Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry,” *Phys. Lett. B*, vol. 76, pp. 575–579, 1978.

[42] H. Goldberg, “Constraint on the Photino Mass from Cosmology,” *Phys. Rev. Lett.*, vol. 50, p. 1419, 1983. [Erratum: Phys. Rev. Lett.103,099905(2009)].

[43] J. R. Ellis, J. S. Hagelin, D. V. Nanopoulos, K. A. Olive, and M. Srednicki, “Supersymmetric Relics from the Big Bang,” *Nucl. Phys. B*, vol. 238, pp. 453–476, 1984.

[44] W. Beenakker, M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas, “Stop production at hadron colliders,” *Nucl. Phys.*, vol. B515, pp. 3–14, 1998, arXiv:hep-ph/9710451.

[45] C. Borschensky, M. Kramer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, “Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13$, 14, 33 and 100 TeV,” *Eur. Phys. J. C*, vol. 74, p. 3174, 2014, arXiv:1407.5066.

[46] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs, Geneva: CERN, 2004.

[47] L. Evans and P. Bryant, “LHC Machine,” *JINST*, vol. 3, p. S08001, 2008.

[48] F. Marcastel, “CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN,” Oct 2013. General Photo.

[49] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, p. S08003, 2008.

[50] S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, 2008.

[51] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3, p. S08002, 2008.

[52] A. A. Alves, Jr. *et al.*, "The LHCb Detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[53] G. Antchev *et al.*, "First measurement of elastic, inelastic and total cross-section at $\sqrt{s} = 13$ TeV by TOTEM and overview of cross-section data at LHC energies," *Eur. Phys. J.*, vol. C79, no. 2, p. 103, 2019, arXiv:1712.06153.

[54] ATLAS Collaboration, "LuminosityPublicResultsRun2," https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2, 2016.

[55] ATLAS Collaboration, "ATLAS Insertable B-Layer Technical Design Report,", ATLAS-TDR-19, 2010.

[56] ATLAS Collaboration, "ATLAS Insertable B-Layer Technical Design Report Addendum,", ATLAS-TDR-19-ADD-1, 2012.

[57] ATLAS Collaboration, "Performance of the ATLAS trigger system in 2015," *Eur. Phys. J. C*, vol. 77, p. 317, 2017, arXiv:1611.09661.

[58] R. Achenbach *et al.*, "The ATLAS level-1 calorimeter trigger," *JINST*, vol. 3, p. P03001, 2008.

[59] A. Collaboration, *ATLAS Computing: technical design report*. Technical Design Report ATLAS, Geneva: CERN, 2005.

[60] P. Calafiura, W. Lavrijsen, C. Leggett, M. Marino, and D. Quarrie, "The Athena Control Framework in Production, New Developments and Lessons Learned," 2005.

[61] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger, "The new ATLAS track reconstruction (NEWT)," *Journal of Physics: Conference Series*, vol. 119, p. 032014, jul 2008.

[62] ATLAS Collaboration, "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2," *Eur. Phys. J. C*, vol. 77, p. 673, 2017, arXiv:1704.07983.

[63] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl. Instrum. Meth.*, vol. A262, pp. 444–450, 1987.

[64] R. O. Duda and P. E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," *Commun. ACM*, vol. 15, pp. 11–15, Jan. 1972.

[65] ATLAS Collaboration, "Vertex Reconstruction Performance of the ATLAS Detector at$\sqrt{s} = 13$ TeV,", ATL-PHYS-PUB-2015-026, 2015.

[66] ATLAS Collaboration, "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC," *Eur. Phys. J. C*, vol. 77, p. 332, 2017, arXiv:1611.10235.

[67] W. Lampl *et al.*, "Calorimeter Clustering Algorithms: Description and Performance,", ATL-LARG-PUB-2008-002, 2008.

[68] ATLAS Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," *Eur. Phys. J. C*, vol. 77, p. 490, 2017, arXiv:1603.02934.

[69] ATLAS Collaboration, "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$," *Eur. Phys. J. C*, vol. 76, p. 292, 2016, arXiv:1603.05598.

[70] ATLAS Collaboration, "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$," *Eur. Phys. J.*, 2019, arXiv:1902.04655.

