Search for scalar and vector leptoquarks decaying into quarks and leptons of different generations

BERGISCHE UNIVERSITÄT WUPPERTAL

Fakultät für Mathematik
und Naturwissenschaften

Dissertation

Volker Andreas Austrup

Erstgutachter: PD Dr. Frank Ellinghaus
Zweitgutachter: Prof. Dr. Erez Etzion
Drittgutachter: Prof. Dr. Wolfgang Wagner

Abgabedatum: 01.07.2022
Prüfungsdatum: 01.12.2022

Abstract

A search for pairs of leptoquarks produced in proton-proton collisions at a centre-of-mass energy of 13 TeV is presented. The data set, recorded with the ATLAS detector at the Large Hadron Collider, corresponds to an integrated luminosity of $139\,\text{fb}^{-1}$. The leptoquarks are assumed to decay only into quarks of the third generation and leptons of the first or the second generation. In particular, final states containing exactly one highly-energetic electron or muon are targeted. Additionally, events of interest contain large amounts of missing transverse momentum. These characteristics allow for an efficient discrimination against background events. Contributions from background processes are further suppressed through the application of neural networks, trained separately for various signal hypotheses. In total, eight different signal models are considered. In the absence of significant deviations between observed data and predictions based on the Standard Model of Particle Physics, exclusion limits on the signal models are derived. Finally, results from a statistical combination with other searches for leptoquarks are presented.

Kurzfassung

Es wird eine Suche nach Leptoquarks präsentiert, die in Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von 13 TeV paarweise produziert werden. Der Datensatz, aufgenommen mit dem ATLAS-Detektor am Large Hadron Collider, entspricht einer integrierten Luminosität von $139\,\mathrm{fb}^{-1}$. Es wird angenommen, dass die Leptoquarks nur in Quarks der dritten Generation und in Leptonen der ersten oder zweiten Generation zerfallen. Von Interesse sind insbesondere Endzustände, die genau ein hochenergetisches Elektron oder Muon enthalten. Entsprechende Ereignisse enthalten außerdem große Mengen an fehlendem transversalem Impuls. Diese Charakteristiken erlauben es, Beiträge von Untergrundprozessen effizient zu unterdrücken. Desweiteren werden neuronale Netze eingesetzt, um übrige Untergrundereignisse bestmöglich zu entfernen. Die neuronalen Netze werden separat für diverse Signalmodelle trainiert. Insgesamt werden acht verschiedene Signalhypothesen betrachtet. Da in den gemessenen Daten keine signifikanten Abweichungen von der Vorhersage des Standardmodells der Teilchenphysik beobachtet werden, werden Ausschlusslimits für die Signalmodelle bestimmt. Im Anschluss werden Ergebnisse einer statistischen Kombination mit anderen Suchen nach Leptoquarks präsentiert.

Contents

Introduction

All known elementary particles and their interactions, except gravity, are described in the Standard Model (SM) of Particle Physics. It provides accurate mathematical descriptions, facilitating precise predictions that have been verified in numerous experiments. However, its many successes notwithstanding, various arguments indicate the necessity of so-called Beyond Standard Model (BSM) theories extending the SM with as-of-yet unknown particles. Such theories can include leptoquarks (LQs), hypothetical bosons simultaneously coupling to leptons and quarks.

In this thesis, a search for pair-produced LQs with the ATLAS experiment at the Large Hadron Collider (LHC) at CERN is presented. The corresponding dataset amounts to an integrated luminosity of $139\,\text{fb}^{-1}$, collected in proton-proton collisions with centre-of-mass energies of $\sqrt{s} = 13\,\text{TeV}$ during the full Run 2 in the years 2015 to 2018. In particular, LQs coupling to quarks of the third generation and to leptons of the first or second generation are targeted. Such LQs introduce a mixing between different generations of quarks and leptons and are of great interest due to their prominence in BSM scenarios providing potential explanations for possible anomalies observed in rare B hadron decays.

The LQs searched for in the analysis described in this dissertation are either of scalar or of vectorial nature. For the former, up-type LQs ($\text{LQ}^{\text{u}}_{\text{mix}}$) with an electric charge of $q = +2/3e$ and down-type LQs ($\text{LQ}^{\text{d}}_{\text{mix}}$) with an electric charge of $q = -1/3e$ are considered. In the case of vector LQs, which carry an electric charge of $q = +2/3$, the Yang-Mills (YM) coupling scenario ($\text{vLQ}^{\text{YM}}_{\text{mix}}$) and the minimal coupling scenario ($\text{vLQ}^{\text{min}}_{\text{mix}}$) are analysed. Both, up-type scalar and vector LQs, decay either into top quarks and neutrinos or into bottom quarks and charged leptons. Conversely, down-type scalar LQs can decay into top quarks and charged leptons or bottom quarks and neutrinos. The branching ratio of the LQ decay is not determined a priori. It is therefore treated as a free model parameter in the analysis and referred to as $\mathcal{B}(\text{LQ} \to q_3\ell)$ with the third-generation quarks q_3 and the light leptons $\ell = e, \mu$. In the following, also the abbreviation $\mathcal{B}$ will be used.

The focus of the analysis is on final states containing a single electron or muon and large values of missing transverse momentum ($E_{\text{T}}^{\text{miss}}$). It is therefore most sensitive for $\mathcal{B} = 0.5$, i.e. when one LQ decays into a charged lepton and the other decays into a neutrino. However, for up-type LQs, the analysis is also sensitive to events with both LQs decaying into a top quark and a neutrino, with one top quark decaying leptonically and the other decaying hadronically. The analysis strategy is based around the application of neural networks (NNs), trained on various signal hypotheses. This approach allows to efficiently incorporate the characteristics

of the different signal models.

The structure of this dissertation is as follows: First, theoretical concepts are discussed, starting with a summary of the SM in Chapter 1. The need for theories extending the SM is motivated, resulting in the introduction of LQs in Chapter 2. Afterwards, the experimental facilities are described in Chapter 3, putting a focus on the different subsystems comprising the ATLAS detector, before detailing the simulation of proton-proton collisions in Chapter 4 and the reconstruction of such collision events in Chapter 5. The dataset and samples of simulated events are presented in Chapter 6, followed by the event categorisation in Chapter 7, including a description of NNs. Sources of systematic uncertainties in the analysis are discussed in Chapter 8. These systemtic uncertainties are used as nuisance parameters in profile likelihood fits, which, together with hypothesis testing techniques utilised for the extraction of exclusion limits on the signal models, are described in Chapter 9. Results from the statistical evaluation are provided in Chapter 10. Finally, in Chapter 11, the prospects of a statistical combination of the analysis presented in this dissertation with complementary searches are discussed.

1 | Standard Model of Particle Physics

There are numerous known particles that are considered to be elementary. Mathematical prescriptions of all these elementary particles and their interactions, except for gravity, are provided by the SM [1–3], a quantum field theory whose development was initiated in the 1960s. It has proven exceptionally successful, facilitating precise descriptions later validated by countless experiments. In addition, it allowed for predictions of particles, which had not yet been observed experimentally, such as the W boson or the top quark. The SM was completed from an experimental perspective with the discovery of the Higgs boson by ATLAS and CMS in 2012 [4, 5], four decades after its prediction.

In this chapter an overview of the particle contents of the SM is given before going into more details about its mathematical formulation. The chapter also includes a representative selection of the experimental results confirming its predictions with discoveries and precision measurements, and ends with a motivation for searches for phenomena predicted by theories extending the SM. In the remainder of this thesis, natural units will be used, setting the reduced Planck constant, $\hbar$, and the speed of light in vacuum, c, to unity. The corresponding constants will thus be omitted from equations and units. Electric charges will implicitly be given in units of the elementary charge.

The constituents of the SM are categorised according to their spin, with spin-1/2 particles referred to as fermions and integer-spin particles as bosons. The former are further split into leptons and quarks. Each of the six leptons carries a lepton number of $L = 1$ and an electric charge of either 0 or -1. The six quarks, in addition to an electric charge, also carry a colour charge, which can have one of three different values, commonly denoted as red, blue, and green. In analogy to the lepton number, quarks are assigned a baryon number of $B = 1/3$.

As shown in Table 1.1, both, leptons and quarks, come in three generations, each consisting of a pair of particles. Every generation of leptons contains an electrically neutral lepton, called neutrino, and a charged lepton. The charged leptons are called electron, muon, and tau, and their corresponding neutrinos are the electron-neutrino, the muon-neutrino, and the tau-neutrino. Quark generations on the other hand consist of an up-type quark with an electric charge of +2/3 and a down-type quark with a charge of -1/3. The up-type quarks are referred to as up, charm, and top, and the down-type quarks are down, strange, and bottom. While neutrinos are assumed to be massless in the SM, the masses of all other fermions increase from generation to generation, ranging over several orders of magnitude from the electron mass of 511 keV to the top-quark mass of 172.8 GeV [6]. Massive fermions of the second or third generation are therefore unstable and decay into lighter particles.

Table 1.1: Overview of the fermions in the SM, divided into the three generations. For each fermion, its mass and electric charge are listed [6], with neutrinos assumed to be massless.

Generation	Quarks			Leptons		
	Name	Charge	Mass	Name	Charge	Mass
I	up (u)	+2/3	2.2 MeV	electron (e^-)	-1	511 keV
	down (d)	-1/3	4.7 MeV	e neutrino (ν_e)	0	0
II	charm (c)	+2/3	1.27 GeV	muon (μ^-)	-1	106 MeV
	strange (s)	-1/3	93 MeV	μ neutrino (ν_μ)	0	0
III	top (t)	+2/3	172.8 GeV	tau (τ^-)	-1	1.78 GeV
	bottom (b)	-1/3	4.18 GeV	τ neutrino (ν_τ)	0	0

Each fermion type is associated with an antiparticle counterpart with the same mass, but opposite quantum numbers. Antiparticles are typically distinguished from particles by adding a bar above the respective symbols. In the case of charged leptons, the sign of the electric charge is usually explicitly noted to distinguish particles from antiparticles.

Fermions interact by way of the electromagnetic, the weak, or the strong interaction, all of which are mediated by spin-1 gauge bosons listed in Table 1.2. The mediator of the electromagnetic interaction is the massless and electrically neutral photon. As photons couple to the electric charge, the electromagnetic force acts on all fermions except neutrinos. The weak interaction is mediated by the neutral Z boson and the charged $W^\pm$ bosons, which couple to the weak isospin quantum number. Since charged leptons and down-type quarks have a weak isospin of -1/2 and neutrinos and up-type quarks of +1/2, the weak interaction is the only fundamental force acting on all fermions, including neutrinos. It is short-ranged because the weak bosons are massive and therefore have short live times. While the W boson decays either into two quarks or a charged lepton and a neutrino, the Z boson decays into two quarks or two oppositely charged leptons or into two neutrinos. Lastly, the strong interaction couples to the colour charge of particles, therefore acting on quarks, but not on leptons. The mediators of the strong force are gluons, which are massless and carry colour charges themselves. The particle content of the SM is completed with the Higgs boson, which is the only known elementary spin-0 particle, giving rise to the masses of all other particles. It is electrically neutral and has a mass of 125.25 GeV [6].

Table 1.2: Overview of the gauge bosons in the SM. For each boson, its mass and electric charge are listed, along with the interaction it mediates [6].

Name	Charge	Mass	Interaction
photon (γ)	0	0	electromagnetic
$W^\pm$ boson	±1	80.4 GeV	weak
Z boson	0	91.2 GeV	weak
gluon (g)	0	0	strong

1.1 Mathematical Formulation

Mathematically, the SM is formulated in the framework of a quantum field theory that combines quantum mechanics with special relativity by representing all elementary particles as excited states of their respective fields. The dynamics of the particles are encoded in a Lagrangian density term, $\mathcal{L}$, which will be referred to as Lagrangian in the following. The SM is a non-abelian gauge theory based on the invariance of the Lagrangian under the tensor product of three gauge groups, namely

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

Here, $SU(2)_L \otimes U(1)_Y$ is the symmetry group of the electroweak interaction, combining the electromagnetic and the weak forces, and $SU(3)_C$ is the symmetry group of the strong interaction.

The dynamics of free fermions are described by the Dirac Lagrangian

$$\mathcal{L} = \overline{\Psi} \left(\mathrm{i}\gamma^\mu \partial_\mu - m \right) \Psi, \tag{1.2}$$

with the fermion field Ψ, the fermion mass m, and the Dirac matrices γ^μ. The Dirac Lagrangian is invariant under global gauge transformations, but to also satisfy local gauge invariance the minimal substitution of the partial derivative ∂_μ with the covariant derivative

$$D_\mu = \partial_\mu - ig A_\mu^a T^a \tag{1.3}$$

is required. As A_μ^a are vector fields introduced to restore local gauge invariance, they are referred to as gauge fields. Their coupling to the scalar fields is governed by the coupling strength parameter g. The number of vector fields necessary corresponds to the number of generators T^a of the underlying symmetry group. The generators follow the commutation relation

$$\left[T^a, T^b \right] = \mathrm{i} f^{abc} T^c \tag{1.4}$$

with the structure constant f^{abc}. The Lagrangian of free vector fields can be written as

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^a F^{a\,\mu\nu}, \tag{1.5}$$

where

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c \tag{1.6}$$

is the field strength tensor. The Lagrangian in Equation (1.5) does not contain any explicit mass terms, as those would break gauge invariance again.

1.1.1 Electroweak Interaction and the Higgs Mechanism

The electromagnetic interaction and the weak interaction are unified in the electroweak theory. It is characterised by the $SU(2)_L \otimes U(1)_Y$ symmetry and describes flavour-conserving neutral currents mediated by the photon and the Z boson as well as flavour-changing charged currents mediated by the $W^\pm$ bosons. While the electromagnetic and the neutral weak interaction act on left- and right-handed fermions, the charged weak interaction exhibits maximal parity violation and exclusively couples to left-handed fermions. As parity violating terms are introduced via the $SU(2)$ symmetry, the subscript L denotes couplings to left-handed fermions only. Consequently, fermion fields are split into a right-handed component, Ψ_R, and a left-handed component, Ψ_L, defined as

$$\Psi_{L/R} = \frac{1}{2} \left(1 \mp \gamma^5 \right) \Psi, \tag{1.7}$$

with the product of the Dirac matrices $\gamma^5 = \gamma^0 \gamma^1 \gamma^2 \gamma^3$. Because γ^5 transforms a vector (V) into an axial vector (A), the coupling structure is termed $V - A$. The left- and the right-handed component then transform differently under $SU(2)$ transformations, the former being a weak isospin doublet, the latter a singlet. Accordingly, the arrangement of fermions into $SU(2)$ doublets and singlets follows as

$$L_L^i = \begin{pmatrix} \nu_L^i \\ \ell_L^i \end{pmatrix}, Q_L^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}, \ell_R^i, u_R^i, d_R^i, \tag{1.8}$$

where L_L^i and Q_L^i are the left-handed lepton and quark doublets of generation i, while ℓ_R^i, u_R^i, and d_R^i denote the right-handed singlets of charged leptons and quarks, respectively. No right-handed neutrinos exist in the SM.

The three generators of $SU(2)$ are

$$T_i = \frac{\sigma_i}{2}, \tag{1.9}$$

with the Pauli matrices σ_i. Conversely, the subscript Y indicates the generator of the $U(1)$ symmetry, the weak hypercharge Y, which is related to the electric charge, Q, and the third component of the weak isospin, T_3, via the Gell-Mann–Nishijima relation

$$\frac{Y}{2} = Q - T_3. \tag{1.10}$$

Following Equation (1.3), the covariant derivative is defined as

$$D_\mu = \partial_\mu - ig W_\mu^i \frac{\sigma^i}{2} - ig' B_\mu \frac{Y}{2} \tag{1.11}$$

to satisfy local gauge invariance, with the three gauge fields W_μ^i associated with $SU(2)_L$ and the gauge field B_μ associated with $U(1)_Y$. The corresponding coupling constants are denoted

as g and g' for $SU(2)_L$ and $U(1)_Y$, respectively. With these definitions, the electroweak Lagrangian can be written as

$$\mathcal{L}_{\text{EW}} = -g'\overline{\Psi}\gamma^\mu B_\mu \frac{Y}{2}\Psi - g\overline{\Psi}\gamma^\mu W_\mu^i \frac{\sigma^i}{2}\Psi - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu}, \tag{1.12}$$

where

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.13}$$

and

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\varepsilon^{ijk}W_\mu^j W_\nu^k \tag{1.14}$$

correspond to the kinematic terms of the gauge bosons, with the totally antisymmetric Levi-Civita tensor ε^{ijk}.

At this point, several problems arise, indicating the necessity for an additional mechanism in the SM. For one, gauge boson masses are not accounted for in the theory, although the weak bosons are known to be massive from experiments. Similarly, while local gauge invariance holds for the Dirac Lagrangian in Equation (1.2), the same is not true for separate left- and right-handed fermion fields. Furthermore, since the $SU(2)$ symmetry group is non-abelian, weak gauge bosons exhibit self-couplings as evidenced by the non-vanishing structure constant in Equation (1.14). These self-couplings can lead to divergences, causing unitarity violation in vector boson scattering for centre-of-mass energies on the TeV scale.

Serving as a remedy for these shortcomings, the Englert–Brout–Higgs mechanism [7–9] introduces an additional isospin doublet of complex scalar fields, consisting of an electrically neutral and a positively charged component,

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix}. \tag{1.15}$$

Denoting the conjugate transpose of a matrix M with complex entries as $M^\dagger$, the corresponding Lagrangian is

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\Phi)^\dagger (D^\mu\Phi) - \overbrace{\left(\mu^2\Phi^\dagger\Phi + \lambda\left(\Phi^\dagger\Phi\right)^2\right)}^{V(\Phi)}, \tag{1.16}$$

with the Higgs potential $V(\Phi)$, whose behaviour depends on the two parameters λ and μ^2. The case of $\lambda < 0$ is considered to be unphysical, as it results in a potential without any stable minima. If both λ and μ^2 are positive, a unique stable global minimum exists for $\Phi = 0$, as shown in Figure 1.1a. Conversely, if $\lambda > 0$, but $\mu^2 < 0$, the minimum is degenerate
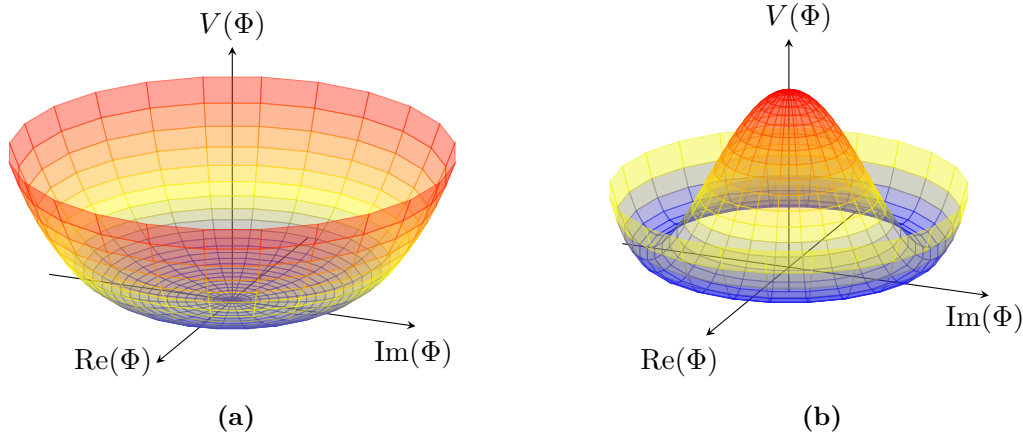
Figure 1.1: Schematic of the Higgs potential, $V(\Phi)$, in two dimensions with the parameters $\lambda > 0$ and (a) $\mu^2 > 0$ or (b) $\mu^2 < 0$.

and continuous in three dimensions, satisfying

$$\Phi^\dagger \Phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}, \tag{1.17}$$

where v is referred to as the vacuum expectation value (VEV). This functional behaviour is depicted schematically for two dimensions in Figure 1.1b.

The non-vanishing VEV leads to spontaneous breaking of the continuous symmetry of the underlying Lagrangian, which according to the Goldstone theorem implies the existence of three massless Goldstone bosons [10]. Choosing the so-called unitary gauge and expanding about the minimum, the Higgs field can be written as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.18}$$

where $H(x)$ is the field of a new massive, scalar boson. In the process, the Goldstone bosons are absorbed, providing the longitudinal polarisations for the weak gauge bosons. The choice of the unitary gauge yields a Higgs potential of the form

$$V(H) = \frac{1}{2} m_H^2 H^2 + \frac{m_H^2}{2v} H^3 + \frac{m_H^2}{8v^2} H^4 - \frac{1}{4} m_H^2 v^2, \tag{1.19}$$

with the Higgs mass $m_H^2 = 2\lambda v^2$. Higher orders of H correspond to the Higgs self-interactions. Identifying the physical W bosons as $W^\pm = \frac{1}{\sqrt{2}} \left(W^1 \mp i W^2 \right)$, the kinematic term in the Higgs Lagrangian in Equation (1.16), evaluated at the VEV, expands to

$$\left. (D_\mu \Phi)^\dagger (D^\mu \Phi) \right|_{\text{VEV}} = \frac{g^2 v^2}{8} W_\mu^+ W^{-\mu} + \frac{v^2}{8} \begin{pmatrix} W_\mu^3 & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix}. \tag{1.20}$$

From the first term in Equation (1.20), the W boson mass can be identified as $m_W = \frac{gv}{2}$. The second term in Equation (1.20) contains off-diagonal matrix elements, coupling together the W^3 and the B field. The mass eigenstates of the corresponding physical bosons are obtained by diagonalising the matrix, such that

$$\frac{v^2}{8} \begin{pmatrix} W_\mu^3 & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix} = \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} m_A^2 & 0 \\ 0 & m_Z^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix}, \tag{1.21}$$

where A_μ and Z_μ represent the photon field and the Z boson field, respectively. Their masses follow as $m_A = 0$ and $m_Z = \frac{v}{2}\sqrt{g^2 + g'^2}$. Through the Higgs mechanism, the boson masses have thus been introduced in a gauge invariant manner. Defining the weak mixing angle as $\tan\theta_W = \frac{g'}{g}$, the photon field and the Z boson field can be expressed as linear combinations of W^3 and B, i.e.

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \sin\theta_W & \cos\theta_W \\ \cos\theta_W & -\sin\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \tag{1.22}$$

Additionally, couplings between gauge bosons and the Higgs boson cancel contributions from vector boson self-interactions, thereby avoiding unitarity violations. Lastly, a mechanism to generate fermion masses, m_f, through Yukawa couplings to the Higgs boson of the form

$$g_f = \sqrt{2}\frac{m_f}{v} \tag{1.23}$$

is provided. These couplings appear in the Yukawa Lagrangian

$$\mathcal{L}_{\text{Yukawa}} = g_f \left(\overline{\Psi}_{\text{f, L}} \Phi \Psi_{\text{f, R}} + \overline{\Psi}_{\text{f, R}} \overline{\Phi} \Psi_{\text{f, L}} \right), \tag{1.24}$$

describing the interactions of left- and right-handed fermions with the Higgs doublet. These terms are gauge invariant as they transform as singlets under $SU(2)_L$. The mass eigenstates obtained through the Yukawa couplings are not necessarily aligned with the electroweak eigenstates. In particular, left-handed up- and down-type quarks are transformed between the eigenstates by different matrices, leading to a mixing between generations, which is quantified in the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix [11, 12]. The CKM matrix is commonly represented as

$$\begin{aligned} V_{\text{CKM}} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}e^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{-i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \end{aligned} \tag{1.25}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. It contains four free parameters, namely a phase, δ, violating the charge conjugation parity (CP) symmetry and three Euler angles, θ_{12}, θ_{13}, and θ_{23}.

1.1.2 Strong Interaction

The strong interaction is not affected by the Higgs mechanism, and the gluons remain massless, since the Higgs boson does not carry colour charge. Instead, interactions between quarks and gluons are described with the theory of quantum chromodynamics (QCD). Its underlying symmetry group, $SU(3)$, has eight generators

$$T^a = \frac{\lambda^a}{2},$$

(1.26)

with the Gell-Mann matrices λ^a, $a = 1 - 8$. Thus, an equal number of gauge fields, G_μ^a, is introduced, corresponding to a set of eight gluons. Denoting the coupling strength as g_S, the covariant derivative is then defined as

$$D_\mu = \partial_\mu - \mathrm{i}g_S \frac{\lambda^a}{2}G_\mu^a$$

(1.27)

and the gluon field strength tensor as

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

(1.28)

where f^{abc} is the group structure constant. With these equations and denoting the fermionic quark fields as Ψ_q, the QCD Lagrangian becomes

$$\mathcal{L}_{QCD} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} + \sum_q \overline{\Psi}_q \mathrm{i}\gamma^\mu D_\mu \Psi_q,$$

(1.29)

summing over all quark flavours. In the Lagrangian, interactions between quarks and gluons are encoded in the covariant derivative. As $SU(3)_C$ is non-abelian and its group structure constant does not vanish, gluon self-couplings are introduced.
The coupling strength is commonly expressed as

$$\alpha_{\mathrm{S}}(Q^2) = \frac{g_S^2(Q^2)}{4\pi} \approx \left(\frac{11n_{\mathrm{C}} - 2n_{\mathrm{f}}}{12\pi}\ln\left(\frac{Q^2}{\Lambda^2}\right)\right)^{-1}$$

(1.30)

as a function of the energy scale, Q, with the infrared cut-off scale, Λ, where $n_{\mathrm{C}} = 3$ is the number of colour charges and n_{f} is the number of quark flavours with masses $m_{q_f} < Q$. Since the SM contains six quark flavours and therefore $n_f \leq 6$, it is evident from Equation (1.30) that the coupling strength becomes small at high energies, but large at low energies. These effects, arising from the gluon self-interactions, are known as asymptotic freedom [13] and

confinement [14], respectively. Due to asymptotic freedom, quarks and gluons can be considered as free particles in high-energy collisions. Conversely, particles carrying a colour charge cannot be observed in isolation and instead form bound colourless states. These bound states either consist of three (anti-)quarks with the three different colour charges, called baryons, or of one quark with a certain colour and one antiquark with its anticolour, referred to as mesons.

1.2 Experimental Validation

Ever since its inception, the SM has been the subject of rigorous tests in order to validate its predictions. These checks include searches for particles predicted years earlier. In particular, theories including the top quark [12] or the Higgs boson [7–9] preceded their eventual experimental observation at Tevatron [15, 16] and at the LHC [4, 5], respectively, by decades. An overview of elementary particles first observed in the year 1974 or later is given in Table 1.3. All of these particles were predicted within a period of ten years between 1964 and 1973.

In addition, the SM has proven immensely successful in providing coherent descriptions of the nature of the interactions between elementary particles. Theoretical expectations have been validated in numerous measurements in physics processes across multiple orders of magnitude. For instance, production cross-sections of various processes have been evaluated at the ATLAS experiment in proton-proton collisions at centre-of-mass energies between 5 TeV and 13 TeV. A concise summary of a representative set of these measurements is given in Figure 1.2, comparing them to predictions and showcasing the excellent agreement between theory and experiment.

Precision measurements are not limited to the evaluation of cross-sections, as illustrated by the selection of plots in Figure 1.3, covering different sectors of the SM. For example, the aforementioned dependence of the strong coupling strength on the energy scale is experimentally well established and has been verified extensively in a wide range of experiments across several orders of magnitude up to energies on the TeV scale. This is demonstrated by the

Table 1.3: Summary of elementary particles predicted in the SM and later discovered experimentally. The table lists the years in which the particles were first predicted and the years in which they were subsequently observed for the first time.

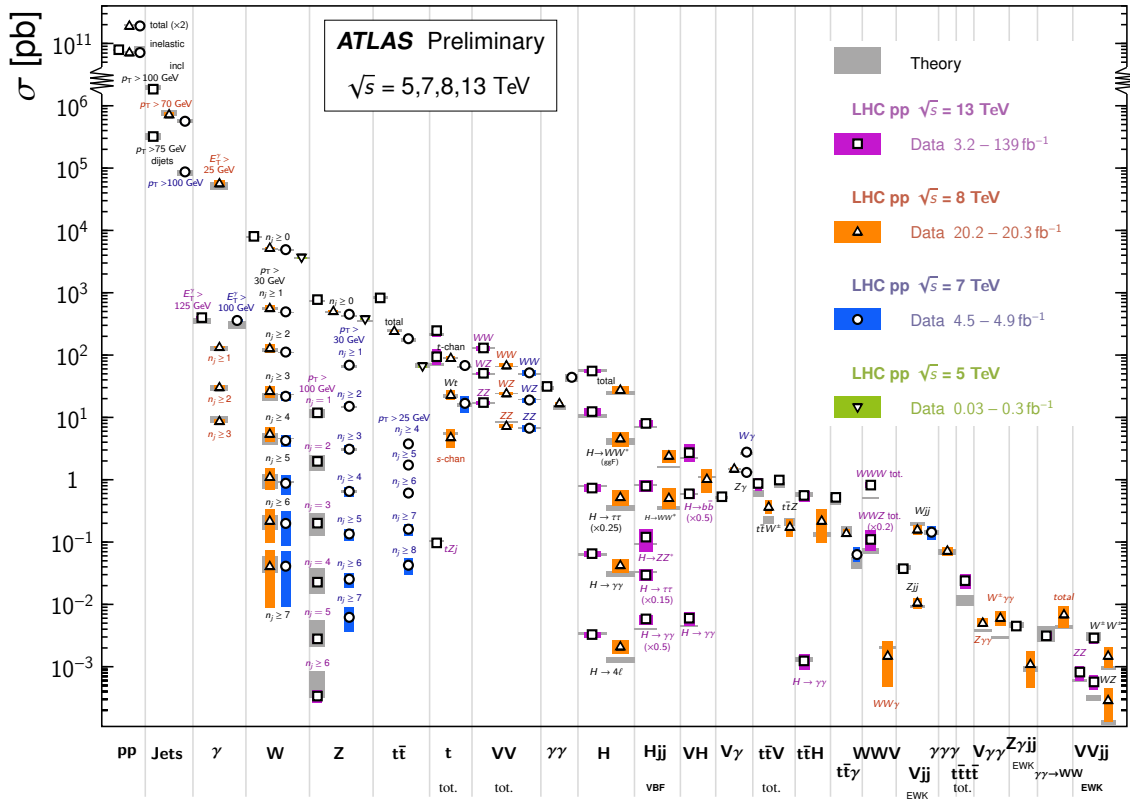| Particle	Theorised	Observed
charm	1970 [17]	1974 [18, 19]
τ	1971 [20]	1975 [21]
bottom	1973 [12]	1977 [22]
$W^{\pm}$	1968 [1–3]	1983 [23, 24]
Z^0	1968 [1–3]	1983 [25, 26]
top	1973 [12]	1995 [15, 16]
Higgs	1964 [7–9]	2012 [4, 5]

Figure 1.2: Representative summary of various total cross-sections measured at the ATLAS experiment in proton-proton collisions at $\sqrt{s} = 5, 7, 8$, and $13\,\text{TeV}$, compared with theory predictions [27]. Theory predictions and their uncertainties are shown as gray bands, experimental uncertainties as coloured areas.

summary plot in Figure 1.3a, which shows the experimental and theoretical behaviour of α_S for energy scales ranging from $Q = 1\,\text{GeV}$ to $2\,\text{TeV}$, confirming its predicted decrease towards high energy scales.

Of particular interest is the accurate determination of the Higgs boson mass, as, due to its unique and prominent role in the SM, the Higgs boson and its properties are highly susceptible to contributions from potential unknown physics processes. An overview of results obtained by the ATLAS collaboration in proton-proton collisions at centre-of-mass energies of $7-8\,\text{TeV}$ and $13\,\text{TeV}$ in several decay channels and a combination with CMS results at $7 - 8\,\text{TeV}$ is shown in Figure 1.3b. An excellent agreement between the measurements is observed, with uncertainties well below the percent level. Such precise knowledge of the Higgs boson mass is invaluable for verifying the internal consistency of the SM through global fits of the electroweak sector. These global fits are performed by the Gfitter group [29]. The plot shown in Figure 1.3c illustrates the deviations between individual measurements and results of the global simultaneous fit, with differences given in units of the experimental uncertainty.
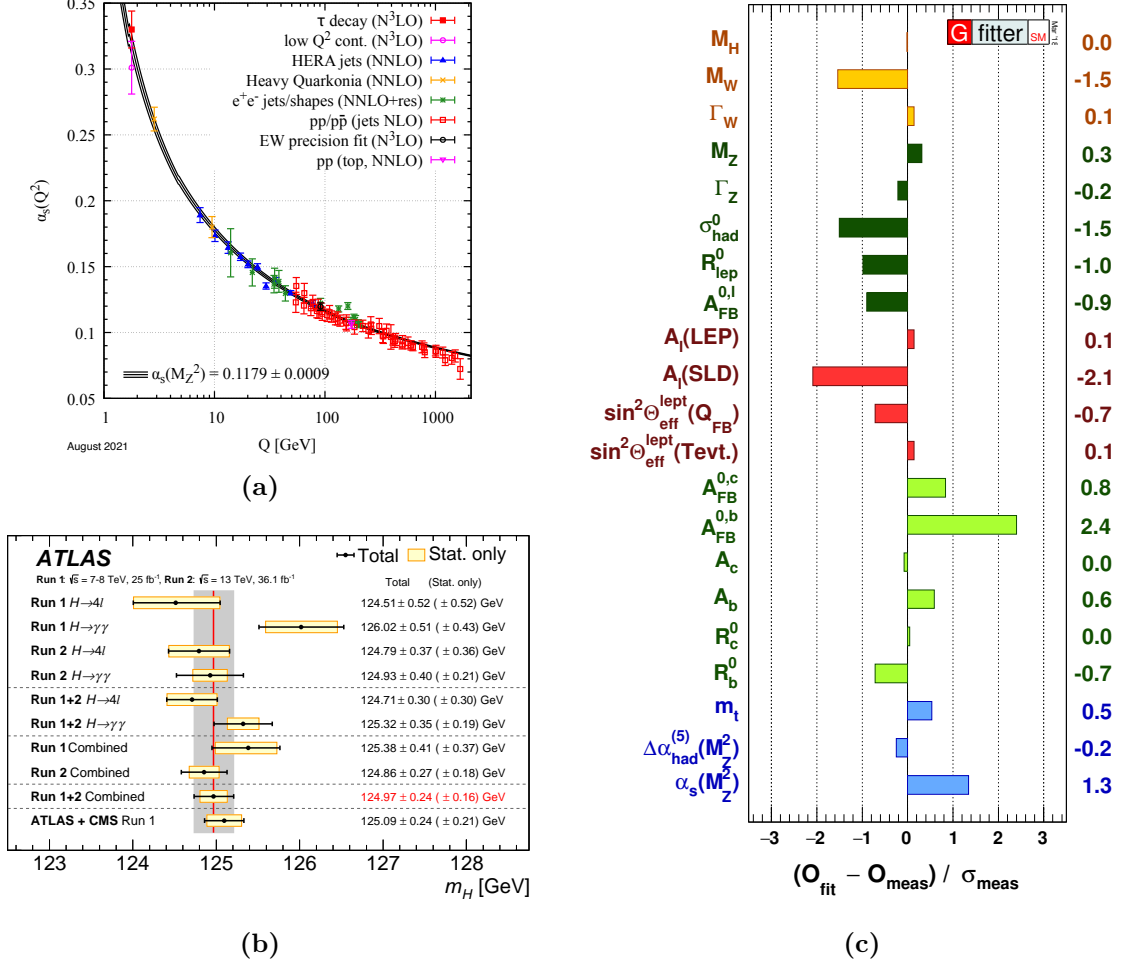
(a)

(b)

(c)

Figure 1.3: Representative selection of various experimental results probing and validating the predictions of the SM. (a) Summary of measurements of the strong coupling strength, α_S, as a function of the energy scale, Q [6]. (b) Overview of Higgs boson mass measurements carried out by the ATLAS experiment in proton-proton collisions at centre-of-mass energies of $7 - 8\,\mathrm{TeV}$ and $13\,\mathrm{TeV}$ [28]. The red vertical line indicates the combined result and the shaded gray area the corresponding total uncertainty. Results from individual channels and their total uncertainties are overlaid, with the statistical uncertainty components highlighted with yellow boxes. For comparison, a combination of ATLAS and CMS results at $7 - 8\,\mathrm{TeV}$ is included. (c) Results of a simultaneous fit of the electroweak SM parameters performed by the Gfitter group [29]. For each parameter, the difference between the measured value and the one obtained from the global fit is shown in units of the experimental uncertainty.

13

Because the fit is overconstrained, it allows for exceedingly precise indirect determinations of observables. As Figure 1.3c indicates, these indirectly measured values are highly consistent with the direct measurements, the largest deviation being 2.4σ. The majority of parameters agree within 1σ, demonstrating their collective compatibility with the SM. It should be noted that this does not yet take the recent W boson mass measurement by the CDF collaboration into account, which exhibits a tension of 7.0σ to SM predictions [30].

1.3 Limitations of the Standard Model

Despite the overwhelming successes of the SM, there are several limitations motivating the search for physics beyond the SM (BSM). An obvious deficiency is the omission of gravity, which is the only of the four fundamental interactions not described in the SM, but in the theory of general relativity instead. While at the electroweak scale of 246 GeV the gravitational force is many orders of magnitude weaker than the weak interaction, quantum gravity effects become sizeable at higher energy scales and lead to a break-down of the present description of the SM in a quantum field theory at the Planck scale of $\mathcal{O}\left(10^{19}\,\mathrm{GeV}\right)$.

In addition, on the theoretical side, aesthetical considerations hint towards potential extensions to the SM. This includes questions about why at least 19 free parameters are needed in the SM, i.e.

- nine fermion masses, six for quarks and three for the charged leptons,

- three mixing angles and a CP-violating phase for a total of four parameters to describe the CKM matrix,

- three gauge couplings, one for each symmetry group,

- two parameters describing the Higgs sector, namely the Higgs boson mass and the VEV,

- a CP-violating phase in the SM QCD Lagrangian.

These parameters can only be determined experimentally, but not calculated from first principles, prompting discussions about potential more fundamental theories containing fewer degrees of freedom. Moreover, despite a potential CP-violating phase in the QCD Lagrangian not breaking renormalisability, no CP-violation in the strong sector has been observed as of yet and the phase has been determined to be effectively zero [31, 32]. This fine-tuning of the phase parameter, which a priori could take non-vanishing values, is commonly referred to as the strong CP problem.

Another fine-tuning problem arises in the Higgs sector as a result of the so-called hierarchy problem, which itself is based on the question why the weak force is stronger than gravity by many orders of magnitude at the electroweak scale. As a consequence of divergent radiative loop corrections to the Higgs boson mass increasing quadratically with the energy cut-off scale, a cancellation between the bare Higgs boson mass and the loop corrections is necessary

to arrive at the experimentally established value. Depending on the size of the cut-off scale, these cancellations can become unnaturally precise, e.g. when the cut-off scale coincides with the Planck scale. The naturalness problem therefore indicates a need for contributions from new physics processes entering at energy scales in the TeV regime. Furthermore, similarities between the quark and the lepton sector, both consisting of three generations of doublets, hint at a more fundamental connection between the two fermion sectors. In particular, these symmetries are required to explain vanishing gauge anomalies, which is necessary for any self-consistent theory. However, ad hoc there is no known reason within the SM why the symmetries required for the contributions from quarks and leptons to the gauge anomalies to cancel are indeed realised in nature.

In addition to these phenomenological concerns, several experimental observations indicate the necessity of extensions of the SM. Besides the aforementioned discrepancies in the W boson mass, this includes measurements of neutrino flavour oscillations, implying misalignment between neutrino flavour eigenstates and mass eigenstates, which must be non-degenerate. The experimental upper limit on the sum of the three neutrino masses is $m_{\mathrm{tot}} = \sum_\nu m_\nu < 0.12\,\mathrm{eV}$ [33], while the lower bound is $m_{\mathrm{tot}} > 0.06\,\mathrm{eV}$ [34]. Moreover, cosmological measurements of gravitational lensing effects [35] and of rotation curves of spiral galaxies [36] show that gravitational forces generated by common matter cannot account for what is measured, therefore hinting at the possible presence of to-date unknown, non-luminous (or *dark*) matter. Furthermore, observations of supernovae reveal an accelerated expansion of the universe, implying the existence of an unknown form of energy counteracting the effects of gravity on cosmological scales [37, 38], termed *dark energy*. According to measurements of the cosmic microwave background, baryonic matter comprises only $5\,\%$ of the energy content of the universe, whereas dark matter and dark energy are significantly more prevalent, constituting $26\,\%$ and $68\,\%$ of the energy content, respectively [33]. Cosmological observations further indicate a predominance of matter over antimatter [39]. Explanations for this baryon asymmetry require sources of CP-violation surpassing those observed in the electroweak sector of the SM [40].

An exceedingly active field in accelerator-based particle physics is the study of the flavour sector. A prime candidate for the observation of BSM effects in the flavour sector are $b \to s\ell\ell$ transitions, as they are forbidden at tree level and heavily suppressed at loop level due to the Glashow–Iliopoulos–Maiani mechanism [17]. Decays facilitated by $b \to s\ell\ell$ transitions are therefore rare and sensitive to contributions from BSM particles. An example is the decay of the neutral B^0 meson into a $K^{\star 0}$ meson and two charged leptons, which is depicted in the Feynman diagram in Figure 1.4a. Under the SM hypothesis, lepton flavour universality is predicted. Except for small phase space corrections due to the non-vanishing lepton masses, equal branching ratios into electrons and muons are thus expected. However, measurements of $R_{K^{\star 0}}$, effectively corresponding to the ratio of branching ratios into electrons and muons,
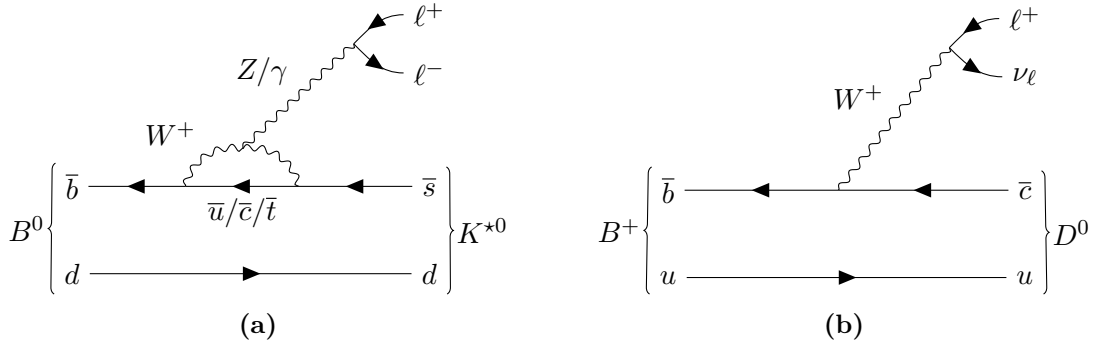
Figure 1.4: Feynman diagram of (a) the lowest SM loop-level contributions to the $B^0 \to K^{\star 0} \ell\, \ell$ decay and of (b) the SM $B^+ \to D^0 \ell^+ \nu_\ell$ decay at tree-level.

i.e.,

$$R_{K^{\star 0}} \hat{=} \frac{\mathcal{B}\left(B^0 \to K^{\star 0} \mu^+ \mu^-\right)}{\mathcal{B}\left(B^0 \to K^{\star 0} e^+ e^-\right)}, \tag{1.31}$$

reported by the LHCb collaboration indicate a tension of $2.5\,\sigma$ with respect to the SM prediction [41].

Figure 1.5a illustrates these findings in comparison to theory expectations in two bins of the momentum transfer, q^2. As evidenced by the small error bars on the predictions, the theory uncertainties on the ratios of electrons and muons are very well controlled, such that the significance of the measurements depends almost entirely on the experimental uncertainties.
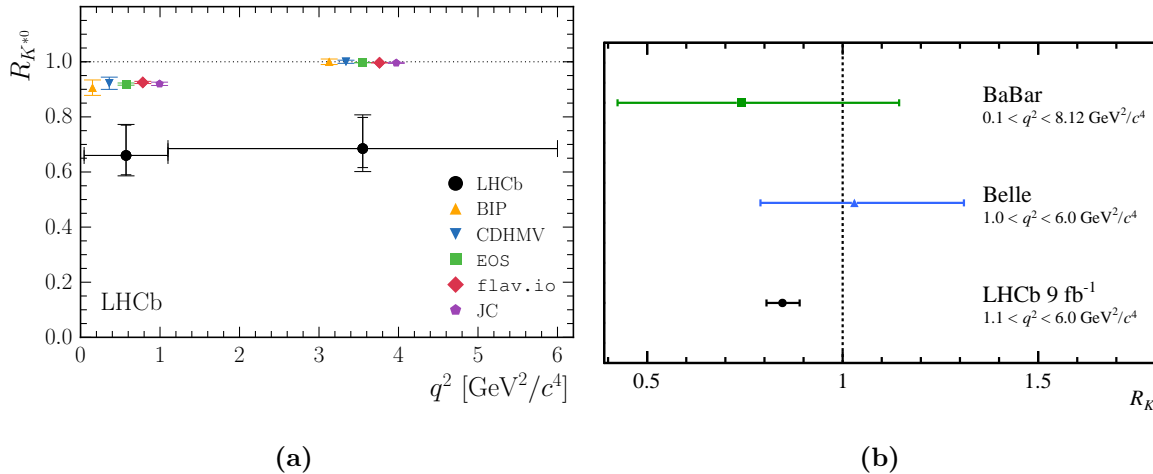

Figure 1.5: Results of measurements of $R_{K^{\star 0}}$ and R_K, with (a) a comparison of $R_{K^{\star 0}}$ as measured by the LHCb collaboration to theoretical SM predictions [41] and (b) a comparison of measurements of R_K at the LHCb, Belle, and BaBar experiments [42].