[71] ATLAS Collaboration, "Photon identification in 2015 ATLAS data,", ATL-PHYS-PUB-2016-014, 2016.

[72] T. G. Cornelissen, M. Elsing, I. Gavrilenko, J. F. Laporte, W. Liebig, M. Limper, K. Nikolopoulos, A. Poppleton, and A. Salzburger, “The global chi**2 track fitter in ATLAS,” *J. Phys. Conf. Ser.*, vol. 119, p. 032013, 2008.

[73] ATLAS Collaboration, “Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung,”, ATLAS-CONF-2012-047, Geneva, May 2012.

[74] ATLAS Collaboration, “Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016,” *Eur. Phys. J. C*, vol. 79, p. 205, 2019, arXiv:1810.05087.

[75] ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data,” *JINST*, vol. 14, p. P03017, 2019, arXiv:1812.03848.

[76] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k_t jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008, arXiv:0802.1189.

[77] ATLAS Collaboration, “Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV,”, ATL-PHYS-PUB-2015-036, 2015.

[78] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual,” *Eur. Phys. J.*, vol. C72, p. 1896, 2012, arXiv:1111.6097.

[79] ATLAS Collaboration, “Tagging and suppression of pileup jets with the ATLAS detector,”, ATLAS-CONF-2014-018, 2014.

[80] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas,” *Phys. Lett.*, vol. B659, pp. 119–126, 2008, arXiv:0707.1378.

[81] ATLAS Collaboration, "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector," *Phys. Rev. D*, vol. 96, p. 072002, 2017, arXiv:1703.09665.

[82] ATLAS Collaboration, "Performance of b-jet identification in the ATLAS experiment," *JINST*, vol. 11, p. P04008, 2016, arXiv:1512.01094.

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

[84] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$," *Eur. Phys. J. C*, vol. 78, p. 903, 2018, arXiv:1802.08168.

[85] ATLAS Collaboration, "$E_{\text{T}}^{\text{miss}}$ performance in the ATLAS detector using 2015–2016 LHC pp collisions,", ATLAS-CONF-2018-023, 2018.

[86] A. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione," *Giornale dell' Istituto Italiano degli Attuari*, vol. 4, pp. 83–91, 1933.

[87] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Ann. Math. Statist.*, vol. 19, pp. 279–281, 06 1948.

[88] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014, arXiv:1405.0301.

[89] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008, arXiv:0710.3820.

[90] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth.*, vol. A462, pp. 152–155, 2001.

[91] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, and I. Niessen, "Supersymmetric top and bottom squark production at hadron colliders," *JHEP.*, vol. 1008, p. 098, 2010, arXiv:arXiv:1006.4771 [hep-ph].

[92] W. Beenakker *et al.*, "Squark and gluino hadroproduction," *Int.J.Mod.Phys.*, vol. A26, pp. 2637–2664, 2011, arXiv:1105.1110.

[93] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," *JHEP*, vol. 02, p. 007, 2009, arXiv:0811.4622.

[94] S. Höche, F. Krauss, S. Schumann, and F. Siegert, "QCD matrix elements and truncated showers," *JHEP*, vol. 05, p. 053, 2009, arXiv:0903.1219.

[95] T. Gleisberg and S. Höche, "Comix, a new matrix element generator," *JHEP*, vol. 12, p. 039, 2008, arXiv:0808.3674.

[96] S. Schumann and F. Krauss, "A Parton shower algorithm based on Catani-Seymour dipole factorisation," *JHEP*, vol. 03, p. 038, 2008, arXiv:0709.1027.

[97] R. D. Ball *et al.*, "Parton distributions with LHC data," *Nucl. Phys.*, vol. B867, pp. 244–289, 2013, arXiv:1207.1303.

[98] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 06, p. 043, 2010, arXiv:1002.2581.

[99] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 05, p. 026, 2006, arXiv:hep-ph/0603175.

[100] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP*, vol. 06, p. 029, 2002, arXiv:hep-ph/0204244.

[101] ATLAS Collaboration, "ATLAS Pythia 8 tunes to 7 TeV data,", ATL-PHYS-PUB-2014-021, 2014.

[102] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, "New parton distributions for collider physics," *Phys. Rev.*, vol. D82, p. 074024, 2010, arXiv:1007.2241.