Analogous measurements of the decay of the charged B^+ meson into a K^+ meson by the LHCb collaboration yield similar deviations of 3.1σ from the SM prediction [42]. A comparison of measurements of the corresponding ratio of branching ratios

$$R_K \widehat{=} \frac{\mathcal{B}\left(B^+ \to K^+ \mu^+ \mu^-\right)}{\mathcal{B}\left(B^+ \to K^+ e^+ e^-\right)} \tag{1.32}$$

by the LHCb, BaBar, and Belle collaborations is shown in Figure 1.5b. Furthermore, measurements of lepton flavour universality in $B^0 \to K_S \ell\ell$ and in $B^+ \to K^{\star+} \ell\ell$ decays by the LHCb collaboration [43] also exhibit minor tensions with the SM, deviating by around 1.5σ each. These measurements are in agreement with those at the Belle experiment [44, 45]. In addition to neutral-current $b \to s\ell\ell$ transitions, decays of B mesons into $D^{(\star)}$ mesons, a charged lepton, and a neutrino via $b \to c\ell\nu$ transitions have been under study at the same three experiments. The corresponding process is illustrated as a Feynman diagram in Figure 1.4b. Here, the ratio of decays into τ leptons over those into light leptons is of particular interest. In accordance with Equation (1.31) and Equation (1.32), it is defined as

$$R_{D^{(\star)}} \widehat{=} \frac{\mathcal{B}\left(B \to D^{(\star)} \tau \nu_\tau\right)}{\mathcal{B}\left(B \to D^{(\star)} \ell \nu_\ell\right)} \tag{1.33}$$

with the light leptons $\ell = e, \mu$. The results from the different experiments and their combination by the Heavy Flavor Averaging Group is shown in Figure 1.6.
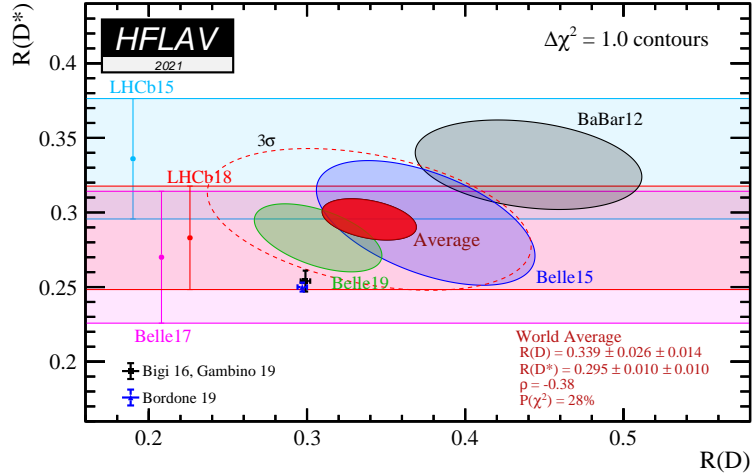
Figure 1.6: Comparison of measurements of theoretical predictions of R_D and $R_{D^\star}$ with measurements at the LHCb, Belle, and BaBar experiments as well as their combination [46].

Deviations from the theory prediction of more than 3σ are evident [46]. These deviations are particularly striking because the charged-current $b \to c\ell\nu$ transitions are allowed already at tree level in the SM. Thus, in order to be detectable, potential new physics effects need to

be sizeable enough to compete with tree-level SM contributions, hinting at the existence of BSM particles at scales directly accessible in proton-proton collisions at the LHC.

Measurements of lepton flavour universality via ratios of branching ratios are advantageous because the observables are theoretically well understood, as uncertainties on the individual branching ratios are highly correlated and hence largely cancel when calculating their ratio. While these ratios are therefore the easiest to interpret in view of potential BSM physics, they are not the only observables exhibiting tensions with SM predictions. For instance, the branching ratio of the rare $B_S^0 \to \phi\mu^+\mu^-$ decay has been measured at the LHCb experiment differentially in bins of the momentum transfer and, as displayed in Figure 1.7a, was consistently found to be 3.6σ below theoretical SM calculations [47]. Furthermore, angular distributions in $B^0 \to K^{\star 0}\ell\ell$ decays have received particular scrutiny from experimentalists and theorists alike, as it is possible to construct observables that in first order are independent of leading theoretical uncertainties [48]. The most signficant differences between SM prediction and measurement arise in the so-called P_5' variable, with a tension of 3.3σ reported by the LHCb collaboration [49] as shown in Figure 1.7b.
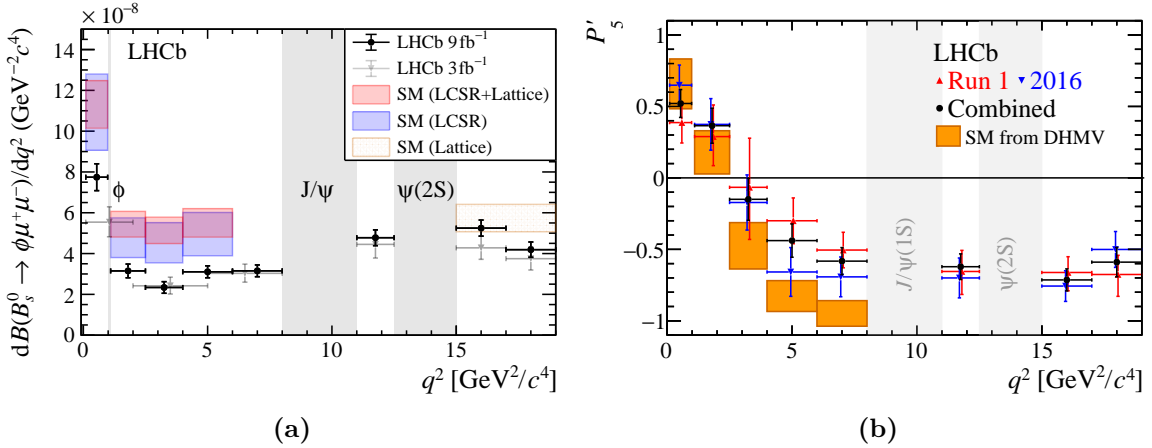
Figure 1.7: Comparison of SM predictions to LHCb measurements for (a) the differential branching ratio of $B_S^0 \to \phi\mu^+\mu^-$ in bins of the momentum transfer, q^2, [47] and (b) the distribution of P_5' depending on the momentum transfer, q^2 [49].

Another probe of SM predictions relies on the magnetic moment, $\vec{M}$, which for elementary spin-1/2 particles with mass m, charge q, and spin $\vec{S}$ is defined as

$$\vec{M} = g\frac{q}{2m}\vec{S}. \tag{1.34}$$

At tree level, the proportionality factor g is expected to be exactly two. However, small deviations from this value emerge due to loop corrections, motivating the definition of the anomalous magnetic moment, a, as

$$a = \frac{g-2}{2}. \tag{1.35}$$

For electrons, the observed anomalous magnetic moment agrees remarkably well with the predicted value [50]. However, in the case of muons, a combination of precise measurements of the anomalous magnetic moment performed at Brookhaven National Laboratory and at Fermilab yields a difference of the measured value, a_μ^{meas}, with respect to the theoretically calculated value, a_μ^{SM}, of

$$a_\mu^{\mathrm{meas}} - a_\mu^{\mathrm{SM}} = (251 \pm 59) \cdot 10^{-11}, \tag{1.36}$$

corresponding to a tension between theory and experiment of 4.2σ [51]. While the largest contributions to the loop corrections are induced by QED, the dominating source of uncertainty on the theoretical predictions stems from strong interactions. In light of this, new, lattice-QCD-based calculations were performed as cross-checks for the previous calculations, resulting in considerably smaller deviations from the measurements [52].

2 | Leptoquarks in Beyond Standard Model Theories

The hints at deviations from the SM prediction observed in experiments as discussed in the previous section point to the need for BSM theories. With indications of lepton flavour universality violations in semileptonic B decays arising both in $b \to s\ell\ell$ and in $b \to c\ell\nu$ transitions, an extension of the SM coherently explaining the observations in both channels would be conceptually very appealing. In many promising BSM models attempting to consistently explain the potential flavour anomalies, hypothetical scalar or vector bosons, referred to as LQs, emerge. Carrying both baryon and lepton number, LQs provide a possible connection between the quark and the lepton sector. Predating observations of hints at lepton flavour universality violation by several decades, their existence was first theorised in Grand Unified Theories (GUTs) like the Georgi-Glashow model [53], in which quarks and leptons emerge from the same multiplets in a spontaneously broken $SU(5)$ symmetry. Similarly, LQs also appear in Pati-Salam models [54], in which the lepton number is treated as a fourth colour charge. In both theories, LQs constitute the gauge bosons associated to the respective $SU(5)$ or $SU(4)$ symmetries. As such, they also appear in GUTs based on $SO(10)$, which include Georgi-Glashow and Pati-Salam models [55, 56]. Additionally, LQs have been discussed in the context of technicolour theories [57–59] and composite models [60, 61].

As the number of potential combinations of lepton fields and quark fields in the SM is small, all valid LQs connecting the quarks and the leptons can easily be listed and classified. Following the Buchmüller-Rückl-Wyler classification [62], there are ten possible LQ multiplets, five of them being scalars and five vectors. They are listed in Table 2.1 along with their couplings to left- and right-handed fermions and their fermion numbers

$$F = 3B + L. \tag{2.1}$$

LQs with $|F| = 0$ couple to fermion-antifermion pairs, whereas those with $|F| = 2$ have couplings exclusively to either fermions or antifermions. Given the couplings to fermions and antifermions, the electric charge that LQs can carry is limited to one of $|q| = 5/3, 4/3, 2/3$, or $1/3$.

As quarks and leptons transform as triplets and singlets under $SU(3)$, respectively, LQs constitute $SU(3)$ triplets, too. In the case of LQs with $|F| = 2$, this gives rise to couplings of the LQ to two quarks, whereas simultaneous couplings of LQs to two leptons are never allowed. However, these di-quark couplings are omitted in Table 2.1 as they induce rapid

proton decays at tree level. In light of the observed lower limits on the proton lifetime, which are on the order of 10^{34} years [63, 64], severe lower bounds on the LQ mass or extremely small couplings to fermions are implied. In either case, LQs coupling to two quarks are not of interest here, as they are not accessible at the LHC.

Because quarks and leptons transform as $SU(2)$ doublets or singlets, the various LQs necessarily transform either as singlets, as doublets, or as triplets under $SU(2)$. The multiplet structure under $SU(2)$ of a given LQ is indicated by the subscript of its symbol. There are two scalar singlets, namely S_1 and $\tilde{S}_1$, two scalar doublets, R_2 and $\tilde{R}_2$, and the scalar triplet S_3. Analogously, the two vector singlets are called U_1 and $\tilde{U}_1$, the two vector doublets are referred to as V_2 and $\tilde{V}_2$, and the vector triplet as U_3. With neutrinos assumed to be massless in the SM, one additional scalar and one additional vector singlet, coupling exclusively to a right-handed neutrino and a quark, are omitted from Table 2.1.

Table 2.1: List of possible scalar and vector LQs according to the Buchmüller-Rückl-Wyler classification. Couplings to left- and right-handed fermions are indicated by the letters L and R in the subscripts. No distinction is made between LQs and their conjugates.

Symbol	Spin	F	$\lvert q \rvert$	Coupling to
S_1	0	-2	1/3	$\ell^-_{\mathrm{L/R}} u_{\mathrm{L/R}},\ \nu_{\mathrm{L}} d_{\mathrm{L}}$
$\widetilde{S}_1$	0	-2	1/3	$\ell^-_{\mathrm{R}} u_{\mathrm{R}}$
S_3	0	-2	4/3	$\ell^-_{\mathrm{L}} d_{\mathrm{L}}$
			1/3	$\ell^-_{\mathrm{L}} u_{\mathrm{L}},\ \nu_{\mathrm{L}} d_{\mathrm{L}}$
			2/3	$\nu_{\mathrm{L}} u_{\mathrm{L}}$
R_2	0	0	5/3	$\ell^-_{\mathrm{L/R}} \overline{u}_{\mathrm{R/L}}$
			2/3	$\ell^-_{\mathrm{L/R}} \overline{d}_{\mathrm{R/L}},\ \nu_{\mathrm{L}} \overline{u}_{\mathrm{R}}$
$\widetilde{R}_2$	0	0	2/3	$\ell^-_{\mathrm{L}} \overline{d}_{\mathrm{R}}$
			1/3	$\nu_{\mathrm{L}} \overline{d}_{\mathrm{R}}$
U_1	1	0	2/3	$\ell^-_{\mathrm{L/R}} \overline{d}_{\mathrm{L/R}},\ \nu_{\mathrm{L}} \overline{u}_{\mathrm{L}}$
$\widetilde{U}_1$	1	0	5/3	$\ell^-_{\mathrm{R}} \overline{u}_{\mathrm{R}}$
U_3	1	0	5/3	$\ell^-_{\mathrm{R}} \overline{u}_{\mathrm{R}}$
			2/3	$\ell^-_{\mathrm{L}} \overline{d}_{\mathrm{L}},\ \nu_{\mathrm{L}} \overline{u}_{\mathrm{L}}$
			1/3	$\nu_{\mathrm{L}} \overline{d}_{\mathrm{L}}$
V_2	1	-2	4/3	$\ell^-_{\mathrm{L/R}} d_{\mathrm{R/L}}$
			1/3	$\ell^-_{\mathrm{R}} u_{\mathrm{L}},\ \nu_{\mathrm{L}} d_{\mathrm{R}}$
$\widetilde{V}_2$	1	-2	4/3	$\nu_{\mathrm{L}} u_{\mathrm{R}}$
			1/3	$\ell^-_{\mathrm{R}} u_{\mathrm{L}}$

Even though already introduced as early as 1974, LQs have recently gained renewed interest with the emergence of flavour anomalies in $b \to s\ell\ell$ and $b \to c\ell\nu$ transitions. While previously the focus was on LQs coupling to quarks and leptons of the same generation, thereby providing

a possible explanation for the similarities in the structures of the quark and the lepton sector, now in particular LQs with flavour off-diagonal couplings are of special interest. Facilitating tree-level decays of the B meson as demonstrated in Figure 2.1, such cross-generational LQs could be a viable explanation for the flavour anomalies [65–67]. In particular, the U_1 vector LQ has been identified as the simplest consistent scenario [68]. Likewise, a model containing S_1 and S_3 scalar LQs has been discussed in the same context [69]. This model, however, requires a higher degree of fine-tuning to comply with constraints from experiments. Similarly, anomalies in $b \to s\ell\ell$ transitions can be accomodated in models with R_2 scalar LQs and $\widetilde{U}_1$ vector LQs [70].

Assuming non-chiral LQs with couplings to left-handed and right-handed fermions, cross-generational LQs can also be an explanation for the deviations observed in the anomalous magnetic moment of the muon because of loop-level modifications of EM interactions of the muon [71, 72]. Moreover, similar contributions have been proposed as the potential origin of neutrino masses [73, 74]. In this context, the corrections appearing only at the one- or two-loop level is seen as a natural justification for the small size of the neutrino masses. In addition, there are proposals to provide an explanation for the tensions in the W boson mass observed with the CDF experiment by introducing models containing scalar LQs and vector-like quarks. Such models incorporate R_2 LQs [75] or a mixture of $\widetilde{R}_2$ and S_3 LQs [76].
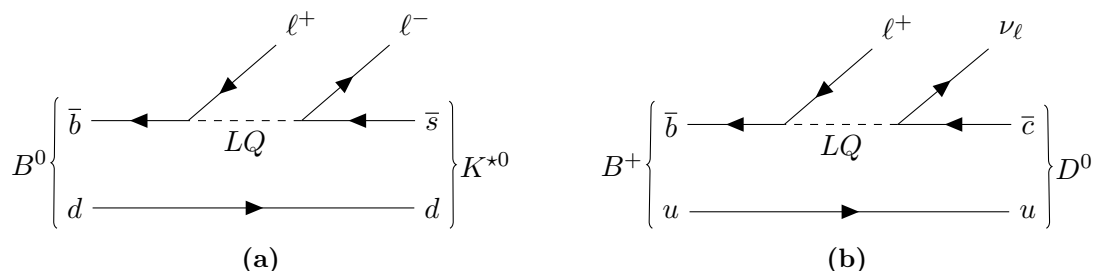

Figure 2.1: Feynman diagram of (a) the $B^0 \to K^{\star 0}\ell\ell$ decay and of (b) the $B^+ \to D^0\ell^+\nu_\ell$ decay. Assuming the existence of cross-generational LQs, both decays are possible at tree level.

2.1 Production and Decay

In proton-proton collisions at the LHC, pairs of LQs are produced primarily through gluon-gluon fusion or quark-antiquark annihilation, with production via a t-channel lepton exchange also being possible. The latter, however, is less prominent and particularly suppressed for LQs coupling to quarks solely of the third generation, because, due to their large mass, top quarks are not part of the proton's parton content, while for similar reasons bottom quarks constitute only a small proportion of the available sea quarks. Feynman diagrams illustrating the different production mechanisms at leading order are displayed in Figure 2.2.
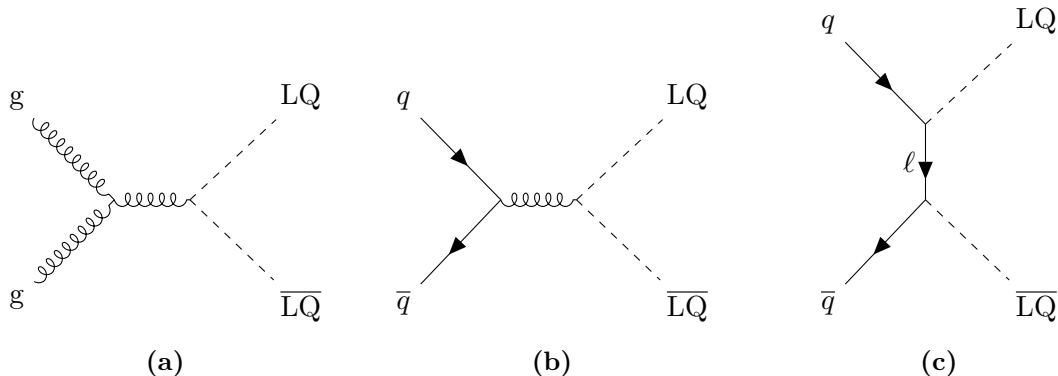
Figure 2.2: Feynman diagrams representing pair-poduction of LQs in proton-proton collisions through (a) gluon-gluon fusion, (b) quark-antiquark annihilation, and (c) a *t*-channel lepton exchange.

Due to the colour-triplet structure, the production cross-section of scalar LQs at the LHC in practice solely depends on their mass when neglecting contributions from *t*-channel lepton exchange. In the case of vector LQs, additional assumptions about the nature of the LQs affect the cross-sections. Typically, two scenarios are considered, with the vector LQ arising either as a massive gauge boson of a spontaneously broken gauge symmetry [77–80] or as the composite state of two elementary fermions [81, 82]. In the former scenario, the LQ couples to SM gauge bosons, thereby significantly enhancing production cross-sections. This is referred to as the YM coupling scenario. On the other hand, such couplings do not emerge if the LQ is a composite state. Consequently, the latter case is referred to as the *minimal* coupling scenario. However, even for the minimal coupling scenario, cross-sections are larger than those for scalar LQs by a factor of around three to seven, depending on the LQ mass, because of additional degrees of freedom due to the spin. In the YM coupling scenario, production cross-sections are further enhanced by a factor of five with respect to the minimal coupling case. The production cross-sections for pairs of scalar LQs at next-to-next-to-leading (NNLO) plus next-to-next-to-leading-log (NNLL) precision [83–86] as well as of vector LQs in the minimal and in the YM coupling scenario at leading-order (LO) precision are visualised in Figure 2.3. The latter are obtained using the MADGRAPH5_AMC@NLO(MG5_AMC@NLO) generator [87], as no higher-order calculations are available.

As various scalar LQs with similar properties are discussed in connection with the flavour anomalies, analyses often rely on a simplified model [88]. This model serves as a benchmark and does not implement the exact LQs according to the Buchmüller-Rückl-Wyler classification, but rather a generic Lagrangian of the form

$$\mathcal{L}_d = \sqrt{\beta}\lambda \left(\sqrt{\eta}\overline{u}_\mathrm{R}\ell_\mathrm{L} + \sqrt{1-\eta}\overline{u}_\mathrm{L}\ell_\mathrm{R}^+ \right) \mathrm{LQ}_\mathrm{mix}^\mathrm{d} + \sqrt{1-\beta}\lambda\overline{d}_\mathrm{R}\nu_\mathrm{L}\mathrm{LQ}_\mathrm{mix}^\mathrm{d} + \mathrm{h.c.} \qquad (2.2)$$
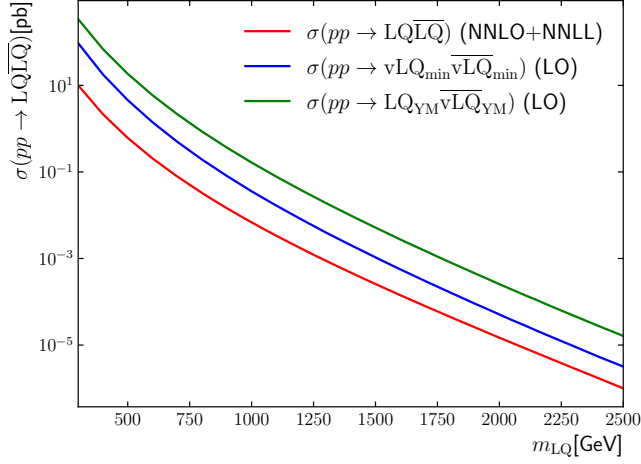
Figure 2.3: Production cross-sections for pairs of scalar LQs at NNLO+NNLL precision and for pairs of vector LQs in the minimal and in the YM coupling scenario at LO precision.

for down-type LQs with an electric charge of $q = -1/3$ and

$$\mathcal{L}_u = \sqrt{\beta}\lambda \left(\sqrt{\eta}\overline{d}_\mathrm{R}\ell_\mathrm{L} + \sqrt{1-\eta}\widetilde{d}_\mathrm{L}\ell_\mathrm{R}^- \right) \mathrm{LQ}_\mathrm{mix}^\mathrm{u} + \sqrt{1-\beta}\lambda\overline{u}_\mathrm{R}\nu_\mathrm{L}\mathrm{LQ}_\mathrm{mix}^\mathrm{u} + \mathrm{h.c.} \qquad (2.3)$$

for up-type LQs with an electric charge of $q = +2/3$, which could in principle be mixtures of the different possible scalar LQs in Table 2.1 with the corresponding electric charges. Here, h.c. refers to the Hermitian conjugate of the preceding terms. Couplings to left- and right-handed fermions are given by the parameter $\eta \in [0, 1]$ and $1-\eta$, respectively. However, in this analysis, only couplings of the scalar LQs to left-handed fermions are considered, i.e. $\eta = 1$, and the parameter is omitted in the following. Another model parameter, $\beta \in [0, 1]$, reflects the coupling of LQs to charged leptons and neutrinos, with the coupling to the former given by $\sqrt{\beta}\lambda$ and to the latter by $\sqrt{1-\beta}\lambda$. A choice of $\beta = 1.0$ therefore corresponds to LQs decaying exclusively into charged leptons and quarks, whereas LQs couple solely to neutrinos and quarks when β is zero. Assuming the bottom quark mass to vanish, the decay widths follow as

$$\Gamma(\mathrm{LQ}_\mathrm{mix}^\mathrm{d}) = \frac{1}{16\pi m_\mathrm{LQ}^3}(\overbrace{\lambda^2\beta(m_\mathrm{LQ}^2 - m_t^2)^2}^{\mathrm{LQ}_\mathrm{mix}^\mathrm{d} \to t\ell} + \overbrace{\lambda^2(1-\beta)m_\mathrm{LQ}^4}^{\mathrm{LQ}_\mathrm{mix}^\mathrm{d} \to b\nu}) \qquad (2.4)$$

and

$$\Gamma(\mathrm{LQ}_\mathrm{mix}^\mathrm{u}) = \frac{1}{16\pi m_\mathrm{LQ}^3}(\overbrace{\lambda^2\beta m_\mathrm{LQ}^4}^{\mathrm{LQ}_\mathrm{mix}^\mathrm{u} \to b\ell} + \overbrace{\lambda^2(1-\beta)(m_\mathrm{LQ}^2 - m_t^2)^2}^{\mathrm{LQ}_\mathrm{mix}^\mathrm{u} \to t\nu}) \qquad (2.5)$$

for down-type and up-type LQs, respectively.

As opposed to the simplified scalar LQ model, the one used for the vector LQs [89] corresponds directly to one of the LQs in the Buchmüller-Rückl-Wyler classification. Specifically, the U_1 LQ is targeted, as it offers a potential direct explanation for the flavour anomalies. Given its electric charge of $q = 2/3$, it is the vector analogue to the scalar $\text{LQ}^{\text{u}}_{\text{mix}}$, decaying either into a top quark and a neutrino or a bottom quark and a charged lepton. The Lagrangian is given as

$$
\begin{aligned}
\mathcal{L} = & -\mathrm{i}g_S(1 - \kappa_U)U_{1\mu}^\dagger \frac{\lambda^a}{2} U_{1\nu} G^{a\mu\nu} - \mathrm{i}g' \frac{2}{3}(1 - \tilde{\kappa}_U)U_{1\mu}^\dagger U_{1\nu} B^{\mu\nu} \\
& + \frac{g_U}{\sqrt{2}} \left[\left(\beta_{\mathrm{L}} \overline{Q}_{\mathrm{L}} \gamma_\mu L_{\mathrm{L}} + \beta_{\mathrm{R}} \overline{d}_{\mathrm{R}} \gamma_\mu \ell_{\mathrm{R}} \right) U_1 + \text{h.c.} \right]
\end{aligned}
\tag{2.6}
$$

where β_{L} and β_{R} represent the coupling to left-handed and right-handed fermions, respectively, and g_U is the overall coupling strength. Unlike the scalar LQs, this Lagrangian contains non-minimal interactions with SM gauge fields, which vanish only when the parameters are chosen as $\kappa_U = \tilde{\kappa}_U = 1$. This choice corresponds to the aforementioned minimal coupling scenario. On the other hand, if $\kappa_U = \tilde{\kappa}_U = 0$ and therefore the non-minimal couplings are allowed, the YM coupling scenario arises. In either case, the decay width follows as

$$
\begin{aligned}
\Gamma(\text{vLQ}_{\text{mix}}) = & \frac{\beta_{\mathrm{L}}^2 g_U^2}{48\pi} m_{\text{LQ}} \big(\overbrace{1}^{\text{vLQ}_{\text{mix}} \to b\ell} + \overbrace{V_{23}^2}^{\text{vLQ}_{\text{mix}} \to c\nu} + \overbrace{V_{13}^2}^{\text{vLQ}_{\text{mix}} \to u\nu} \big) \\
& + \underbrace{\frac{\beta_{\mathrm{L}}^2 g_U^2}{48\pi} \frac{m_{\text{LQ}}^2 - m_t^2}{m_{\text{LQ}}^3} V_{33}^2 \left(m_{\text{LQ}}^2 - \frac{m_t^2}{2} - \frac{m_t^4}{2m_{\text{LQ}}^2} \right)}_{\text{vLQ}_{\text{mix}} \to t\nu},
\end{aligned}
\tag{2.7}
$$

assuming couplings to quarks of the third generation and leptons of either the first or the second generation. As becomes evident from Equation (2.7), couplings of the vector LQ to up quarks and charm quarks are in principle also possible. These contributions, however, are severely suppressed with the off-diagonal CKM matrix elements and are neglected in the following.

Since the model aims at a coherent explanation for the B flavour anomalies, it is desirable from a theoretical standpoint to ensure it is ultraviolet (UV) complete and therefore renormalisable. Furthermore, it should comply with existing constraints from experiments. To achieve this, two additional massive spin-1 mediators are introduced in the model, namely a heavy gluon, G', and a Z', both of which are assumed to have masses on a similar scale as the vector LQ [89].

2.2 Properties of Signal Processes

As visualised in Figure 2.4, the mass peaks of scalar LQs are pronounced and have a width below 0.2%, corresponding to just a few GeV. Vector LQs on the other hand have a significantly larger width slightly above 10%. While in Figure 2.4 only vector LQs in the YM coupling scenario are shown, those in the minimal coupling scenario exhibit virtually the same mass distributions. The large differences in behaviour between scalar and vector LQs are caused by the choice of the respective coupling parameter. For scalar LQs, it is set to $\lambda = 0.3$ in agreement with previous LQ searches at the ATLAS experiment [90], since in the case of pair production it does not affect the production cross-section, but only the decay width. For vector LQs, a larger value of $g_U = 3.0$ is chosen in order to suppress the production cross-section of the other mediators arising in the UV-complete model, since these could lead to strong constraints from existing searches at the LHC, for example in the dileptonic final state. The observed widths are in agreement with those calculated in Equation (2.5) and Equation (2.7).

In general, the signal processes are expected to be characterised by large transverse momenta of the particles produced in the LQ decays. However, the kinematic behaviour of the decay products strongly depends on the LQ mass, as shown exemplarily in Figure 2.5 for the transverse momenta of top quarks and charged leptons emerging from the decays $\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to t\nu$ and $\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to b\ell$, respectively. Since the LQs are massive and therefore produced approximately at rest, the maxima in these distributions are reached for transverse momenta slightly below half the LQ's mass in all cases. Phase space constraints from the top-quark mass are noticeable as small corrections in Figure 2.5a, in which the distributions are slightly shifted to smaller transverse momenta when compared to the distributions in Figure 2.5b. This effect is particularly visible for $m_{\mathrm{LQ}} = 500\,\mathrm{GeV}$, where the top-quark mass is sizeable compared to the LQ mass, while for larger LQ masses the top-quark mass quickly becomes negligible. Equivalent behaviour is also observed for vector LQs.
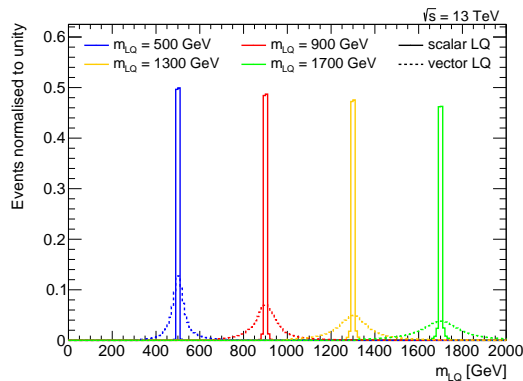

Figure 2.4: Mass distributions of simulated scalar and vector LQs with $m_{\mathrm{LQ}} = 500, 900, 1300, 1700\,\mathrm{GeV}$. In the scalar case up-type LQs are shown, and in the vector case those in the YM coupling scenario.
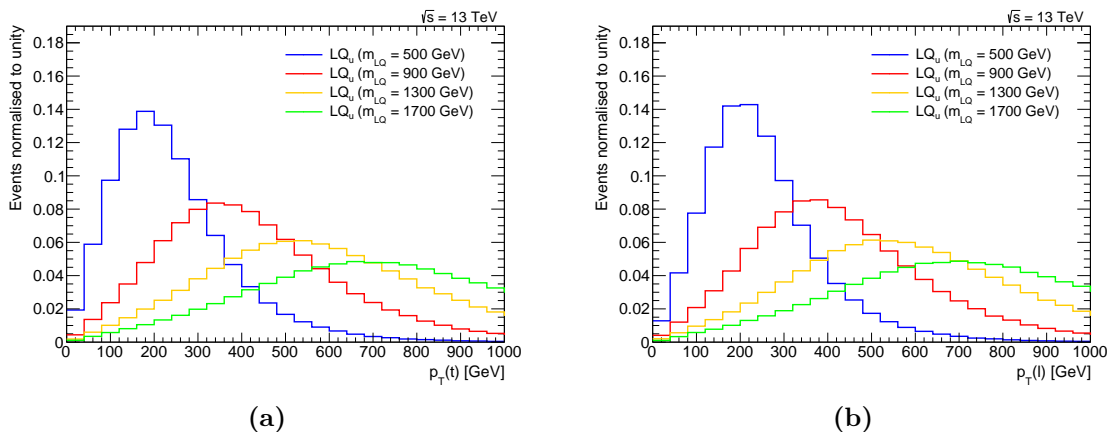
Figure 2.5: Distributions of the transverse momenta of (a) top quarks and of (b) charged leptons originating directly from simulated decays of scalar up-type LQs with masses of $m_{\mathrm{LQ}} = 500$, 900, 1300, and $1700\,\mathrm{GeV}$.

While scalar and vector LQs are very similar in their dependence of kinematic behaviours on the LQ mass, certain differences can be observed. In particular, decay products of top quarks are affected by differences in spin correlations propagating from the different spins of scalar and vector LQs. The impact of these effects on the transverse momenta of charged leptons and bottom quarks originating from the top-quark decays is shown in Figure 2.6 for a LQ mass of $m_{\mathrm{LQ}} = 1300\,\mathrm{GeV}$. In the case of scalar LQs, the distribution of the transverse momentum of the charged lepton reaches its maximum around $p_{\mathrm{T}}(\ell) = 100\,\mathrm{GeV}$ before decreasing again. On the other hand, the same distribution is monotonically falling for vector LQs. The opposite behaviour can be observed for the transverse momenta of the bottom quarks, for which the distribution reaches its maximum later in the case of vector LQs than in the case of scalar LQs and consequently has a larger tail towards higher transverse momenta.
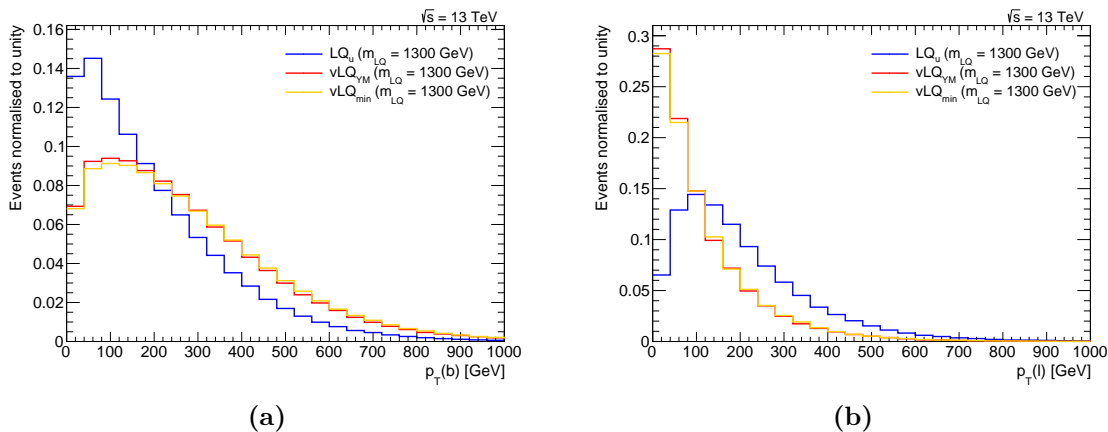

Figure 2.6: Distributions of the transverse momenta of (a) bottom quarks and of (b) charged leptons originating from top-quarks decays. The top quarks are decay products of up-type scalar or vector LQs with $m_{\mathrm{LQ}} = 1300\,\mathrm{GeV}$. The distributions are obtained from simulation.

The deviations can be understood by retracing the decay of the LQ into an antineutrino and a top quark and the subsequent decay of the top quark. A schematic illustration is given in Figure 2.7.
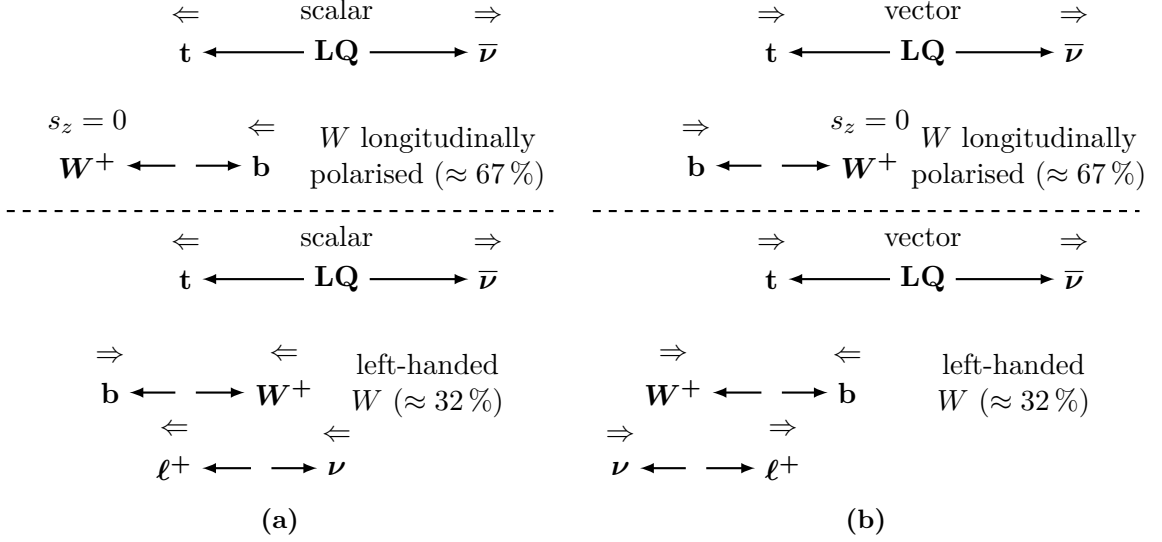

Figure 2.7: Schematic illustration of the decay of (a) a scalar LQ and (b) a vector LQ into an antineutrino and a leptonically decaying top quark. The direction of flight of the daughter particles in the rest frame of the mother particle is indicated by single arrows ($\rightarrow$). The direction of spin is given by the double arrows above each particle's symbol ($\Rightarrow$).

Since the antineutrino necessarily is right-handed, its spin is determined unambiguously. From this, the spin of the top quark can be deduced by imposing angular momentum conservation. It has to be in the opposite direction of the spin of the antineutrino in the case of scalar LQs and in the same direction as the spin of the antineutrino in the case of vector LQs. Following this, the top-quark decays into a bottom quark and a W boson, the latter of which, being a spin-1 boson, can be either transversely (i.e. left-handed or right-handed) or longitudinally polarised. At around 67%, the majority of W bosons is longitudinally polarised, with left-handed W bosons amounting to approximately 32% [91]. The decay of top quarks into right-handed W bosons is heavily suppressed. The differences between left-handed and longitudinally polarised W bosons and their implications are considered separately in the following:

1. The W boson is longitudinally polarised. Consequently, the bottom quark has the same spin as the top quark. As it is comparably light, it is mostly left-handed. Hence, the direction of its momentum is predominantly opposite to the top quark's momentum in the case of scalar LQs and in the same direction as the top quark's momentum in the case of a vector LQ decay. In the rest frame of the top quark, the W boson and the bottom quark are produced in a back-to-back topology, such that they have momenta in opposite directions. The decay products of the W boson do not have a

preferred direction due to the longitudinal polarisation. On average, they are therefore characterised only by the boost due to the W boson momentum.

2. The W boson is left-handed. This only allows for the direction of its momentum to be opposite to the top quark's momentum in the case of scalar LQs and vice versa in the case of vector LQs. Again, in the top quark's rest frame, bottom quark and W boson have opposite momenta. The bottom quark thus moves in the same direction as the top quark did originally in the case of scalar LQs and in the opposite direction to the top quark in the case of vector LQs. Unlike for longitudinally polarised W bosons, the decay products of left-handed W bosons have preferred directions, because the neutrino necessarily is left-handed. Due to angular momentum conservation, it therefore moves in the same direction as the W boson before, whereas the charged lepton is produced in the opposite direction in the rest frame of the W boson.

Based on these considerations, the behaviour observed in Figure 2.6a can be easily explained, because in the case of scalar LQs the majority of bottom quarks is produced in the opposite direction of the top quark and therefore has less transverse momentum than in the case of a vector LQ. The opposite argument holds true for the charged leptons in Figure 2.6b, which hence have higher transverse momenta in scalar LQ decays.

Further small differences between scalar and vector LQs arise in angular distributions as exemplarily shown for the azimuthal separation between the charged lepton and the bottom quark stemming directly from the LQ decay in Figure 2.8. These differences however are considerably less significant than the ones observed in the decay products of the top quarks, with the distributions generally increasing towards larger angular distances as expected from a decay of the LQ approximately at rest. While the increase is slightly more pronounced for scalar LQs, deviations between vector LQs in the YM scenario and in the minimal coupling scenario are found to be negligble in Figure 2.8.
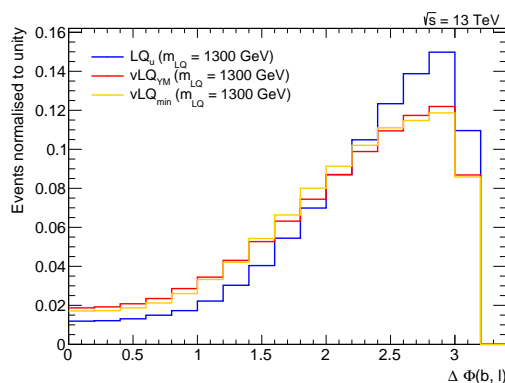

Figure 2.8: Distributions of the azimuthal distance between the charged lepton and the bottom quark originating from decays of up-type scalar and vector LQs with a mass of $m_{\mathrm{LQ}} = 1300 \, \mathrm{GeV}$. The distributions are obtained from simulation.

Nonetheless, small kinematic differences also exist between the YM and the minimal coupling scenario. This is demonstrated in Figure 2.9, where deviations are particularly visible for low LQ masses of $m_{\mathrm{LQ}} = 500 \, \mathrm{GeV}$. The transverse momenta of vector LQs in the minimal coupling scenario tend to higher values than those of vector LQs in the YM coupling scenario. The effect is less significant for higher masses, where the distributions are generally broader. To a certain extent, the discrepancies propagate also to the kinematic behaviour of the decay products, as shown for the example of the transverse momenta of neutrinos originating directly from LQ decays. While the distributions for a LQ mass of $m_{\mathrm{LQ}} = 1300 \, \mathrm{GeV}$ are practically identical, residual differences are still evident for a lower LQ mass.
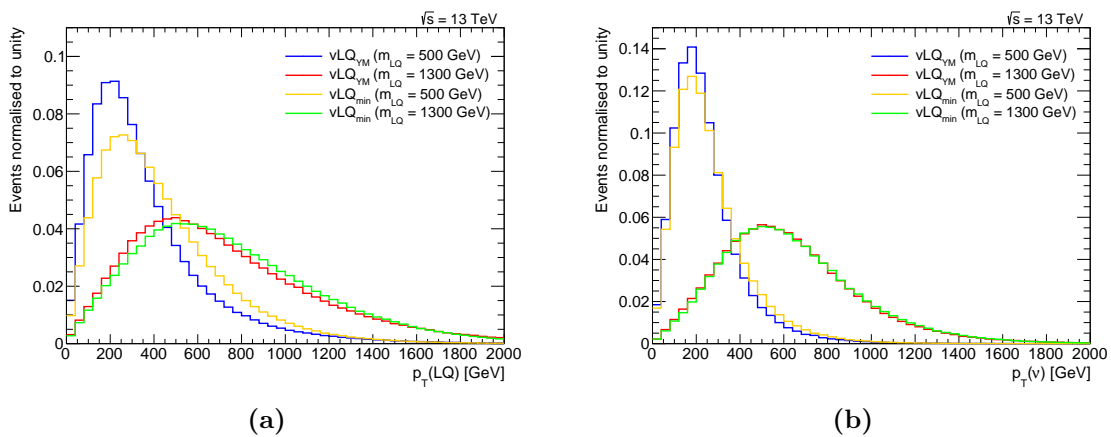

Figure 2.9: Distributions of the transverse momentum of (a) vector LQs and of (b) neutrinos originating directly from vector LQ decays with masses of $m_{\mathrm{LQ}} = 500$ or $1300 \, \mathrm{GeV}$. The distributions are obtained from simulation.

2.3 Direct Searches for Leptoquarks at the LHC

Leptoquarks are searched for at the LHC in broad programmes by both the ATLAS and the CMS experiment, covering single and pair production in a multitude of potential final states. Due to phase space considerations, single production is expected to exceed pair production at masses around the TeV scale, leading to a better sensitivity for the former processes. However, this strongly depends on the choice of the LQ-lepton-quark coupling λ, because the single production is proportional to λ^2.

For very large coupling strenghts, the sensitivity can be further improved through analyses targeting non-resonant production in dileptonic final states emerging due to the t-channel exchange of LQs, which is proportional to λ^4. A schematic illustration of the regions in the plane of the LQ mass and the coupling strength that can be excluded by searches for the different production channels is given in Figure 2.10. While searches for single production and for non-resonant production of LQs target somewhat similar parts of the parameter space, searches for pair-produced LQs, being independent of the choice of coupling strength, are
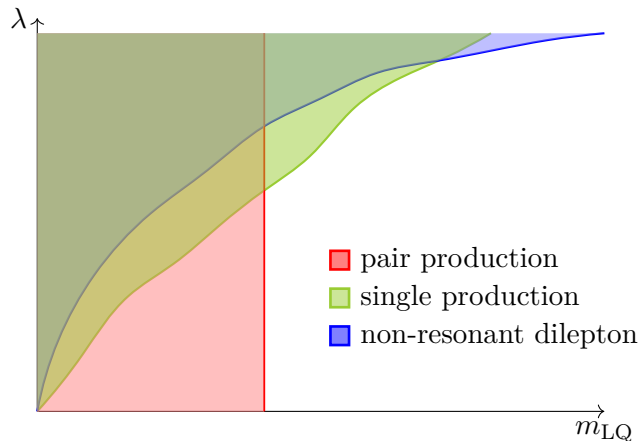
Figure 2.10: Schematic illustration of the exclusion reach of searches for LQ pair production, LQ single production, and non-resonant dilepton production as a function of the LQ mass, m_{LQ}, and the coupling strength, λ.

complementary.

Given the absence of experimental evidence for the existence of LQs as of yet, lower mass exclusion limits have been set on the various signal hypotheses. These range from 1.0 TeV in the case of scalar LQs decaying exclusively into bottom quarks and tau leptons [90, 92] to 1.8 TeV for scalar LQs decaying solely into light quarks and electrons [93]. In addition, limits on vector LQs have been derived for certain couplings to quarks and leptons, reaching mass exclusions as high as 2 TeV [94]. A representative selection of the most stringent limits from direct searches for scalar and vector LQs by the ATLAS and the CMS collaborations is given in Table 2.2.

Most of these results stem from searches for pair-produced LQs using the full or a partial dataset obtained during Run 2. However, in certain cases searches for single production are competitive or lead to improvements when combined with pair production, as done for example in a search by the CMS collaboration for LQs decaying into third-generation quarks and leptons [100]. Results presented in subsequent sections of this thesis are not included in Table 2.2, but signal hypotheses for which significant contributions to the sensitivity are foreseen are highlighted to indicate where improvements due to the analysis at hand can be expected.

Table 2.2: Lower mass exclusion limits in searches for scalar and vector LQs at the LHC. For the branching ratios of the LQs, $\mathcal{B}(LQ \to q\ell) = 1 - \mathcal{B}(LQ \to q'\nu)$ with q and q' of the same generation is assumed. The light quarks are denoted as $q = u, d, s$ in the table. Results from searches for singly-produced LQs are included when competitive with limits on pair production. Results obtained with the analysis presented in this thesis are not included. However, channels in which a significant contribution to the sensitivity is expected are highlighted with a bold font.

Signal Process		$\sqrt{s}$ [TeV]	$\mathcal{L}$ [fb^{-1}]	Limit [TeV] scalar	vector min	YM
$LQ\overline{LQ}, \mathcal{B}(LQ \to qe) = 1.0$	ATLAS [93]	13	139	1.79		
	CMS [95]	13	36	1.44		
$LQ\overline{LQ}, \mathcal{B}(LQ \to qe) = 0.5$	ATLAS [93]	13	139	1.38		
	CMS [95]	13	36	1.27		
$eLQ, \lambda = 1.0, \mathcal{B}(LQ \to qe) = 1.0$	CMS [96]	8	20	1.73		
$LQ\overline{LQ}, \mathcal{B}(LQ \to ce) = 1.0$	ATLAS [93]	13	139	1.75		
$LQ\overline{LQ}, \mathcal{B}(LQ \to ce) = 0.5$	ATLAS [93]	13	139	1.35		
$LQ\overline{LQ}, \mathcal{B}(LQ \to be) = 1.0$	ATLAS [93]	13	139	1.74		
LQ$\overline{\textbf{LQ}}, \mathcal{B}(\textbf{LQ} \to \textbf{be}) = \textbf{0.5}$	ATLAS [93]	13	139	1.40		
$LQ\overline{LQ}, \mathcal{B}(LQ \to te) = 1.0$	ATLAS [97]	13	139	1.48		
	CMS [94]	13	137	1.34		
LQ$\overline{\textbf{LQ}}, \mathcal{B}(\textbf{LQ} \to \textbf{te}) = \textbf{0.5}$	ATLAS [97]	13	139	1.29		
$LQ\overline{LQ}, \mathcal{B}(LQ \to q\mu) = 1.0$	ATLAS [93]	13	139	1.73		
$LQ\overline{LQ}, \mathcal{B}(LQ \to c\mu) = 1.0$	ATLAS [93]	13	139	1.68		
	CMS [95]	13	36	1.53		
$LQ\overline{LQ}, \mathcal{B}(LQ \to c\mu) = 0.5$	ATLAS [93]	13	139	1.41		
	CMS [95]	13	36	1.27		
$LQ\overline{LQ}, \mathcal{B}(LQ \to b\mu) = 1.0$	ATLAS [93]	13	139	1.77		
LQ$\overline{\textbf{LQ}}, \mathcal{B}(\textbf{LQ} \to \textbf{b}\mu) = \textbf{0.5}$	ATLAS [93]	13	139	1.48		
$LQ\overline{LQ}, \mathcal{B}(LQ \to t\mu) = 1.0$	ATLAS [97]	13	139	1.47		
	CMS [94]	13	36	1.42	1.70	2.05
LQ$\overline{\textbf{LQ}}, \mathcal{B}(\textbf{LQ} \to \textbf{t}\mu) = \textbf{0.5}$	ATLAS [97]	13	139	1.25		
	CMS [94]	13	36	1.15	1.46	1.76
$LQ\overline{LQ}, \mathcal{B}(LQ \to b\tau) = 1.0$	ATLAS [98]	13	139	1.19	1.47	1.76
	CMS [92]	13	36	1.02		
$LQ\overline{LQ}, \mathcal{B}(LQ \to b\tau) = 0.5$	ATLAS [98]	13	139	1.24	1.50	1.77
$LQ\overline{LQ}, \mathcal{B}(LQ \to t\tau) = 1.0$	ATLAS [99]	13	139	1.43		
	CMS [94]	13	137	1.12		
$LQ\overline{LQ}, \mathcal{B}(LQ \to t\tau) = 0.5$	ATLAS [99]	13	139	1.22		
	CMS [100]	13	137	0.95		
$LQ\overline{LQ} + \nu LQ, \lambda = 2.5, \mathcal{B}(LQ \to t\tau) = 0.5$	CMS [100]	13	137	1.02	1.41	1.73
$LQ\overline{LQ}, \mathcal{B}(LQ \to q/c\nu) = 1.0$	CMS [101]	13	137	1.14	1.55	1.98
$LQ\overline{LQ}, \mathcal{B}(LQ \to b\nu) = 1.0$	ATLAS [102]	13	139	1.26		
	CMS [101]	13	137	1.20	1.55	1.93
LQ$\overline{\textbf{LQ}}, \mathcal{B}(\textbf{LQ} \to \textbf{t}\nu) = \textbf{1.0}$	ATLAS [103]	13	139	1.24		
	CMS [101]	13	137	1.14	1.47	1.81

3 | The ATLAS Detector at the Large Hadron Collider

3.1 The Large Hadron Collider

With a circumference of $27\,\mathrm{km}$, the LHC [104], located at CERN near Geneva, Switzerland, is the largest and most powerful particle accelerator in the world. It houses the four major detectors ALICE [105], ATLAS [106], CMS [107], and LHCb [108] at its interaction points. In the interaction points, protons can be brought to collision at very high center-of-mass energies $\sqrt{s}$. The LHC is designed for $\sqrt{s}$ of up to $14\,\mathrm{TeV}$. During Run 2 from 2015 until 2018 energies of $13\,\mathrm{TeV}$ were reached. In addition, heavy-ion collisions are possible, mostly lead-lead, but also proton-lead and xenon-xenon. Such heavy-ion collisions are the main physics interest for the ALICE experiment, whereas the other three detectors primarily concentrate on proton-proton collisions. While LHCb is designed specifically for high-precision b-hadron experiments, ATLAS and CMS are general-purpose detectors built to cover nearly the whole solid angle. Their physics programmes include precision measurements of SM parameters as well as direct searches for BSM physics.

The LHC is embedded in a complex of accelerators as depicted in Figure 3.1. After production from ionised hydrogen gas, protons are first accelerated to an energy of $50\,\mathrm{MeV}$ in a linear accelerator (LINAC2). Next, they are injected into the circular Proton Synchrotron Booster (BOOSTER), in which the energy is further increased to $1.4\,\mathrm{GeV}$, before reaching energies of $25\,\mathrm{GeV}$ in the Proton Synchrotron (PS). In the last step before injection into the LHC, protons are accelerated to energies of $450\,\mathrm{GeV}$ in the Super Proton Synchrotron (SPS). Then, protons are inserted in opposite directions into the two beam pipes of the LHC, grouped into bunches, and are kept on circular paths by helium-cooled superconducting dipole magnets. The distance between two bunches is $25\,\mathrm{ns}$, equivalent to a collision rate of $40\,\mathrm{MHz}$.

The rate of a specific process, $pp \to X$, is directly proportional to its cross-section, $\sigma_{pp \to X}$, which generally is a function of the collision energy, such that

$$\frac{dN_{pp \to X}}{dt} = \sigma_{pp \to X}\mathcal{L},\tag{3.1}$$

where the proportionality factor is the instantaneous luminosity, $\mathcal{L}$. A large instantaneous luminosity is therefore beneficial, especially in searches for rare processes, which are primarily limited by the available data statistics rather than systematic uncertainties. The luminosity
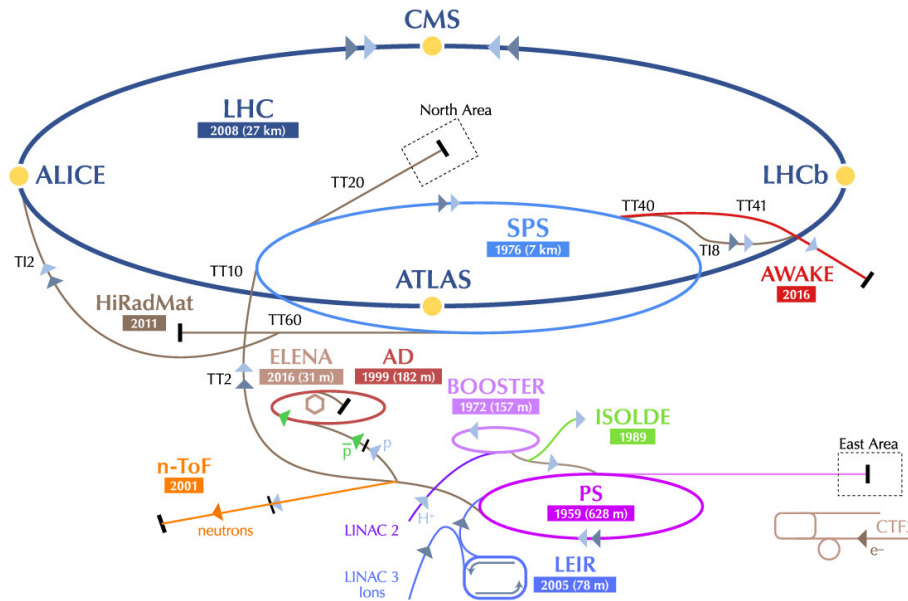
Figure 3.1: Schematic overview of CERN's accelerator complex. Protons are gradually accelarated in the chain before injection into the LHC [109].

is determined by various LHC machine parameters and can be expressed as

$$\mathcal{L} = \frac{n_{\mathrm{b}} n_p^2 f_{\mathrm{r}}}{2\pi \sigma_x \sigma_y}, \tag{3.2}$$

where f_{r} is the revolution frequency and $\sigma_{x,y}$ are the convoluted transverse beam sizes. The number of bunches, n_{b}, in a typical LHC fill is above 2000, and each bunch contains $n_p \approx 10^{11}$ protons. With these settings, peak luminosities of up to $2.1 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ are reached [110]. The exact values depend on the filling scheme and vary per year. They are summarised in Table 3.1 along with other key performance markers. The table also includes the beam focus parameter, $\beta^\star$, which provides a measure of the transverse beam size.

Due to the large number of protons per bunch, multiple collisions occur in a typical bunch crossing, an effect which is termed in-time pile-up. The average number of collisions, μ, follows from the cross-section for inelastic scattering, σ_{pp}, as

$$\mu = \frac{\mathcal{L} \sigma_{pp}}{n_{\mathrm{b}} f_{\mathrm{r}}}. \tag{3.3}$$

An additional contribution comes from out-of-time pile-up, caused by remnants from collisions in previous bunch crossings. The mean number of pile-up events increased during Run 2 from 13.4 in 2015 to 36.1 in 2018, leading to challenging conditions in the detectors. The pile-up profile in the different years, weighted by luminosity, is shown in Figure 3.2.

Table 3.1: Values of key LHC performance markers during Run 2. The design values are shown in a separate column for comparison [110].

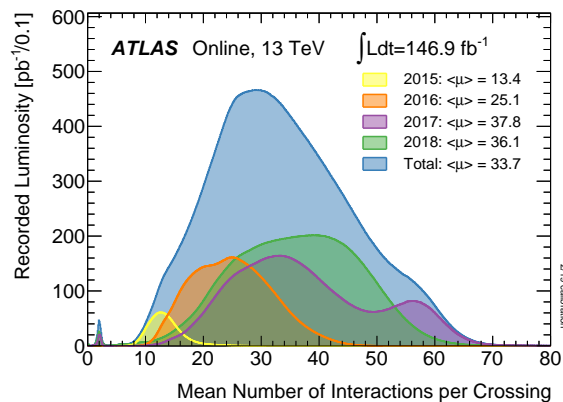| Parameter	2015	2016	2017	2018	Design
Beam energy [TeV]	6.5	6.5	6.5	6.5	7.0
Bunch spacing [ns]	25	25	25	25	25
Bunch population $[10^{11}p]$	1.2	1.25	1.25	1.1	1.15
Max. no. of bunches	2244	2200	2556	2556	2808
Beam focus $\beta^\star$ [cm]	80	40	40/30	30/27/25	55
Peak luminosity $[10^{34}\,/\mathrm{cm}^2\,/\mathrm{s}]$	0.5	1.5	2.0	2.1	1.0

Figure 3.2: Distribution of the number of interactions per bunch crossing during Run 2 [111].

3.2 The ATLAS Detector

The ATLAS detector is the largest detector at the LHC, measuring 44 m in length with a diameter of 25 m and weighing 7000 t. It contains several subsystems, installed in concentric layers around the interaction point, facilitating the detection, identification, and measurement of different types of particles. The layout of the detector, a schematic of which is shown in Figure 3.3, is designed to provide almost complete coverage across the full solid angle.

After introducing the coordinate system used in ATLAS, the various subsystems are presented in this section, highlighting the different purposes and functional principles, starting with the innermost layer, the Inner Detector (ID), and moving outwards via the calorimeter system to the muon spectrometer (MS).

3.2.1 Coordinate System

ATLAS employs a right-handed coordinate system centered about the nominal interaction point. The x-axis is defined towards the centre of the LHC ring and the y-axis upwards, while the z-axis points in the anti-clockwise beam direction. Instead of Cartesian coordinates, the azimuthal angle, ϕ, measured in the transverse plane is commonly used to better reflect the rotational symmetry of the detector. Accordingly, the polar angle, θ, is defined as the angle
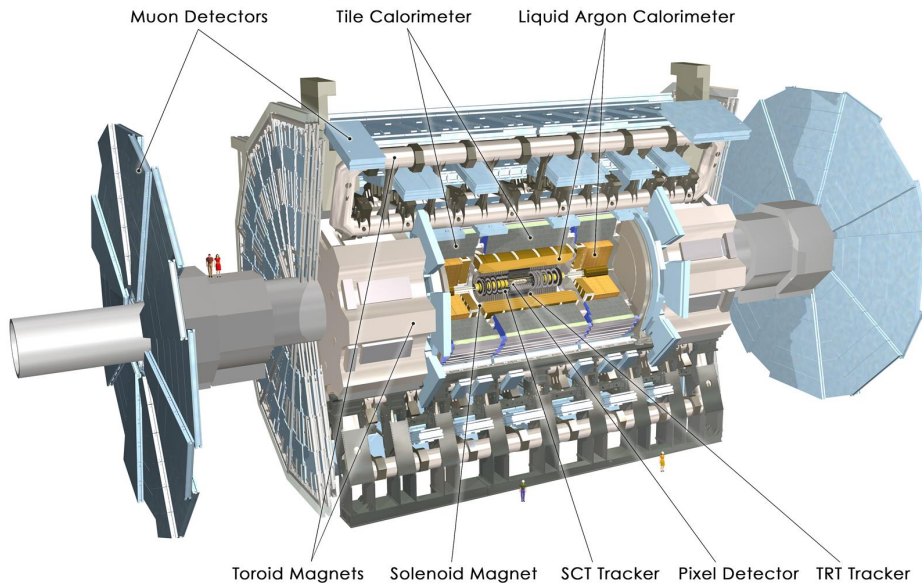
Figure 3.3: Cut-away view of the ATLAS detector. The various layers of subsystems as well as the magnet system are shown and labelled [112].

with respect to the z-axis. From this, the pseudorapidity is calculated as

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

which in the relativistic limit coincides with the rapidity, y, determined from a particle's energy, E, and its momentum along the beam axis, p_z, as

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

Differences in rapidity are invariant under Lorentz boosts along the z-axis. Angular distances are calculated as

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{3.6}$$

3.2.2 Inner Detector

Closest to the beam line, the ID enables precise tracking of charged particles, covering the region $|\eta| < 2.5$ [113, 114]. Its layout is shown in Figure 3.4.

The ID is submersed in a homogeneous $2\,\mathrm{T}$ magnetic field parallel to the beam provided by the superconducting solenoid surrounding it. Thus, trajectories are bent, thereby allowing for the determination of charge and momentum of the charged particle. The ID itself is composed of three subsystems, the innermost of which is the pixel detector. The pixel detector originally consisted of three cylindrical layers of high-granularity silicon semiconductor sensors with
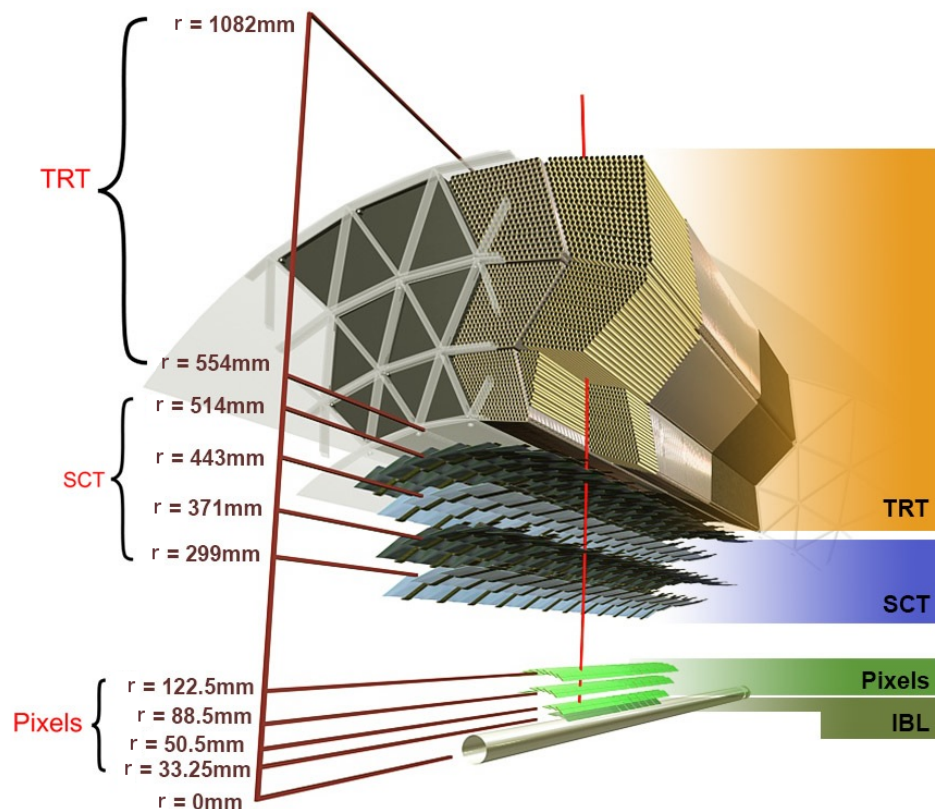
r = 1082mm

TRT

r = 554mm
r = 514mm
r = 443mm

SCT

r = 371mm
r = 299mm

TRT

SCT

r = 122.5mm
r = 88.5mm

Pixels

r = 50.5mm
r = 33.25mm
r = 0mm

Pixels

IBL

Figure 3.4: Layout of the Inner Detector [115].

pixels of size 50 μm x 400 μm and three additional concentric discs at each end of the barrel. The spatial resolution measured with these sensors is 10 μm in the R-ϕ plane and 115 μm in the z-direction. During Long Shutdown 1, the Insertable B-layer (IBL) was added [116], decreasing the distance to the interaction point to 33 mm. Smaller pixel areas of 50 μm x 250 μm allow for an improved resolution of 70 μm in the z-direction. This accuracy is essential for precisely reconstructing vertices, separating primary vertices from pile-up vertices, and identifying tracks originating from the decay of B hadrons.

Built around the pixel detector, the Semiconductor Tracker (SCT), comprising four layers in the barrel region and nine discs in each end-cap, consists of silicon micro-strip sensors with areas of 80 μm x 64 mm. In each layer, two sets of sensors are mounted back-to-back with a stereo angle of 40 mrad between them, thus allowing for the measurement of the hit location along the strip axis. The resolution offered by the SCT is 17 μm in R-ϕ and 580 μm in z.

The outermost layer of the ID is the Transition Radiation Tracker (TRT), extending the tracking capabilities to radii of 1 m. As opposed to the pixel detector and the SCT, it provides a coverage limited to $|\eta| < 2.0$. The TRT consists of 73 layers of straw tubes in the barrel region and 160 layers in the end-caps. Each straw tube has a diameter of 4 mm and is filled with a xenon- or argon-based gas mixture, which is ionised by traversing particles. By making use of the drift time of the free electrons thus created, a resolution of 130 μm can be achieved.

Additionally, the straw tubes are interleaved with transition radiation material. Since the amount of transition radiation depends on the particle mass, this allows for distinguishing electrons from heavier particles such as pions.

The combined information collected by all subsystems results in a transverse momentum resolution of

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 0.05\,\%\,p_{\mathrm{T}}[\mathrm{GeV}] \oplus 1\,\%, \tag{3.7}$$

with the symbol $\oplus$ denoting addition in quadrature.

3.2.3 Calorimeter System

Outside the ID, the calorimeter system is installed [117], providing coverage up to $|\eta| < 4.9$. Its purpose is to identify particles as well as to measure their energy and determine their direction of flight. The ATLAS calorimeter system comprises several subsystems as shown in Figure 3.5, all of which are based on the sampling principle. Calorimeters of this class are constructed from alternating layers of active and passive material. The latter is designed to ensure the energy is contained within the calorimeter structure by inducing particle showers, which cause measurable energy deposits in the active medium. The nature of energy loss differs between electrons and photons on the one hand and hadrons on the other. While for the former it is caused primarily by bremsstrahlung and pair production, respectively, hadrons predominantly deposit energy due to inelastic hadronic interactions. To account for these different behaviours, the system is split into the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL).
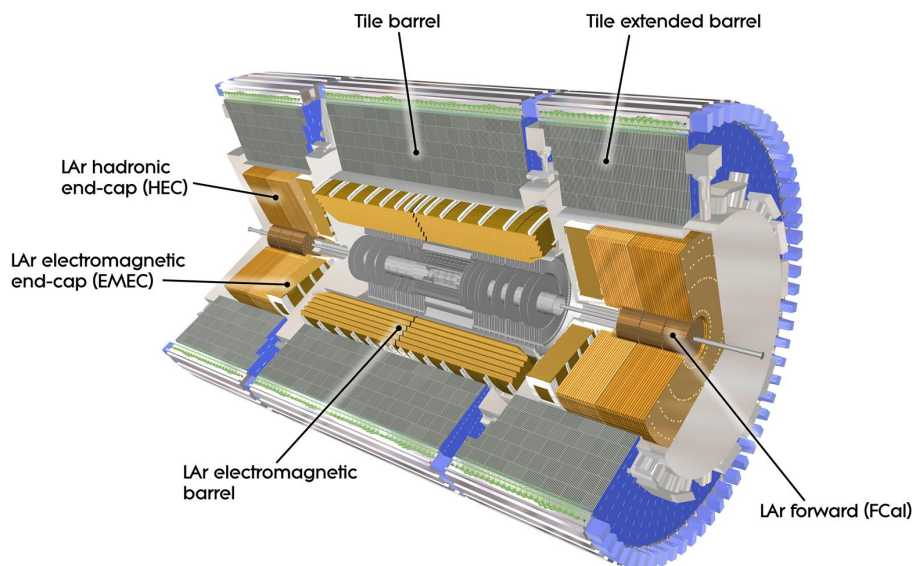

Figure 3.5: Cut-away view of the ATLAS calorimeter system [118].

The ECAL is located closer to the beamline than the HCAL and consists of a barrel region within $|\eta| < 1.475$ and one end-cap on either side, extending the coverage in the region of $1.4 < |\eta| < 3.2$. It relies on the liquid argon technique [119] with three layers of accordion shaped electrodes, using lead as absorber material. An additional layer of purely active material is utilised as a presampler to correct for energy losses occuring before a particle reaches the calorimeter. The depth of the calorimeter is typically given in units of the radiation length X_0, which is the mean distance after which an electrons' energy is reduced to 1/e of its initial value. The ECAL has a total depth of 22-33 X_0 in the barrel and 24-38 X_0 in the end-cap regions, respectively, and a relative energy resolution of

$$\frac{\sigma(E)}{E} = \frac{10\,\%}{\sqrt{E[\text{GeV}]}} \oplus 0.7\,\%. \tag{3.8}$$

Besides precise energy measurements, an important task of the ECAL is the accurate determination of a particle's direction of flight. To facilitate this, the first layer is finely segmented in η, with a granularity in η-ϕ as low as 0.003×0.1, whereas the outer layers are coarser in η but finer in ϕ, with granularities of 0.025×0.025 and 0.05×0.025, respectively.

The HCAL is divided into a barrel region and end-caps on either side, with granularities of 0.1×0.1. The end-caps, composed of two disks each, provide coverage for $1.5 < |\eta| < 3.2$ and follow similar principles as the ECAL, utilising liquid argon as the active medium, but copper as absorber. In the barrel region up to $|\eta| = 1.7$, the tile calorimeter is installed. Its three layers of steel absorbers and scintillating tiles have a total depth of $7.4\,\lambda$, where λ is the hadronic equivalent to X_0 and represents the characteristic interaction length. The relative energy resolution of the HCAL is

$$\frac{\sigma(E)}{E} = \frac{50\,\%}{\sqrt{E[\text{GeV}]}} \oplus 3\,\%. \tag{3.9}$$

In order to detect and measure very forward particles, the range of the calorimeter system is extended by the liquid argon-based Forward Calorimeter in $3.1 < |\eta| < 4.9$, providing an energy resolution of

$$\frac{\sigma(E)}{E} = \frac{100\,\%}{\sqrt{E[\text{GeV}]}} \oplus 10\,\%. \tag{3.10}$$

Its first layer uses copper as passive material, ensuring containment for electromagnetic showers, the other two layers are made of tungsten. The extension of the coverage towards very high pseudorapidities is crucial for the precise determination of the missing transverse momentum of a collision, as it reduces the amount of very forward particles escaping ATLAS undetected.

3.2.4 Muon Spectrometer

Muons are low-ionising and therefore barely interact with the material in the calorimeter system. They are detected and reconstructed primarily using information from the ID and from the outermost subsystem of the ATLAS detector, the MS [120]. Immersed in a strong magnetic field produced by superconducting toroids, it provides tracking capabilities with high momentum resolution of 10% for $1\,\mathrm{TeV}$ muons. There are three toroids, one in the barrel ($|\eta| < 1.2$) and one in each end-cap ($1.0 < |\eta| < 2.7$), with a maximum magnetic field strength of $3.5\,\mathrm{T}$ in the end-caps and $2.5\,\mathrm{T}$ in the barrel region.

The MS is divided into four subsystems with the layout as shown in the cut-away view in Figure 3.6. Monitored Drift Tube (MDT) chambers, filled with a mixture of argon and CO_2, are installed in three layers and offer coverage across the whole range of the MS with spatial resolutions of $50\,\mu\mathrm{m}$. MDTs have a maximum drift time of $700\,\mathrm{ns}$. This is a limiting factor in regions with high rates. Due to their smaller drift time of $40\,\mathrm{ns}$, the innermost layer relies on Cathode Strips Chambers (CSCs) for $|\eta| > 2.0$ instead, owing to the increased background rates in this region. In addition to the detection chambers, triggering capabilities are added through Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs). Both have much higher time resolutions and can distinguish subsequent bunch crossings. As the MDT only provides measurements of the pseudorapidity, information on the azimuthal angle is taken from hits in the RPCs.
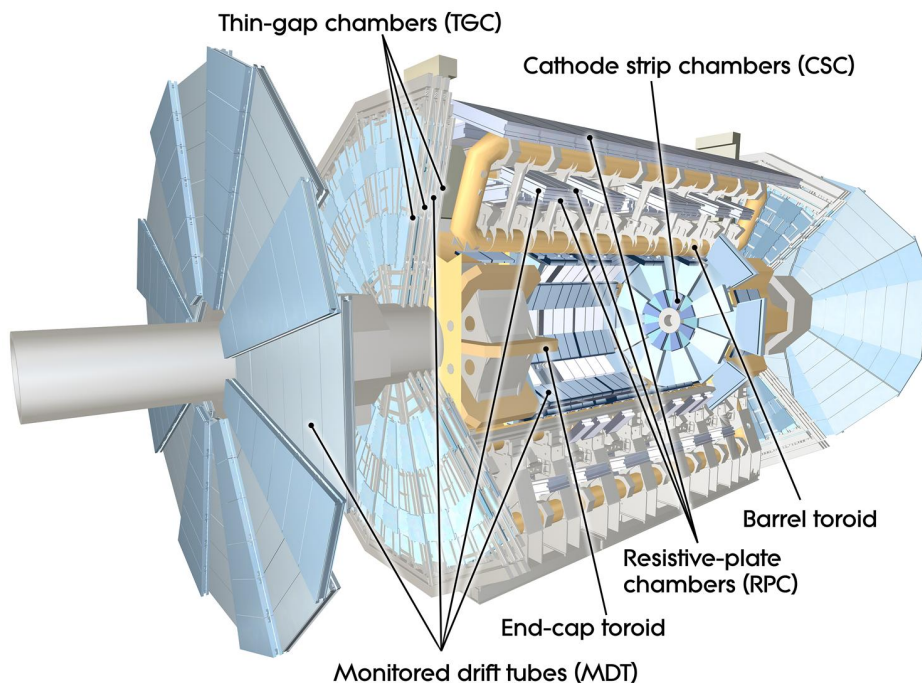

Figure 3.6: Cut-away view of the ATLAS muon spectrometer [121].

3.2.5 Luminosity Measurement

As in most analyses measured data are compared to simulated events, a precise knowledge of the luminosity is essential for the correct normalisation of the prediction. Accurately measuring the luminosity delivered by the LHC and recorded by ATLAS is therefore paramount. To this end, ATLAS employs two dedicated subdetectors, which are placed around the beam line at high pseudorapidities. The primary luminosity measurement is delivered by LUCID2 (LUminosity measurements using Cherenkov Integrated Detector) [122], a Cherenkov detector located approximately 17 m from the interaction point. A second, independent measurement is conducted with the Beam Conditions Monitor [123], a diamond detector placed 184 cm from the interaction point. Both are able to resolve single bunch crossings. Further cross-checks on these measurements come from comparisons with measurements of subsystems of the ATLAS detector, such as the counts of primary vertices detected in the ID.

The analysis presented in this dissertation uses the full Run 2 dataset, corresponding to an integrated luminosity of $139\,\text{fb}^{-1}$. This number has been measured with an uncertainty of $1.7\,\%$ [124]. The luminosities recorded by ATLAS in the different years are listed in Table 3.2 and compared to the luminosities delivered by the LHC. Depending on the year, around $90\,\%$ to $95\,\%$ of delivered luminosity has been recorded by the ATLAS detector. Across the full Run 2, the share of recorded events deemed good for physics analyses amounts to almost $90\,\%$ of the delivered luminosity, containing only data taken during periods in which all detector systems were fully operational.

Table 3.2: Luminosities delivered by the LHC and recorded by ATLAS in the years 2015 to 2018. The last row shows the luminosity corresponding to data taking periods deemed good for physics analysis, after applying data quality requirements.

	2015	2016	2017	2018	total
Delivered by LHC	$4.0\,\text{fb}^{-1}$	$38.5\,\text{fb}^{-1}$	$50.2\,\text{fb}^{-1}$	$63.4\,\text{fb}^{-1}$	$156\,\text{fb}^{-1}$
Recorded by ATLAS	$3.6\,\text{fb}^{-1}$	$35.5\,\text{fb}^{-1}$	$46.4\,\text{fb}^{-1}$	$60.0\,\text{fb}^{-1}$	$147\,\text{fb}^{-1}$
Good for physics	$3.2\,\text{fb}^{-1}$	$33.0\,\text{fb}^{-1}$	$44.3\,\text{fb}^{-1}$	$58.5\,\text{fb}^{-1}$	$139\,\text{fb}^{-1}$

3.2.6 Trigger System

The high luminosity at the LHC renders storing and processing all collision events impossible, due to limitations in computing resources and disk space. To reduce the rate of events stored for future analysis, a real-time trigger system [125] is employed, targeting events of interest, which generally feature particles with high transverse momenta and have a comparably low cross-section.

The system consists of two stages, the first one of which, called first-level trigger (Level-1, L1), is implemented in purpose-built hardware. To account for the need for rapid evaluations and ensure low latency, it utilises limited information provided by the calorimeters and the MS,

thereby allowing for a reduction of the event rate to a maximum of approximately 100 kHz across all trigger signatures with a decision time not exceeding 2.5 µs.

Events accepted by the L1-trigger are further processed by the software-based high-level trigger (HLT). Here, depending on the trigger signature, a refined object reconstruction can be either executed for the full calorimeter or limited to regions of interest identified by the L1-trigger. Utilising finer-granularity information from the calorimeter system and the MS as well as tracking information from the ID, the event rate is reduced to about 1.5 kHz with a decision time of $\mathcal{O}(500\,\mathrm{ms})$.

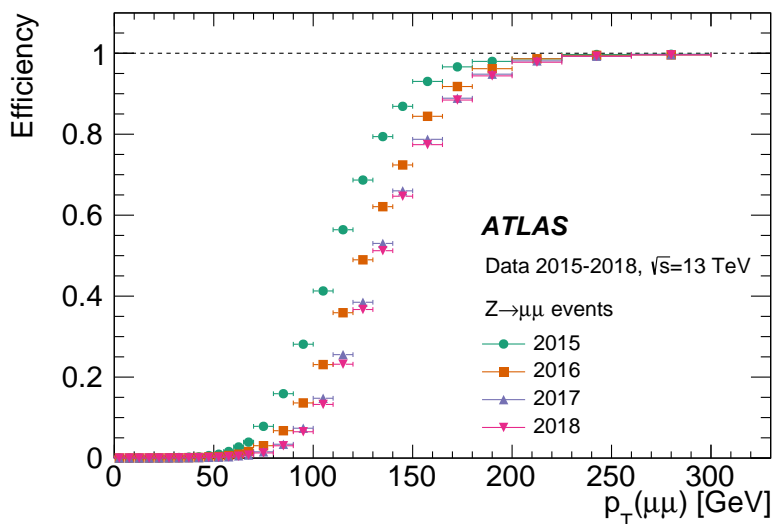**Missing Transverse Momentum Triggers**

Events of interest in this analysis are characterised by large missing transverse momenta. Specialised $E_\mathrm{T}^\mathrm{miss}$ triggers [126] accept events with a rate of up to 10 kHz at L1 level and up to 230 Hz at HLT level, varying over the course of Run 2. The L1 trigger decision is formed based on projective towers built from calorimeter cells, removing low-energy contributions. Accepted events are transferred to the HLT, which employs several algorithms. The *cell* algorithm is used to calculate $E_\mathrm{T}^\mathrm{miss}$ as the sum across all calorimeter cells with energy deposits above a pre-determined threshold that takes into account electronic noise and pile-up contributions. Another approach is based on topological clusters of calorimeter cells. This offers the advantage of being able to determine the origin of a cluster as either electromagnetic or hadronic and therefore to apply the corresponding calibration after clustering. The clusters can be used to evaluate $E_\mathrm{T}^\mathrm{miss}$ directly or as input to more sophisticated algorithms, such as *mht* $E_\mathrm{T}^\mathrm{miss}$. This type of $E_\mathrm{T}^\mathrm{miss}$ is calculated only from calibrated jets built from the topological clusters, because hadronic activity usually represents the largest contributions to the visible momentum of an event, and was used in the HLT trigger decisions during the years 2015 and 2016. In another approach, applied in 2017 and 2018, pile-up effects are suppressed by building patches with the size of a typical $R = 0.4$ jet from topological clusters. Contributions from pile-up are subtracted by performing a fit under the assumption that energy deposits due to pile-up correspond to a total vectorial momentum sum of zero in the x- and the y-direction.

An overview of which trigger was used during which period of data taking is given in Tab. 3.3. The trigger names indicate the required thresholds on HLT and L1 level, respectively, as well as the algorithm used on HLT level. During 2015 and 2016, the thresholds on *mht* $E_\mathrm{T}^\mathrm{miss}$ increased from 70 GeV to 110 GeV to account for rising pile-up. The threshold for the *pufit* algorithm is a constant 110 GeV, with an additional requirement on the *cell* $E_\mathrm{T}^\mathrm{miss}$ of 50 GeV (2017) and 65-70 GeV (2018), respectively.

Since energy deposits in the MS are not considered in the calculation of the missing transverse momentum on trigger level, the performance of the $E_\mathrm{T}^\mathrm{miss}$ triggers can be evaluated in data on $Z \to \mu\mu$ events, with the transverse momentum of the dimuon system constituting a proxy for real $E_\mathrm{T}^\mathrm{miss}$. As Figure 3.7 shows, the efficiencies increase substantially for transverse

Table 3.3: List of E_T^{miss} triggers required to record the dataset used in this analysis.

Year	Period	Trigger
2015		HLT_xe70_mht
2016	A-D3	HLT_xe90_mht_L1XE50
2016	D4-F1	HLT_xe100_mht_L1XE50
2016	F2-	HLT_xe110_mht_L1XE50
2017	B1-D5	HLT_xe110_pufit_L1XE55
2017	D6-	HLT_xe110_pufit_L1XE50
2018	B-C5	HLT_xe110_pufit_xe70_L1XE50
2018	C6-	HLT_xe110_pufit_xe65_L1XE50

Figure 3.7: Efficiencies for the full E_T^{miss} trigger chain, separately evaluated for the years 2015 to 2018 as a function of the transverse momentum of the dimuon system, $p_T(\mu\mu)$ [126].

momenta between $100\,\mathrm{GeV}$ and $200\,\mathrm{GeV}$, before reaching a plateau slightly below $250\,\mathrm{GeV}$. Even though differences between the years due to varying trigger thresholds are evident, the general behaviour is the same in each year. The trigger chain can thus be assumed to be fully efficient for missing transverse momenta above $250\,\mathrm{GeV}$.

4 | Simulation of Proton-Proton Collisions

The general strategy for direct searches of BSM processes is based on comparisons of SM and BSM predictions to experimental data. In order to assess the consistency between observations and predictions, analyses such as the one presented in this thesis therefore rely on an accurate description of the underlying physics processes and on a realistic modelling of the interactions of particles with the detector. As the calculation of these processes involves complicated integrals, usually Monte Carlo (MC) methods are employed, in which pseudo-random numbers are used to simulate events from a given process and generate the predicted distributions.

A typical proton-proton collision of interest in many analyses is illustrated schematically in Figure 4.1, which in this case depicts the production of a Higgs boson in association with a pair of top quarks. It is characterised by the interaction of two partons in what is commonly termed the *hard scattering process*, shown here in red. Massive particles produced in the hard scattering process subsequently decay into lighter particles. QCD radiation is emitted from the partons both before and after the scattering process, and is referred to as initial-state radiation (ISR) or final-state radiation (FSR), respectively, resulting in parton showers (PS). Due to confinement, the partons eventually hadronise and form bound, colourless states, which are shown in light green in Figure 4.1 and can decay further into lighter, stable particles. These collimated sprays of hadrons are reconstructed as jets. Further activity, arising from the remaining partons due to colour connections and beam remnants, is predominantly restricted to lower energies and referred to as the *underlying event* (UE). Electromagnetic radiation from charged particles is possible in any of the different stages and is shown in yellow.

When calculating cross-sections of specific processes, the compositeness of the proton must be accounted for. According to the factorisation theorem, the cross-section for a generic process, $pp \to X$, can be factorised into the hard scattering process, in which the partons are asymptotically free, and the soft, non-perturbative regime, i.e.

$$\sigma(pp \to X) = \sum_{i,j} \int \mathrm{d}x_i \mathrm{d}x_j f_i(x_i, \mu_\mathrm{F}^2) f_j(x_j, \mu_\mathrm{F}^2) \hat{\sigma}_{ij \to X}(x_i x_j s, \mu_\mathrm{R}^2, \mu_\mathrm{F}^2), \tag{4.1}$$

where the sum runs over all partons i and j that contribute to the final state X with the partonic cross-section $\hat{\sigma}_{ij \to X}$. The factorisation scale, μ_F, indicates the threshold between the soft and the hard interaction. When calculating the partonic cross-section at fixed order, the renormalisation scale, μ_R, is introduced, because virtual loops of order n lead to corrections
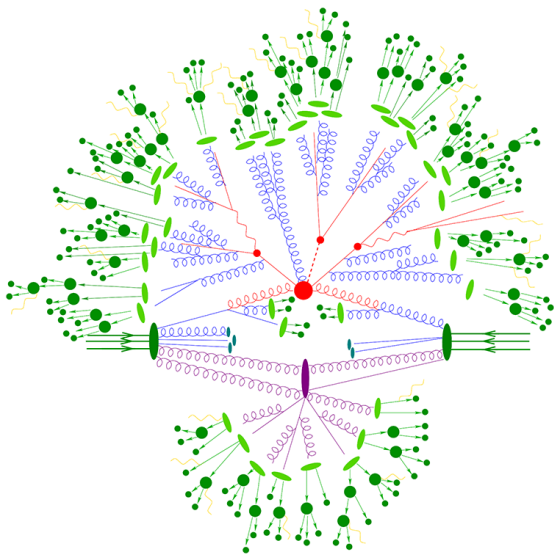
Figure 4.1: Schematic overview of the different processes during a proton-proton collision. The hard scattering process is shown in red, the UE in violet. The PS evolves from the products of the hard scattering process, its contents eventually hadronising as indicated by the light green ellipses. Electromagnetic radiation can occur at any stage and is shown in yellow [127].

of order $\alpha_{\mathrm{S}}^n(\mu_{\mathrm{R}})$.

As the actual collisions happen between two partons, each carrying a fraction x of the respective proton's momentum, the quantity relevant for the calculation of the partonic cross-section is not the centre-of-mass energy, $\sqrt{s}$, of the two colliding protons, but rather the partonic centre-of-mass energy, $\sqrt{\hat{s}} = \sqrt{x_i x_j s}$. The probability for a parton to carry the fraction x_i of the total proton momentum is described by the parton density function (PDF) $f_i(x_i, Q^2)$ as a function of the momentum transfer, Q^2. It cannot be calculated analytically from first principles, but has to be measured experimentally instead, separately for each quark flavour and gluons. PDFs are provided by several groups, each assuming different parameterisations of the underlying functional behaviour and statistically combining inputs from various sources, such as deep-inelastic scattering, fixed target experiments, or hadron colliders. Once the PDF has been determined at a certain energy scale, it can be calculated for other energy scales using the DGLAP evolution equations [128–130]. An example of such a set of PDFs for the different partons, as provided by the NNPDF collatoration, is shown in Figure 4.2 for a momentum transfer of $Q^2 = 10^4\,\mathrm{GeV}^2$. The dominance of the valence quarks u and d is clearly visible for high values of x, whereas sea quarks and gluons dominate for low x.

The partonic cross-section is calculated as

$$\hat{\sigma}_{ij \to X} = \frac{1}{2\hat{s}} \int \mathrm{d}\Phi_X \left| \sum_l \mathcal{M}_{ij \to X}^l(\Phi_X; \mu_{\mathrm{R}}, \mu_{\mathrm{F}}) \right|^2 , \tag{4.2}$$

where Φ_X is the phase space available for final state X and $\mathcal{M}_{ij \to X}^l$ is the matrix element (ME) for the process $ij \to X$ containing l virtual loops. The inclusive partonic cross-section can be evaluated by repeating this calculation for the processes $X + k$, with k indicating the number of real emissions. In this way, higher-order loop corrections to arbitrary precision can
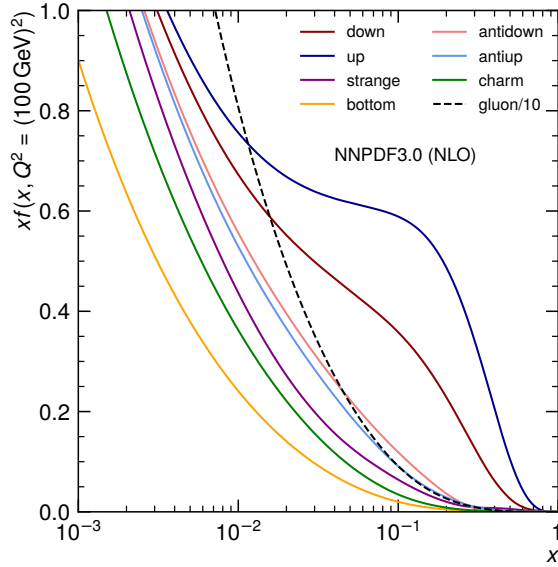
Figure 4.2: PDFs of the NNPDF3.0 PDF set, provided by the NNPDF collaboration, for a momentum transfer of $Q^2 = 10^4\,\mathrm{GeV}^2$ [131, 132].

in principle be added to the LO tree-level calculation, for which $k = l = 0$. The corresponding calculations, however, quickly become computationally prohibitive and are therefore typically only carried out to next-to-leading (NLO, $k + l \leq 1$) or NNLO ($k + l \leq 2$), resulting in a residual dependence on the renormalisation and factorisation scales, which are set to typical scales of the process and commonly chosen to be equal. While in theory real emissions can be evaluated to arbitrary order in the ME, the approach has two drawbacks. For one, final states with numerous real emissions require extensive calculations. Secondly, emissions in the limit of low momenta or under very small angles can lead to divergences. Such soft or collinear divergences are avoided by not integrating over the problematic phase spaces. Instead, PS generators interfaced to the ME generators are utilised to approximate corrections due to higher-order real emissions by simulating parton splittings both in the initial and in the final state. This procedure is applied until energy scales of around $1\,\mathrm{GeV}$ are reached, at which point the partons are effectively confined in colourless hadrons and can no longer be treated perturbatively, requiring the hadronisation to be modelled with phenomenological approaches, such as the Lund string model [133] or the cluster hadronisation model [134].

As the last step of the MC toolchain, the generated events are passed through a simulation of the ATLAS detector, in which interactions of the final state particles with the detector material are simulated. The results are subsequently digitised in the same way as data. MC samples are produced either with a detailed detector model fully implemented in GEANT4 [135] or with parameterised calorimeter responses [136] for a less detailed, but also less computationally expensive simulation. The former is known as Full Simulation and is used for MC samples of the SM backgrounds, while the latter is referred to as Fast Simulation or AFII and is used for signal samples and systematic variations of the SM background samples.

5 | Event Reconstruction

Considering the final state of interest, physics objects used in this analysis are jets, electrons, muons, and the missing transverse momentum. They are each identified and reconstructed from energy deposits throughout the detector. Objects have to fulfil certain quality criteria in order to be considered further. In this section, first, the reconstruction of tracks in the ID caused by charged particles is detailed. Subsequently, the reconstruction procedures for the different objects relevant in the analysis presented in this thesis are discussed, for which the ID tracks are an important ingredient. In addition, the quality criteria imposed on the reconstructed objects are summarised. For electrons and muons, two quality levels are defined, referred to as baseline and signal in the following, with the objects passing the signal level requirements being a subset of the objects fulfilling the baseline level requirements. The signal level is required in the event selection, while the baseline level is used to veto against additional leptons, to calculate the missing transverse momentum, and to resolve ambiguities in the reconstruction through the overlap removal procedure, which is described at the end of this section.