[103] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes," *Phys. Rev.*, vol. D82, p. 074018, 2010, arXiv:1005.3457.

[104] ATLAS Collaboration, "Validation of Monte Carlo event generators in the ATLAS Collaboration for LHC Run 2,", ATL-PHYS-PUB-2016-001, 2016.

[105] ATLAS Collaboration, "Multi-boson simulation for 13 TeV ATLAS analyses,", ATL-PHYS-PUB-2016-002, 2016.

[106] ATLAS Collaboration, "Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2,", ATL-PHYS-PUB-2016-003, 2016.

[107] ATLAS Collaboration, "Simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV,", ATL-PHYS-PUB-2016-004, 2016.

[108] ATLAS Collaboration, "Modelling of the $t\bar{t}H$ and $t\bar{t}V (V = W, Z)$ processes for $\sqrt{s} =$ 13 TeV ATLAS analyses,", ATL-PHYS-PUB-2016-005, 2016.

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

[110] ATLAS Collaboration, "The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim,", ATL-PHYS-PUB-2010-013, 2010.

[111] ATLAS Collaboration, "Summary of ATLAS Pythia 8 tunes,", ATL-PHYS-PUB-2012-003, 2012.

[112] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC," *The European Physical Journal C*, vol. 63, pp. 189–285, Sep 2009.

[113] ATLAS Collaboration, "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector," *JHEP*, vol. 09, p. 076, 2013, arXiv:1306.4945.

[114] C. G. Lester and D. J. Summers, "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders," *Phys. Lett.*, vol. B463, pp. 99–103, 1999, arXiv:hep-ph/9906349.

[115] C. Lüdtke, *The search for direct top-squark production in Hadronic final states with the ATLAS detector at* $\sqrt{s} = 13$ *TeV*. Dissertation, University of Freiburg, 2018.

[116] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J.*, vol. C71, p. 1554, 2011, arXiv:1007.1727. [Erratum: Eur. Phys. J.C73,2501(2013)].

[117] G. Cowan, "Discovery sensitivity for a counting experiment with background uncertainty," http://www.pp.rhul.ac.uk/ cowan/stat/medsig/medsigNote.pdf, 05 2012.

[118] P. Jackson, C. Rogan, and M. Santoni, "Sparticles in motion: Analyzing compressed SUSY scenarios with a new method of event reconstruction," *Phys. Rev.*, vol. D95, no. 3, p. 035031, 2017, arXiv:1607.08307.

[119] ATLAS Collaboration, "ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider," *Eur. Phys. J. C*, vol. 75, p. 510, 2015, arXiv:1506.08616. [Erratum: Eur. Phys. J.C76,no.3,153(2016)].

[120] ATLAS Collaboration, "Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s}$ =8 TeV pp collisions with the ATLAS detector," *JHEP*, vol. 11, p. 118, 2014, arXiv:1407.0583.

[121] ATLAS Collaboration, "Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using $4.7\,\mathrm{fb}^{-1}$ of $\sqrt{s} = 7\,\mathrm{TeV}$ proton–proton collision data," *Phys. Rev. D*, vol. 87, p. 012008, 2013, arXiv:1208.0949.

[122] ATLAS Collaboration, "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ with the ATLAS detector," *Phys. Rev. D*, vol. 96, p. 072002, 2017, arXiv:1703.09665.

[123] 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.

[124] ATLAS Collaboration, "Performance of b-jet identification in the ATLAS experiment," *JINST*, vol. 11, p. P04008, 2016, arXiv:1512.01094.

[125] ATLAS Collaboration, "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC," *Eur. Phys. J. C*, vol. 77, p. 332, 2017, arXiv:1611.10235.

[126] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$," *Eur. Phys. J. C*, vol. 78, p. 903, 2018, arXiv:1802.08168.

[127] ATLAS Collaboration, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data," *Eur. Phys. J. C*, vol. 74, p. 3071, 2014, arXiv:1407.5063.

[128] ATLAS Collaboration, "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$," *Eur. Phys. J. C*, vol. 76, p. 292, 2016, arXiv:1603.05598.

[129] ATLAS Collaboration, "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data,", ATLAS-CONF-2016-024, 2016.