5.1 Tracks and Vertices

Trajectories of charged particles traversing the magnetic field inside the ID are bent. The curvature of the helix is inversely proportional to the particles' momenta, which, together with the sign of charge, can therefore be determined by fitting trajectory models to the energy deposits observed in the ID [115].

The tracks are reconstructed from measurements in the different layers of the ID, starting with clusters built from energy deposits in the pixel detector and the SCT. The results from the clusterisation serve as input to an iterative track finding algorithm, that relies on a pattern recognition approach with several stages. The algorithm employs an inside-out strategy, first defining a set of track seeds from collections of three clusters that are compatible with an additional cluster. Quality criteria are imposed to ensure only well-defined seeds are considered in the computationally more expensive step, which consists of a simplified Kalman filter [137]. Using this filter, the tracks are extended both in- and outwards by adding compatible clusters along the projected trajectory. Track candidates obtained with this technique can in principle have overlaps, which are resolved by applying a stringent ambiguity solver. It assigns a score to the tracks based on quality indicators and resolves overlaps by iteratively removing tracks with smaller scores. Finally, tracks are extended into

the TRT and refitted with the full ID information, if they fulfil quality requirements.

Each reconstructed track is characterised by a set of five parameters. These include their azimuthal angle, ϕ, and their polar angle, θ, the ratio of charge and momentum, q/p, as well as the two impact parameters d_0 and z_0, specifying the minimal distance to the nominal interaction point in the transverse and the longitudinal direction, respectively.

A primary vertex is the point of collision of two protons. Since typically numerous collisions occur in one bunch crossing, multiple vertices can be reconstructed from the set of tracks. They are determined by applying an adaptive vertex finding algorithm [138]. Fits to the tracks and a seed vertex are performed iteratively, gradually weighing down and eventually removing incompatible tracks. Once a vertex is found, the procedure is repeated with all tracks not yet associated with any vertex. From all reconstructed vertices, those with at least two associated tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ are retained. The primary vertex is chosen as the one with the highest sum of squared transverse momenta of associated tracks. Vertices outside of the estimated region of transverse overlap between the two proton beams constitute the set of secondary vertices and are used to identify B hadrons as explained in more detail in Section 5.5.

5.2 Electrons

Electrons are reconstructed from ID tracks and from energy deposits in the calorimeter system. The reconstruction procedure relies on a dynamic clusterisation approach, resulting in so-called superclusters with variable sizes [139] as depicted schematically in Figure 5.1. As opposed to fixed-size clusters, this method has the advantage of recovering energy from bremsstrahlung photons. In a first step, ECAL and HCAL cells with energy deposits above a noise threshold are used as seeds for topo-clusters consisting of topologically connected calorimeter cells. Topo-clusters with only minor energy deposits in the ECAL are removed. The remaining topo-clusters are matched to tracks in the ID, which are refitted to account for bremsstrahlung effects on the trajectories.
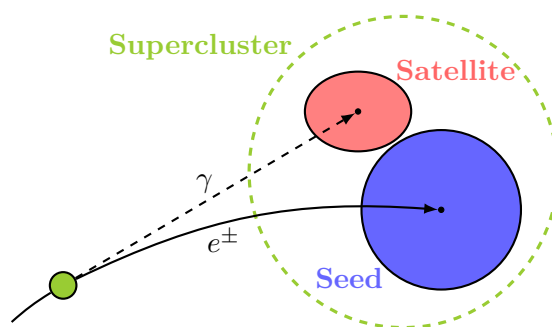

Figure 5.1: Schematic visualisation of a supercluster of calorimeter cells. Satellite clusters are combined with seeds to capture energy losses due to bremsstrahlung photons.

Topo-clusters containing significant energy deposits matched to well-reconstructed tracks serve as seeds and are combined with satellite clusters in close proximity. The resulting superclusters are matched to tracks again. To account for energy losses due to material in front of the calorimeter, the electron energy scale and resolution are adjusted in a MC-based procedure [140]. Utilising a multivariate regression scheme, the electron energy is estimated from shower shape information of the reconstructed cluster. Corrections are applied for intercalibration effects due to energy scale differences between calorimeter layers as well as for local detector non-uniformities not modelled in simulation. In the last step, simulated $Z \to ee$ events are used for the final calibration in a binned-likelihood fit.

In order to suppress contributions from hadronic activity misidentified as electrons and from non-prompt electrons from photon conversions and decays of heavy-flavour hadrons, identification criteria are imposed [139]. These rely on a likelihood discriminant, which combines calorimeter information such as the shower shape and the amount of energy deposited in the various layers of the ECAL and the HCAL with information on the track quality and the track-cluster matching. Based on the likelihood discriminant, several operational points with varying background rejection rates and signal efficiencies are defined, optimised in bins of the pseudorapidity and the transverse energy. In this analysis, two of these levels are of interest. On baseline level, requirements are placed on the likelihood discriminant corresponding to the *LooseLH* operating point and on the number of hits in the ID, in particular including a hit in the IBL. For the *Tight* operational point, which is used for signal electrons, additional requirements on both tracking and calorimeter information are included to further increase the background rejection.

Moreover, contributions from electrons from heavy-flavour decays are reduced by requiring the electron candidates to be isolated from other activity in the ID or the calorimeter. In the analysis, electron candidates are excluded if the scalar sum of transverse momenta of tracks within a cone of $\Delta R = \min(10\,\text{GeV}/p_\text{T}(e), 0.2)$, excluding the electron itself, is larger than $15\,\%$ of the electron's p_T. Analogously, an electron is only considered further if the transverse energy deposited within $\Delta R = 0.2$ around the electron does not exceed $20\,\%$ of the transverse energy of the electron itself.

Efficiencies for electrons to pass the reconstruction, identification, and isolation criteria, respectively, are determined in $Z \to ee$ and $J/\Psi \to ee$ events. For these measurements, the tag-and-probe method is used, in which tag electrons have to pass strong selection criteria, whereas probe electrons are selected by constraining the reconstructed masses to the resonance. From this, efficiencies for probe electrons to pass the respective reconstruction, identification, and isolation requirements are calculated. This is done separately for data and MC events, such that scale factors can be derived to correct for differences in efficiency between data and simulation. These scale factors are applied to all MC simulated events in the analysis.

Electrons are selected only within $|\eta| < 2.47$, additionally excluding the calorimeter transition region $1.37 < |\eta| < 1.52$. To ensure track-to-vertex-association (TTVA), electrons are

removed if their d_0 significance, $|d_0|/\sigma(d_0)$, is above five or if $|z_0 \sin\theta| > 0.5\,\mathrm{mm}$. For an electron to be considered as baseline, it has to fulfil $p_\mathrm{T}(e) > 10\,\mathrm{GeV}$, while for signal electrons the lower threshold is $p_\mathrm{T}(e) > 30\,\mathrm{GeV}$.

5.3 Muons

At the typical energy scales relevant for the ATLAS experiment, muons constitute minimum-ionising particles and therefore exhibit only minor energy deposits in the calorimeter system. As such, their characteristic signature exploited by reconstruction algorithms primarily consists of combinations of hits in the ID and in the MS [141, 142]. Tracks in the ID are reconstructed as described in Section 5.1, while tracks in the MS are reconstructed by building segments in individual stations, which are combined into track candidates, assuming a parabolic trajectory due to the bending power of the magnetic field in the MS. Finally, a global χ^2 fitting procedure is performed, extrapolating the tracks to the beam line and taking into account energy losses in the calorimeter system. Different muon types are defined, based on their reconstruction strategy:

- A combined fit is applied on MS tracks matched to ID tracks, resulting in combined muons (CB).

- Muons reconstructed by extrapolating ID tracks outwards and matching them to at least three loosely-aligned MS hits are referred to as inside-out muons (IO). Their trajectories are refitted with the combined ID, calorimeter, and MS information.

- Segment-tagged muons (ST) are reconstructed from ID tracks, extrapolated to the MS and tightly matched to at least one reconstructed MS segment.

- For calorimeter-tagged muons (CT), ID tracks are extrapolated through the calorimeter system in a search for energy deposits compatible with a minimum-ionising particle.

A schematic visualisation of the different muon types is given in Figure 5.2.

Identification criteria based on the number of hits in the ID and the MS and the measurement quality are imposed to reject non-prompt muons, particularly those resulting from decays of light hadrons. Different operating points with varying signal efficiencies and background rejection rates are available to cover different analysis use cases. In this analysis, signal muons are required to pass the *Medium* operating point, accepting only CB and IO muons with consistent individual measurements in the ID and the MS. In accordance with other searches for pair-produced LQs [93, 97], for $p_\mathrm{T}(\mu) > 800\,\mathrm{GeV}$ the *High-p_T* operating point is prerequisite. As the trajectories of muons with high transverse momenta are barely bent in the magnetic field, stricter requirements on the number of hits in the MS are imposed to ensure good resolution. In addition, a threshold on the uncertainty of the measurement of the transverse momentum is introduced to remove inaccurately measured muons. For baseline muons, the Loose operating point is used, which includes all Medium muons and also
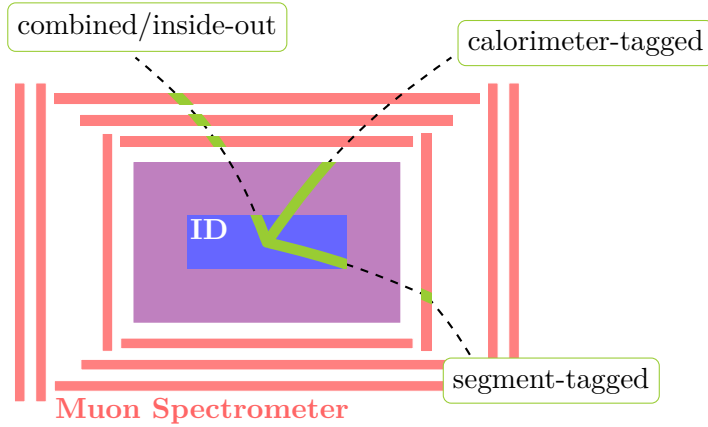
Figure 5.2: Overview of the four muon types used in this analysis and their different interactions with the detector material. Energy deposits in the ID and the MS are shown in light green.

accepts CT and ST muons for $|\eta| < 0.1$, where the CB muon efficiency is low due to limited coverage in the MS. Contributions from muons originating from hadron decays are further removed through selections applied on the impact parameters of the muon tracks. The relevant TTVA criteria are placed on the d_0 significance, which is required to be smaller than three, and on $|z_0| \sin \theta < 0.5$ mm. Track-based or calorimeter-based isolation criteria aid in suppressing contributions from muons from heavy-flavour decays. In this analysis, a signal muon is removed if the scalar sum of transverse track momenta within a cone with radius $\Delta R = \min(10 \, \text{GeV}/p_\text{T}(\mu), 0.3)$ around the muon exceeds $6\,\%$ of its transverse momentum. This criterion corresponds to the *TightTrackOnly* operating point. No isolation criteria are imposed on baseline muons.

Analogously to electrons, the muon momentum scale and resolution is calibrated with a binned likelihood fit of the invariant mass distributions of the two muons in $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ events in data and MC. Efficiencies for muons to pass reconstruction, identification, TTVA, and isolation criteria are determined with the tag-and-probe method, using the same decay processes. Again, scale factors for the simulation, binned in the transverse momentum and the pseudorapidity, are derived by performing the procedure separately for data and MC events.

Muons considered in this analysis are constrained to $|\eta| < 2.5$. The p_T requirements are 10 GeV and 30 GeV for baseline and signal muons, respectively.

5.4 Jets

Due to colour confinement, partons produced in the initial scattering event hadronise, forming collimated sprays of charged and neutral hadrons, referred to as jets. They can be detected and subsequently reconstructed. From the characteristics of these jets, conclusions can be

drawn on the properties of the initial partons.

To ensure good momentum resolution, information from the ID and the calorimeter system are combined into so-called particle-flow (PFlow) objects [143]. Similar to the electron reconstruction, calorimeter cells are combined into topo-clusters, to which an origin correction is applied to account for the position of the primary vertex. Afterwards, tracks are matched to the corrected topo-clusters. Topo-clusters without any associated tracks are assumed to stem from uncharged particles. In topo-clusters with associated tracks, contributions of charged particles are subtracted, depending on the track-p_T and on the energy deposited in the clusters. For this removal procedure, energy deposits are estimated from the topo-cluster position and track kinematics. Afterwards, relying only on momentum information obtained from tracks, contributions from charged particles are added. The subtraction scheme is primarily applied for softer particles, whereas a reconstruction from topo-clusters alone is preferred at high energies and momenta, thereby combining the advantages of great tracking capabilities for low-p_T trajectories in the ID with the excellent resolution in the calorimeters at high energies. Compared to an approach relying solely on calorimeter information, this procedure also offers improved pile-up stability. After applying the subtraction scheme, the collection of PFlow objects consists of tracks as well as modified and unmodified topo-clusters. It serves as input to the jet reconstruction method, which typically is the anti-k_T algorithm [144] as implemented in the FASTJET software package [145, 146]. The algorithm is based on measures of the distance d_{ij} between two PFlow objects i and j defined as

$$d_{ij} = \min\left(\frac{1}{k_{Ti}^2}, \frac{1}{k_{Tj}^2}\right) \frac{\Delta R_{ij}^2}{R^2} \tag{5.1}$$

and of the distance of a PFlow object i to the LHC beam

$$d_{iB} = \frac{1}{k_{Ti}^2}, \tag{5.2}$$

where k_T is the transverse momentum of the object and R is the radius parameter, which controls the size of the reconstructed jets. In an iterative approach, the distances d_{ij} and d_{iB} are compared, combining objects i and j if the former is smaller and declaring i a jet if the latter is. The procedure is repeated until all PFlow objects are clustered. With the anti-k_T algorithm, closeby particles and those with large momenta are assigned small distances. These therefore tend to be clustered first, forming seeds for the jet. Particles carrying less energy are added to these seeds, creating roughly circular shapes in the η-ϕ-plane, as illustrated in Figure 5.3 for an example with $R = 1.0$. This approach is advantageous, as it is independent of soft or collinear emissions and therefore inherently offers infrared and collinear safety. In the following, jets formed with a radius parameter of $R = 0.4$ are used, referred to as small-R jets.

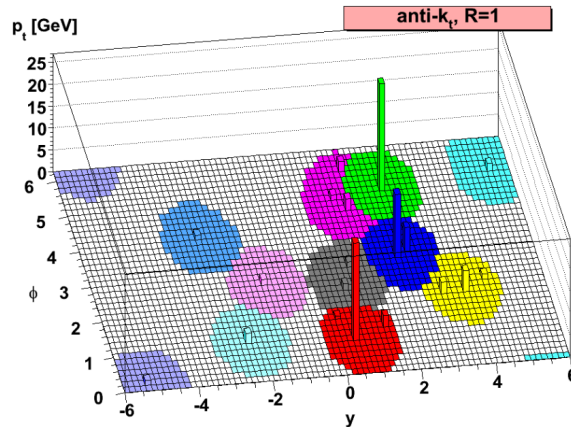After reconstruction, the jet energy scale (JES) is calibrated to compensate for incorrect

Figure 5.3: Illustration of the jet shapes when applying the anti-k_T algorithm with a radius parameter of $R = 1.0$ [144].

physics modelling and various detector effects [147]. This is done in multiple stages, all but the last of which are applied to data and simulated events, while the last one is applied to data only. The calibration procedure is illustrated in Figure 5.4. In the first stage, excess energy from pile-up contributions is removed in a two-step approach. A subtraction scheme is applied, based on the jet area A and the jet-p_T density, $\langle p_\mathrm{T}/A \rangle$. Residual pile-up dependencies are corrected for as a linear function of the mean number of interactions per bunch crossing and of the number of reconstructed primary vertices in the event.

In the next stage, the absolute JES is calibrated using dijet MC events to compensate for mismeasurements, energy losses in passive material, and incorrectly clustered particles. Furthermore, biases in the pseudorapidity of the reconstructed jets due to transitions between detector geometries and technologies are reduced.
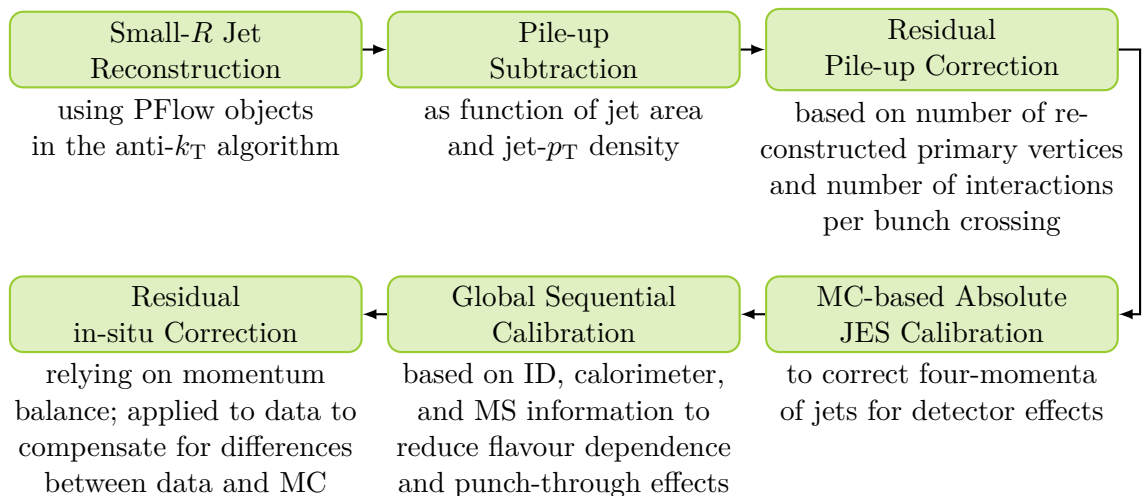

Figure 5.4: Schematic overview of the different stages of the jet energy scale calibration procedure applied to reconstructed jets. Graphic adapted from Ref. [147].

The third stage consists of several multiplicative corrections, referred to as Global Sequential Calibration (GSC), and is derived from six observables covering ID, calorimeter, and MS information. It is constructed to compensate for fluctuations in the properties of jet constituents, especially due to the type of initiating particles, as well as for variations in the shower development and to reduce the impact from jets not fully contained within the calorimeter system, an effect called punch-through.

Remaining differences between data and simulation are caused by imperfect modelling of the detector as well as of the underlying physics processes. They are mitigated in the final, in-situ calibration stage, itself consisting of three sub-stages, all of which rely on the transverse momentum balance in events with hadronic activity and at least one well-calibrated object. First, the η intercalibration is derived in dijet systems to match the energy scale of jets with $|\eta| > 0.8$ to that of jets within $|\eta| < 0.8$, which is assumed to be well understood. Subsequently, $Z(\to \ell\ell)/\gamma+$jet events are utilised for calibrations up to a jet-p_T threshold of approximately $1.2\,\mathrm{TeV}$. Jets with higher transverse momenta are calibrated in the last step, based on systems in which a single high-p_T jet recoils against multiple lower-p_T jets, which themselves have been calibrated in the previous step.

An additional calibration of the jet mass scale (JMS) is derived using dijet MC events, analogously to the absolute JES calibration.

To be considered in this analysis, jets need to have $p_\mathrm{T} > 25\,\mathrm{GeV}$ and to be within $|\eta| < 2.5$. Additionally, jets with $p_\mathrm{T} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ have to pass the tight Jet Vertex Tagger (JVT) working point [148], which suppresses contributions from pile-up jets by imposing requirements on the multivariate combination of information on pile-up tracks and on those matched to the primary vertex. Scale factors to compensate for efficiency differences of the JVT requirement in data and simulation are derived with a tag-and-probe method in $Z(\to \mu\mu)+$jets events.

5.5 Flavour Tagging

Hadrons originating from bottom quarks have mean life times of approximately $1.5\,\mathrm{ps}$ and, at typical transverse momenta, can therefore travel through the detector for several millimeters before decaying, resulting in a secondary vertex. The procedure of identifying such jets initiated by bottom quarks is called b-tagging. To ensure good selection efficiencies and low misidentification rates, the characteristics of b-jets are exploited in several complementary low-level tagging algorithms, later combined into one multivariate high-level observable, called DL1r [149, 150].

One class of low-level taggers, specifically IP2D and IP3D, relies solely on impact parameter information, resulting in log-likelihood discriminants to separate between different jet categories. To improve sensitivity to correlations of impact parameters across tracks, which are not fully taken into account in the aforementioned taggers, a recurrent NN (RNN) is employed for the RNNIP tagger, trained on track impact parameters as well as kinematic

and angular information.

These methods are complemented by vertex-finding algorithms, which are used to reconstruct the vertex structures of B hadron decays. The secondary vertex finding algorithm, SV1, searches for displaced secondary vertices within associated tracks. It provides observables such as the number of tracks associated with the secondary vertex or its distance to the primary vertex. Lastly, the full decay chain of B hadrons is reconstructed, using an approach based on a Kalman filter to approximate the hadron flight path and evaluate vertex positions. From this, conclusions are drawn about properties of the identified displaced vertices, particularly the secondary and tertiary ones.

Information from the low-level algorithms is used as input for DL1r, combining it in a deep feed-forward NN and including the transverse momenta and pseudorapidities of jets in order to exploit correlations of kinematic information and the low-level taggers. The NN has three output nodes, one each for the flavour hypothesis light, charm, or bottom. These can be interpreted as the probabilites p_{light}, p_c, and p_b for a jet to originate from the corresponding flavour. The final DL1r discriminant is defined as

$$D_{\text{DL1r}} = \ln \left(\frac{p_b}{f_c p_c + (1 - f_c) p_{\text{light}}} \right) \tag{5.3}$$

where $f_c = 0.018$ is the effective c-jet fraction in the background training sample. Different b-tagging efficiencies and c-/light-jet rejection rates are achieved by varying the requirement on the DL1r score. It is important to note that the tagging rate depends on the event kinematics, in particular on the transverse momentum of the b-jet. While in this analysis the 77 % working point is used, corresponding to a requirement of $D_{\text{DL1r}} > 2.195$, this efficiency only represents an average across the whole phase space. Tagging rates decrease significantly towards higher transverse b-jet momenta, as indicated in Figure 5.5, which shows b-jet identification efficiencies measured in $t\bar{t}$ events using a tag-and-probe method [151]. Differences between data and MC due to mismodelling are apparent in the plot. These deviations are corrected for through scale factors applied on MC events, which are derived in bins of b-jet-p_{T}.

5.6 Missing Transverse Momentum

Due to their nature as solely weakly interacting particles, neutrinos cannot be measured directly in the ATLAS detector, resulting in an imbalance in the observed sum of transverse momenta in events containing neutrinos. Instead, the missing transverse momentum is used as proxy for the transverse momentum carried by neutrinos. It is calculated as the negative vectorial sum of transverse momenta of all visible objects, comprising contributions from fully reconstructed and calibrated particles and jets, referred to as hard objects, as well as from a *soft term* built solely from tracking information [152, 153]. Hence, $\vec{E}_{\text{T}}^{\text{miss}}$ in this analysis is
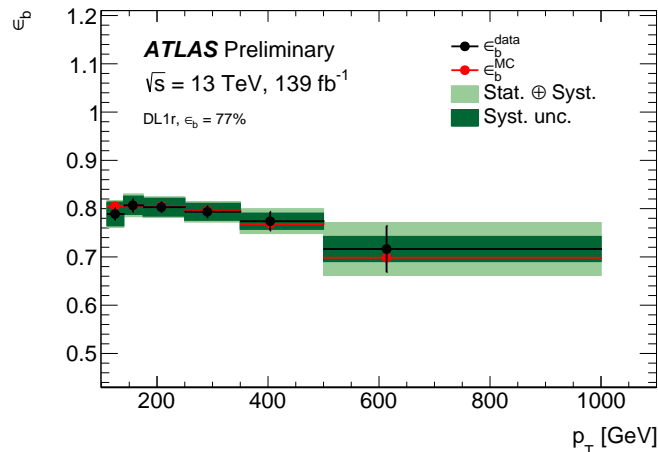
Figure 5.5: Efficiencies to identify a b-jet using the DL1r algorithm at a working point of 77 % as a function of the transverse b-jet momentum. The tagging rates are measured in $t\bar{t}$ events in data and MC [151].

defined as

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum \vec{p}_{\mathrm{T}}^{\,e} - \sum \vec{p}_{\mathrm{T}}^{\,\mu} - \sum \vec{p}_{\mathrm{T}}^{\,\mathrm{jets}} - \sum \vec{p}_{\mathrm{T}}^{\,\mathrm{soft}}. \qquad (5.4)$$

The order of the terms of the sum is important, as in case of ambiguities in the reconstruction lower-priority particles are discarded in favor of ones entering the sum earlier. For electrons and muons, the analysis-specific object definitions are applied. In case of the analysis at hand, baseline objects are used. For jets, the $E_{\mathrm{T}}^{\mathrm{miss}}$ working point determines the criteria imposed. Here, the *Tight* working point is chosen, which includes jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. Jets within $|\eta| < 2.4$ and with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ need to also pass the JVT criteria. A dedicated ΔR-based overlap removal procedure is applied to resolve ambiguities between reconstructed objects. The track-based soft term is calculated from tracks not associated with any previously reconstructed object. Tracks are included if they have $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ and fulfil $|z_0 \sin\theta| < 2.0\,\mathrm{mm}$. This approach cannot account for soft neutral particles, but is chosen over a calorimeter-based approach due to its robustness against pile-up effects.

5.7 Overlap Removal

With the procedures described above, the same energy deposit could potentially be part of multiple reconstructed objects. To avoid ambiguities, a procedure is used which resolves these overlaps deterministically. This overlap removal algorithm consists of several stages and is applied on the baseline objects as defined above. At first, overlaps between leptons are handled. Electrons are rejected when they share an ID track with a muon, unless the muon in question is only calorimeter-tagged, in which case the muon is removed instead. Jet-lepton overlap is determined from the distance between objects in the η-ϕ plane, calculated

as $\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$ using the rapidity instead of the pseudorapidity. Jets within $\Delta R < 0.2$ to an electron are removed, as are electrons within $\Delta R < 0.4$ to remaining jets. Jet-muon overlap is resolved by discarding jets with less than three associated tracks if they are within $\Delta R < 0.2$ to a muon or ghost-associated [154]. Otherwise, muons are rejected if their distance to a jet is smaller than $\Delta R < \min\left(0.4, 0.04 + 10\,\mathrm{GeV}/p_\mathrm{T}(\mu)\right)$.

5.8 Jet Reclustering

As all signal events of interest contain a hadronically decaying top quark, reconstructing that top quark from the individual small-R jets is of great interest. To this end, two different reclustering approaches, one with a fixed, the other with a variable radius parameter, are described in the following, both relying on calibrated small-R jets as input. Reclustering from calibrated objects is beneficial, because properties can simply be propagated to the resulting reclustered jets and no dedicated calibration is required.

5.8.1 Large-R jet reclustering

Since the hadronic top-quark decay results in three individual quarks, it is beneficial to not limit the jet reconstruction to small radii of $R = 0.4$, but to increase the size of the reconstructed jet cone to capture all decay products. For this purpose, large-R jets with $R = 1.0$ are a common choice. These large-R jets are reclustered from small-R jets using the anti-k_T algorithm. In order to improve pile-up stability and mass resolution, a trimming procedure is applied, removing constituent jets carrying less than $5\,\%$ of a reclustered jet's transverse momentum. Large-R jet candidates are discarded if they have only one constituent, a transverse momentum of less than $150\,\mathrm{GeV}$, or a mass of less than $50\,\mathrm{GeV}$.

5.8.2 Hadronic Top Quark Reconstruction

A second approach to reconstruct hadronically decaying top quarks relies on an iterative anti-k_T reclustering method with a variable radius parameter R [155], allowing for a higher reconstruction efficiency than fixed-radius methods. The algorithm makes use of the relation between the maximum distance, $R(p_\mathrm{T})$, between two massless daughter particles originating from a two-body decay of a mother particle with mass m and transverse momentum p_T, which can be approximated by

$$R(p_\mathrm{T}) = \frac{2m}{p_\mathrm{T}}. \tag{5.5}$$

The maximum distance decreases for higher momenta of the mother particle, as its decay products are increasingly collimated. Even though the W boson mass cannot be assumed to be negligible, the relation in Equation (5.5) holds reasonably well as an approximation for the radius of reconstructed jets containing the products from top-quark decays. Particularly

for transverse top-quark momenta above $400\,\mathrm{GeV}$, the spatial separation decreases to values considerably smaller than $R = 1.0$ used for the fixed-radius large-R jets introduced above. Therefore, in order to retain high selection efficiencies for top quarks with high transverse momenta and as the true transverse momentum of the hadronically decaying top quark is unknown, a recursive approach to the reclustering is used, which proceeds as follows:

1. Small-R jets are reclustered with an initial radius parameter of $R_0 = 3.0$, corresponding to a transverse top-quark momentum of $p_\mathrm{T}(t_\mathrm{had}) \approx 115\,\mathrm{GeV}$. As the radius is adapted iteratively, the exact choice of the starting parameter is not critical.

2. Given the transverse momentum of a reclustered variable-radius jet, the estimated radius is calculated following Equation (5.5) with $m = 173\,\mathrm{GeV}$ and compared to the radius hypothesis from the previous iteration, R_{i-1}. Three different cases need to be handled:

   a) If the current estimated radius is within a certain range to the previous one, i.e. if

   $$R_{i-1} - R_\mathrm{down} \leq R(p_\mathrm{T}) \leq R_{i-1} + R_\mathrm{up} \tag{5.6}$$

   with $R_\mathrm{down} = 0.5$ and $R_\mathrm{up} = 0.3$, then the reclustered large-R jet is classified as a hadronic top candidate, unless its mass is below $150\,\mathrm{GeV}$, in which case it is discarded.

   b) If, however, $R(p_\mathrm{T})$ is above the range given in Equation (5.6) instead, the candidate is discarded. This is necessary to remove low-p_T jets not originating from top-quark decays.

   c) Finally, if $R(p_\mathrm{T})$ is below the range in Equation (5.6), the second step is repeated recursively with the reduced radius parameter of

   $$R_i = \max\left(R(p_\mathrm{T}) + \frac{1}{2}R_\mathrm{down}, \frac{1}{2}R_{i-1} \right). \tag{5.7}$$

   The adapted radius is required to be at least half as large as the previous one in order to prevent it from shrinking too rapidly and potentially splitting a true hadronically decaying top quark.

3. The hadronic top quark-candidate is discarded if no stop condition is reached within ten iterations.

In each event, only the hadronic top-quark candidate with the highest transverse momentum is considered.

5.9 Derived Variables

In addition to the reclustered jets, a number of variables is derived by combining various kinematic observables. Instead of solely relying on simple kinematic behaviour of individual objects such as their transverse momentum, this allows for making use of knowledge about the underlying physics processes and about detector effects. Four important variables are introduced in the following.

Transverse mass m_T: Calculated from the transverse energies of two particles, p_1 and p_2 with masses $m(p_1)$ and $m(p_2)$ and transverse momenta $\vec{p}_\mathrm{T}(p_1)$ and $\vec{p}_\mathrm{T}(p_2)$, the transverse mass m_T is defined as

$$m_\mathrm{T}^2\left(\vec{p}_\mathrm{T}(p_1),\vec{p}_\mathrm{T}(p_2)\right) \quad = \quad \left(E_\mathrm{T}(p_1)+E_\mathrm{T}(p_2)\right)^2 - \left(\vec{p}_\mathrm{T}(p_1)+\vec{p}_\mathrm{T}(p_2)\right)^2 \tag{5.8}$$

$$\overset{\substack{m(p_1)\to 0 \\ m(p_2)\to 0}}{=} 2 p_\mathrm{T}(p_1) p_\mathrm{T}(p_2)\left(1-\cos\Delta\phi\left(\vec{p}_\mathrm{T}(p_1),\vec{p}_\mathrm{T}(p_2)\right)\right) \tag{5.9}$$

with the azimuthal angle $\Delta\phi$ between the two particles. It is commonly used to reconstruct particles with a weakly interacting particle among their decay products, because in the case of $E_\mathrm{T}^\mathrm{miss}$ only information about the transverse, but not the longitudinal four-vector components is available. As such, the observable $m_T(\ell, E_\mathrm{T}^\mathrm{miss})$ is particularly relevant for background contributions from events in which a W boson is produced in association with a number of jets (W+jets), because it represents a lower bound on the reconstructed W boson mass. Conversely, down-type LQs can decay into a bottom quark and a neutrino. For these decays, the transverse mass of the leading-p_T b-jet and $E_\mathrm{T}^\mathrm{miss}$ is useful.

$E_\mathsf{T}^\mathsf{miss}$ significance: The missing transverse momentum is calculated from the calibrated reconstructed jets and leptons. However, limited reconstruction efficiencies and resolutions are inherent to all these calibrated objects, depending on their types and kinematics. An observed transverse momentum imbalance is therefore not necessarily caused by non-interacting particles, but can also arise through the inaccurate determination of transverse momenta. These inaccuracies are taken into account in the object-based $E_\mathrm{T}^\mathrm{miss}$ significance, which thus is a measure for the compatibility of the observed value of $E_\mathrm{T}^\mathrm{miss}$ with real $E_\mathrm{T}^\mathrm{miss}$ from a particle not depositing energy in the detector. Assuming independent measurements of individual particles with Gaussian p_T resolutions, parameterised using MC simulation, the $E_\mathrm{T}^\mathrm{miss}$ significance can be calculated as

$$S = \frac{E_\mathrm{T}^\mathrm{miss}}{\sqrt{\sigma_\mathrm{L}^2(1-\rho_\mathrm{LT}^2)}}. \tag{5.10}$$

The total variance in parallel to the measured $\vec{E}_\mathrm{T}^\mathrm{miss}$, σ_L^2, and the correlation factor between momentum measurements in parallel and perpendicularly to $\vec{E}_\mathrm{T}^\mathrm{miss}$, ρ_LT, are obtained event-by-event based on the reconstructed objects. High values of the $E_\mathrm{T}^\mathrm{miss}$ significance indicate the presence of weakly-interacting particles as the source of real $E_\mathrm{T}^\mathrm{miss}$, whereas lower values correspond to missing transverse momenta compatible with zero within resolution effects.

Asymmetric transverse mass $am_\mathbf{T2}$:   As discussed above in the context of the $E_\mathrm{T}^\mathrm{miss}$ significance, missing transverse energy is not necessarily caused solely by weakly interacting particles traversing the detector without depositing energy or by mismeasurements of calibrated objects. It can, for instance, also be the result of the failure to identify a charged lepton. Such is the case with dileptonic $t\bar{t}$ events, in which one of the charged leptons is not detected. For the purpose of suppressing contributions from such events, the asymmetric transverse mass is introduced, which for dileptonic $t\bar{t}$ events is designed to exhibit a kinematic endpoint at the top-quark mass. This is achieved by calculating the transverse masses of the decay products of the top quark and the antitop quark following Equation (5.8).

However, since in the scenario at hand one of the charged leptons remains undetected, one branch of the decay consists of a bottom quark and the undetected W boson, while the other one contains a bottom quark, the reconstructed lepton, and an invisible neutrino. The former branch is termed *branch a* in the following, the latter *branch b*. For visualisation purposes, a Feynman diagram of a dileptonic $t\bar{t}$ decay is shown in Figure 5.6, also indicating the two branches and highlighting reconstructed and undetected final-state particles.
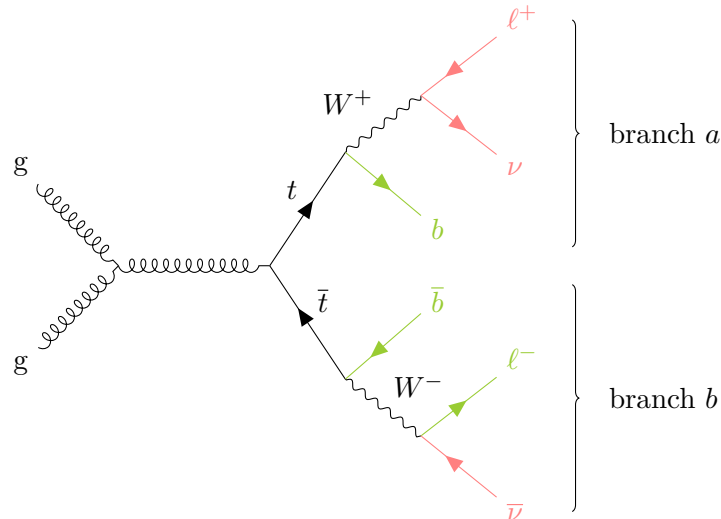

Figure 5.6: Feynman diagram of the dileptonic $t\bar{t}$ decay. Reconstructed final-state particles are shown in green, undetected ones in red. While in branch *b* both, the charged lepton and the bottom quark, are identified and reconstructed, in branch *a* only the bottom quark is detected.

Assuming the mass of the charged lepton to be negligible, the transverse masses are given as

$$m_{T,a}^2 = \left(\sqrt{p_T^2(b_a) + m^2(b_a)} + \sqrt{p_T^2(W_a) + m^2(W_a)} \right)^2 - (\vec{p}_T(b_a) + \vec{p}_T(W_a))^2 \qquad (5.11)$$

for branch a and

$$m_{T,b}^2 = \left(\sqrt{p_T^2(b_b) + m^2(b_b)} + p_T(\ell_b) + p_T(\nu_b) \right)^2 - (\vec{p}_T(b_b) + \vec{p}_T(\ell_b) + \vec{p}_T(\nu_b))^2 \qquad (5.12)$$

for branch b, respectively, with the bottom quarks b_a and b_b, the invisible W boson W_a, the reconstructed lepton ℓ_b, and the associated neutrino ν_b. The indices indicate the respective branch. Both transverse masses are bound from above by the top-quark mass. Since the individual contributions from the transverse momenta of W_a and ν_b to the missing transverse momentum are unknown, a minimisation is performed constraining the sum of invisible transverse momenta to E_T^{miss}, i.e.

$$am_{T2} = \min_{\vec{p}_T(\nu_b) + \vec{p}_T(W_a) = \vec{E}_T^{\mathrm{miss}}} \left(\max(m_{T,a}, m_{T,b}) \right). \qquad (5.13)$$

Since the b-jets cannot be assigned to the branches unambigiously, either combination is evaluated and am_{T2} is taken to be the minimum of the two alternatives. Unlike dileptonic $t\bar{t}$ events, other processes do not fulfil the underlying assumptions entering the computation of am_{T2} and are therefore not limited by the top-quark mass, making this observable a powerful tool to distinguish dileptonic $t\bar{t}$ from other processes.

Effective mass m_{eff}:   The effective mass is a measure for the total energy of the initial collision. It is calculated as the scalar sum of the transverse momenta of all calibrated objects and E_T^{miss}. In the scope of this analysis, it is therefore defined as

$$m_{\mathrm{eff}} = p_T(\ell) + \sum_{i=1}^{n_j} p_T(j_i) + E_T^{\mathrm{miss}} \qquad (5.14)$$

with the jet multiplicity n_j and the individual jets j_i.

6 | Data and Simulated Event Samples

In the analysis, data from proton-proton collisions at centre-of-mass energies of $\sqrt{s} = 13\,\text{TeV}$, collected by the ATLAS experiment in the years 2015 to 2018, is utilised, corresponding to an integrated luminosity of $139\,\text{fb}^{-1}$. Only data collected during stable beam conditions and with all detector subsystems operational are considered [156].

Samples of MC simulated events are utilised to model distributions of BSM signal and SM background processes. The set of available MC samples comprises scalar and vector LQ signals and numerous background processes. The various generator setups for the different physics processes are described in the following, detailing the ME generators and the PS generators they are interfaced to. As discussed in Chapter 4, PS and hadronisation cannot be calculated analytically and are simulated with phenomenological approaches introducing sets of free parameters, that affect the modelling of UE, ISR, and FSR, and are tuned to data. The specific choices of parameters are termed *tunes* in the following, the most common of which is the A14 tune [157], used for all samples in this analysis unless indicated otherwise. Furthermore, in all MC samples, except those generated with SHERPA [158], B hadron decays are handled by EVTGEN [159] to fully take spin correlations in heavy-flavour decays into account.

Top-antitop-quark pair production, referred to as $t\bar{t}$, and single top-quark events as well as the production of a Higgs boson in association with a $t\bar{t}$ pair ($t\bar{t} + H$) were all modelled with the POWHEG BOX v2 [160–163] generator and the `NNPDF3.0NLO` PDF set [164] at NLO in QCD, interfaced to PYTHIA 8.230 [165] for the simulation of PS, hadronisation, and UE, using the A14 tune. In the case of $t\bar{t}$ pair production, the samples were normalised to the theoretical prediction at NNLO, which includes resummation of NNLL soft-gluon terms, calculated using `Top++ 2.0` [166]. Events containing a $t\bar{t}$ pair and a vector boson ($t\bar{t} + V$) were simulated using the MG5_AMC@NLO 2.3.3 generator with the `NNPDF3.0NLO` PDF set and MEs at NLO accuracy. They were matched to the PYTHIA 8.210 PS generator using the A14 tune.

The production of vector bosons, $V = W, Z$, in association with jets was modelled with the SHERPA 2.2.1 generator, with MEs accurate to NLO for up to two jets and to LO for up to four jets, as calculated with the Comix [167] and the OPENLOOPS [168–170] libraries. The samples were normalised to the theoretical NNLO expectation [171]. Similarly, depending on the process, the production of diboson (VV) events was simulated with the SHERPA 2.2.1 generator or the SHERPA 2.2.2 generator, using NLO-accurate MEs for up to one additional jet and LO-accurate MEs for up to three additional jets. In both cases, the `NNPDF3.0NNLO` PDF set was used and simulated hard scattering events were matched to the SHERPA PS

with the MEPS@NLO [172–175] prescription and a SHERPA-internal tune provided by the authors.

Samples of pair-produced scalar or vector LQs were generated with MG5_AMC@NLO 2.6.0 and MG5_AMC@NLO 2.8.1, respectively, using the NNPDF3.0NLO PDF set with $\alpha_S = 0.118$. For the purpose of retaining spin correlations, decays of LQs and top quarks were handled with MADSPIN, while UE, PS, and hadronisation were simulated with PYTHIA 8.230 for scalar LQs and with PYTHIA 8.244 for vector LQs.

As discussed in Section 2.1, the model used for the simulation of scalar LQs, is an extension of the one described in Ref. [88], allowing for flavour off-diagonal couplings to quarks and leptons in addition. It is based on previous fixed-order NLO QCD calculations [176, 177]. The coupling strength to quarks and leptons is set to $\lambda = 0.3$. The value of β differs from the branching ratio of the LQ into charged leptons because of limitations of the phase space originating primarily from the large top-quark mass and was set to 0.5 in the simulation. The corresponding branching ratios differ depending on the LQ mass, as the decay into top quarks is restricted more strictly for smaller signal masses. Using Equation (2.4) and Equation (2.5), they are calculated analytically for down-type scalar LQs as

$$\mathcal{B}(LQ_d \to t\ell) = \frac{\beta \left(m_{\mathrm{LQ}}^2 - m_t^2\right)^2}{\beta \left(m_{\mathrm{LQ}}^2 - m_t^2\right)^2 + (1-\beta)m_{\mathrm{LQ}}^4} \tag{6.1}$$

and for up-type scalar LQs as

$$\mathcal{B}(LQ_u \to b\ell) = \frac{\beta m_{\mathrm{LQ}}^4}{(1-\beta)\left(m_{\mathrm{LQ}}^2 - m_t^2\right)^2 + \beta m_{\mathrm{LQ}}^4}, \tag{6.2}$$

with the mass of the top quark, m_t, and assuming the mass of the bottom quark to vanish. The dependence of the branching ratio on the LQ mass is visualised in Figure 6.1, assuming $\beta = 0.5$. As the figure indicates, the branching ratio significantly deviates from $\mathcal{B} = 0.5$ for masses below $m_{\mathrm{LQ}} = 500\,\mathrm{GeV}$ and asymptotically approaches $\mathcal{B} = 0.5$ towards higher masses.

Samples can be derived for all branching ratios through reweighting of the simulated events based on the LQ decays. For this, per-event weights are calculated based on the number $n_\ell \in [0, 2]$ of LQs decaying into charged leptons as

$$w\left(\mathcal{B}(LQ \to q_3\ell)\right) = \left(\frac{\mathcal{B}(LQ \to q_3\ell)}{\hat{\mathcal{B}}(LQ \to q_3\ell)}\right)^{n_\ell} \left(\frac{1 - \mathcal{B}(LQ \to q_3\ell)}{1 - \hat{\mathcal{B}}(LQ \to q_3\ell)}\right)^{2-n_\ell}, \tag{6.3}$$

with the signal-mass-dependent branching ratio the sample was produced with, $\hat{\mathcal{B}}(LQ \to q_3\ell)$, and the branching ratio it is reweighted to, $\mathcal{B}(LQ \to q_3\ell)$.
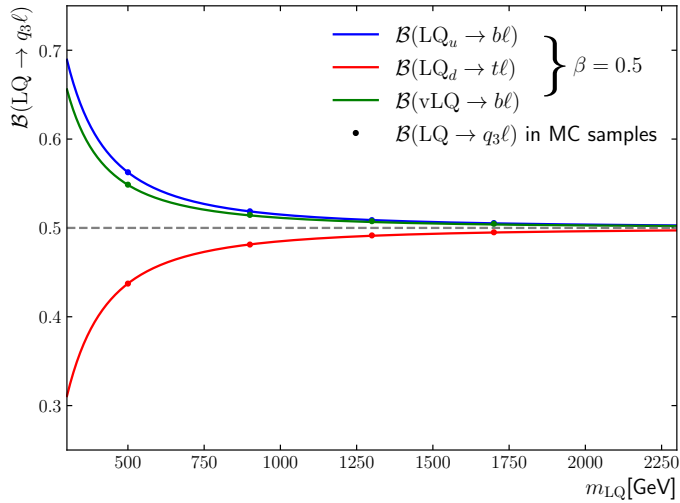
Figure 6.1: Branching ratios of up- and down-type scalar LQs and up-type vector LQs into charged leptons assuming $\beta = 0.5$, depending on the signal mass. Solid lines indicate branching ratios as calculated from the equations given in the text, while branching ratios obtained from MC simulated samples for signal masses of $500\,\mathrm{GeV}$, $900\,\mathrm{GeV}$, $1300\,\mathrm{GeV}$, and $1700\,\mathrm{GeV}$ are overlaid as circles.

Cross-sections for the pair production of scalar LQs were taken from the calculation of top-squark pair production, which has identical production modes. They are computed at approximate NNLO in QCD including a resummation of NNLL soft gluon terms [83–86]. Both, in the model used for the simulation of the events and in the calculation of the cross-sections, lepton t-channel contributions, which could lead to corrections at the percent level [178], are neglected. Relative uncertainties due to the choice of PDF and of renormalisation and factorisation scales amount to less than $7\,\%$ for a LQ mass of $300\,\mathrm{GeV}$ up to around $25\,\%$ for a LQ mass of $2000\,\mathrm{GeV}$. Separate signal samples are available for up- and down-type LQs decaying into electrons or muons, covering the LQ mass range from $300\,\mathrm{GeV}$ to $800\,\mathrm{GeV}$ and from $1600\,\mathrm{GeV}$ to $2000\,\mathrm{GeV}$ in steps of $100\,\mathrm{GeV}$. For the range from $800\,\mathrm{GeV}$ to $1600\,\mathrm{GeV}$ simulated samples are produced in steps of $50\,\mathrm{GeV}$ in order to increase the resolution around the expected sensitivity limit.

The model used for the simulation of pair-produced vector LQs is based on the one described in Ref. [89]. It includes an extension to allow for couplings of the LQs to electrons, which were prohibited in the original model. Here, a coupling strength of the vector LQ to quarks and leptons of $g_U = 3.0$ is chosen. The model allows for left- and right-handed couplings to fermions through the adjustable parameters β_L and β_R. Samples for this analysis have been generated with only left-handed couplings, i.e. equal coupling strength to charged leptons and neutrinos ($\beta = 0.5$). However, by applying event weights calculated from Equation (6.3), the same reweighting approach for MC simulated vector LQ events is used as described above for scalar LQs to probe different values of the branching ratio into charged leptons. As is evident

from Figure 6.1, the dependence of the branching ratio on the signal mass for a particular value of β, while qualitatively similar, is slightly different from the one for scalar LQs. It is given as

$$\mathcal{B}(\text{vLQ}_{\text{mix}} \rightarrow b\ell) = \frac{m_{\text{LQ}}^4 \left(\beta_{\text{L}}^2 + \beta_{\text{R}}^2 \right)}{\beta_{\text{L}}^2 \left(m_{\text{LQ}}^2 - m_t^2 \right) \left(m_{\text{LQ}}^2 - \frac{m_t^4}{2m_{\text{LQ}}^2} - \frac{m_t^2}{2} \right) + m_{\text{LQ}}^4 \left(\beta_{\text{L}}^2 + \beta_{\text{R}}^2 \right)}, \tag{6.4}$$

where different choices of β_{L}^2 and β_{R}^2 effectively result in different values of β.

The model allows for either the minimal ($\text{vLQ}_{\text{mix}}^{\text{min}}$) or the YM ($\text{vLQ}_{\text{mix}}^{\text{YM}}$) coupling scenario. Samples have been produced separately for both coupling scenarios and for muons and electrons with a mass spacing of $100\,\text{GeV}$ from $300\,\text{GeV}$ to $1400\,\text{GeV}$ and from $2300\,\text{GeV}$ to $2500\,\text{GeV}$. Between $1400\,\text{GeV}$ and $2300\,\text{GeV}$, a finer mass spacing of $50\,\text{GeV}$ has been used. Since no higher-order cross-section computations are available for this model, the cross-sections used in this analysis are computed at leading order by MG5_aMC@NLO instead, with relative uncertainties of around $25\,\%$ for a vector LQ mass of $300\,\text{GeV}$ up to $65\,\%$ for a vector LQ mass of $2500\,\text{GeV}$.

An overview over the different generator setups used for the various signal and background processes is given in Table 6.1. In all cases, pile-up was modelled by overlaying minimum-bias events produced with Pythia 8.186 [179] and the A3 tune [180] onto the hard scattering events.

Table 6.1: List of ME generator and the order of the strong coupling constant α_{S} in the perturbative calculation, PDF set, PS generator and tune for the different signal and background processes.

Process	ME Generator	ME order	PDF set	PS, UE, and hadronisation	Tune
$\text{LQ}_{\text{mix}}^{\text{u}}/\text{LQ}_{\text{mix}}^{\text{d}}$	MG5_aMC@NLO 2.6.0	NLO	NNPDF3.0NLO	Pythia 8.230	A14
$\text{vLQ}_{\text{mix}}^{\text{YM}}/\text{vLQ}_{\text{mix}}^{\text{min}}$	MG5_aMC@NLO 2.8.1	LO	NNPDF3.0NLO	Pythia 8.244	A14
$t\bar{t}$/single top	Powheg Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14
V+jets	Sherpa 2.2.1	MEPS@NLO	NNPDF3.0NNLO	Sherpa	Sherpa-internal
Diboson	Sherpa 2.2.1-2.2.2	MEPS@NLO	NNPDF3.0NNLO	Sherpa	Sherpa-internal
$t\bar{t} + V$	MG5_aMC@NLO 2.3.3	NLO	NNPDF3.0NLO	Pythia 8.210	A14
$t\bar{t} + H$	Powheg Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14

7 | Event Categorisation

In Chapter 2, LQs were introduced as a potential element of BSM theories, in particular to provide a possible explanation for the apparent anomalies observed in flavour physics. These LQs can come with a variety of different characteristics, such as their spin or their couplings to different generations of quarks and charged or uncharged leptons.

For the analysis described in this thesis, a special focus is put on scalar and vector LQs decaying into quarks of the third generation and charged leptons or neutrinos of the first or second generation. In particular, the search targets a final state containing a single charged lepton, which is expected to be highly energetic. It is therefore primarily sensitive to events with one LQ decaying into an electron or a muon and the other LQ decaying into the respective neutrino. However, in the case of up-type LQs, single-lepton final states can also be the result of both LQs decaying into a top quark and a neutrino, if one top-quark decays hadronically and the other one leptonically. Necessarily, the chosen final state therefore in general coincides with large missing transverse momenta due to at least one neutrino stemming directly from a LQ decay. Events of interest are further characterised by containing exactly one hadronically decaying top quark and one additional b-jet.

Based on these properties, a baseline event selection is defined in Section 7.1. Building on the baseline event selection, the analysis strategy discussed in detail in the following sections relies on several orthogonal regions of phase space defined by placing a number of selection criteria on various observables. A kinematic reweighting procedure of $t\bar{t}$ and single top-quark events, referred to as top reweighting, is derived in a dedicated region to combat modelling issues associated with the corresponding processes. The reweighting procedure is described in Section 7.2.

After applying the reweighting factors, the separation between signal and background events is studied in Section 7.3. Several discriminating variables are examined, and artificial NNs are trained for various signal hypotheses in a dedicated training region. Based on the output of the NNs, the training region is further split into a signal-enriched signal region (SR) and a region mainly dominated by $t\bar{t}$, which is utilised as a control region (CR) to constrain the normalisation of the $t\bar{t}$ background. An additional CR for single top-quark processes is defined in Section 7.4, as is a CR for W+jets events.

7.1 Baseline Event Selection

Based on the targeted signal event topology, criteria for a baseline event selection are defined, constituting the foundation for subsequent selections described in later sections. While the baseline selection, a summary of which is given in Table 7.1, only provides relatively loose constraints, it can already be used to gain first insights into the data-MC modelling in the phase space of interest.

Initial selection requirements consist of quality criteria ensuring all sub-detectors to be fully functional. Events are discarded in case of noise bursts in the calorimeter systems. Furthermore, events containing a jet failing loose jet cleaning criteria are removed as well, as are those with a poorly measured muon.

In view of the expectation of large missing transverse momenta, only events passing the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are considered in the following. An additional requirement on $E_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$ is applied, at which point the triggers have reached full efficiency. As the focus of this analysis is put on single-lepton final states, events have to contain exactly one signal electron or muon with $p_{\mathrm{T}}(\ell) \geq 30\,\mathrm{GeV}$. Events containing additional baseline leptons are vetoed, as are events with less than four reconstructed small-R jets, because any signal event with exactly one lepton in the final state necessarily contains four jets. At least one of the small-R jets has to be b-tagged. Lastly, contributions from events with large missing transverse momentum caused by mismeasurements of leptons or jets are suppressed by requirements on the transverse mass of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ of $m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) > 30\,\mathrm{GeV}$ and on the azimuthal distance between the two leading-p_{T} jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ of $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_{1,2})| > 0.4$, respectively. A Feynman diagram motivating the central baseline event selection criteria using the example of up-type LQs is shown in Figure 7.1. The same arguments also hold for down-type LQs.

The composition of SM background processes after the baseline event selection is shown in Figure 7.2a. With above 70 %, major contributions stem primarily from $t\bar{t}$. Additionally, significant numbers of background events can be attributed to W+jets and single top-quark processes. Further contributions from $t\bar{t} + V$, $t\bar{t} + H$, diboson, and Z+jets processes are

Table 7.1: Overview of the baseline event selection criteria.

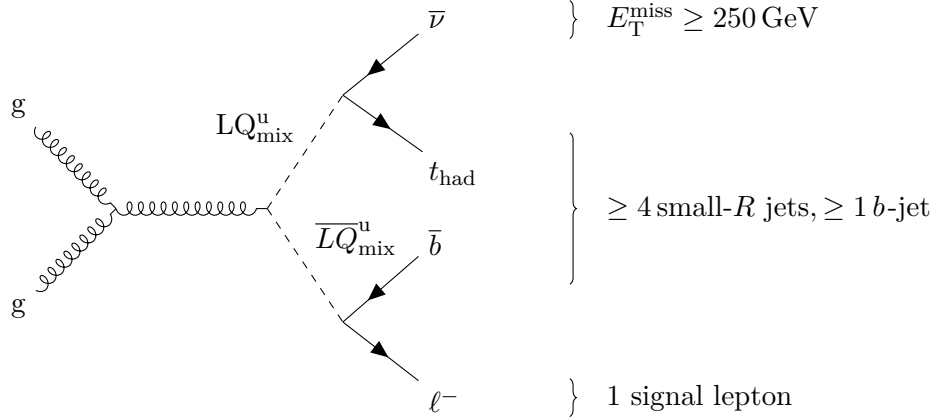| Baseline Event Selection |
| :---: |
| Event Cleaning Criteria |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers |
| $= 1$ signal lepton |
| veto on additional baseline leptons |
| $E_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$ |
| ≥ 4 small-R jets |
| ≥ 1 b-jet |
| $m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) > 30\,\mathrm{GeV}$ |
| $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_{1,2})| > 0.4$ |

Figure 7.1: Exemplary Feynman diagram of pair-produced up-type LQs, with one LQ decaying into a charged lepton and a bottom quark, the other into a neutrino and a hadronic top quark, motivating the central baseline event selection criteria as indicated by the braces.

subsumed in the category *others*. Contributions from $t\bar{t}$ events with both top quarks decaying hadronically are found to be negligble, indicating that events containing objects incorrectly identified as a charged lepton do generally not pass the baseline event selection. Kinematic distributions are shown in Figure 7.2b and Figure 7.2c for m_{eff} and $p_{\mathrm{T}}(\ell)$, respectively. Deviations between MC expectation and observed data are evident, particularly towards higher transverse momenta, with the ratio of data and prediction gradually decreasing below unity.
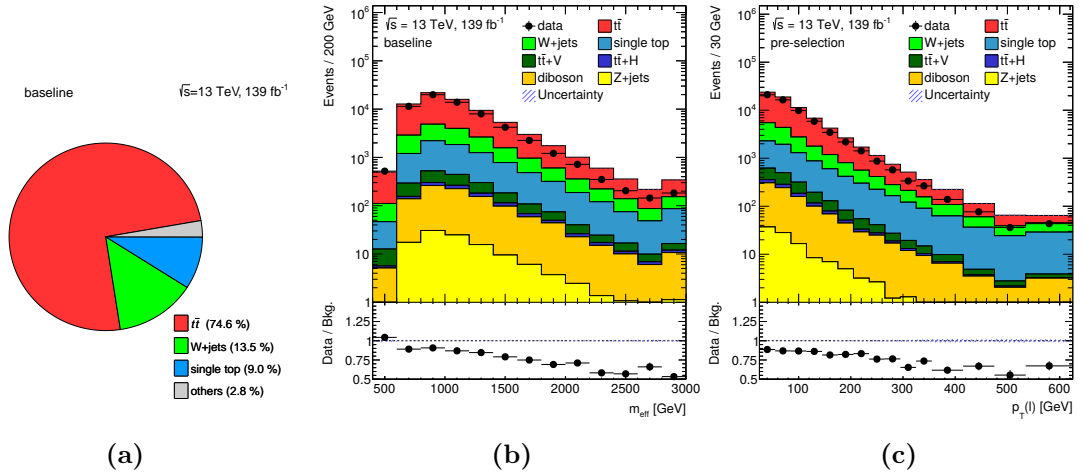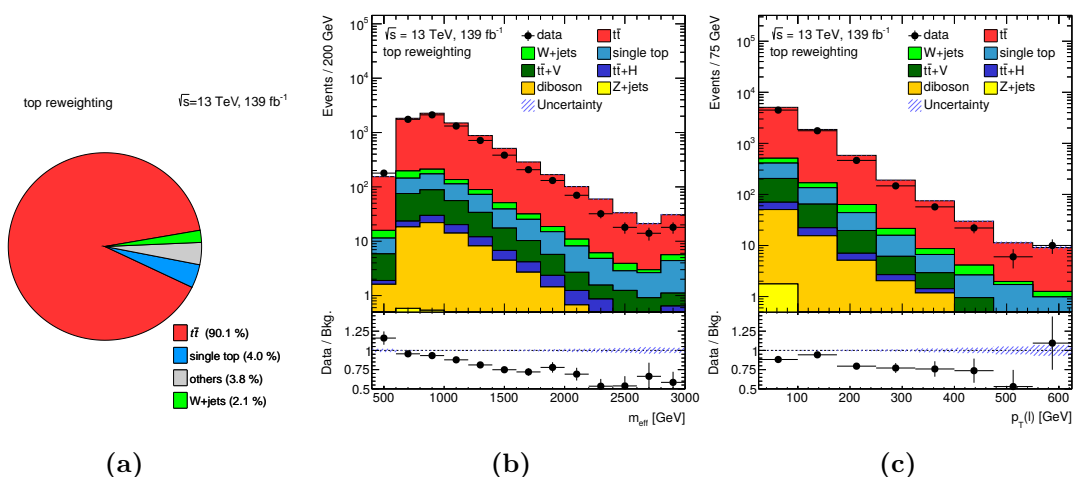

Figure 7.2: (a) Contributions from different SM processes after the baseline event selection. Contributions from $t\bar{t}+V$, $t\bar{t}+H$, diboson, and Z+jets production are summarised as *others*. Distributions of (b) the effective mass and of (c) the transverse momentum of the lepton. The uncertainty bands contains MC statistical uncertainties only.

7.2 Kinematic Top Reweighting

Differences between data and MC simulation after the baseline event selection are caused by mismodelling at high top-quark momenta, which is a known effect also observed in other analyses. The underlying causes are missing higher-order QCD corrections not accounted for in simulation. In order to compensate for the mismodelling, a reweighting scheme depending on kinematic properties of the affected events is developed. Reweighting factors in the form of event weights to be applied on $t\bar{t}$ and single top-quark MC simulated events are derived as a function of two observables sensitive to the missing corrections, namely the jet multiplicity, n_j, and m_{eff}.

For the derivation of the reweighting factors, a dedicated top reweighting region, mainly enriched in $t\bar{t}$, is defined. Due to strong interference effects between $t\bar{t}$ and tW and because the underlying MC generator setups used for the production of the MC simulated samples are equivalent, the same reweighting factors are applied on $t\bar{t}$ and single top-quark events. To ensure kinematic similarity to the training region, the top reweighting region is defined by requiring at least one b-jet and a transverse mass of the lepton and $E_{\text{T}}^{\text{miss}}$ of $m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}}) \geq 120\,\text{GeV}$. Orthogonality to training and control regions is ensured with an asymmetric transverse mass requirement of $am_{\text{T2}} < 200\,\text{GeV}$. As am_{T2} is computed from two b-jets despite only one b-jet being required in the selection, the two jets with the highest b-tagging score are used for the calculation here and in the following sections. With these criteria, a region of phase space is selected that is very pure in $t\bar{t}$ events. As shown in Figure 7.3a, the purity is about 90 %, depending on the jet multiplicity, with single top-quark events adding another 4 %. Other SM processes therefore account for only approximately 7 % of events in the top reweighting region.

(a)   **(b)**   **(c)**

Figure 7.3: (a) Contributions from different SM processes to the top reweighting region. Contributions from $t\bar{t}+V$, $t\bar{t}+H$, diboson, and Z+jets production are summarised as *others*. Distributions of (b) the effective mass and of (c) the transverse momentum of the lepton. The uncertainty band contains MC statistical uncertainties only.

In addition, data and MC distributions of m_{eff} and $p_{\text{T}}(\ell)$ are shown in Figure 7.3b and Figure 7.3c, respectively, before applying reweighting factors. Both distributions qualitatively exhibit the same mismodelling as was observed on baseline event selection level.

In total, four separate bins of the jet multiplicity are considered, containing four, five, six, or at least seven jets, respectively. For jet multiplicity bin i, the reweighting factors required to scale the $t\bar{t}$ and single top-quark event yields, n_{top}^i, to those expected from the difference between observed data, n_{data}^i, and the sum of all other SM background processes, n_{other}^i, are then calculated as

$$f_i(m_{\text{eff}}) = \frac{n_{\text{data}}^i(m_{\text{eff}}) - n_{\text{other}}^i(m_{\text{eff}})}{n_{\text{top}}^i(m_{\text{eff}})}. \tag{7.1}$$

The resulting reweighting factors are shown in Figure 7.4. The error bars contain statistical uncertainties and an additional uncertainty on the MC estimate of the non-top backgrounds of 50 %. Due to the purity of the top reweighting region, the latter uncertainty component is negligible and does not noticeably affect the total size of the error bars.
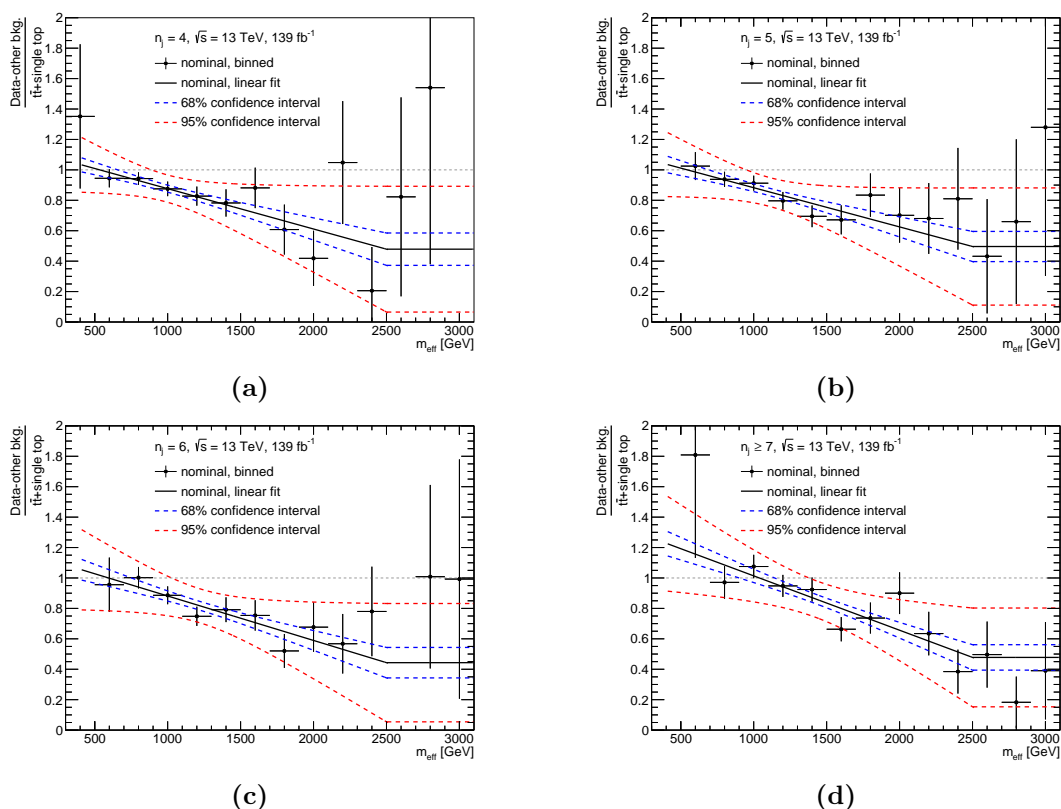
Figure 7.4: Kinematic top reweighting factors as a function of m_{eff} in four bins of jet multiplicity, with (a) four jets, (b) five jets, (c) six jets, and (d) at least seven jets. The result of the linear fit is shown as the solid black line, the 68 % and 95 % confidence intervals obtained by propagating uncertainties on the fit parameters as a dashed blue and red line, respectively.

The reweighting factors are evaluated in bins of m_{eff} and generally decrease towards higher energies. They are fitted with a first-order polynomial for effective masses up to $2500\,\text{GeV}$, with the factors being assumed to be constant from that point onwards to avoid artificially downweighting events too much. The uncertainty on the reweighting procedure in each jet multiplicity bin, σ_i, is evaluated by propagating the uncertainties on the fit parameters. Defining the linear fit function as $f(m_{\text{eff}}) = a \cdot m_{\text{eff}} + b$, error propagation yields

$$\sigma_i^2 = m_{\text{eff}}^2 \sigma_{a,i}^2 + \sigma_b^2 + 2m_{\text{eff}}\rho_{ab,i}\sigma_{a,i}\sigma_{b,i} \tag{7.2}$$

with the uncertainties on the slope, σ_a, and on the offset, σ_b, and the correlation coefficient ρ_{ab} between slope and offset. Potential non-linearities in the dependence of the reweighting factors on m_{eff} are accounted for by using the $95\,\%$ confidence interval as uncertainty. This choice for the systematic uncertainty covers the majority of bin-by-bin fluctuations, as demonstrated in Figure 7.4.

After deriving the factors, they are applied on all $t\bar{t}$ and single top-quark events. Two exemplary distributions displaying the data-MC agreement in the top reweighting region post reweighting are shown in Figure 7.5, which, like all figures displaying data-MC ratios in the following, includes systematic uncertainties as described in Chapter 8. The data-MC ratio is close to unity across the whole phase space and well within the statistical and systematic uncertainties. For comparison, the total background sum before reweighting is included as a dashed black line, highlighting the increased effects of the reweighting procedure on highly energetic events.

To ensure that the reweighting procedure derived in this dedicated region of phase space is applicable also in the other regions considered in the analysis and indeed fixes the mismodelling, distributions of the effective mass and of the transverse momentum of the lepton after applying the top reweighting to the events passing the baseline event selection are shown in Figure 7.6. Here, too, the agreement between observation and expectation is markedly improved. The top reweighting factors are applied in all distributions shown in the remaining sections of this thesis, unless explicitly stated otherwise.
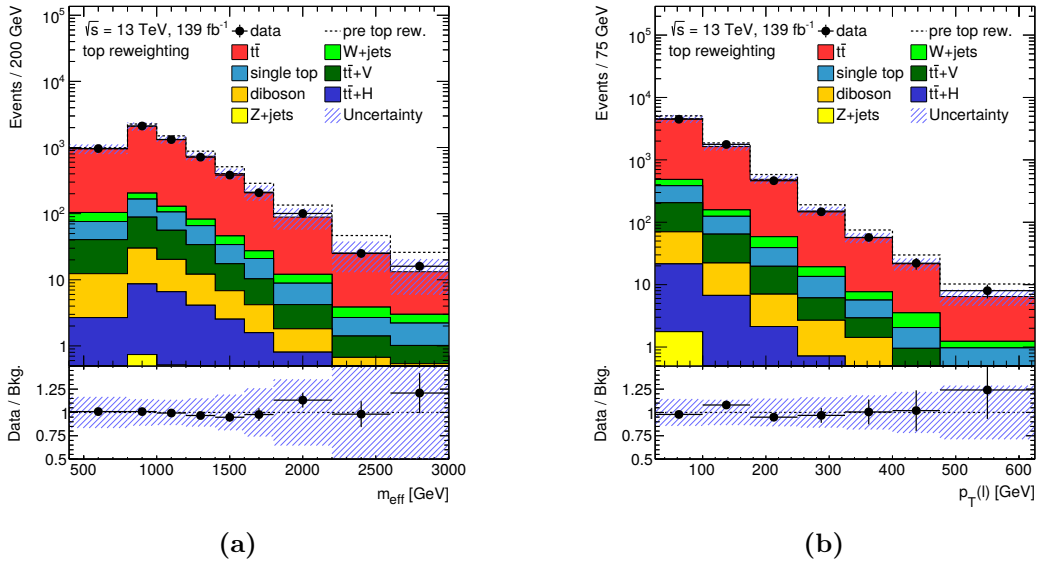
Figure 7.5: Distributions of (a) m_{eff} and (b) $p_{\text{T}}(\ell)$ in the top reweighting region after applying the kinematic top reweighting. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the kinematic top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.
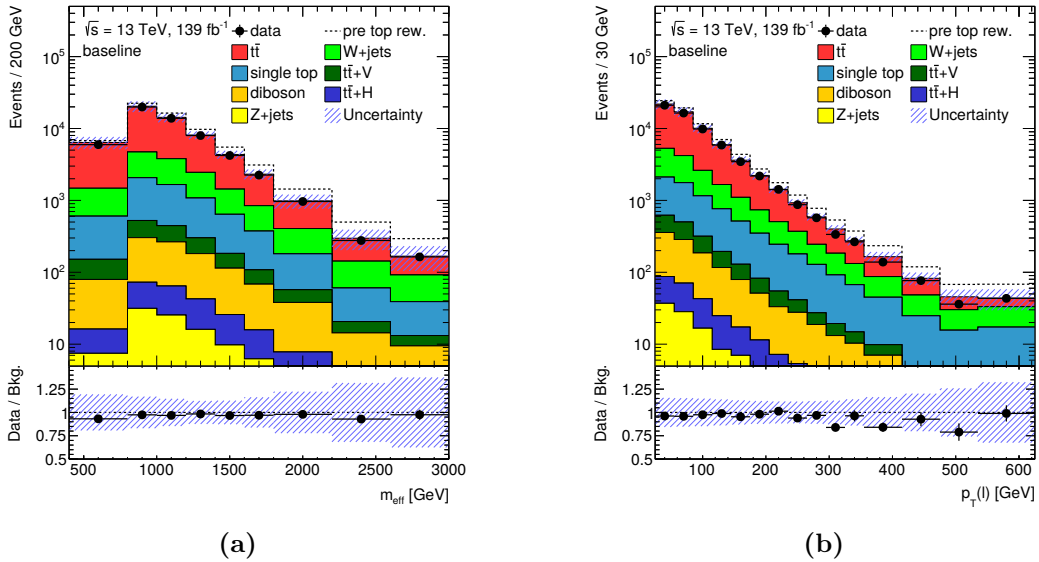

Figure 7.6: Distributions of (a) m_{eff} and (b) $p_{\text{T}}(\ell)$ after the baseline event selection after applying the kinematic top reweighting. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the kinematic top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.

7.3 Separation Between Signal and Background Processes

After ensuring the observed data to be modelled well by the MC prediction following the baseline event selection, further requirements are applied to reduce contributions from background events, resulting in the definition of the training region. The selection criteria for this region are chosen to be relatively loose in order to retain high signal selection efficiencies for a variety of signal hypotheses across a wide range of LQ masses. As such, only requirements on the transverse mass of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ of $m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) > 120\,\mathrm{GeV}$ and on the asymmetric transverse mass of $am_{\mathrm{T2}} > 200\,\mathrm{GeV}$ are imposed on top of the baseline event selection, with the former criterion targeting W+jets background and the latter suppressing contributions from $t\bar{t}$ events. The background composition after applying this selection is shown in Figure 7.7. Main contributions stem from $t\bar{t}$, single top-quark, and W+jets events, with $t\bar{t} + V$ processes following in fourth position.

The selection efficiencies for signal events primarily depend on the mass of the LQs. A visualisation is given in Figure 7.8, showing the product of the signal acceptance and the selection efficiency for six representative signal hypotheses as a function of m_{LQ}. As expected, all models exhibit selection efficiencies that rapidly increase with the LQ mass, before reaching plateaus around $m_{\mathrm{LQ}} = 1200\,\mathrm{GeV}$. After the plateau, a decrease can be observed. The reason for the decline are kinematic dependences of the reconstruction of certain objects. This was discussed in the context of the b-tagging working point in Section 5.5. An additional factor is the switch to the High-p_{T} muon identification criterion for muons with transverse momenta above $800\,\mathrm{GeV}$, which is more relevant the higher the LQ mass is. Discrepancies between up-type scalar and vector LQs, the former reaching a maximum selection efficiency of $17\,\%$ and the latter of $20\,\%$, are caused by differences in kinematics as discussed in Section 2.2. For low LQ masses, differences can also be observed between vector LQs in the YM coupling scenario and vector LQs in the minimal coupling scenario. These, too, were previously examined in Section 2.2, where deviations in the transverse momentum of the neutrino were seen. These differences propagate to the reconstructed missing transverse momentum, and, since
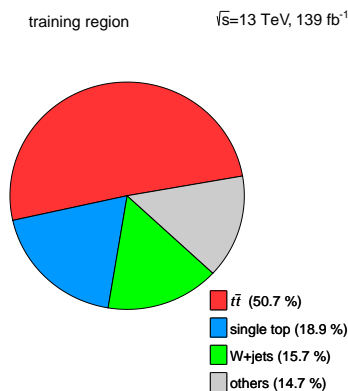

Figure 7.7: Pie chart visualising the contributions from the different background processes in the training region. Minor contributions from $t\bar{t}+V$, $t\bar{t}+H$, diboson, and Z+jets processes are summarised as *others*.
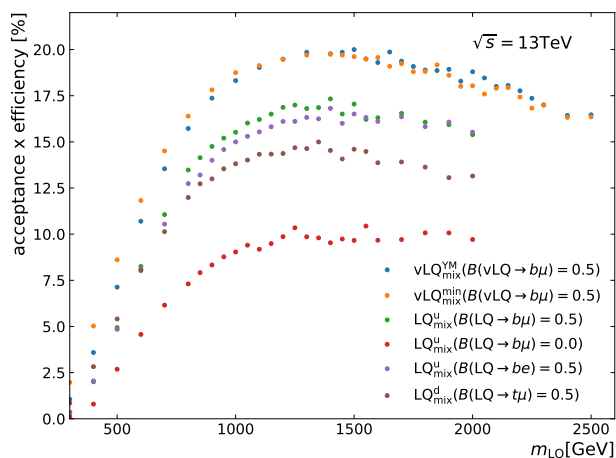
Figure 7.8: Signal Acceptance times efficiency for six signal models as a function of the LQ mass.

for low LQ masses only events in the high-$E_\mathrm{T}^\mathrm{miss}$ tail are selected, lead to the discrepancies in the efficiencies. Differences between scalar up-type and down-type LQs arise due to the additional contributions from events with both up-type LQs decaying into a top quark and a neutrino. Finally, selection efficiencies are significantly lower and reach their maximum at approximately 10 % when assuming both LQs to exclusively decay into top quarks and neutrinos, because only events with a leptonically decaying top quark are selected.

7.3.1 Discriminating Variables

Building on the findings from Section 2.1, various observables distinguishing between signal and background processes are considered to establish potential avenues towards a more refined event selection allowing for higher signal purity. A representative subset of discriminating variables is discussed in this section, containing distributions of kinematic observables in the training region that provide good separation power between signal and background processes. Further distributions can be found in Appendix A.1.

In Figure 7.9, the effective mass, the transverse momentum of the lepton, and the invariant mass of the leading-p_T b-jet and the lepton are presented. The total SM background prediction is compared to data, showing agreement within the statistical and systematic uncertainties. In particular, the ratio of data over the MC prediction is observed to be near unity across the whole range of the effective mass, indicating the effectiveness of the top reweighting procedure also in the training region. Similarly, the distribution of the transverse momentum of the lepton does not demonstrates large deviations from a constant data-MC ratio. Sizeable shape effects only arise in the distribution of the invariant mass of the leading-p_T b-jet and the lepton, which is among the least well-modelled variables considered in the analysis. However, even in this case, data are generally within the uncertainy band.

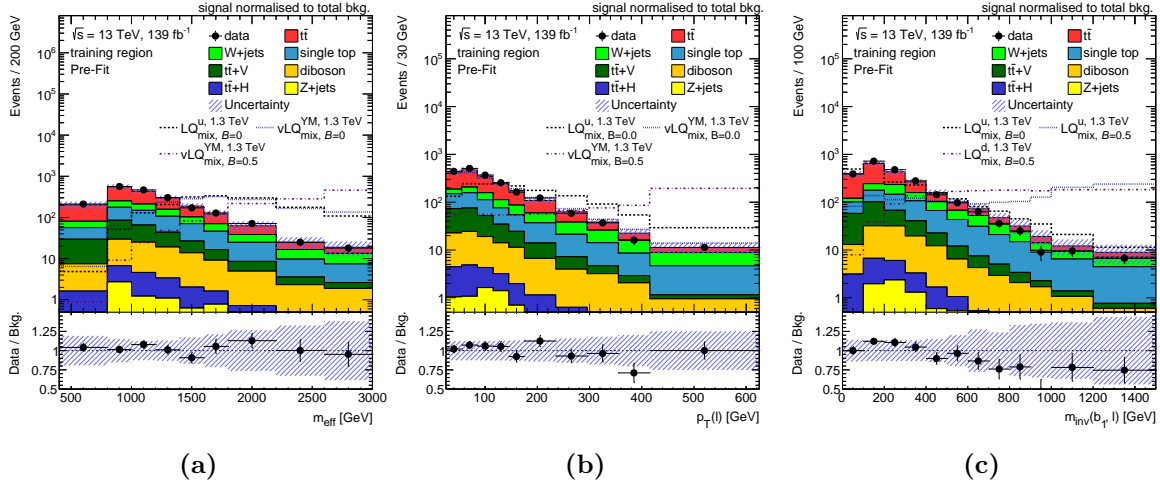**Figure 7.9:** Distributions of exemplary discriminating variables, namely (a) m_{eff}, (b) $p_{\mathrm{T}}(\ell)$, and (c) $m_{\mathrm{inv}}(b_1, \ell)$. The plots show the SM background expectation in comparison with the observed data. The hatched bands include statistical and systematic uncertainties. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow. Three exemplary signal processes, normalised to the background expectation, are overlaid as dashed lines.

In addition to providing the means to ensure good modelling of the data in the phase space of interest, the figures include information about the shape of three exemplary signal distributions with a LQ mass of $m_{\mathrm{LQ}} = 1300\,\mathrm{GeV}$ as well. These are overlaid on top of the data-MC comparison, normalised to the total yields of the background processes in order to highlight differences in their shapes. It becomes evident that signal events tend to contain more energy than background events, which is required to produce pairs of massive bosons. This effect is particularly striking for the effective mass, shown in Figure 7.9a, which in essence reflects the total energy released in the hard scattering process. The distribution reaches its maximum below about $3.0\,\mathrm{TeV}$ in the case of signal with a branching ratio into charged leptons of $\mathcal{B}(\mathrm{LQ} \to b\ell) = 0.5$, corresponding to approximately twice the LQ mass. The slight shift towards higher effective masses is due to the momentum carried by the LQs themselves. For events with both LQs decaying into a top quark and a neutrino, the behaviour is different as the maximum of the distribution is already observed around $1.7\,\mathrm{TeV}$. This is because the missing transverse momentum in such events does not represent only one neutrino, but rather three, with two of them originating directly from the LQ decays and one from a leptonic top-quark decay. The behaviour is the same for scalar and vector LQs.

Differences between scalar and vector LQs emerge, however, in the distributions of the transverse momentum of the lepton, specifically in events with both LQs decaying into top quarks and neutrinos. Displayed in Figure 7.9b and previously discussed in Section 2.1, these deviations at low transverse momenta are caused by spin correlations propagating through the complete decay chain. As a result, the distribution is very similar in shape to the total SM background for vector LQs with $\mathcal{B}(\mathrm{vLQ}_{\mathrm{mix}} \to b\ell) = 0.0$, but shifted towards higher transverse

momenta in the case of scalar LQs with $\mathcal{B}(\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to b\ell) = 0.0$. On the other hand, if the charged lepton originates directly from the LQ decay, its transverse momentum is considerably larger, therefore providing significantly more discriminating power for such events.

Shape differences between a signal hypothesis with $\mathcal{B}(\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to b\ell) = 0.0$ and a signal hypothesis with $\mathcal{B}(\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to b\ell) = 0.5$ also arise in the invariant mass of the leading-p_{T} b-jet and the lepton, as evidenced by Figure 7.9c. This observable is particularly useful for events with one LQ decaying into a bottom quark and a charged lepton, as the variable allows for the reconstruction of its mass by assuming the relevant bottom quark to correspond to the leading-p_{T} b-jet, thereby providing strong separation between signal and background processes. When neither LQ decays directly into a charged lepton, however, the observable loses its effectiveness, and the signal distributions become very similar to the background, because the missing transverse momentum is not taken into account and because the association between b-jet and lepton is ambiguous. Moreover, Figure 7.9c indicates that, due to the highly energetic nature of the signal event, discrimination power is also provided for down-type LQs, even though the observable cannot be used to directly reconstruct the LQ mass in this case. In Figure 7.10, distributions of the transverse mass of the lepton and the missing transverse momentum, of the asymmetric transverse mass, and of the large-R jet multiplicity are shown.
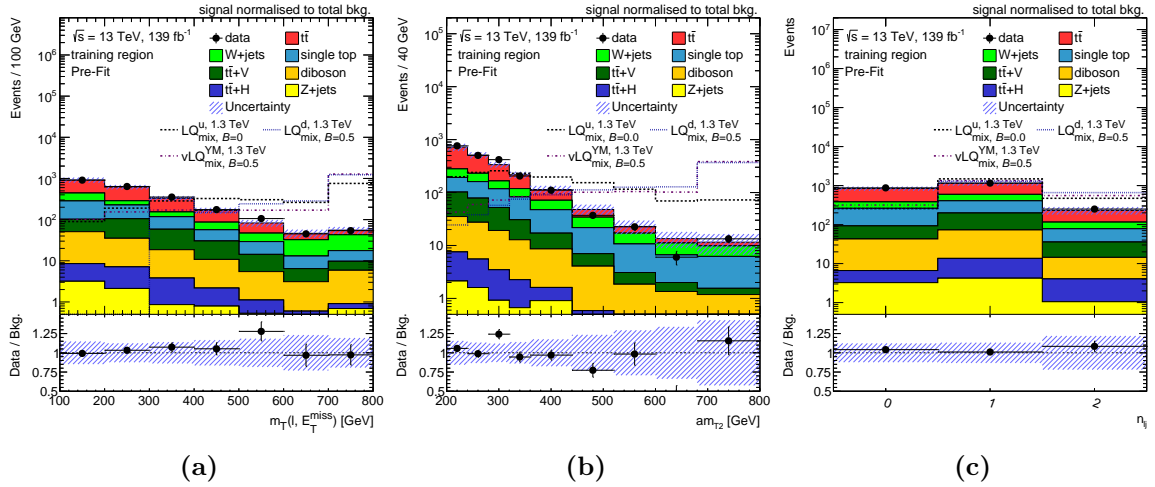
Figure 7.10: Distributions of exemplary discriminating variables, namely (a) $m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$, (b) am_{T2}, and (c) the large-R jet multiplicity. The plots show the SM background expectation in comparison with the observed data. The hatched bands include statistical and systematic uncertainties. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow. Three exemplary signal processes, normalised to the background expectation, are overlaid as dashed lines.

Again, the observed data agree well with the MC prediction within uncertainties. Only minor differences between the different signal hypotheses are visible in the distributions of the transverse mass, with the discrimination between signal and background processes being primarily due to the higher energy in signal events. Events with both LQs decaying into a top-quark and a neutrino tend to slightly lower values, due to the lepton originating from the

top-quark decay and due to the missing transverse momentum capturing the sum of three neutrinos. Similarly, such signal events are more background-like in am_{T2} than events in which the charged lepton and a bottom quarks originate directly from LQ decays. Finally, all signal hypotheses under consideration tend to higher multiplicities of reconstructed large-R jets than the background processes.

7.3.2 Neural Networks

The problem space under study in this analysis is characterised by a high dimensionality. Not only is information discriminating between signal and background processes contained in many different observables, but, depending on the LQ mass and its decay topology, signal characteristics can differ considerably. To allow for exploitation of as much of the available discriminating information as possible and to efficiently cover a wide range of signal parameters, a multivariate approach in the form of NNs is employed. Utilising a NN, a set of input variables is mapped to the NN output, which can be used to distinguish between signal and background processes. The dimensionality is thus reduced substantially. Moreover, an important advantage of NNs is their capability of incorporating information about correlations between variables, as opposed to a purely sequential cut-based analysis. In this thesis, three-layer feed-forward NNs implemented in the NeuroBayes package [181, 182] are used. The general structure of such a NN is shown in Figure 7.11. It consists of an input layer containing a set of $n + 1$ nodes, with one node, v_i, for each of the n input variables and one additional bias node, the latter increasing the network's flexibility.
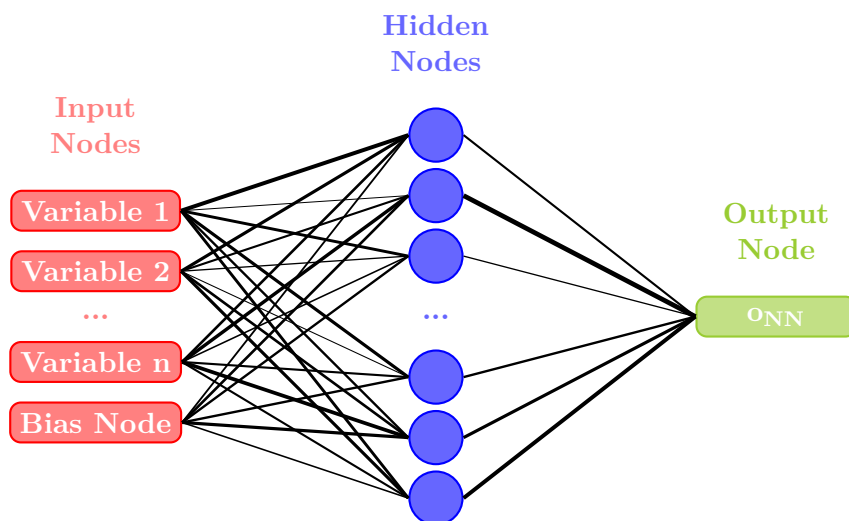
Figure 7.11: Schematic depiction of the structure of a three-layer feed-forward neural network with n input variables and a bias node in the input layer, a certain number of nodes in the hidden layer, and one node in the output layer, yielding the final discriminant, NN_{out}. Connections between nodes of different layers are weighted as represented in the figure by the widths of the corresponding edges.

Referred to as the hidden layer, the second layer consists of an a-priori arbitrary number of nodes, m. The network's performance depends only slightly on the exact number of hidden nodes, which is chosen as $m = 15$ in this case. Very large numbers of nodes become computationally prohibitive, whereas very small numbers can limit the network's capabilities. The value of each hidden node h_j is calculated as the weighted sum of the input nodes v_i, with the weights w_{ij} represented by edges connecting all input nodes with all hidden nodes. Similarly, the output node is connected to every hidden node via edges with weight w_j. In each hidden node and in the output node, the weighted sums are transformed using the sigmoid function as activation function, which is defined as

$$S(x) = \frac{2}{1 + e^{-x}} - 1. \tag{7.3}$$

The sigmoid function is symmetric and transforms the interval $(-\infty, +\infty)$ to the interval $[-1, +1]$, with an approximately linear response around zero. The NN output is thus given by

$$o_{\mathrm{NN}} = S\left(\sum_{j=1}^{m} w_j \cdot S\left(\sum_{i=1}^{n} w_{ij} x_i\right)\right). \tag{7.4}$$

The final discriminant is obtained by transforming o_{NN} to the range $[0, 1]$ and is referred to as NN_{out} in the following.

The weights of the connections between nodes of different layers are determined in a training procedure, utilising a training dataset consisting of MC simulated signal and background events. For a simulated event k, the corresponding target value of the NN, denoted as t_k, is known. It is $t_k = 1$ for signal events and $t_k = -1$ for background events. Using this, the classification accuracy of the NN can be quantified by means of an entropy loss function, which is to be minimised by iteratively adapting the weights w_{ij} and w_j through backpropagation of classification errors. Here, the loss function is defined as

$$E = \sum_{k} \log\left(\frac{1 + t_k \cdot o_k + \epsilon}{2}\right), \tag{7.5}$$

where o_k is the NN output for a given event and ϵ is a regularisation constant, included to avoid numerical instabilities for events with $t_k \approx -o_k$. With the choice of a logarithmic loss function, particularly large differences between true category and predicted one are penalised. The training procedure is stopped once the loss function is below a certain threshold.

As multivariate methods contain many adjustable parameters, they can be prone to overtraining issues, i.e. models being adapted very well to the observations they are trained on, but failing to generalise to additional data points. To avoid these effects, a Bayesian regularisation scheme is employed, which, by way of applying Bayesian inference on the weights,

downweights and thus effectively eliminates irrelevant connections between nodes. In practice, insignificant weights are set to zero, reducing the model's degrees of freedom and thereby avoiding potential overtraining.