[130] ATLAS Collaboration, "Luminosity determination in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ using the ATLAS detector at the LHC," *Eur. Phys. J. C*, vol. 76, p. 653, 2016, arXiv:1608.03953.

[131] M. Bähr *et al.*, "Herwig++ physics and manual," *Eur. Phys. J. C*, vol. 58, p. 639, 2008, arXiv:0803.0883.

[132] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, "Single-top hadroproduction in association with a W boson," *JHEP*, vol. 07, p. 029, 2008, arXiv:0805.3067.

[133] A. L. Read, "Presentation of search results: the CL_S technique," *J. Phys. G*, vol. 28, p. 2693, 2002.

[134] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933.

[135] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *Ann. Math. Statist.*, vol. 9, pp. 60–62, 03 1938.

[136] G. Ranucci, "The Profile likelihood ratio and the look elsewhere effect in high energy physics," *Nucl. Instrum. Meth.*, vol. A661, pp. 77–85, 2012, arXiv:1201.4604.

[137] M. Baak, G. J. Besjes, D. Côte, A. Koutsman, J. Lorenz, and D. Short, "HistFitter software framework for statistical data analysis," *Eur. Phys. J.*, vol. C75, p. 153, 2015, arXiv:1410.1280.

[138] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural SUSY Endures," *JHEP*, vol. 09, p. 035, 2012, arXiv:1110.6926.

[139] ATLAS Collaboration, “Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 13\,\text{TeV}$ pp collisions with the ATLAS detector,” *Phys. Rev. D*, vol. 94, p. 052009, 2016, arXiv:1606.03903.

[140] ATLAS Collaboration, “Search for squarks and gluinos in events with isolated leptons, jets and missing transverse momentum at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector,” *JHEP*, vol. 04, p. 116, 2015, arXiv:1501.03555.

[141] ATLAS Collaboration, “Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector,” *Phys. Rev. D*, vol. 94, p. 032005, 2016, arXiv:1604.07773.

[142] A. Djouadi *et al.*, “The Minimal supersymmetric standard model: Group summary report,” in *GDR (Groupement De Recherche) - Supersymetrie Montpellier, France, April 15-17, 1998*, 1998, arXiv:hep-ph/9901246.

[143] C. F. Berger, J. S. Gainer, J. L. Hewett, and T. G. Rizzo, “Supersymmetry Without Prejudice,” *JHEP*, vol. 02, p. 023, 2009, arXiv:0812.0980.

[144] N. Arkani-Hamed, A. Delgado, and G. F. Giudice, “The Well-tempered neutralino,” *Nucl. Phys.*, vol. B741, pp. 108–130, 2006, arXiv:hep-ph/0601041.

[145] ATLAS Collaboration, “Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC,” *Eur. Phys. J.*, vol. C79, no. 5, p. 375, 2019, arXiv:1808.07858.

[146] D. Krohn, J. Thaler, and L.-T. Wang, “Jet Trimming,” *JHEP*, vol. 02, p. 084, 2010, arXiv:0912.1342.

[147] ATLAS Collaboration, “Jet mass reconstruction with the ATLAS Detector in early Run 2 data,”, ATLAS-CONF-2016-035, 2016.

[148] J. Thaler and K. Van Tilburg, “Identifying Boosted Objects with N-subjettiness,” *JHEP*, vol. 03, p. 015, 2011, arXiv:1011.2268.

[149] ATLAS Collaboration, "Measurement of kT splitting scales in W->lv events at sqrt(s)=7 TeV with the ATLAS detector," *Eur. Phys. J.*, vol. C73, no. 5, p. 2432, 2013, arXiv:1302.1415.

[150] J. Thaler and L.-T. Wang, "Strategies to Identify Boosted Tops," *JHEP*, vol. 07, p. 092, 2008, arXiv:0806.0023.

[151] ATLAS Collaboration, "Top and Boson Tagger Signal Performance - Moriond2017," https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2017-004, 2017.

[152] A. J. Larkoski, G. P. Salam, and J. Thaler, "Energy Correlation Functions for Jet Substructure," *JHEP*, vol. 06, p. 108, 2013, arXiv:1305.0007.

[153] ATLAS Collaboration, "Object-based missing transverse momentum significance in the ATLAS detector,", ATLAS-CONF-2018-038, Geneva, 2018.