The input dataset consists of MC simulated signal and SM background events, which are weighted such that the ratio of signal and background is unity. This is done to avoid problems when the nominal signal yields are very small, resulting in the entropy loss being dominated by fluctuations in the background distributions. Even though Bayesian NNs as implemented in NeuroBayes are not prone to overtraining, the input dataset is split into a training set and a validation set for further cross-checks, with 80 % of events belonging to the former and the remaining 20 % to the latter category. For the training procedure only the training dataset is utilised, while for the validation set the entropy loss is calculated in each iteration. As long as the NN performance has been improved with respect to the previous iteration, the test error is expected to decrease. Once an increase above a certain threshold is detected, however, the training process is stopped and the final NN is determined from the iteration before the increase. This procedure is illustrated in Figure 7.12 for one exemplary NN training. Both, training and test error, are generally decreasing, thereby indicating that no overtraining is occuring.
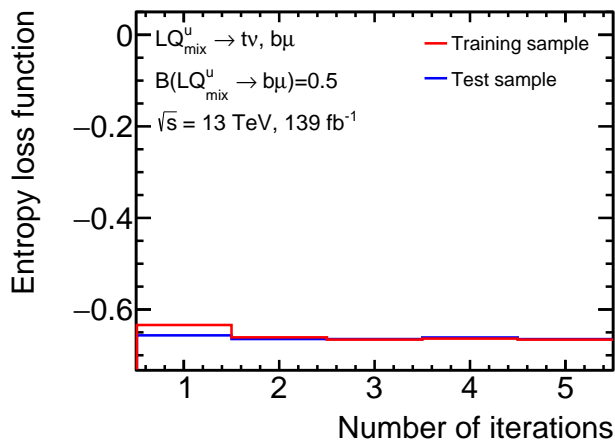
Figure 7.12: Entropy loss in training and test dataset as a function of the iteration for an exemplary NN training. The loss in the training dataset is shown in red, the loss in the test dataset in blue.

Before including the input variables in the training procedure, a preprocessing is applied to facilitate a fast and robust training. First, signal and background are transformed to uniform distributions, containing the same number of events in each bin. Using these equalised distributions, the signal purity is calculated and fitted with a spline function in order to suppress statistical fluctuations. Finally, the numerical stability is further improved by converting the spline function output into a distribution with a mean of zero and a standard deviation of one. In addition, the preprocessing procedure also allows ranking the input variables by their separation power. To obtain this ranking, first, the correlation of each transformed input variable to the target function is calculated. The target function assumes a value of one for signal events and of zero for background events. In the next step, the loss in total correlation to target when removing an observable from the input is computed. The variable resulting

in the smallest loss is discarded, and the procedure is repeated iteratively until only the most significant variable remains. In the interest of limiting the number of input variables, only those observables providing significant additional information are considered in the training by requiring their added significance to exceed a certain threshold. For the purpose of the analysis at hand, the threshold is chosen as 8σ in order to retain approximately 10 variables in each training. Including more variables in the training has been found to not increase the performance. As summarised in Table 7.2, 18 individual NNs are trained in total. In each of the trainings, several signal masses are combined to simplify the analysis, but also to enhance the separation performance when the NN is applied on signal masses it was not trained on. For this purpose, all trainings are based on MC samples of simulated signal events with LQ masses of $m_{\mathrm{LQ}} = 500\,\mathrm{GeV}, 900\,\mathrm{GeV}$, and $1300\,\mathrm{GeV}$. For scalar LQs, this set of masses is sufficient to ensure good performance across the whole range of LQ masses studied in this analysis, without significant losses in sensitivity. An additional higher mass point is omitted from the training, as its inclusion would lead to degraded sensitivity for low masses. In trainings involving vector LQs on the other hand, the $m_{\mathrm{LQ}} = 1700\,\mathrm{GeV}$ mass point is included as well to account for significantly higher production cross-sections and therefore higher expected exclusion sensitivity. All signal samples included in the trainings are scaled to the same cross-sections, such that, bar acceptance effects, all mass points have the same share. Separate NNs for LQ decays into muons or into electrons are trained for scalar up- and down-type LQs and for vector LQs in the YM coupling scenario. As discussed in Chapter 2, vector LQs in the minimal coupling scenario show kinematically largely identical behaviour to those in the YM scenario, apart from very low masses below around $500\,\mathrm{GeV}$. Dedicated NNs for the minimal coupling scenario are therefore not necessary.

NNs for down-type LQs are solely trained for $\mathcal{B} = 0.5$, because only events with one LQ decaying into a charged lepton and the other decaying into a neutrino contribute to the final state of interest. In the case of up-type LQs, events with both LQs decaying into neutrinos are relevant, too, because the charged lepton can be produced in the top-quark decay. However, these decays have different kinematic behaviour than $\mathrm{LQLQ} \to t\nu b\ell$, as

Table 7.2: Summary of signal hypotheses for which a NN is trained. For scalar LQs, MC samples with $m_{\mathrm{LQ}} = 500\,\mathrm{GeV}, 900\,\mathrm{GeV}$, and $1300\,\mathrm{GeV}$ are combined. For vector LQs, $m_{\mathrm{LQ}} = 1700\,\mathrm{GeV}$ is included in addition.

LQ Type	$\mathcal{B}(\mathrm{LQ} \to q\ell)$
$\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to t\nu/b\mu$	0.0, 0.25, 0.5, 0.9
$\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to t\nu/be$	0.0, 0.25, 0.5, 0.9
$\mathrm{LQ}^{\mathrm{d}}_{\mathrm{mix}} \to b\nu/t\mu$	0.5
$\mathrm{LQ}^{\mathrm{d}}_{\mathrm{mix}} \to b\nu/te$	0.5
$\mathrm{vLQ}^{\mathrm{YM}}_{\mathrm{mix}} \to t\nu/b\mu$	0.0, 0.25, 0.5, 0.9
$\mathrm{vLQ}^{\mathrm{YM}}_{\mathrm{mix}} \to t\nu/be$	0.0, 0.25, 0.5, 0.9

discussed in Section 7.3.1. To capture these differences, four NNs are trained for each up-type LQ hypothesis, with $\mathcal{B} = 0.0, 0.25, 0.5, 0.9$. For each of the 18 NNs, the preprocessing and training procedures are performed. The set of input variables considered in the preprocessing consists of a total of 15 observables, as listed in Table 7.3. The list includes kinematic and angular variables such as transverse momenta of various reconstructed particles or azimuthal angular separations between different objects. Furthermore, the flavour of the signal lepton is used as an input. This allows for an efficient separation of the signal process from backgrounds containing a lepton of the opposite flavour. Instead of applying a stringent requirement on the lepton flavour, however, the decision is left to the NN in order to also capture events with both LQs decaying into top quarks, with one of the top quarks decaying hadronically and the other leptonically. These top-quark decays are independent of the signal hypothesis and can therefore result in an opposite-flavour lepton as well. Contributions from such events are primarily relevant for up-type LQs, particularly for low values of $\mathcal{B}$.

Table 7.3: Input variables to the NN training, approximately sorted in descending discriminating power between signal and background. The order is not strict as there is some dependence on the signal model and $\mathcal{B}$.

Variable	Description
$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$	transverse mass of lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$
m_{eff}	scalar sum of the transverse momenta of leptons, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$
lepton flavour	flavour of the signal lepton
$p_{\mathrm{T}}(\ell)$	transverse momentum of the lepton
$m_{\mathrm{inv}}(b_1, \ell)$	invariant mass of leading-p_{T} b-jet and lepton
n_{lj}	reclustered large-R jet multiplicity
am_{T2}	asymmetric transverse mass
$E_{\mathrm{T}}^{\mathrm{miss}}$ significance	measure for the compatibility of the observed $E_{\mathrm{T}}^{\mathrm{miss}}$ with zero, taking resolutions of reconstructed objects into account
$m_T(b_1, E_{\mathrm{T}}^{\mathrm{miss}})$	transverse mass of leading-p_{T} b-jet and $E_{\mathrm{T}}^{\mathrm{miss}}$
$p_{\mathrm{T}}(t_{\mathrm{had}})$	transverse momentum of t_{had}
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, b_2)$	azimuthal angle separation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and subleading-p_{T} b-jet
$m_{\mathrm{inv}}(b_2, \ell)$	invariant mass of subleading-p_{T} b-jet and lepton
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, b_1)$	azimuthal angle separation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and leading-p_{T} b-jet
$\Delta\phi(t_{\mathrm{had}}, \ell)$	azimuthal angle separation between t_{had} and lepton
$p_{\mathrm{T}}(b_1)$	transverse momentum of leading-p_{T} b-jet

While the modelling of individual variables has been examined in Section 7.3.1, NNs, as discussed before, not only take individual variables into account, but also their correlations among each other. It is therefore paramount to not only ensure good modelling of the input variables to the training themselves, but to also compare their correlations in data and MC simulation. This is done with a correlation plot, which, for the sake of legibility, is shown in Figure 7.13 only for a subset of the variables. The corresponding plot for the full set of NN input variables is included in Appendix A.2. For each combination of two variables, v_i and v_j, the correlation is visualised with two profile plots, displaying $\langle v_j \rangle$ as a function of v_i and vice versa. Generally, good agreement between data and the SM background expectation is found,
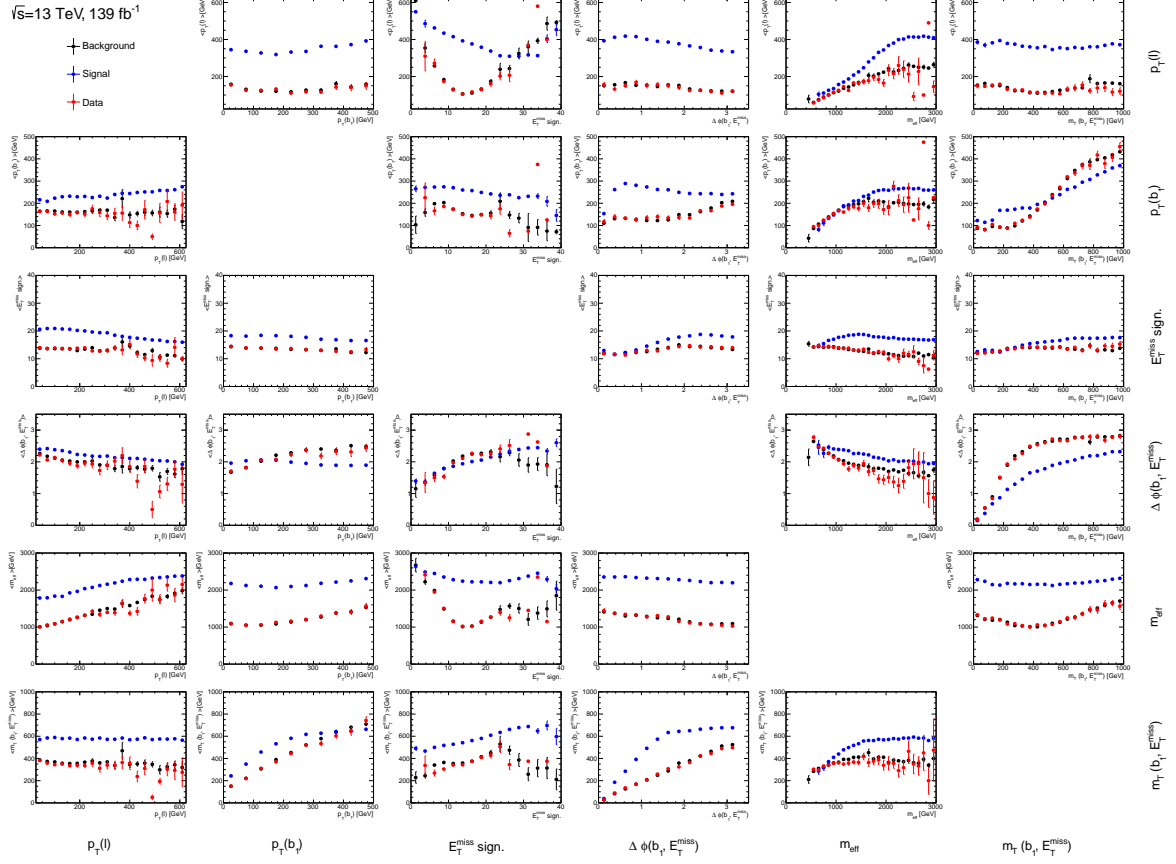
Figure 7.13: Profile plots visualising the correlations between a subset of NN input variables. MC simulated SM background events are shown in black, signal events in blue, and data in red. Here, scalar up-type LQs with a mass of $m_{\mathrm{LQ}} = 1300\,\mathrm{GeV}$ and $\mathcal{B}(\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to b\mu) = 0.5$ are shown as signal.

with data and MC behaving very similarly in the majority of the phase space. Deviations such as those observed in events with large values of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance can be ascribed to a lack of available statistics, particularly in data. The associated statistical uncertainties are not captured in the error bars.

In addition to serving as a cross-check for the data-MC agreement, Figure 7.13 also provides interesting information pertaining the differences in correlation for signal and background processes. To highlight these differences, the correlations for one exemplary signal process are overlaid, showing scalar up-type LQs with a mass of $m_{\mathrm{LQ}} = 1300\,\mathrm{GeV}$ decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. In many cases, the higher energy content of signal events is immediately evident, e.g. when the signal profile is shifted to higher values. This effect can, for instance, be seen in the correlations between the transverse momentum of the lepton and the transverse mass of the leading-p_{T} b-jet and the missing transverse momentum. The corresponding profile plots exhibit a flat profile, both for the signal and for the background processes, because the two observables are not

correlated. Discrepancies hence originate only from the signal events containing objects with higher momenta. Differences in correlations are not useful for the NN training in such cases. The situation is different for example for correlations between the transverse mass and the azimuthal separation of the leading-p_T b-jet and the missing transverse momentum. While a strong correlation between the two variables can be observed for signal and background, the transverse mass increases much faster for signal events. In events with azimuthal separations between the leading-p_T b-jet and E_T^{miss} of $\Delta\phi(b_1, E_T^{miss}) \approx \pi/2$, this leads to the average transverse mass to be more than twice as high for the signal process as for background, with smaller differences for other angles. This shows that taking such correlations into account can further improve the separation power of the NNs.

As mentioned before, overtraining of the NNs is exceedingly unlikely by construction and due to the calculation of the entropy loss on the test sample. However, to further substantiate this claim, an additional cross-check is performed a posteriori by comparing the NN_{out} distributions of the training and the validation samples separately for the four major SM background processes used in the training process, namely $t\bar{t}$, single top-quark, W+jets, and $t\bar{t}+V$ production. Exemplary plots for the NN trained with scalar up-type LQs with $\mathcal{B} = 0.5$ are shown in Figure 7.14. No significant deviations between training sample and test sample are visible, providing further assurance that no overtraining is occuring.

Distributions of a representative subset of the NN_{out} values obtained from the different NNs are shown in Figure 7.15, with the distributions corresponding to the remaining NNs included in Appendix A.3. In all cases, good agreement between observation and prediction is evident. The largest deviation between data and MC is found in the bin $0.9 < NN_{out} < 1.0$ in Figure 7.15a. This fluctuation, however, is not significant. In addition to serving as a check for the data-MC modelling, each of the figures also includes a signal distribution, scaled to the total background expectation, which indicates how well the NN is able to separate between signal and background processes in each case. Due to the excellent separation performance of the NNs, the signal is almost exclusively found at $NN_{out} > 0.5$, whereas the majority of the SM background events tends to low values of NN_{out}. The behaviour of the individual background processes differs slightly. This is shown exemplarily in Figure 7.16a for a NN trained on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$.
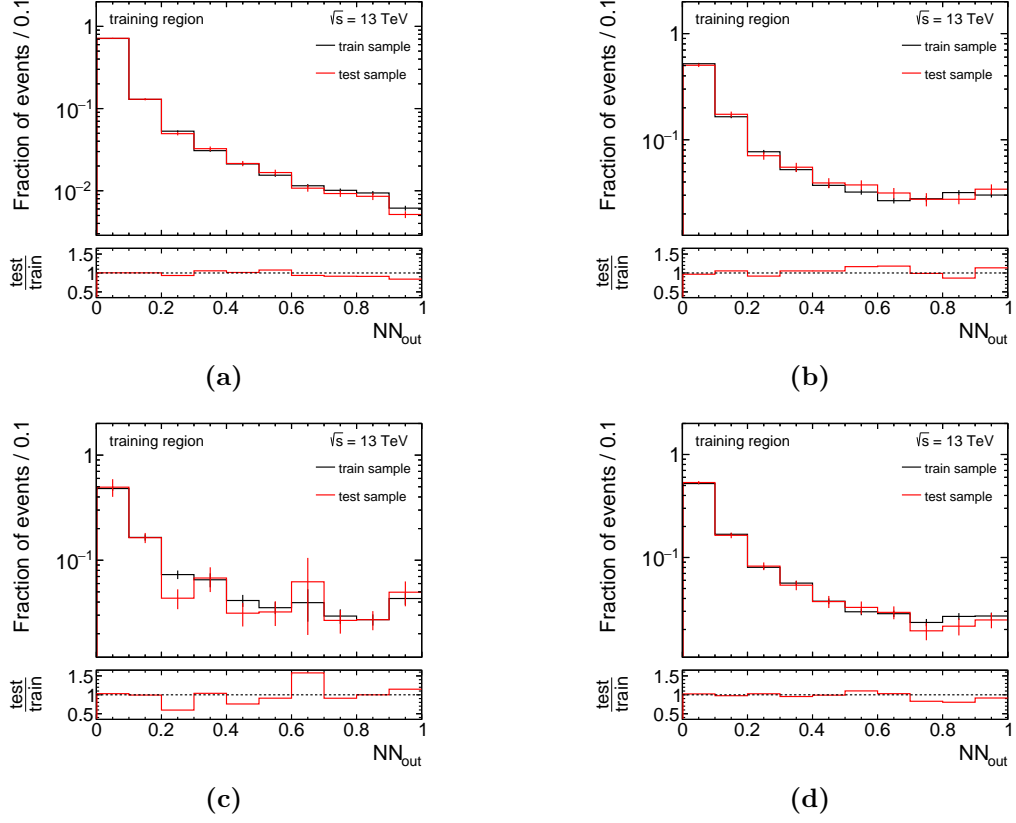
Figure 7.14: Comparison of NN_out distributions of the training and the test samples for the NN trained with scalar up-type LQs with $\mathcal{B} = 0.5$. The figure includes the four main SM background processes, i.e. (a) $t\bar{t}$, (b) single top-production, (c) W+jets, and (d) $t\bar{t} + V$ production.
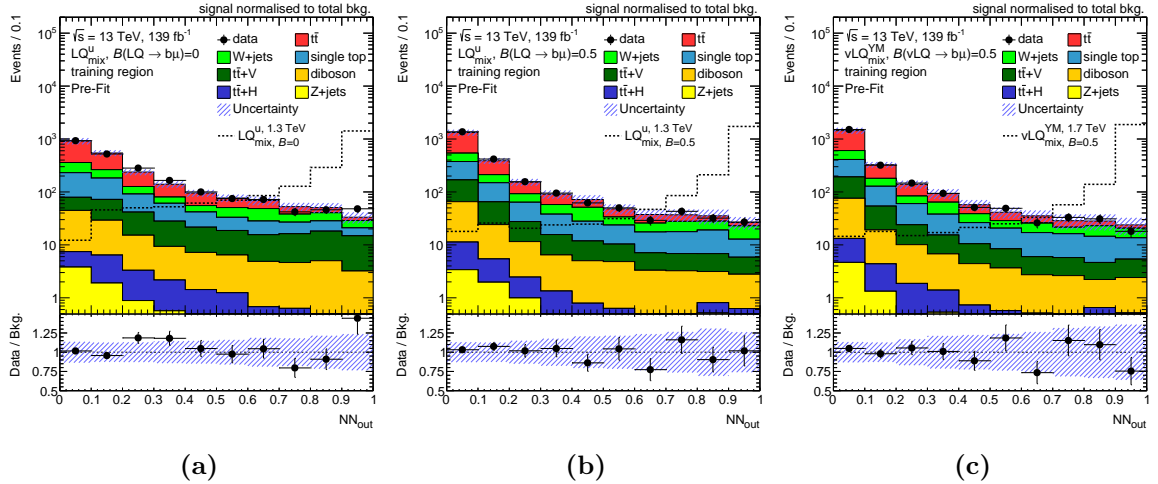

Figure 7.15: Distributions of NN_out obtained from NNs trained (a) on scalar up-type LQs decaying into top quarks and neutrinos, (b) on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, and (c) on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$.
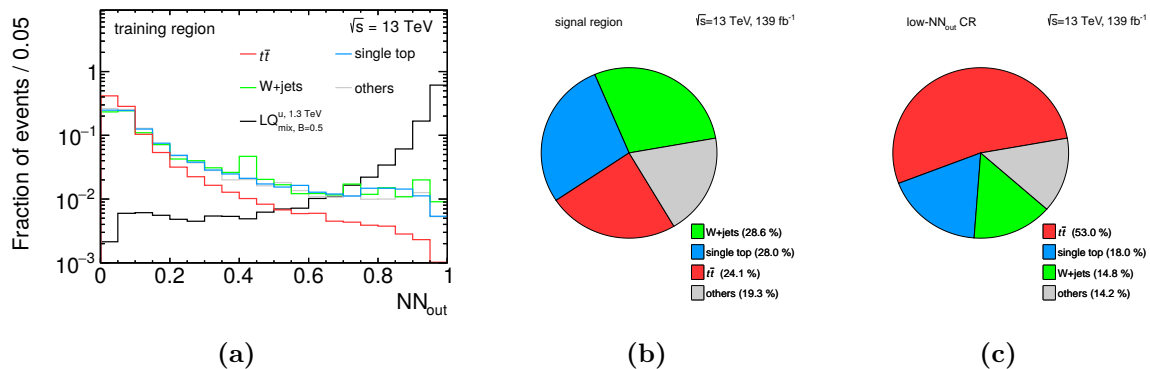
Figure 7.16: (a) NN_{out} distributions of background and signal processes obtained from a NN trained on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. The distributions are normalised to unit area. Pie charts visualising the background composition in (b) the SR and in (c) the low-NN_{out} CR. Contributions from $t\bar{t} + V$, $t\bar{t} + H$, diboson, and Z+jets processes are summarised as *others*.

In particular, the $t\bar{t}$ background exhibits a steeply falling distribution. Thus, the SR is defined as the range of NN_{out} above 0.5. As shown in Figure 7.16b, it contains nearly equal contributions from $t\bar{t}$, single top-quark, and W+jets events. The remainder, summarised as *others* in the figure, is dominated by $t\bar{t}+V$ processes. The lower half of the NN_{out} distribution on the other hand is enriched primarily in $t\bar{t}$ events, as illustrated in Figure 7.16c, and is therefore used to constrain the $t\bar{t}$ normalisation.

7.4 Background Estimation

While the background-like part of the training region is mostly enriched in $t\bar{t}$ events and therefore used as a CR to determine the normalisation of the $t\bar{t}$ process, dedicated orthogonal CRs are defined for the two other major background processes, W+jets and single top-quark production. Their definition is guided by the goal to obtain regions close in phase space to the signal region, enriched in each specific background, but without significant signal contamination.

As such, the W+jets CR is defined by selecting only events within a certain window in the transverse mass of the lepton and the missing transverse energy around the Jacobian peak. The window is chosen as $50\,\text{GeV} \leq m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}}) \leq 120\,\text{GeV}$, ensuring orthogonality to the other regions considered in this analysis. Furthermore, the number of b-tagged jets is limited to exactly one. While adding events with zero b-jets would increase the purity considerably, the additional events would primarily contain only light jets. Thus, in order to not introduce large deviations in flavour composition between the CR and the SR, the lower purity is tolerated. Additionally, to suppress contamination from $t\bar{t}$ events, a veto is placed on reconstructed hadronic top candidates.

The purity is further increased by exploiting the charge asymmetry caused by the prevalence

of up quarks over down quarks in the colliding protons, which leads to higher production cross-sections for W^++jets events than for W^-+jets events. This effect is reinforced by the fact that W bosons produced in proton-proton collisions with large transverse momenta are predominantly left-handed [183]. In a manner similar to the spin correlations discussed in the context of LQ decays in Section 2.2, these spins propagate to the decay products and lead to the neutrino originating from a W^+ to receive a boost in the direction of flight of the W boson, resulting in larger missing transverse momenta. Antineutrinos stemming from the decay of a W^- boson, on the other hand, are boosted in the opposite direction, thus leading to a reduced missing transverse momentum. The charge asymmetry is hence increased in the phase space with large missing transverse momenta. For these reasons, other background processes, that do not exhibit such a charge asymmetry, can be efficiently suppressed by requiring the lepton to be positively charged.

Finally, the same selection criterion as for the training region is applied on the asymmetric transverse mass, i.e. $am_{T2} > 200\,\text{GeV}$. After applying these requirements, W+jets events constitute around $58\,\%$ of the yields in the W+jets CR, with major additional contributions of $29\,\%$ and $9\,\%$ from $t\bar{t}$ and single top-quark events, respectively, as illustrated in Figure 7.17a.
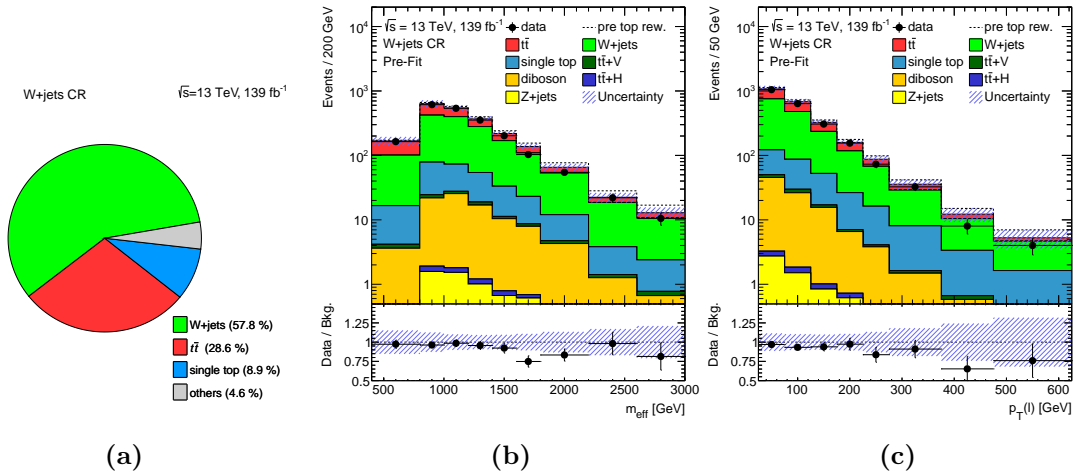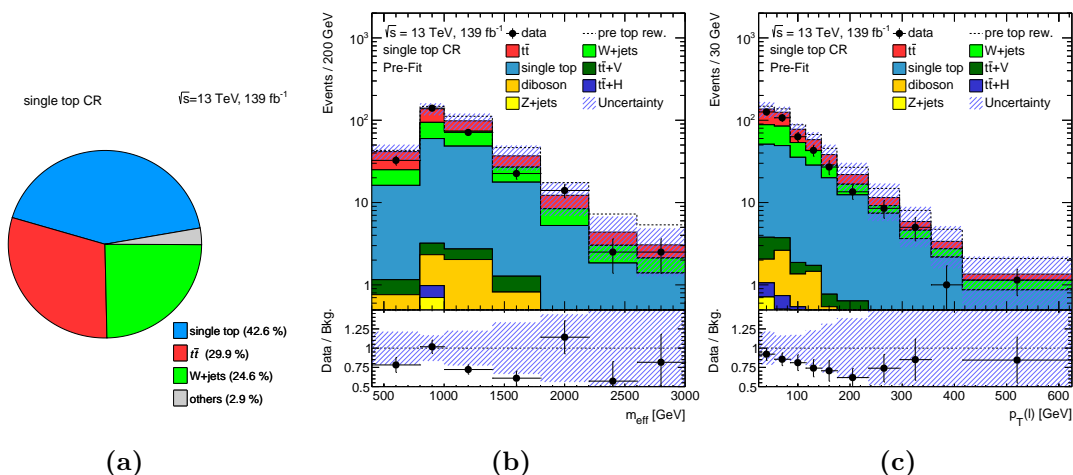

Figure 7.17: (a) Contributions from the different SM background processes in the W+jets CR. Distributions of m_{eff} (b) and of $p_T(\ell)$ (c) in the W+jets CR after applying the top reweighting factors, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.

The signal contamination amounts to below $1\,\%$ for scalar LQs with a mass of $m_{\text{LQ}} = 1000\,\text{GeV}$ and $\mathcal{B}(LQ \to q\ell) = 0.5$. The modelling in the W+jets CR is validated in various observables, all of which indicate good agreement between data and MC within statistical and systematic uncertainties. Exemplary plots are shown in Figure 7.17b and Figure 7.17c for m_{eff} and $p_T(\ell)$. Plots of the remaining NN input variables can be found in Appendix A.4. Again, the total background expectation prior to applying the top reweighting factors is overlaid

as a dashed black line, demonstrating the improvement in the ratio between expected and observed events.

Another control region is defined for the single top-quark background. Here, the requirement on the transverse mass between lepton and E_T^{miss} is inverted with respect to the training region, and hence only events with $m_T(\ell, E_T^{miss}) < 120\,\text{GeV}$ are selected. Orthogonality to the W+jets CR is guaranteed by requiring all events to contain exactly two b-tagged jets, which in addition are required to have an angular separation of $\Delta R(b_1, b_2) > 1.2$. In line with the training region definition, a selection criterion of $am_{T2} > 200\,\text{GeV}$ is applied. Lastly, contributions from $t\bar{t}$ processes are suppressed by selecting only events without any reclustered large-R jets. As illustrated in the pie chart in Figure 7.18a, the CR consists of about 40 % single top-quark background, and further contributions arise mainly from $t\bar{t}$ and W+jets events. The signal contamination amounts to roughly 1 % for scalar LQs with a mass of $m_{LQ} = 1000\,\text{GeV}$ and $\mathcal{B}(LQ \to q\ell) = 0.5$. As before, the modelling of the data in the single top CR is validated for different kinematic variables, with the effective mass and the transverse momentum of the lepton being shown in Figure 7.18b and Figure 7.18c, respectively. Plots of the remaining NN input variables can be found in Appendix A.4. While the agreement between data and prediction is improved through applying the top reweighting factors as evidenced by the dashed line indicating the total sum of background prior to the top reweighting, a general tension between the two is still evident. The W+jets background being modelled well, this points to issues in the description of the interference between single top-quark and $t\bar{t}$ production, which will be discussed in Section 8.1. However, the differences are covered by large systematic uncertainties associated to the interference terms.

(a)   **(b)**   **(c)**

Figure 7.18: (a) Contributions from the different SM background processes in the single top CR. Distributions of m_{eff} (b) and of $p_T(\ell)$ (c) in the single top CR after applying the top reweighting factors, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.

8 | Systematic Uncertainties

Besides uncertainties arising due to the inherently statistical nature of the experiment, analyses need to account for systematic deviations between the expectation and the actual measured values. Sources of these systematic discrepancies can either be of experimental nature, emerging due to limited knowledge of the interactions of particles with the detector material, or stem from theoretical considerations, as assumptions have to be made on parameters that cannot be derived from first principles. Depending on the source of uncertainty, the parameter values, therefore, have to be chosen based on previous measurements or even ad hoc, introducing ambiguities as the predictions become dependent on how the values for the different parameters are chosen. Both, experimental and theoretical systematic uncertainties, can impact the overall normalisation of the background and signal processes, but also the shape, i.e. the bin-by-bin behaviour. To ensure stability, the systematic uncertainties are subject to a preprocessing procedure. It consists of the four steps detailed in the following, some of which are only applied for certain processes or systematic variations.

Rescaling   Theoretical modelling uncertainties are rescaled, such that for each variation the sum of event yields in CRs and SR is the same as that of the nominal event yields. This is done in order to retain only shape and acceptance effects and to not introduce double-counting effects of the normalisation uncertainties in the profile likelihood fit. Due to this treatment, normalisation components of certain systematic uncertainties are not included in pre-fit uncertainty bands either.

Smoothing   The comparison between the nominal prediction and systematic variations can introduce fluctuations due to limited statistics, for example when systematic uncertainties are evaluated from alternative MC simulated samples. These fluctuations can lead to unphysically small or large estimations of the uncertainty. Such effects are avoided by averaging bin contents, a method referred to as smoothing.

Symmetrisation   Systematic uncertainties can be evaluated either from one or from two variations, the former being referred to as one-sided and the latter as two-sided variation. To increase the fit stability, one-sided systematics are symmetrised by mirroring the variation about the nominal prediction. Similarly, for two-sided systematics the up- and down-variations are shifted such that the nominal expectation is in the center of the two variations.

Pruning   When considering many different sources of systematic uncertainty, the statistical model can become very complex. In order to limit the number of insignificant systematic uncertainties, a pruning procedure is performed separately for every individual CR and SR. For each uncertainty, the normalisation component in a given region is discarded if it is smaller than a threshold of 1 %. Analogously, shape components are removed if in every bin of a region the systematic uncertainty deviates no more than 1 % from the nominal distribution.

8.1 Theoretical Systematic Uncertainties

Theoretical uncertainties are typically assessed by studying variations of particular parameter settings in order to evaluate their impact on the predictions. Various variations are considered for the SM background processes and for the signal hypothesis.

Theoretical Background Uncertainties

A typical example for the procedure is the uncertainty on renormalisation and factorisation scale, which are evaluated by varying each scale independently by a factor of two up and down with respect to its nominal value. These scale variations are calculated for all background processes. An additional uncertainty of 50 % is introduced for $W + b$-jets events to cover differences in flavour composition between the training region and the W+jets CR observed in MC as shown in Figure 8.1.
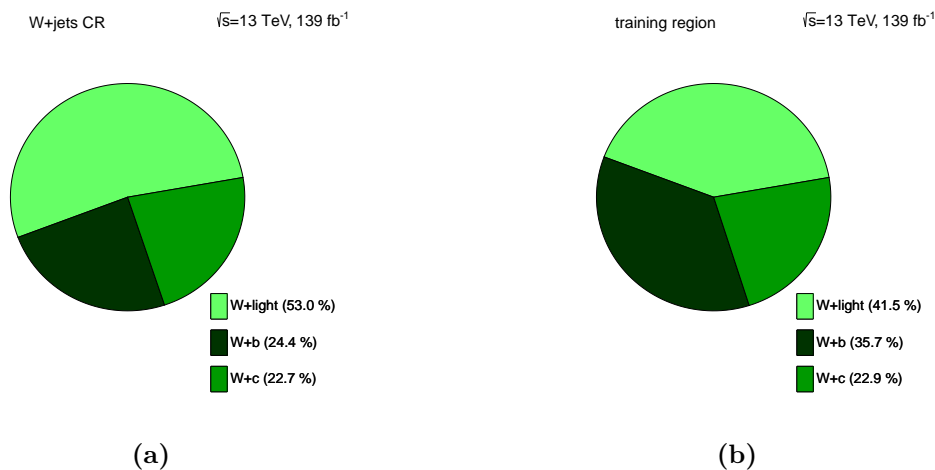
Figure 8.1: Flavour composition of the W+jets background process in (a) the W+jets CR and (b) the training region, based on the truth flavour of the reconstructed jets.

Further modelling uncertainties are included for $t\bar{t}$ and single top-quark processes. Alternative MC samples produced with MG5_AMC@NLO as ME generator and interfaced to PYTHIA 8 are utilised to assess the uncertainty on the NLO matching between ME and PS. Similarly, the effects of the choice of fragmentation and hadronisation models are evaluated by comparing

MC samples produced with POWHEG BOX as ME generator and interfaced to HERWIG 7 [184, 185] for the shower simulation. Additionally, the choice of a particular PDF set has an influence on the calculation of the hard scattering process. Uncertainties due to this choice are evaluated by comparing the nominal sample to 30 variations of the PDF4LHC15 set [186], consisting of a statistical combination of independent PDF sets. Furthermore, interference effects between the $t\bar{t}$ process and the production of a single top quark in association with a W boson and a bottom quark at NLO need to be accounted for. Corresponding Feynman diagrams reflecting these interferences are shown in Figure 8.2. For the nominal tW samples, these effects are handled with the diagram removal (DR) scheme [187], in which interfering diagrams are removed from the ME calculation, thereby avoiding overlaps between the two processes. However, this results in a ME that is not gauge invariant. An alternative approach is based on the diagram subtraction (DS) scheme [188], which in addition to the interfering contributions to tWb includes a gauge-invariant subtraction term to the ME. While this latter procedure is gauge invariant, the subtraction term is arbitrary and therefore adds ambiguity. Differences between the two schemes can be significant and constitute another source of modelling uncertainty.
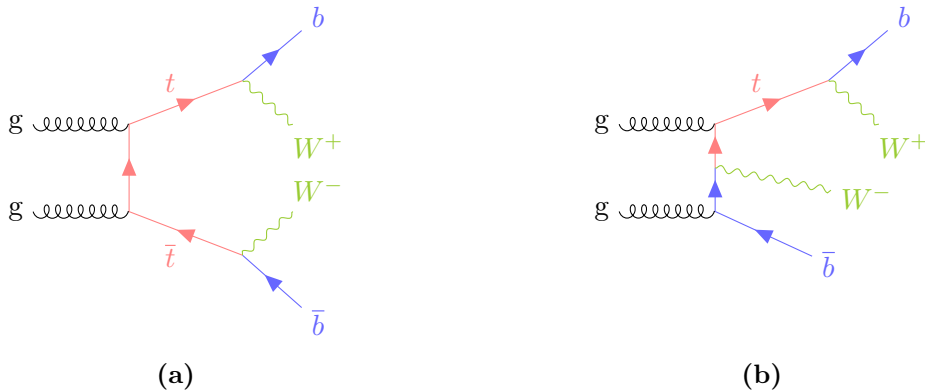
Figure 8.2: Feynman diagrams representing production of (a) $t\bar{t}$ events and (b) tWb events, leading to interference effects.

Variations of initial-state radiation (ISR) are estimated by varying α_S in the A14 tune from its nominal value of $\alpha_S^{\text{nom}} = 0.118$ to $\alpha_S^{\text{down}} = 0.117$ and $\alpha_S^{\text{up}} = 0.119$, respectively. Similarly, the uncertainty on final-state radiation (FSR) is assessed by varying the renormalisation scale for final-state PS emissions by a factor of two up and down. As discussed in Section 7.2, an additional uncertainty is applied on $t\bar{t}$ and single top-quark events to account for statistical uncertainties and potential non-linearities in the kinematic top reweighting procedure. For this, each of the four jet multiplicity bins is treated independently. In the case of $t\bar{t}$, the choice of the Powheg-dependent parameter hdamp, which regulates the radiation of jets with large transverse momenta, is free. The parameter is chosen to be hdamp$= 1.5m_t$ in the nominal sample, and a systematic uncertainty is evaluated by doubling this value.

While the normalisations of the signal and the major background processes $t\bar{t}$, single top-quark, and W+jets production are free parameters in the likelihood fit, uncertainties on the cross-sections of the minor backgrounds are considered. These amount to 11 % [189] in the case of $t\bar{t} + H$ and 15 % [189] for $t\bar{t} + Z$. A larger cross-section uncertainty of 50 % is used in the case of $t\bar{t}+W$, accounting for potential differences between the measured values and those predicted theoretically [190]. Lastly, the cross-section uncertainties on diboson and Z+jets processes are taken as 6 % [191] and 5 % [192], respectively.

Theoretical Signal Uncertainties

Theoretical uncertainties on the signal prediction are evaluated in a similar way as those on the SM backgrounds. As such, variations of the factorisation and the renormalisation scale by factors of two are used to estimate the impact of scale variations. Furthermore, uncertainties related to the choice of PDF are evaluated using 100 replicas of the NNPDF set, taking their root-mean-square deviation as the uncertainty. Uncertainties on ISR and FSR are assessed by varying α_S in the A14 tune and the renormalisation scale for final-state emissions, respectively, as explained for the theoretical background systematics. Taking into account only acceptance differences in the phase space under study, the quadratic sum of the signal modelling uncertainties is not strongly correlated with the LQ mass and does not exceed 3 % for LQ masses up to 1700 GeV, with leading contributions from PDF+α_S and from ISR variations. As the evaluation of each systematic variation for every individual signal hypothesis is infeasible due to computational limitations, a conservative choice of 5 % is made for the combined theoretical systematic uncertainties on the signal, thereby ensuring the uncertainties to be covered also for higher LQ masses up to 2.5 TeV.

8.2 Experimental Systematic Uncertainties

Experimental systematic uncertainties capture the uncertainties arising in the reconstruction of events and individual objects. As such, they cover the identification of objects as well as their calibration.

Lepton Calibration

The measurement of the energy and momentum of electrons and muons is affected by various systematic uncertainties. Sources of these uncertainties include noise due to pile up and electronics, sampling and shower shape fluctuations, and imprecise knowledge of the detector material. Potential misalignment between the ID and the MS and between the different layers of the MS itself are another source of uncertainty in the measurement of muons. In the analysis, simplified correlation models are used, combining contributions from the different sources into one systematic uncertainty on the energy scale and one systematic uncertainty on the energy resolution for electrons and muons, respectively. For muons, an additional

uncertainty accounting for the observed charge-dependent sagitta bias is applied.
Systematic uncertainties also arise in the determination of the identification, reconstruction, and isolation efficiencies. They are assessed by varying the selection criteria employed in the tag-and-probe method, resulting in systematic variations of the lepton SFs. An uncertainty accounting for the observed charge-dependent sagitta bias is applied.

Small-R Jet Calibration

Uncertainties on the JES and the jet energy resolution (JER) as well as on the JMS and the jet mass resolution (JMR) are assessed for small-R jets only, but not for reclustered jets, because uncertainties on the former are propagated to the latter. Various sources of systematic uncertainties on the JES are considered. These cover pile-up and flavour dependences, but also include effects from the in situ calibrations such as selection criteria, MC mismodelling, and statistical limitations. For this analysis, a simplified uncertainty scheme is used, combining all effects into 30 uncorrelated components.

Systematic uncertainties on the JER are obtained by propagating uncertainties from the JES. Further uncertainties stem from non-closure of the dijet balance. Lastly, effects of selection criteria are estimated by varying the measurement requirements and the MC modelling. As for JES, a reduced set of systematic uncertainties is utilised, corresponding to eight uncorrelated components in total.

Uncertainties on the JMS are obtained by varying the reconstructed small-R jet mass by 10% up and down, whereas systematic uncertainties on the JMR are derived by randomly varying the reconstructed small-R jet mass according to a p_T-dependend resolution map. A relative 20% JMR uncertainty is obtained by smearing the small-R jet mass with a Gaussian with width $0.66\,\sigma_{nominal}$. The nominal resolution $\sigma_{nominal}$ as a function of the jet-p_T is shown in Fig. 8.3.
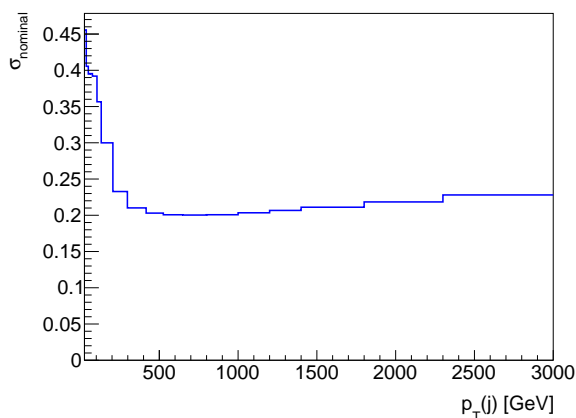
Figure 8.3: Nominal resolution of the small-R jet mass $\sigma_{nominal}$ depending on the transverse momentum of the jet, $p_T(j)$.

$E_\mathrm{T}^\mathrm{miss}$ Soft Term

As uncertainties are propagated through the calculation of the missing transverse momentum, no special systematic variations are required for the terms in the $E_\mathrm{T}^\mathrm{miss}$ calculation related to calibrated objects. For the soft term, however, additional uncertainties are included on the overall scale and on the resolution in parallel and perpendicularly to the direction of the soft term.

Flavor-Tagging Uncertainties

Uncertainties on the determination of b-tagging efficiencies of jets from different sources are propagated to the SFs in the form of eigenvariations. In total, there are nine eigenvariations for jets originating from bottom quarks and four each for jets originating from charm or light quarks. Furthermore, uncertainties on the extrapolation to high-p_T topologies above $400\,\mathrm{GeV}$ and on the extrapolation from charm quarks to τ-jets are considered. Moreover, as the calculation of am_T2 relies on the b-tagging score, which is uncalibrated, an additional $20\,\%$ uncertainty is introduced for MC events in which exactly one jet is b-tagged if the jet with the second-highest DL1r score does not originate from a bottom quark.

Other Instrumental Uncertainties

As mentioned in the description of the luminosity measurement in Section 3.2.5, the integrated luminosity of the full Run 2 dataset has been evaluated with an accuracy of $1.7\,\%$ [124]. This number is therefore applied as a systematic uncertainty on the total MC normalisation. Uncertainties on the pile-up reweighting procedure are derived by varying the factor used to shift the pile-up distribution prior to calculating the pile-up reweighting SFs. Its nominal value is 1.03, the corresponding up- and down-variations are 1.07 and 0.99, respectively.

9 | Statistical Techniques

The statistical analyses consists of two separate steps, the first one of which is a simultaneous profile likelihood fit to the observed data in the CRs and the SR. Building on the definition of the likelihood, hypothesis testing methods are introduced.

9.1 Profile Likelihood Fit

For binned analyses, like the one at hand, the expected number of events, N_i, in each bin i entering the fit can we written as

$$N_i = \mu s_i + \sum_j \mu_j b_{ji} \tag{9.1}$$

based on the predicted yields of the signal process, s_i, and of each background process j, $\mu_j b_{ji}$. Here, the signal strength μ is a multiplicative factor on the signal yields, with $\mu = 1$ corresponding to the nominal signal hypothesis. Similarly, μ_j represents the background normalisation factor for background j, which for certain background processes is a free parameter and for others is set to one. Information from CRs is used to constrain nuisance parameters and background normalisation factors in order to reduce the impact of systematic uncertainties on the sensitivity. From this, the likelihood function L can be formed as a product of the Poisson probability for each bin i of the N bins and of the constraint term $\rho(\Theta_j | \Theta_j^0, \sigma_{\Theta_j})$ for each of the M nuisance parameters, Θ_j, such that

$$L(\mu, \Theta) = \prod_{i=1}^{N} \frac{\left(\mu s_i + \sum_j \mu_j b_{ji}\right)^{n_i}}{n_i!} e^{-(\mu s_i + \sum_j \mu_j b_{ji})} \prod_{j=1}^{M} \rho(\Theta_j | \Theta_j^0, \sigma_{\Theta_j}), \tag{9.2}$$

with the number of observed events, n_i, and the vector of nuisance parameters, Θ. The constraint terms for the nuisance parameters corresponding to the systematic uncertainties discussed in Chapter 8 follow a Gaussian prior with mean Θ_j^0 and width σ_{Θ_j}, typically obtained from auxiliary measurements. Each of these nuisance parameters is divided into one component affecting solely the overall normalisation and one affecting the shape of the distribution only. The effects of the latter component are interpolated linearly from the 1σ up and down variations, while for the former an exponential interpolation is used in order to preclude negative values for the normalisation. The exponential interpolation is approximately linear only close to the nominal value, therefore leading to potentially asymmetric

effects for larger deviations from the prior. However, despite using different interpolation schemes for the normalisation and for the shape component, the same underlying constraint term is used for both by default, and the components are thus not independent. Profile likelihood fits are then performed by maximising the likelihood function within the constraints of the nuisance parameters. In addition to the systematic uncertainties, the MC statistical uncertainties, γ, are considered in the likelihood function, using a separate Poisson constraint term for each individual bin. In the following, fits under the assumption $\mu = 0$ are referred to as *background-only* fits, whereas the signal strength is a free parameter in *signal-plus-background* fits. Deviations of the best-fit value of a nuisance parameter from the nominal value are termed *pulls*, while a *constraint* occurs when its variance is limited by the fit. Such information can serve as a cross check, because due to limited statistics deviations from the auxiliary measurements are generally not expected in analyses directly searching for new physics contributions. Severe pulls or constraints could therefore point to instabilities in the fitting procedure.

9.2 Hypothesis Tests

In order to make statements about the compatibility of the data observed in the experiment with either the SM expectation or with BSM physics phenomena, two hypotheses are introduced. Here, H_0 refers to the background-only hypothesis, in which only the known SM processes are considered, whereas H_μ contains the additional contributions from the signal process as well and is referred to as signal-plus-background hypothesis. The goal of the statistical analysis is to find the parameters, for which with reasonable certainty either one or the other hypothesis can be excluded. As such, for each hypothesis H, a p-value can be computed, which represents the probability to measure data at least as incompatible with the predictions of H as the given observation. A hypothesis can be excluded if its p-value is below a certain threshold. Of particular interest in this analysis is the lowest signal production cross-section for which the p-value of H_μ is still below the threshold. For such an upper exclusion limit on the cross-section, the threshold is typically chosen as $5\,\%$. It is customary to convert the p-value into a significance $Z = \Phi^{-1}(1 - p)$, where Φ^{-1} refers to the quantile, i.e. the inverse of the cumulative distribution of the standard Gaussian. The aforementioned threshold then is $1.64\,\sigma$. On the other hand, to claim discovery of a signal process, a $5\,\sigma$ threshold is commonly used in particle physics, corresponding to a p-value of $2.87 \cdot 10^{-7}$.

In order to determine the p-value, the test statistic $\lambda(\mu)$ is used, which is defined as the ratio between the profile likelihood function $L(\mu, \hat{\hat{\Theta}})$ and the maximum likelihood $L(\hat{\mu}, \hat{\Theta})$, i.e.

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\Theta}})}{L(\hat{\mu}, \hat{\Theta})},$$ (9.3)

where $\hat{\hat{\Theta}}$ is the set of values of Θ that maximises the likelihood for a given μ. Similarly, $\hat{\mu}$ and $\hat{\Theta}$ are the values maximising the likelihood overall. By definition, $\lambda(\mu)$ is always between zero and one. According to the Neyman-Pearson lemma [193], the likelihood ratio constitutes the most powerful test statistic to distinguish between two hypotheses at a certain significance level.

Further, the test statistic is tranformed to

$$t_\mu = -2\ln(\lambda(\mu)), \tag{9.4}$$

converting the range from $[0, 1]$ to $[\infty, 0]$, with higher values of t_μ corresponding to greater incompatibility to data. Assuming the signal process to only ever increase event yields with respect to the SM expectation, the signal strength is bound from below to be at least zero. Best fit values of $\hat{\mu} < 0$ are considered to be non-physical in this context and the modified test function $\tilde{\lambda}(\mu)$ is thus defined as

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu,\hat{\hat{\Theta}}(\mu))}{L(\hat{\mu},\hat{\Theta})}, & \hat{\mu} \geq 0 \\ \frac{L(\mu,\hat{\hat{\Theta}}(\mu))}{L(0,\hat{\hat{\Theta}}(0))}, & \hat{\mu} < 0. \end{cases} \tag{9.5}$$

When deriving exclusion limits by checking the compatibility of a signal hypothesis with data, cases where $\hat{\mu} > \mu$ should not be regarded as increasing incompatibility and are therefore not taken into account further. Hence, the definition of the test statistic from Equation (9.4) is extended to

$$\tilde{q}_\mu = \begin{cases} -2\ln\left(\tilde{\lambda}(\mu)\right), & \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L\left(\mu,\hat{\hat{\Theta}}(\mu)\right)}{L\left(0,\hat{\hat{\Theta}}(0)\right)}, & \hat{\mu} < 0 \\ -2\ln\frac{L\left(\mu,\hat{\hat{\Theta}}(\mu)\right)}{L\left(\hat{\mu},\hat{\Theta}\right)}, & 0 \leq \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu. \end{cases} \tag{9.6}$$

From this, the p-value can be calculated as

$$p_\mu = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu)d\tilde{q}_\mu, \tag{9.7}$$

where $\tilde{q}_\mu^{\text{obs}}$ is the value of $\tilde{q}_\mu$ observed in data. The probability density function $f(\tilde{q}_\mu|\mu)$ can be evaluated either using pseudo-experiments or, in case of sufficient statistics, with asymptotic approximations [194]. For the former method, $f(\tilde{q}_\mu|\mu)$ is sampled by randomising the expected background and signal distributions within their uncertainties, requiring large sets of random numbers both for the background-only and for the signal-plus-background hypotheses, covering multiple values of μ. Alternatively, to avoid this computationally expensive

sampling of the pdf, the latter method is utilised. Using the asymptotic formulae, ultimately valid in the limit of large sample sizes, but resulting in good approximations even for samples with few events, $\widetilde{q}_\mu$ is calculated as

$$
\tilde{q}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2}, & \hat{\mu} < 0 \\ \frac{(\mu-\hat{\mu})^2}{\sigma^2}, & 0 \leq \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu. \end{cases}
\tag{9.8}
$$

This approach is used in this thesis to obtain exclusion limits for a multitude of signal hypotheses.

The methodology of calculating p-values of signal-plus-background hypotheses in order to accept or reject hypotheses with a certain level of significance suffers from certain drawbacks. This becomes evident when the signal yields are considerably smaller than the background expectation. Even though no sensitivity is expected in such cases, the signal can still be excluded when a downwards fluctuation in data is observed. A more refined approach, avoiding these pitfalls, is given with the CL_S method [195]. Here, the confidence level is calculated as

$$
CL_S = \frac{CL_{S+B}}{1 - CL_B},
\tag{9.9}
$$

where CL_{s+b} corresponds to the p-value of the signal-plus-background hypothesis, and CL_B represents the p-value of the background-only hypothesis. In the limit of vanishing overlaps between signal and background pdfs this yields CL_{S+B} as CL_b goes to zero, but for more similar signal and background distributions and therefore larger CL_B, CL_S becomes larger than CL_{S+B}. For CL_S to be below the significance threshold a larger signal strength is therefore necessary than it would be for the CL_{S+B} method.

10 | Results of the Statistical Analysis

The final results of the analysis are obtained from simultaneous profile likelihood fits to CRs and SR, using the yields in each CR and six bins in the NN_{out} distributions in the SR. The bin edges are chosen as $[0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 1.0]$, motivated primarily by the separation power of the NN causing a majority of signal events to have $NN_{\text{out}} > 0.9$. While a finer binning at high values of NN_{out} could in principle increase the sensitivity of the analysis, low statistics, both in measured data and in MC simulation, can introduce instabilities in the fit, with systematic variations often having particularly limited statistics. To avoid such issues, the distribution is not split further.

Given the similarities in the fitting procedures for the various signal hypotheses, intermediate steps ensuring the stability of the statistical model are discussed with a focus on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons, which serves as a benchmark scenario. These discussions are supplemented with insights from other signal hypotheses when doing so is beneficial as a comparison.

10.1 Background Normalisation

The normalisations of the three major backgrounds, $t\bar{t}$, W+jets, and single top-quark production, are unconstrained parameters in the fit, with corresponding normalisation factors referred to as $\mu_{t\bar{t}}$, $\mu_{W+\text{jets}}$, and $\mu_{\text{single top}}$, respectively. In a first step, profile likelihood fits are performed taking only SM background contributions into account and neglecting any potential signal contributions. Such background-only fits are used to check the stability of the statistical model and the agreement of the SM prediction with the observed data.

In this analysis, the background normalisation factors obtained from background-only fits are consistent across all SRs, fluctuating slightly depending on the training. They vary between 1.09 ± 0.22 and 1.30 ± 0.25 in the case of $t\bar{t}$, between 0.84 ± 0.12 and 0.93 ± 0.12 for W+jets, and between 0.44 ± 0.28 and 0.54 ± 0.26 for single top-quark production. The normalisation factors for each of the individual fits are listed in Table 10.1, together with the correlations between the normalisation factors for the different backgrounds. Large correlations can introduce ambiguities to the fit, as the values of the free-floating parameters become partially arbitrary, leading to fluctuations and to increased uncertainties. However, as is evident from the table, the largest anti-correlation between $t\bar{t}$ and single top-quark production is limited to 36 % at most, whereas correlations between $t\bar{t}$ and W+jets production or single top-quark production and W+jets exceed 10 % only occasionally. The effects of the correlations can be

Table 10.1: Normalisation factors on $t\bar{t}$, W+jets, and single top-quark production obtained from the background-only fit in W+jets CR, the low-NN_{out} training region, the single top CR, and the signal region. Additionally, the correlations, ρ, between $\mu_{t\bar{t}}$, $\mu_{W+\mathrm{jets}}$, and $\mu_{\mathrm{single\ top}}$ are shown.

	$\mu_{t\bar{t}}$	$\mu_{W+\mathrm{jets}}$	$\mu_{\mathrm{single\ top}}$	$\rho_{t\bar{t},W+\mathrm{jets}}$	$\rho_{t\bar{t},\mathrm{single\ top}}$	$\rho_{W+\mathrm{jets},\mathrm{single\ top}}$
$\mathrm{LQ}_{u,\mu}^{\mathcal{B}=0.0}$	1.23 ± 0.24	0.84 ± 0.12	0.46 ± 0.27	$-17.0\,\%$	$-25.6\,\%$	$-4.8\,\%$
$\mathrm{LQ}_{u,\mu}^{\mathcal{B}=0.25}$	1.25 ± 0.22	0.85 ± 0.12	0.47 ± 0.24	$-16.4\,\%$	$-19.9\,\%$	$1.4\,\%$
$\mathrm{LQ}_{u,\mu}^{\mathcal{B}=0.5}$	1.20 ± 0.23	0.88 ± 0.12	0.48 ± 0.26	$-6.9\,\%$	$-28.8\,\%$	$-8.6\,\%$
$\mathrm{LQ}_{u,\mu}^{\mathcal{B}=0.9}$	1.30 ± 0.25	0.84 ± 0.12	0.48 ± 0.30	$-9.7\,\%$	$-35.4\,\%$	$-6.4\,\%$
$\mathrm{LQ}_{u,e}^{\mathcal{B}=0.0}$	1.25 ± 0.24	0.85 ± 0.12	0.47 ± 0.27	$-15.9\,\%$	$-23.2\,\%$	$-8.5\,\%$
$\mathrm{LQ}_{u,e}^{\mathcal{B}=0.25}$	1.09 ± 0.22	0.93 ± 0.12	0.53 ± 0.26	$-0.8\,\%$	$-30.5\,\%$	$-9.2\,\%$
$\mathrm{LQ}_{u,e}^{\mathcal{B}=0.5}$	1.12 ± 0.22	0.91 ± 0.12	0.53 ± 0.25	$-0.4\,\%$	$-28.9\,\%$	$-8.7\,\%$
$\mathrm{LQ}_{u,e}^{\mathcal{B}=0.9}$	1.15 ± 0.22	0.91 ± 0.12	0.50 ± 0.27	$1.1\,\%$	$-31.5\,\%$	$-12.3\,\%$
$\mathrm{LQ}_{d,\mu}^{\mathcal{B}=0.5}$	1.29 ± 0.23	0.84 ± 0.12	0.47 ± 0.29	$-10.3\,\%$	$-35.6\,\%$	$-3.9\,\%$
$\mathrm{LQ}_{d,e}^{\mathcal{B}=0.5}$	1.09 ± 0.22	0.92 ± 0.12	0.54 ± 0.26	$-2.3\,\%$	$-29.9\,\%$	$-11.0\,\%$
$\mathrm{vLQ}_{\mathrm{YM},\mu}^{\mathcal{B}=0.0}$	1.17 ± 0.23	0.87 ± 0.12	0.52 ± 0.24	$-13.7\,\%$	$-22.6\,\%$	$-4.1\,\%$
$\mathrm{vLQ}_{\mathrm{YM},\mu}^{\mathcal{B}=0.25}$	1.25 ± 0.21	0.86 ± 0.12	0.52 ± 0.24	$-13.1\,\%$	$-20.9\,\%$	$0.5\,\%$
$\mathrm{vLQ}_{\mathrm{YM},\mu}^{\mathcal{B}=0.5}$	1.23 ± 0.24	0.87 ± 0.12	0.52 ± 0.27	$-3.4\,\%$	$-30.3\,\%$	$-9.9\,\%$
$\mathrm{vLQ}_{\mathrm{YM},\mu}^{\mathcal{B}=0.9}$	1.26 ± 0.24	0.86 ± 0.12	0.46 ± 0.28	$-8.2\,\%$	$-30.9\,\%$	$-8.6\,\%$
$\mathrm{vLQ}_{\mathrm{YM},e}^{\mathcal{B}=0.0}$	1.18 ± 0.23	0.90 ± 0.13	0.52 ± 0.25	$-7.9\,\%$	$-23.7\,\%$	$-8.3\,\%$
$\mathrm{vLQ}_{\mathrm{YM},e}^{\mathcal{B}=0.25}$	1.14 ± 0.22	0.93 ± 0.13	0.47 ± 0.25	$0.0\,\%$	$-30.1\,\%$	$-9.4\,\%$
$\mathrm{vLQ}_{\mathrm{YM},e}^{\mathcal{B}=0.5}$	1.17 ± 0.22	0.90 ± 0.13	0.49 ± 0.24	$-5.1\,\%$	$-22.7\,\%$	$-9.4\,\%$
$\mathrm{vLQ}_{\mathrm{YM},e}^{\mathcal{B}=0.9}$	1.20 ± 0.24	0.91 ± 0.14	0.44 ± 0.28	$-5.9\,\%$	$-34.1\,\%$	$-11.8\,\%$

observed in the contrasting fluctuations of the normalisation parameters, as when $\mu_{t\bar{t}}$ is at its lowest, $\mu_{W+\mathrm{jets}}$ and $\mu_{\mathrm{single\ top}}$ tend to be close to their maxima. In general, due to the smaller correlation to the other normalisation factors, $\mu_{W+\mathrm{jets}}$ exhibits fewer fluctuations than $\mu_{t\bar{t}}$ or $\mu_{\mathrm{single\ top}}$.

The compatibility of normalisation factors obtained from fits to different NN_{out} distributions is validated with a bootstrapping procedure involving pseudo-experiments. This method is necessary, because the yields in the single top CR and the W+jets CR are the same for each NN. In addition, distributions in the low-NN_{out} CR and the SR differ between the trainings, but the set of events contained in the training region is always the same. The normalisation factors are therefore correlated across the different background-only fits. In fact, normalisation factors obtained from the individual fits are expected to agree within uncertainties. To prove this is indeed the case, random Poisson distributed weights are applied to each data event and the background-only fit is rerun for two different trainings using the reweighted data. This process is repeated 1000 times. The corresponding differences between the normalisation factors obtained from the fits are shown in Figure 10.1 for the
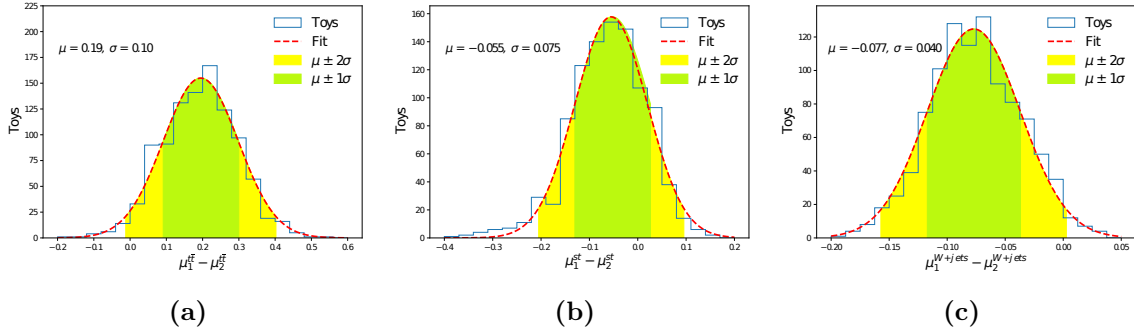
Figure 10.1: Differences between normalisation factors on (a) $t\bar{t}$, (b) single top-quark, and (c) W+jets production obtained from background-only fits of NN_{out} distributions from a training on down-type LQs coupling to muons and from a training on down-type LQs coupling to electrons. For the former training, the normalisation factors are labelled μ_1, for the latter μ_2. The results of the pseudo-experiments (*toys*) are fitted with a Gaussian curve with mean μ and standard deviation σ.

NN trained on down-type LQs coupling to muons and the NN trained on down-type LQs coupling to electrons. These NNs are chosen for the cross-check because they exhibit the largest variations in normalisation factors. Consequently, the distributions of the pseudo-experiments are fitted with a Gaussian curve, indicating compatibility to zero within $1\,\sigma$ for the single top normalisation factors and within $2\,\sigma$ for the $t\bar{t}$ and the W+jets normalisation factors. Moreover, with the two most different fits being consistent within $2\,\sigma$, a general compatibility of the fit results across all 18 NNs can be concluded.

10.2 Nuisance Parameters

To validate the results from the likelihood fit, the behaviour of nuisance parameters is checked. First, correlations of systematic uncertainties and free-floating normalisation parameters are illustrated in Figure 10.2.

As discussed in the context of the normalisation factors, such correlations can introduce ambiguities in the statistical model, leading to an unstable fitting procedure. Among the systematic uncertainties only small correlations occur, which are always below 20 %. However, correlations between systematic uncertainties and normalisation factors can be considerably higher, with the largest correlation arising between $\mu_{\text{single top}}$ and the single top DS uncertainty. Even though the overall normalisation effects of the modelling uncertainties are not considered in the fit, the correlation remains high due to the compositions of the different regions. The corresponding systematic variations in the different regions are displayed in Figure 10.3. The two regions purest in single top-quark events and therefore providing the most constraining power for the normalisation factor, are the single top CR and the SR. In both regions, a variation by $+1\,\sigma$ is associated with a strong reduction in overall yields. Conversely, the same variation results in an enhancement of the overall yields in the W+jets

Figure 10.2: Correlations between nuisance parameters in a background-only fit in control regions and the signal region obtained from the training for up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. Only nuisance parameters which have a correlation of at least 20 % with a least one other nuisance parameter are included in the matrix.
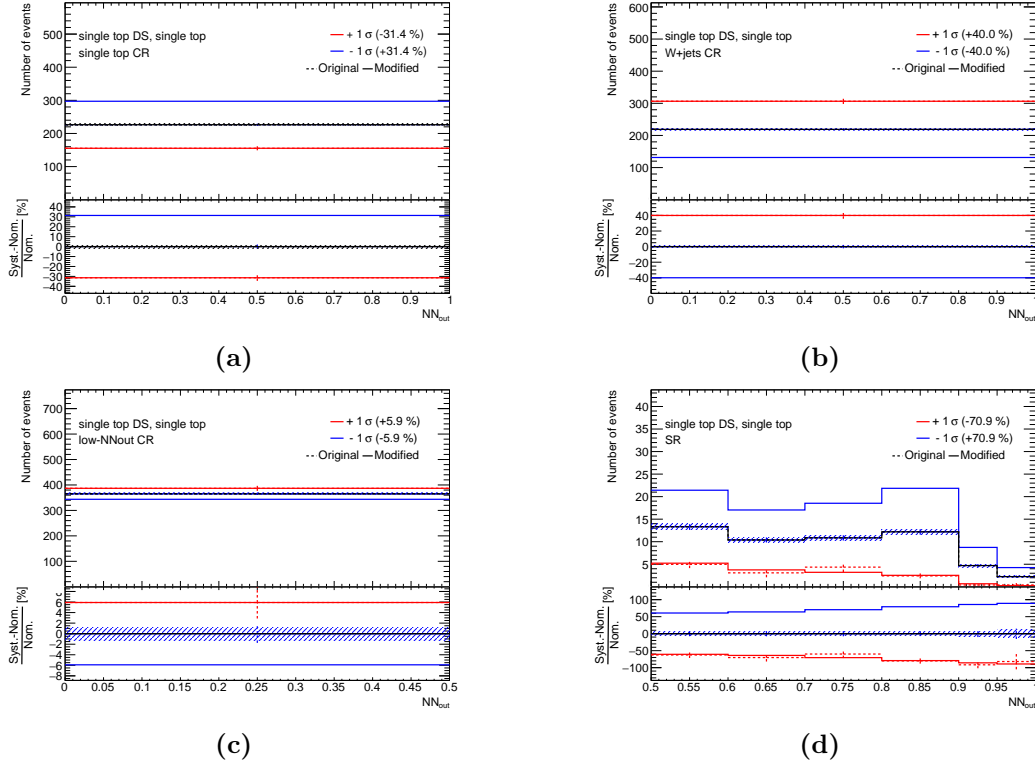
Figure 10.3: Systematic variations corresponding to the single top DS nuisance parameter in (a) the single top CR, in (b) the W+jets CR, in (c) the low-NN_{out} CR, and in (d) the SR. In the lower panel, the relative differences between variation and nominal prediction are shown. The original variation is represented by a dashed line, the modified variation after smoothing and symmetrisation is depicted as a solid line.

CR and in the low-NN_{out} CR. However, as in these regions single top-quark events are not as prevalent, the normalisation factor is primarily affected by the single top CR and the SR, resulting in considerable correlations.

Pulls and constraints of nuisance parameters obtained in background-only fits are visualised in Figure 10.4 for NNs trained on up-type scalar LQs decaying into bottom quarks and muons or top quarks and neutrinos. The largest constraints are found for single top DS, which is consistently constrained by about 40 % in all fits. The estimate of the uncertainty prior to the fit is known to be conservative, since the full difference between the DR and the DS scheme is used and symmetrised with respect to the nominal prediction. A constraint for this particular uncertainty is therefore expected. Pulls are primarily observed for top modelling systematics such as single top hard scatter. They are, however, limited to less than 0.5σ in most cases. Small, but consistent pulls of around 0.1σ emerge also for the JER, because all of the associated variations show very similar behaviour across the CRs and the SR. Equivalent behaviour of the nuisance parameters as discussed above in the context of up-type scalar LQs coupling to muons is also observed for background-only fits using NN_{out} distributions obtained from trainings for other signal hypotheses.
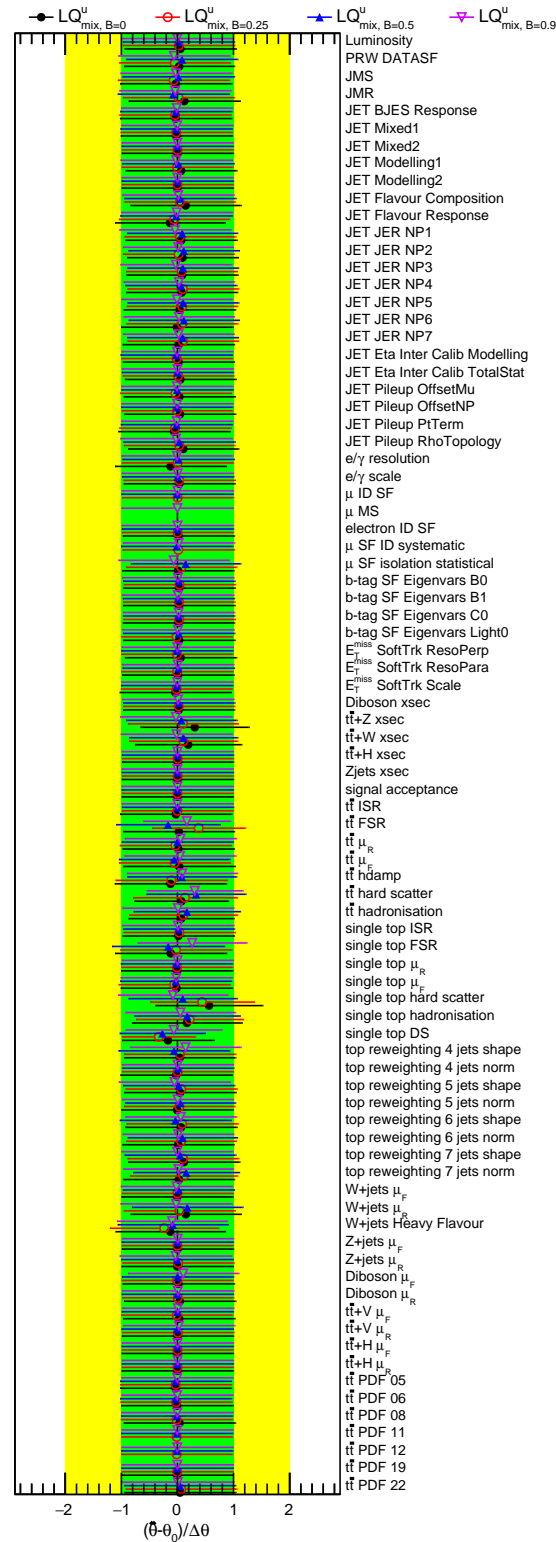
Figure 10.4: Pulls and constraints of nuisance parameters obtained in background-only fits to data for NNs trained on up-type scalar LQs decaying into bottom quarks and muons or top quarks and neutrinos at various branching ratios.

It can be concluded from the correlation matrix and the pull plot that the statistical model is sufficiently stable. In the next step, it is extended to also include contributions from signal processes. By adding the signal to the likelihood fit and thus the signal strength as a parameter, it becomes possible to evaluate the impact of individual nuisance parameters on the signal expectation.

A ranking of the nuisance parameters in terms of their effect on the fitted signal strength is shown in Figure 10.5 for a signal scenario with scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons, assuming a LQ mass of $m_{\mathrm{LQ}} = 1.3\,\mathrm{TeV}$ and a branching ratio of $\mathcal{B} = 0.5$. It is obtained by fixing the nuisance parameter in question to its $\pm 1\sigma$ variations and repeating the fit. For each nuisance parameter, the best fit value of the signal strength derived in this procedure is compared to the one obtained from the complete fit. Deviations are considered as the impact of said nuisance parameter. The pre-fit impact is calculated from variations within the pre-fit uncertainties, while the post-fit impact is evaluated from variations within the post-fit uncertainties, therefore taking constraints into account. Nuisance parameters are then ranked according to their post-fit impact.
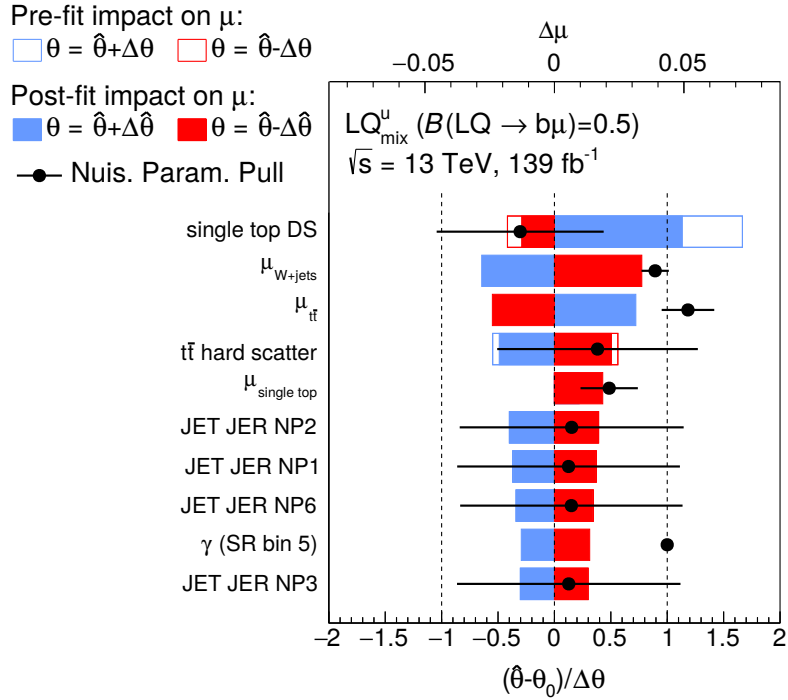
Figure 10.5: Ranking of nuisance parameters included in signal-plus-background fits for scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $m_{\mathrm{LQ}} = 1.3\,\mathrm{TeV}$ and $\mathcal{B} = 0.5$. Nuisance parameters are ranked according to their impact on the observed signal strength, $\Delta\mu$, which is evaluated by fixing each respective nuisance parameter to a value shifted by its uncertainty, repeating the fit, and comparing the signal strength to the nominal fit result. Empty boxes with blue borders indicate pre-fit impacts, whereas post-fit impacts are shown as filled boxes. Pulls of the nuisance parameters and their post-fit uncertainties are shown as black points and error bars.

The most impactful nuisance parameters differ slightly between fits, but are generally related to top modelling uncertainties. In particular, the single top DS uncertainty frequently appears at or near the top of the ranking, with the effects of the aforementioned constraints clearly visible in the differences between pre- and post-fit impacts. As explained above, the sizeable impact is expected due to the magnitude of the uncertainty. Other highly ranked nuisance parameters for instance include the single top hard scatter uncertainty, but also the γ-factor associated with the bin highest in NN_{out}, reflecting the limitations of the available MC statistics. Moreover, the background normalisation factors can have considerable impact on the signal strength, primarily depending on the background composition in the bins containing most signal events. The most impactful experimental systematic uncertainty is associated with the small-R jet energy resolution.

In addition, Figure 10.5 also shows pulls and constraints on the different nuisance parameters as done previously in the pull plot in the case of the background-only fits. This allows for a cross-check to ensure that adding the signal process does not significantly alter the fit results. In fact, a comparison of the ranking plot with Figure 10.4 indicates essentially the same pulls and constraints. Hence, the conclusions drawn from the background-only fits still hold for the signal-plus-background fit.

Another way of quantifying the impact of systematic uncertainties on the analysis is to form groups of nuisance parameters. The fit is then repeated multiple times, each time fixing the nuisance parameters in one such group to their best fit values. This allows to determine the impact of the nuisance parameters in a given group on the uncertainty of the fitted signal strength as

$$\Delta \mu_i^2 = \Delta \mu_{\mathrm{orig}}^2 - \Delta \mu_{\mathrm{not}\,i}^2. \tag{10.1}$$

Here, $\Delta \mu_{\mathrm{orig}}$ is the uncertainty on the signal strength in the original fit and $\Delta \mu_{\mathrm{not}\,i}$ is the uncertainty on the signal strength in the fit where the nuisance parameters of group i are fixed to their best fit values. Accordingly, the impact of the statistical uncertainty is evaluated by fixing all nuisance parameters associated with systematic uncertainties to their best fit values. The results of this test are shown in Table 10.2 for scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B}(\mathrm{LQ}_{\mathrm{mix}}^{\mathrm{u}} \to b\mu) = 0.5$ and $m_{\mathrm{LQ}} = 500\,\mathrm{GeV}$ or $1300\,\mathrm{GeV}$.

For a comparably large mass of $1300\,\mathrm{GeV}$, the statistical uncertainty is dominant and the impact of the systematic variations on the uncertainty of the fitted signal strength is small. The groups with the largest impacts are theory systematics on the $t\bar{t}$ and single top-quark background, uncertainties on the normalisation factors, and experimental uncertainties in the small-R jet calibration, all of which also appear highly ranked in Figure 10.5. The impact of the normalisation factors is caused by correlations to the modelling uncertainties. For lower masses, the importance of systematic uncertainties increases, because due to the higher cross-sections at low masses the signal contamination in the CRs is not negligible anymore.

Table 10.2: Impacts of groups of nuisance parameters on the uncertainty of the fitted signal strength for LQ^u_{mix} decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$ and $m_{LQ} = 500\,\text{GeV}$ or $1300\,\text{GeV}$.

	$\Delta\mu_i$	
Uncertainty Group	$LQ^{u,0.5\,\text{TeV}}_{mix,B=0.5}$	$LQ^{u,1.3\,\text{TeV}}_{mix,B=0.5}$
Lepton Calibration	$0.57 \cdot 10^{-3}$	$0.06 \cdot 10^{-1}$
Jet Calibration	$1.75 \cdot 10^{-3}$	$0.35 \cdot 10^{-1}$
Flavour Tagging	$0.23 \cdot 10^{-3}$	$0.03 \cdot 10^{-1}$
E_T^{miss} Soft Term	$0.08 \cdot 10^{-3}$	$0.02 \cdot 10^{-1}$
Other Experimental	$< 0.01 \cdot 10^{-3}$	$0.01 \cdot 10^{-1}$
Top Modelling	$4.06 \cdot 10^{-3}$	$0.46 \cdot 10^{-1}$
Other Modelling	$0.76 \cdot 10^{-3}$	$0.14 \cdot 10^{-1}$
Normalisation Factors	$2.30 \cdot 10^{-3}$	$0.43 \cdot 10^{-1}$
γ-Factors	$0.50 \cdot 10^{-3}$	$0.15 \cdot 10^{-1}$
Statistical	$3.70 \cdot 10^{-3}$	$1.58 \cdot 10^{-1}$
Total	$6.12 \cdot 10^{-3}$	$1.75 \cdot 10^{-1}$

Constraints on the signal strength are hence not solely driven by the SR, but also from the CRs, which are not statistically limited. Therefore, the impact of systematic uncertainties can exceed that of the statistical uncertainty for low LQ masses as shown here exemplarily for $m_{LQ} = 500\,\text{GeV}$.

10.3 Post-Fit Distributions

Having carefully established the stability of the statistical model in the previous sections, the resulting post-fit yields and distributions are examined in the following. First, the overall background yields in the CRs and the SR obtained from a background-only fit are listed in Table 10.3. For the fit, the low-NN_{out} CR and the SR obtained from the NN trained on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B}(LQ^u_{mix} \to b\mu) = 0.5$ are used. A comparison with the number of observed events in data indicates very good agreement.

Since in the fit the SR is split into several bins, it is instructive to not only examine the overall yields, but to also look at the post-fit distributions. In a first step, the CRs are used to validate the NN_{out} distributions after the profile likelihood fit. As such, the post-fit distributions in the three CRs are shown in Figure 10.6. Differences between data and MC are generally small in the W+jets CR and in the low-NN_{out} CR, where the ratio is compatible with unity across the whole range of NN_{out}. On the other hand, in the single top CR, a downward trend is apparent. While in the likelihood fit the normalisation is extracted correctly, it cannot correct for the differences in the shape, because this information is not included in the fit. However, the observed data still agree with the prediction within the

Table 10.3: Observed and expected event yields in the control and signal regions using a NN trained on $LQ_{mix}^u \to b\mu/t\nu$ and $\mathcal{B}(LQ_{mix}^u \to b\mu) = 0.5$ after the background-only fit. The uncertainties on the background predictions include the statistical and the systematic components. For comparison, expected event yields are shown for a LQ_{mix}^u signal at a mass point of $1300\,\text{GeV}$ and $\mathcal{B}(LQ_{mix}^u \to b\mu) = 0.5$ including its pre-fit uncertainties.

	W+jets CR	single top CR	low-NN_{out} CR	SR
$t\bar{t}$	850 ± 130	188 ± 35	1370 ± 150	61 ± 15
single top	94 ± 74	118 ± 49	200 ± 110	34 ± 15
W+jets	1260 ± 130	112 ± 29	265 ± 55	43.3 ± 8.5
$t\bar{t} + V$	12.4 ± 3.1	4.6 ± 1.1	180 ± 28	20.3 ± 3.7
diboson	95.1 ± 9.8	7.9 ± 2.0	94 ± 11	14.7 ± 1.7
$t\bar{t} + H$	1.27 ± 0.16	1.00 ± 0.12	14.4 ± 1.7	1.58 ± 0.23
Z+jets	6.46 ± 0.32	2.18 ± 0.11	7.20 ± 0.36	1.409 ± 0.078
Total background	2312 ± 48	434 ± 20	2099 ± 45	176 ± 12
Observed events	2310	430	2100	181
LQ_{mix}^u ($m_{LQ} = 1.3\,\text{TeV}$, $\mathcal{B}(LQ_{mix}^u \to b\mu) = 0.5$)	0.0741 ± 0.0059	0.038 ± 0.0040	1.05 ± 0.06	19.7 ± 1.1

(a)   (b)   (c)
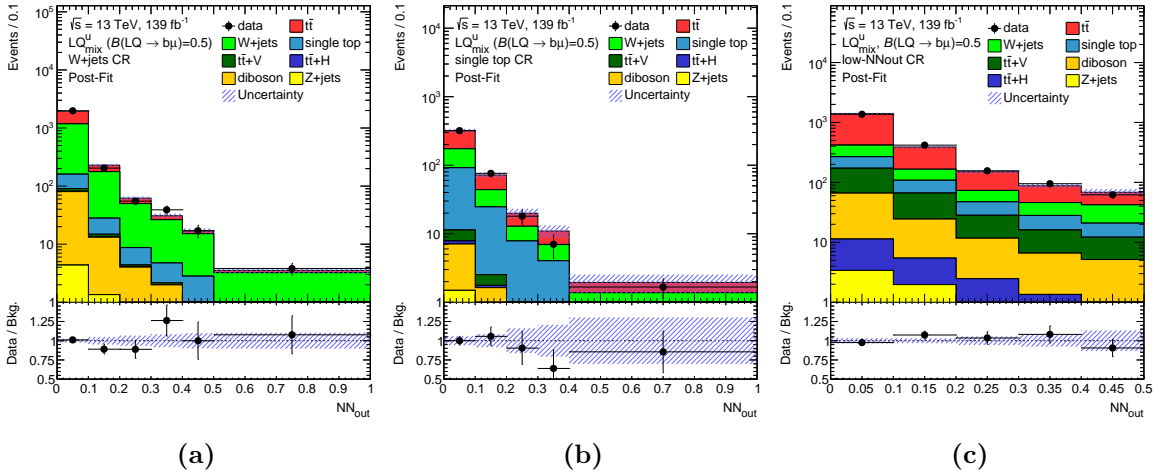
Figure 10.6: Post-fit distributions of NN_{out} in (a) the W+jets CR, (b) the single top CR, and (c) the low-NN_{out} CR for an NN trained for up-type scalar LQs decaying into bottom quarks and muons or top quarks and neutrinos with $\mathcal{B} = 0.5$.

statistical and systematic uncertainties, which are fully propagated to the SR.

Having observed generally good agreement between observation and prediction in the CRs, post-fit distributions for a representative subset of SRs are shown in Figure 10.7, while the corresponding figures for the remaining SRs can be found in Appendix A.5. In the figures, the pre-fit signal expectations are added on top of the post-fit background predictions to indicate which bins contribute to the sensitivity of the analysis the most. As discussed previously, the most sensitive bins are the ones corresponding to $NN_{\text{out}} > 0.9$. This is particularly relevant in the case of Figure 10.7a, because here upwards fluctuations in data can be seen in those two bins. While not significant per se, these represent the largest differences with respect to the SM prediction observed in any of the SRs considered. Interestingly, similar deviations of around $2\,\sigma$ have been reported previously in a search for supersymmetric top partners on the same dataset [196], which targets the same phase space as is relevant for events with both LQs decaying into top quarks and neutrinos.

Despite the small fluctuations in certain SRs, the data are in general described very well by the background prediction. The existence of a certain degree of deviations is not surprising given the large number of SRs.

(a)   **(b)**   **(c)**
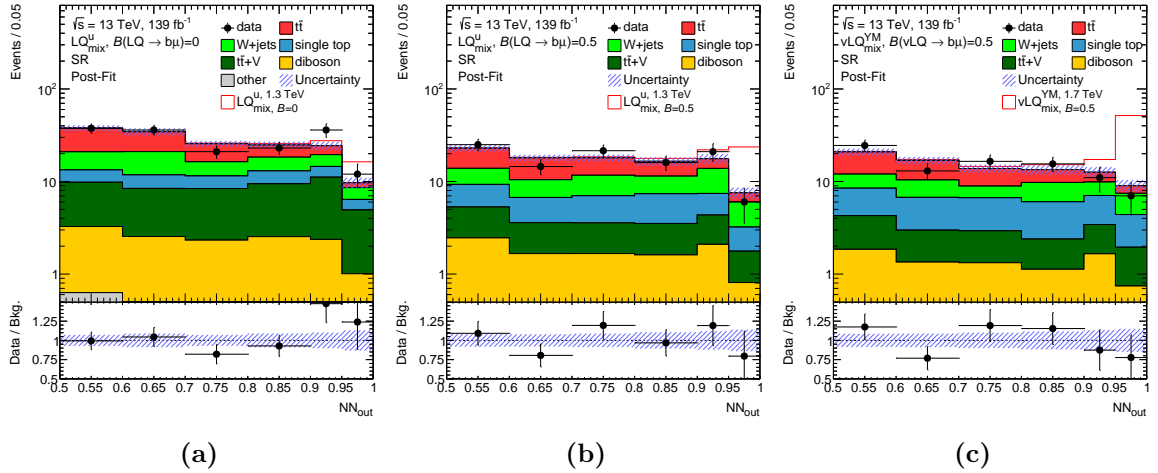
Figure 10.7: Post-fit distributions of NN_{out} in the signal region (a) for an NN trained on up-type scalar LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B} = 0.0$, (b) for an NN trained on up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, and (c) for an NN trained on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. The pre-fit signal expectation at $m_{\text{LQ}} = 1.3\,\text{TeV}$ for scalar LQs and at $m_{\text{LQ}} = 1.7\,\text{TeV}$ for vector LQs is added on top of the post-fit background yields.

10.4 Exclusion Limits

In the absence of significant excesses above the SM expectation, upper 95% CL exclusion limits on the signal cross-section as a function of the LQ mass and $\mathcal{B}$ are determined following the methodology outlined in Section 9.1. Limits are set for all signal masses for which MC samples are available and for branching ratios in steps of 0.05, assuming the LQs can only decay into the third generation of quarks and one generation of leptons, such that $\mathcal{B}(LQ \to q_3 \ell) = 1 - \mathcal{B}(LQ \to q_3' \nu)$, where q_3 and q_3' are the quarks of the third generation. However, limits are not evaluated for $\mathcal{B} = 1.0$, as corresponding signal events are not selected in the analysis. Similarly, for down-type LQs, no limits are computed for $\mathcal{B} = 0.0$. For up-type scalar and vector LQs, for which multiple NNs are trained at various values of the branching ratio into charged leptons, the exclusion limit at a certain point in the plane of signal mass and $\mathcal{B}$ is evaluated by first calculating the expected limits for all relevant NNs and then computing the observed limit for the NN with the best expected limit. The expected limit is determined by assuming the observed data in the SR to correspond to the post-fit prediction, whereas for the calculation of the observed limit the actual, measured data is used. The choice of NN at each point in the parameter space is illustrated in Figure 10.8 using up-type scalar LQs decaying into bottom quarks and muons or top quarks and neutrinos as an example.
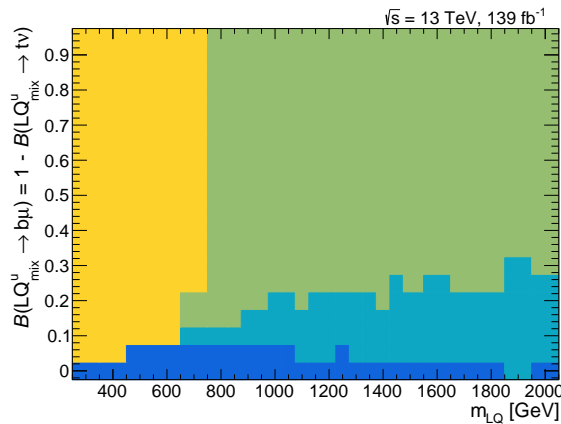
Figure 10.8: Map indicating the NN resulting in the best expected limit for up-type scalar LQs decaying into bottom quarks and muons or top quarks and neutrinos as a function of the signal mass and $\mathcal{B}$. Dark blue corresponds to the NN trained for $\mathcal{B} = 0.0$, light blue to the NN trained for $\mathcal{B} = 0.25$, green to the NN trained for $\mathcal{B} = 0.5$, and yellow to the NN trained for $\mathcal{B} = 0.9$.

The observed behaviour is similar for the other signal hypotheses. As anticipated, the training for $\mathcal{B} = 0.0$ results in the best expected limits for low branching ratios of $\mathcal{B} \leq 0.05$ across the whole LQ mass range. The training for $\mathcal{B} = 0.25$ is preferred for a range of branching ratios of $0.05 \leq \mathcal{B} \leq 0.3$, depending on the signal mass. For higher values of the branching ratio, the NN trained for $\mathcal{B} = 0.5$ yields the best expected limits at LQ masses above $800\,\text{GeV}$,

whereas below this threshold the NN trained for $\mathcal{B} = 0.9$ performs better. The choice of NN depends on the LQ mass because for high LQ masses events with both LQs decaying into a top quark and a neutrino can be separated from background contributions with relative ease due to the high energy content of signal events. It is therefore beneficial in this mass regime to include such events in the NN training. For lower LQ masses, however, only signal events with one LQ decaying into a bottom quark and a charged lepton can be easily distinguished from background processes. Hence, the NN utilised for low masses is trained primarily on such events.

Resulting upper cross-section limits for $\mathcal{B} = 0.5$ are shown in Figure 10.9 for scalar up- and down-type LQs for LQ masses between $300\,\mathrm{GeV}$ and $2000\,\mathrm{GeV}$. They decrease quickly with rising signal masses before reaching a plateau starting from around $m_{\mathrm{LQ}} = 1000\,\mathrm{GeV}$. This is due to kinematic differences between signal and background increasing towards higher LQ masses. This effect, however, becomes less significant for high masses as the separation between signal and background is near optimal above $1\,\mathrm{TeV}$. In fact, as discussed in Section 7.3, the signal selection efficiency begins to decline for high masses, resulting in approximately constant exclusion limits for high LQ masses. Overall, observed limits agree very well with the expectation and generally do not deviate by more than 1σ. Small kinks visible in the observed limits around $m_{\mathrm{LQ}} = 800\,\mathrm{GeV}$ for up-type LQs can be attributed to a switch of the underlying NN and therefore changes in the distributions from which the limits are obtained. From the upper cross-section limits a lower limit on the LQ mass can be deduced. For this purpose, the cross-section limit is compared to the theory cross-section, calculated at NNLO+NNLL and overlaid in the plots as a blue line. All signal masses to the left of the intersection of the theory curve and the cross-section limit are excluded. The observed lower mass limits therefore reach approximately $1460\,\mathrm{GeV}$ for up-type scalar LQs coupling to muons and $1440\,\mathrm{GeV}$ for up-type LQs coupling to electrons. For down-type LQs, slightly lower mass limits are obtained, with around $1370\,\mathrm{GeV}$ for those coupling to muons and $1390\,\mathrm{GeV}$ for those coupling to electrons. The slightly better results for up-type LQs can be attributed to two different factors. For one, events with both LQs decaying into a neutrino can still contribute to the SR in case of up-type LQs, as a charged lepton can stem from a top-quark decay. Secondly, up-type LQs decaying into a bottom quark and a lepton can be reconstructed with relative ease, whereas the reconstruction of down-type LQs decaying into a top quark and a lepton is less efficient, as it first necessitates the successful reclustering of the hadronically decaying top quark.

Analogously to the scalar case, upper cross-section limits for $\mathcal{B} = 0.5$ for vector LQs are shown in Figure 10.10 for the YM coupling scenario and for the minimal coupling scenario. Observed and expected limits agree well within the uncertainties. Again, small kinks can be observed at around $m_{\mathrm{LQ}} = 700\,\mathrm{GeV}$, where a switch between NNs takes place. For high signal masses, the cross-section limits for the two coupling scenarios are practically identical, as kinematic differences become negligble. Conversely, at low signal masses significantly

better cross-section limits are obtained for the minimal coupling scenario due to the discrepancies discussed in Section 2.2, which propagate through the reconstruction and cause large differences in signal acceptance. Observed lower mass limits, evaluated by comparing to the theory prediction at LO, reach 1980 GeV and 1900 GeV for vector LQs coupling to muons and electrons, respectively, assuming YM couplings. In the minimal coupling scenario, the observed limits are 1710 GeV for decays into muons and 1620 GeV for decays into electrons.
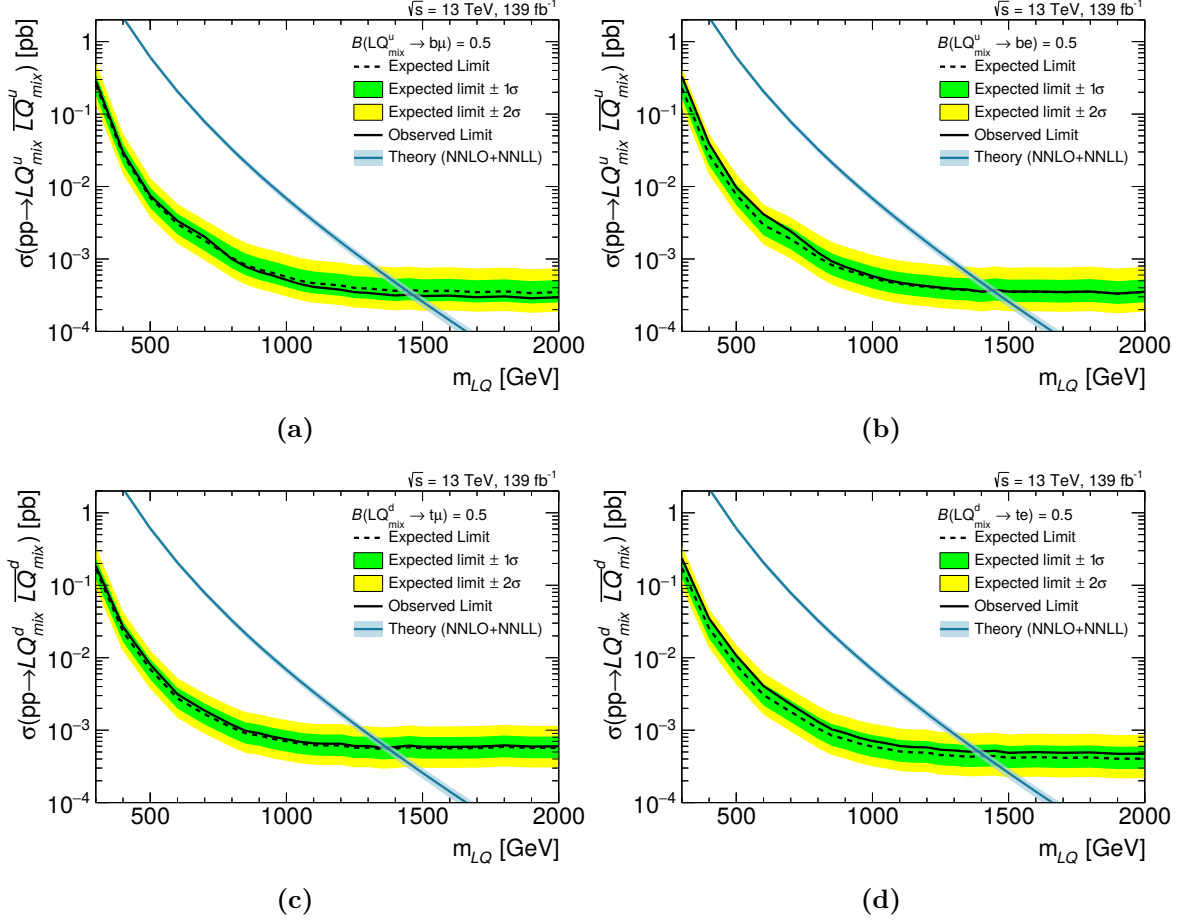
Figure 10.9: Expected (dashed black) and observed (solid black) upper 95 % CL limits on the cross-section of pair-produced scalar LQs, assuming $\mathcal{B} = 0.5$. The green (yellow) band shows the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainty region around the expected limit. The theoretical prediction and its $\pm 1\sigma$ uncertainty band, calculated from PDF+α_{S}, ISR, FSR, and scale variations, are shown in blue. Limits are presented for (a) up-type scalar LQs coupling to muons, for (b) up-type scalar LQs coupling to electrons, for (c) down-type scalar LQs coupling to muons, and for (d) down-type scalar LQs coupling to electrons.
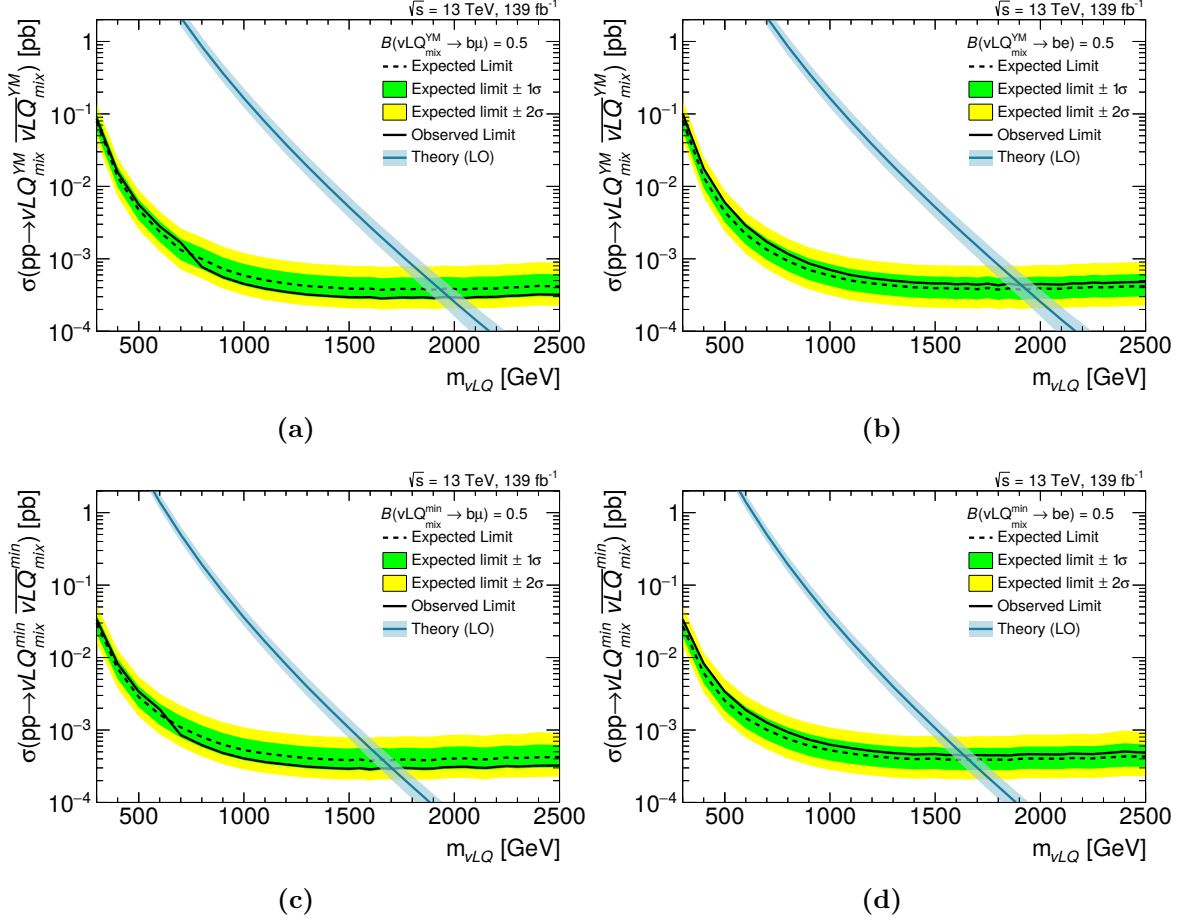
Figure 10.10: Expected (dashed black) and observed (solid black) upper 95 % CL limits on the cross-section of pair-produced vector LQs, assuming $\mathcal{B} = 0.5$. The green (yellow) band shows the $\pm 1\sigma\,(\pm 2\sigma)$ uncertainty region around the expected limit. The theoretical prediction and its $\pm 1\sigma$ uncertainty band, calculated from PDF+α_{S}, ISR, FSR, and scale variations, are shown in blue. Limits are presented for (a) vector LQs in the YM coupling scenario coupling to muons, for (b) vector LQs in the YM coupling scenario coupling to electrons, for (c) vector LQs in the minimal coupling scenario coupling to muons, and for (d) vector LQs in the minimal coupling scenario coupling to electrons.

Table 10.4: Expected and observed lower $95\,\%$ CL limits on the LQ mass at $\mathcal{B} = 0.5$ for the eight signal hypotheses considered in this analysis.

	Exp. limit [GeV]	Obs. limit [GeV]
$\text{LQ}^{\text{u}}_{\text{mix}} \to t\nu/b\mu$	1440^{+60}_{-60}	1460
$\text{LQ}^{\text{u}}_{\text{mix}} \to t\nu/be$	1440^{+60}_{-60}	1440
$\text{LQ}^{\text{d}}_{\text{mix}} \to t\mu/b\nu$	1380^{+50}_{-60}	1370
$\text{LQ}^{\text{d}}_{\text{mix}} \to te/b\nu$	1410^{+60}_{-60}	1390
$\text{vLQ}^{\text{YM}}_{\text{mix}} \to t\nu/b\mu$	1930^{+50}_{-60}	1980
$\text{vLQ}^{\text{YM}}_{\text{mix}} \to t\nu/be$	1930^{+50}_{-70}	1900
$\text{vLQ}^{\text{min}}_{\text{mix}} \to t\nu/b\mu$	1660^{+50}_{-50}	1710
$\text{vLQ}^{\text{min}}_{\text{mix}} \to t\nu/be$	1650^{+50}_{-60}	1620

The expected and observed lower exclusion limits on the signal mass assuming $\mathcal{B} = 0.5$ are summarised in Table 10.4 for all eight signal hypotheses considered in this analysis.

However, while the highest sensitivity is reached for $\mathcal{B} = 0.5$, exclusion limits are obtained for the whole range of $\mathcal{B}$. The observed two-dimensional cross-section exclusion limits for up- and down-type scalar LQs are shown in Figure 10.11. Similarly, the observed cross-section exclusion limits for vector LQs as a function of signal mass and $\mathcal{B}$ are illustrated in Figure 10.12. Curves indicating the corresponding mass limits are overlaid. Again, expected and observed limits agree very well. The largest deviations are obtained for scalar up-type LQs with $\mathcal{B} = 0.0$, for which already in the NN_{out} distributions upwards fluctuations in data were seen, as discussed in Section 10.3.
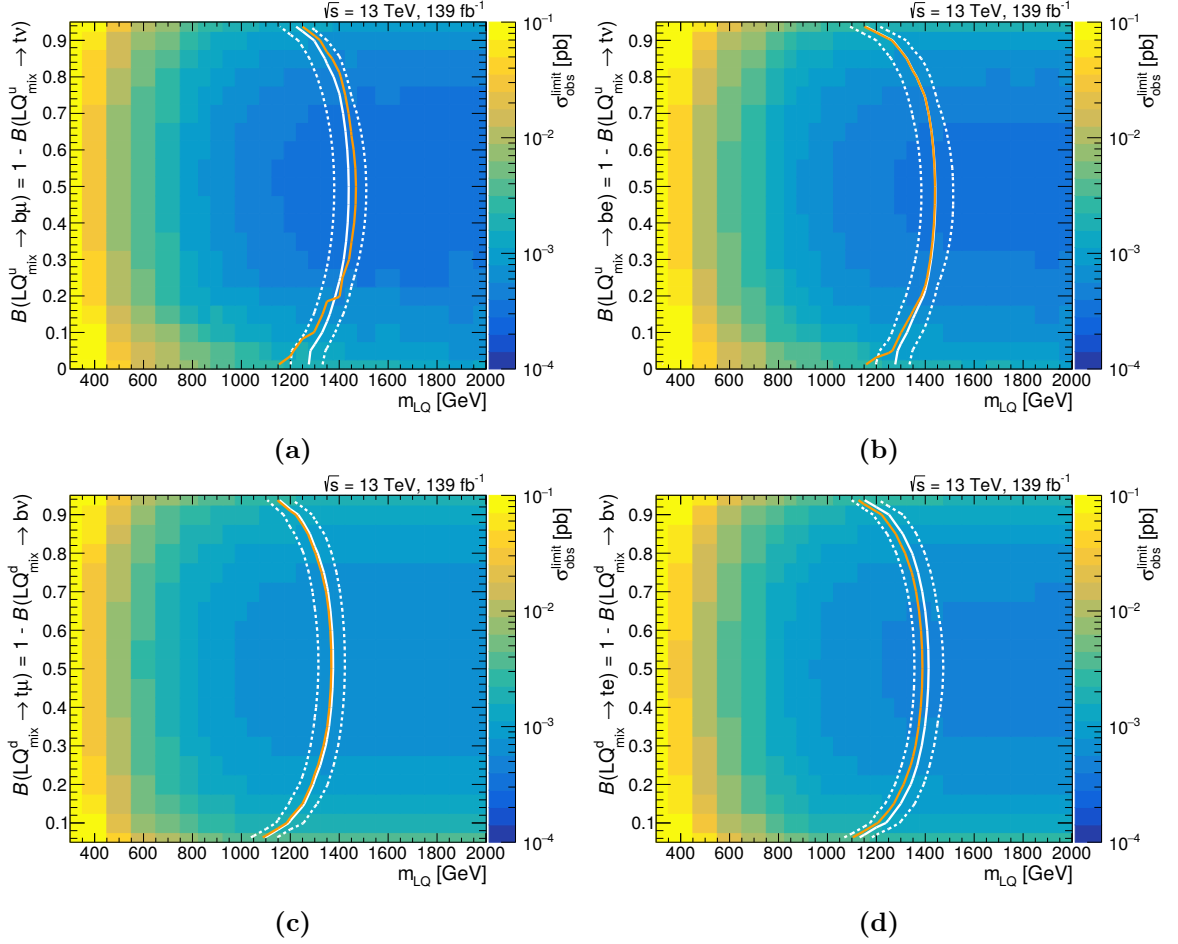
Figure 10.11: Expected (solid white, $\pm 1\sigma$ ranges dashed) and observed (solid red) exclusion limits on the leptoquark mass as a function of the branching ratio into charged leptons at 95 % CL. The observed upper limit on the signal cross-section in each bin is shown on the z-axis. Limits are presented for (a) up-type scalar LQs coupling to muons, for (b) up-type scalar LQs coupling to electrons, for (c) down-type scalar LQs coupling to muons, and for (d) down-type scalar LQs coupling to electrons. For up-type LQs the range in $\mathcal{B}$ is $0 - 0.95$, for down-type it is $0.05 - 0.95$.
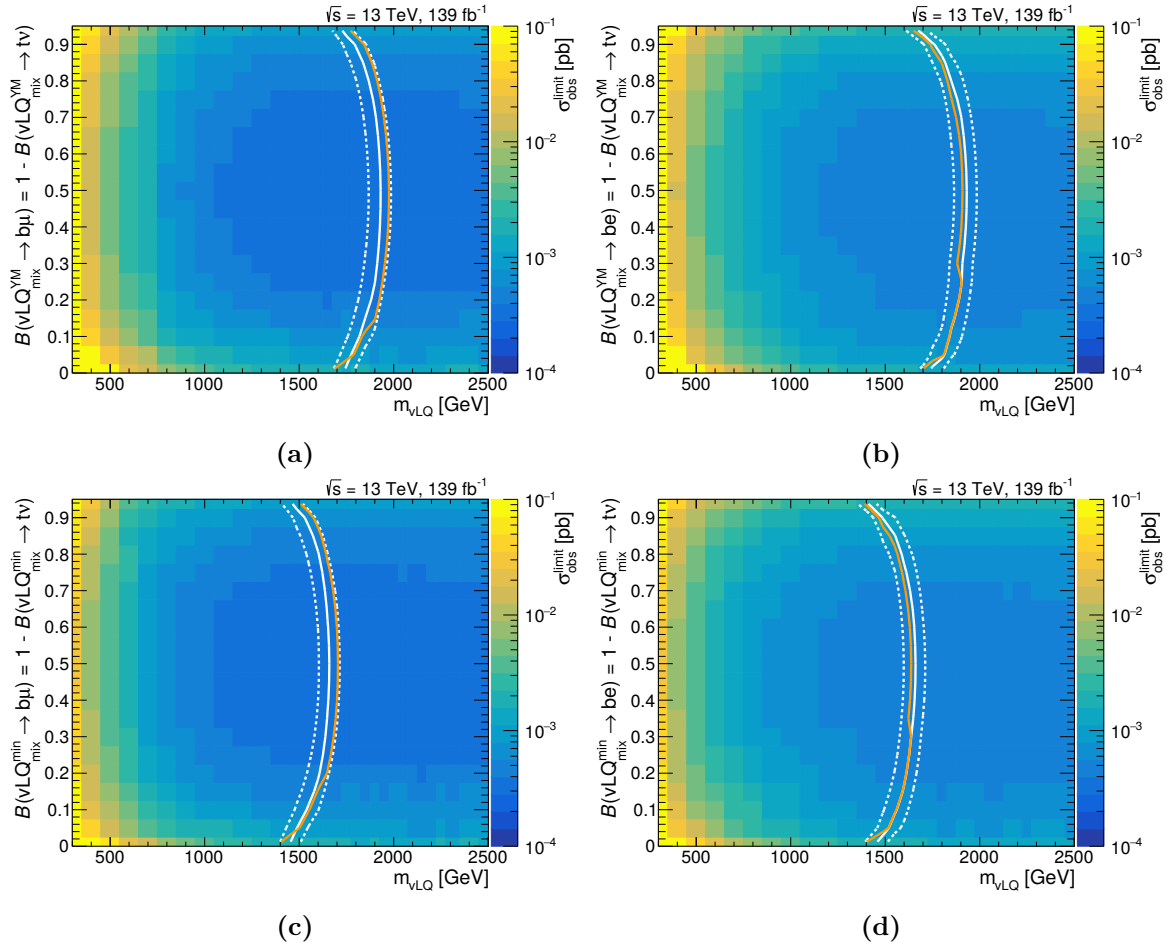
Figure 10.12: Expected (solid white, $\pm 1\sigma$ ranges dashed) and observed (solid red) exclusion limits on the leptoquark mass as a function of the branching ratio into charged leptons at 95 % CL. The observed upper limit on the signal cross-section in each bin is shown on the z-axis. Limits are presented for (a) vector LQs in the YM coupling scenario coupling to muons, for (b) vector LQs in the YM coupling scenario coupling to electrons, for (c) vector LQs in the minimal coupling scenario coupling to muons, and for (d) vector LQs in the minimal coupling scenario coupling to electrons. The range in $\mathcal{B}$ is $0 - 0.95$.

Finally, the limit setting procedure can also be performed without taking systematic uncertainties into account. This allows to quantify the impact of systematic uncertainties on the exclusion limits. The relative differences between exclusion limits with and without accounting for systematic uncertainties are shown in Figure 10.13 exemplarily for up-type scalar LQs coupling to top quarks and neutrinos or bottom quarks and muons. The differences are smaller than around 15 % for LQ masses above 1000 GeV. For lower signal masses, the impact of systematic uncertainties increases significantly. This behaviour is in agreement with the observations made in the context of the grouped impact in Section 10.2.
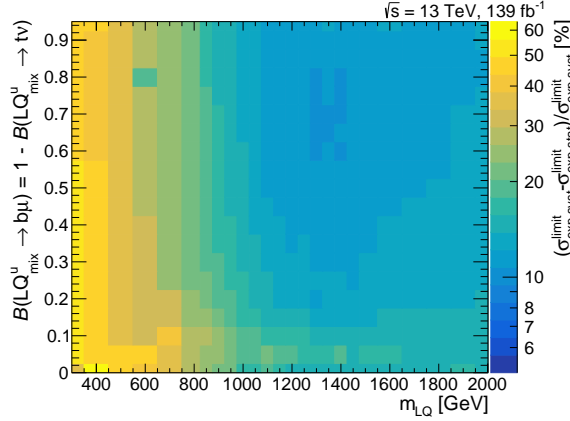
Figure 10.13: Relative differences between exclusion limits on scalar up-type LQs coupling to top quarks and neutrinos or bottom quarks and muons taking statistical and systematic uncertainties into account and exclusion limits taking only statistical uncertainties into account.

10.5 Comparison with Other Leptoquark Searches

Exclusion limits for LQs have been previously obtained in other analyses by the ATLAS and the CMS collaborations. In addition, results obtained for other signal hypotheses can be interpreted as limits for LQs. For instance, as illustrated in Figure 10.14, the search for pair-produced supersymmetric top partners, $\widetilde{t}$, decaying into top quarks and massless neutralinos, $\widetilde{\chi}_1^0$, in the single-lepton channel [196] is focussed on the same decay topology as the one presented here for $\mathcal{B} = 0.0$. In Ref. [196], an expected and observed exclusion limit of 1200 GeV is reported, based on the same dataset as used in this thesis, which is around 50 GeV better than the one observed in the analysis discussed in this thesis, but 50 GeV worse than the expected one. The 2σ excesses in data, discussed previously in Section 10.3, do not propagate to the limits obtained in the search for pair-produced supersymmetric partners of the top quark because the corresponding signal region is not used in the extraction of the exclusion limits.

Other searches using the same LQ hypotheses as the analysis at hand target scalar up-type LQs decaying exclusively into bottom quarks and electrons or muons, i.e. $\mathcal{B} = 1.0$ [93], as well as a search for supersymmetric top partners in the zero-lepton channel, which is reinterpreted in the context of up-type LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B}(\mathrm{LQ}_{\mathrm{mix}}^{\mathrm{u}} \to b\ell) = 0.0$ [103]. The former offers a high sensitivity for $\mathcal{B} \approx 1.0$, but results in competitive exclusion limits also for intermediate values of $\mathcal{B}$, due to the relatively clean
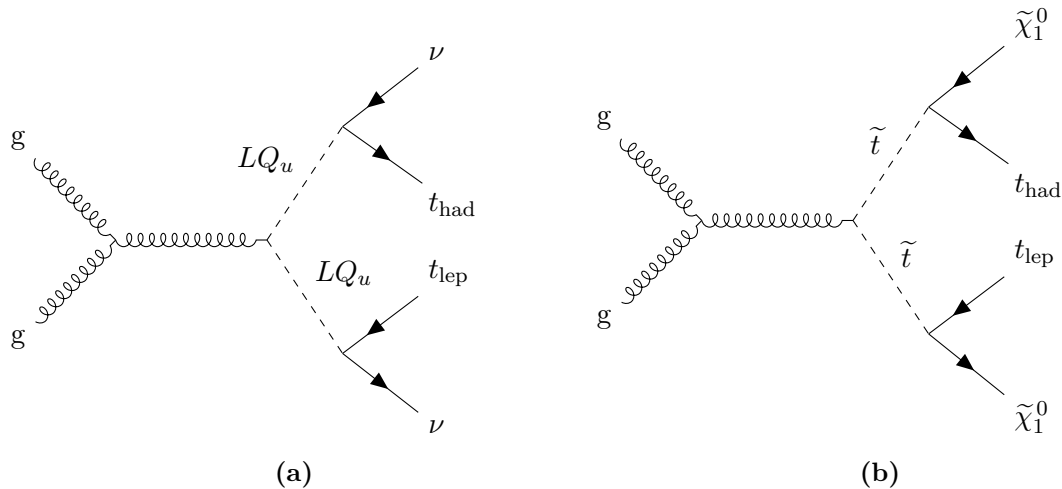
Figure 10.14: Feynman diagram of (a) pair-produced up-type LQs decaying into top quarks and neutrinos and of (b) pair-produced supersymmetric top partners decaying into top quarks and massless neutralinos.

dileptonic final state. As such, with lower mass limits of $1.48\,\mathrm{TeV}$ and $1.40\,\mathrm{TeV}$ in the case of scalar LQs decaying into bottom quarks and muons or electrons, respectively, the exclusion reach for $\mathcal{B} = 0.5$ is comparable to the one of the analysis presented in this dissertation. However, mass limits rapidly decrease towards lower values of the branching ratio into charged leptons. At very low values of $\mathcal{B}$, where the majority of LQs decays into a top quark and a neutrino, the sensitivity is driven both by events with one leptonically decaying and one hadronically decaying top quark and by events with both top quarks decaying hadronically. Limits competitive with the ones presented in the previous section can therefore be extracted from the search for supersymmetric top partners in the zero-lepton channel at $\mathcal{B} = 0$, resulting in an exclusion up to $m_{\mathrm{LQ}} = 1.24\,\mathrm{TeV}$. However, the sensitivity quickly declines towards higher values of $\mathcal{B}$. The analysis presented in this thesis therefore successfully covers a region of the parameter space not previously covered by complementary searches. In addition, it is the only analysis by the ATLAS collaboration providing exclusion limits on vector LQs decaying into third-generation quarks and first- or second-generation leptons.

Similarly, for down-type LQs, a search for LQs in the $t\ell t\ell$ channel with both top quarks decaying hadronically has been made public by the ATLAS collaboration [97]. It is most sensitive to high values of $\mathcal{B}$ and in the intermediate range around $\mathcal{B} \approx 0.5$ results in exclusion limits approximately $100\,\mathrm{GeV}$ to $150\,\mathrm{GeV}$ worse than those reported in Section 10.4. Conversely, for low values of $\mathcal{B}$, sensitivity is provided through a search in the final state containing two bottom quarks and two neutrinos [102]. Again, however, the sensitivity decreases rapidly towards higher branching ratios into charged leptons, such that the exclusion reach is significantly increased by the analysis presented here.

11 | Towards a Statistical Combination of Leptoquark Searches

Building on the findings presented in this dissertation and the comparison with complementary analyses searching for the same signal processes, in this chapter the prospects of a statistical combination are explored. The sensitivity can be improved by combining analyses targeting orthogonal final states in a simultaneous fit. Such a combination is described in the following for scalar up-type LQs, combining the statistical model of the search presented in this thesis with those from searches in final states containing two bottom quarks and two light leptons and in final states with two hadronically decaying top quarks and two neutrinos. The underlying implementation is based on the pyhf [197, 198] and the cabinetry [199] Python packages. First, a short introduction of the individual searches is given, focussing on differences to the analysis at hand, before discussing the combined results.

11.1 Searches for Leptoquarks in Final States with Zero or Two Leptons

The analysis described in Ref. [93] is dedicated to pairs of LQs decaying into an electron or a muon in association with a non-top quark. For the combination discussed here, the relevant final state constitutes two bottom quarks and two charged leptons of the same flavour. The analysis is therefore referred to as $b\ell b\ell$ in the following. Selected events must contain exactly two electrons or muons with opposite electric charges and transverse momenta of $p_\mathrm{T}(\ell) > 27\,\mathrm{GeV}$. Here, electrons have to fulfil the requirements of the *MediumLH* identification working point, which places looser restrictions on the likelihood discriminant than the TightLH working point mentioned in Section 5.2. Muons have to pass the Medium working point for transverse momenta below $800\,\mathrm{GeV}$ and the High-p_T working point above. In addition, events are required to contain at least two reconstructed small-R jets with transverse momenta of $p_\mathrm{T}(j) > 45\,\mathrm{GeV}$, reconstructed from energy deposits in the calorimeter system. Such jets are referred to as *EMTopo* jets and differ from the PFlow jets used in the analysis presented in this thesis, as they do not take tracking information into account. Following the selection described in Ref. [93], events are split into three categories based on their b-jet multiplicity, vetoing events with more than two b-jets. Importantly, here the *MV2c10 b*-tagging algorithm is used [149, 200]. Subsequently, in each event, two LQ candidates are reconstructed from the two leptons and two b-jets, using the leading-p_T jets as substitute

in cases where fewer than two b-jets are available. As the assignment of leptons and b-jets is ambiguous, the combination resulting in the smallest difference in the invariant mass of lepton and jet, $m_{\ell,j}$, is chosen. Events in the SR have to have $m_{\mathrm{asym}} < 0.2$, where the mass asymmetry is defined as

$$m_{\mathrm{asym}} = \frac{m_{\ell,j}^{\max} - m_{\ell,j}^{\min}}{m_{\ell,j}^{\max} + m_{\ell,j}^{\min}}. \tag{11.1}$$

Here, the reconstructed masses of the two LQ candidates are sorted such that $m_{\ell,j}^{\max} > m_{\ell,j}^{\min}$. The sideband region with $0.2 < m_{\mathrm{asym}} < 0.4$ is used as a Z+jets CR (*CRDY*) to constrain the corresponding normalisation factor. An additional $t\bar{t}$ CR (*CRtop*) is defined by requiring opposite lepton flavours and disregarding events without a reconstructed b-jet. The event yields in data and MC after a background-only fit are presented in Figure 11.1 and reproduce the results presented in Ref. [93] well. A small, but negligible overlap on the order of $1\,\%$ with respect to the analysis discussed in this thesis is observed in the selected events, primarily caused by the fact that in the $b\ell b\ell$ search muons are reconstructed for pseudorapidities of $|\eta| < 2.7$ instead of 2.5.
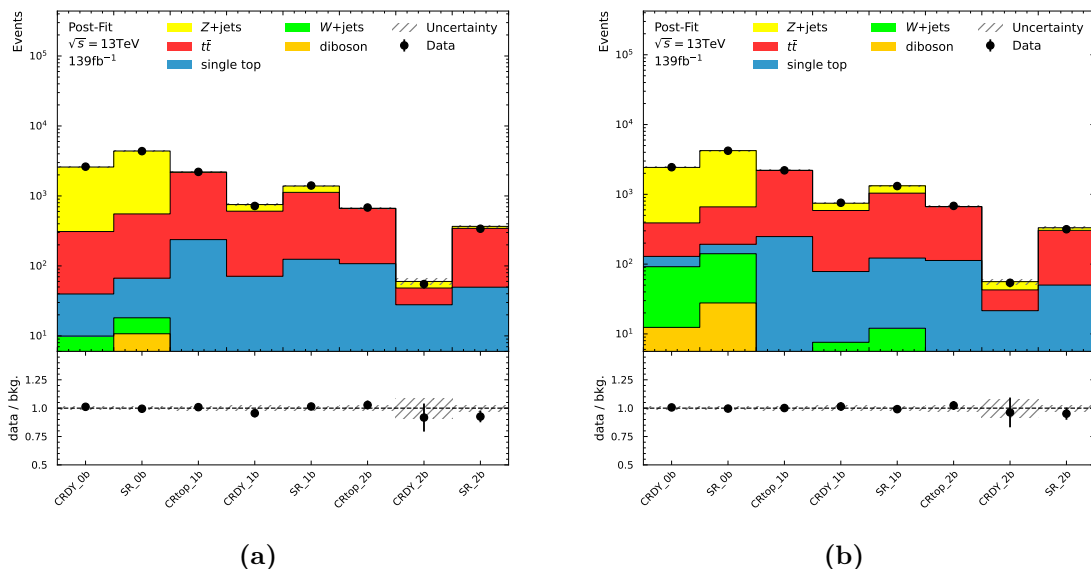
Figure 11.1: Data and MC yields after a simultaneous background-only fit in the CRs and SRs of the $b\ell b\ell$ analysis for (a) the $b\mu b\mu$ channel and for (b) the *bebe* channel. The regions are split by b-jet multiplicity as indicated by the respective suffix.

The analysis targeting events with both LQs decaying into a top quark and a neutrino is described in Ref. [103]. It is referred to as stop-0ℓ in the following. Both top quarks are assumed to decay hadronically, resulting in a final state without any charged leptons. Again, small-R jets are reconstructed as EMTopo jets from energy deposits in the calorimeters and jets originating from B hadron decays are identified with the MV2c10 algorithm. A set of six orthogonal SRs (*SRATT, SRATW, SRAT0, SRBTT, SRBTW, SRBT0*) is defined, for all of

which a veto on reconstructed leptons and a requirement of $E_T^{miss} > 250\,\mathrm{GeV}$ are imposed. Among other criteria on the jet content, the leading-p_T reclustered $R = 1.2$ jet is required to have a mass of $m_1^{R=1.2} > 120\,\mathrm{GeV}$. The SRs are then divided into two categories according to the m_{T2}-based m_{T2,χ^2} variable, which is calculated from top-quark and W-boson candidates obtained by minimising

$$\chi^2 = \frac{(m_{cand} - m_{true})^2}{m_{true}}, \tag{11.2}$$

with the true top-quark or W-boson mass, respectively, m_{true}, and with the candidate jet's mass, m_{cand}. For SRA, a requirement of $m_{T2,\chi^2} > 450\,\mathrm{GeV}$ is applied, whereas for SRB the criterion is inverted. Both sets of events are further divided into three subcategories each, based on the mass of the subleading-p_T reclustered $R = 1.2$ jet, $m_2^{R=1.2}$. The so-called TT category contains events with $m_2^{R=1.2} > 120\,\mathrm{GeV}$, i.e. with a second reconstructed top-quark. Similarly, events in the TW category fulfil $60\,\mathrm{GeV} < m_2^{R=1.2} < 120\,\mathrm{GeV}$, corresponding to a reconstructed W boson. Finally, events in the $T0$ category contain no second reconstructed heavy particle, i.e. $m_2^{R=1.2} < 60\,\mathrm{GeV}$.

The normalisations of the major background processes are constrained in dedicated CRs. For this purpose, separate CRs are defined for the $t\bar{t}$ ($CRttbarAB$), the single top-quark ($CRSTAB$), the W+jets ($CRWAB$), the Z+jets ($CRZABT0$, $CRZABTTW$). and the $t\bar{t} + Z$ processes ($CRttZ$). Due to limited statistics, shape information is not taken into consideration, and each region consists of a single bin only. The corresponding event yields in SRs and CRs after a background-only fit to data are shown in Figure 11.2. Slight, but negligible overlaps of approximately $1\,\%$ with respect to the analysis discussed in this thesis arise due to differences in the handling of the overlap removal procedure.
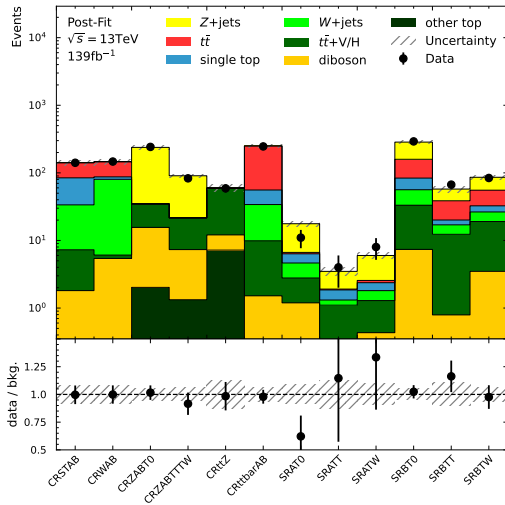
Figure 11.2: Data and MC yields after a simultaneous background-only fit in the CRs and SRs of the stop-0ℓ analysis. Minor background processes involving top quarks are summarised as *other top*.

11.2 Combined Exclusion Limits

The combined statistical model of the three analyses consists of 14 CRs and 10 SRs. When combining the statistical models corresponding to the individual analyses, nuisance parameters can be treated either as uncorrelated or they can be correlated across the analyses. While the former corresponds to a conservative approach, not taking the correlations into account can also mean not making use of the full power of the profile likelihood fit. However, because the analyses cover vastly different regions of phase space, it is not sensible to blindly correlate all nuisance parameters. In particular, theoretical modelling uncertainties can exhibit pulls or constraints in one region of phase space that are not applicable to other regions. Thus, only the impact of correlating or not correlating experimental systematic uncertainties on the JER, the E_T^{miss} soft term, the lepton calibration, and the pile-up reweighting is studied. As the corresponding nuisance parameters are not strongly pulled or constrained by any of the three analyses, no significant effects are expected. Small changes in the normalisation factors are observed in the combined background-only fit, as summarised in Table 11.1. Separate normalisation factors for the same background processes are used across the different analyses, because the targeted phase spaces differ. However, very similar behaviour is observed in particular for normalisation factors obtained in the stop-0ℓ analysis and in the single-lepton analysis, termed $t\nu b\ell$ in the following. Both analyses cover very similar regions of phase space in which the effects of interference between $t\bar{t}$ and single top-quark production are important and mismodelled. Therefore, the normalisation factor of the single top-quark background is significantly smaller than unity in both channels, whereas the $t\bar{t}$ normalisation is above unity. On the other hand, the dilepton analysis is not affected by the interference, and thus the normalisation of the $t\bar{t}$ background is very close to unity.

As Table 11.1 shows, the the majority of changes between treating experimental systematics as correlated or uncorrelated across analyses is smaller than $1\,\%$, with the largest deviations of up to $5\,\%$ observed in the normalisation of the $t\bar{t}$ and the W+jets background in the $t\nu b\ell$ analysis. While the slight changes in normalisation propagate to the upper cross-section limits, the deviations are small, and for all LQ masses considered the results differ by around $2\,\%$ or less when correlating or not correlating the experimental systematics. This is exemplified in Figure 11.3 for scalar up-type LQs with $\mathcal{B} = 0.5$ in the muon channel and in the electron channel. In particular, for electrons the discrepancies can be attributed to statistical fluctuations, as the ratio between the two approaches is close to unity for all masspoints.

Table 11.1: Normalisation factors obtained in a background-only fit of the combined statistical model with correlated and with uncorrelated experimental systematic uncertainties. For the fit, the SRs associated with scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$ are used. The normalisation factors are denoted as μ, with the index indicating the corresponding process and the exponent the associated analysis.

Normalisation factor	uncorrelated	correlated
$\mu_{t\bar{t}}^{b\ell b\ell}$	1.019 ± 0.057	1.020 ± 0.057
$\mu_{Z+\text{jets}}^{b\ell b\ell}$	1.061 ± 0.095	1.062 ± 0.095
$\mu_{t\bar{t}}^{t\nu b\ell}$	1.20 ± 0.23	1.13 ± 0.22
$\mu_{W+\text{jets}}^{t\nu b\ell}$	0.88 ± 0.12	0.91 ± 0.11
$\mu_{\text{single top}}^{t\nu b\ell}$	0.48 ± 0.26	0.49 ± 0.26
$\mu_{t\bar{t}}^{\text{stop-0}\ell}$	1.15 ± 0.20	1.14 ± 0.20
$\mu_{t\bar{t}+Z}^{\text{stop-0}\ell}$	0.85 ± 0.15	0.84 ± 0.14
$\mu_{Z+\text{jets}}^{\text{stop-0}\ell}$	1.054 ± 0.085	1.054 ± 0.085
$\mu_{W+\text{jets}}^{\text{stop-0}\ell}$	0.72 ± 0.24	0.72 ± 0.24
$\mu_{\text{single top}}^{\text{stop-0}\ell}$	0.66 ± 0.30	0.68 ± 0.30
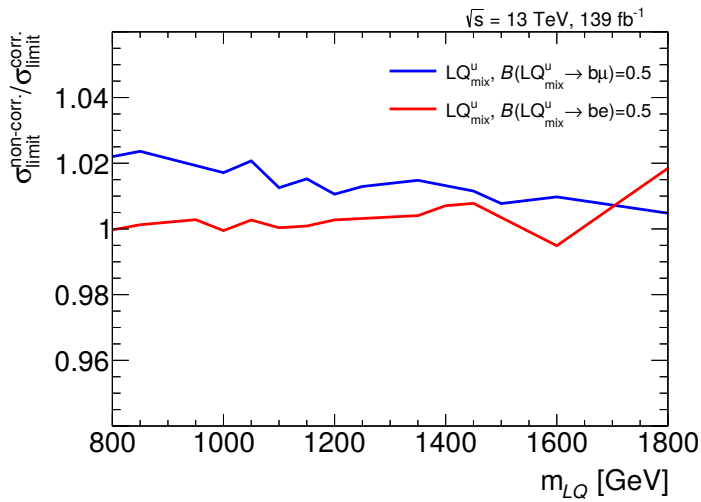
Figure 11.3: Ratio of observed upper limits on the signal production cross-section when correlating or not correlating nuisance parameters associated with experimental systematic uncertainties. Results for LQs decaying into electrons or muons and bottom quarks or into top quarks and neutrinos with $\mathcal{B} = 0.5$ are shown.

Since the observed differences between the approaches are small, in what follows limits are evaluated without taking correlations between nuisance parameters of different analyses into consideration. Observed and predicted event yields after a combined background-only fit are presented in Figure 11.4 for the electron and the muon channel. No relevant deviations from the individual background-only fits are found, validating the stability of the statistical model.
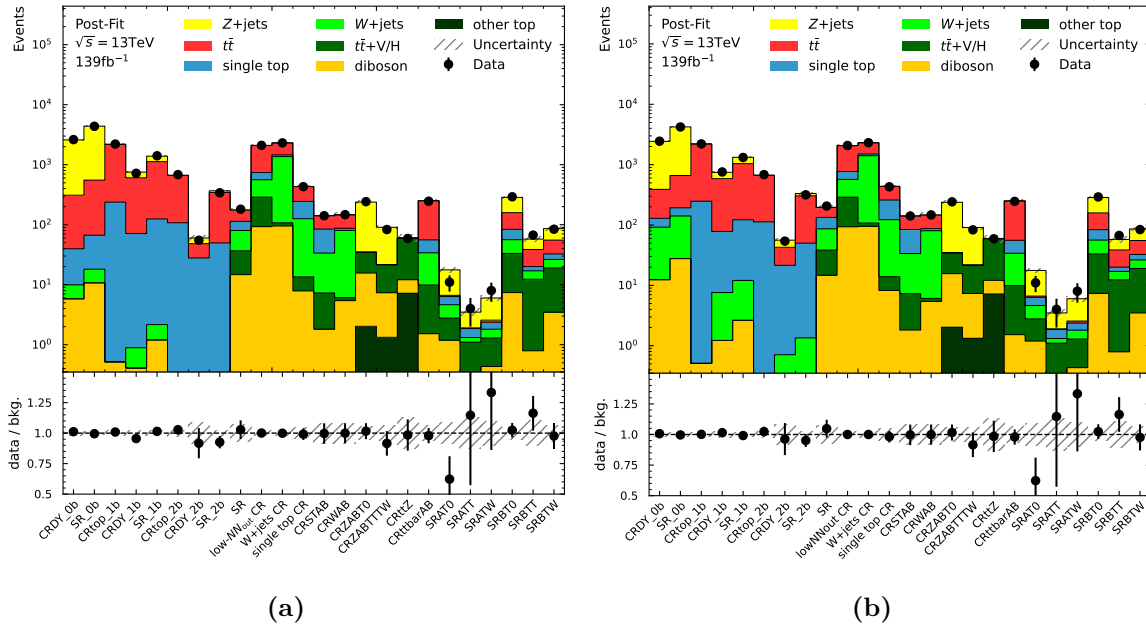
Figure 11.4: Comparison of data and MC yields after a simultaneous background-only fit in the CRs and SRs in (a) the muon channel and in (b) the electron channel. Minor background processes involving top quarks are summarised as *other top*.

Having confirmed the validity of the combined model, exclusion limits are derived. Improvements compared to the individual analyses are primarily expected for values of the branching ratio of leptoquarks decaying into charged leptons for which two analyses provide sensitivity. Such is the case for $\mathcal{B} \approx 0.5$, where the $b\ell b\ell$ analysis and the $t\nu b\ell$ analysis offer similar exclusion reaches. As Figure 11.5 shows, the combined analysis results in expected lower mass limits of above $1500\,\text{GeV}$ both for LQs coupling to muons and LQs coupling to electrons. The observed limits are improved by around $40\,\text{GeV}$ to $100\,\text{GeV}$ and reach $1560\,\text{GeV}$ in the muon channel and $1480\,\text{GeV}$ in the electron channel.

Finally, limits are evaluated for the whole plane of m_{LQ} and $\mathcal{B}$. The results are displayed in Figure 11.6 and compared to the limits obtained by the individual analyses. Small differences, potentially caused by different interpolation schemes, arise between the combined results and the individual analyses even in ranges of $\mathcal{B}$ where only one analysis dominates the sensitivity. However, qualitatively the general behaviour of the combined exclusion limit is as expected. The enhanced sensitivity around $\mathcal{B} = 0.5$ is clearly visible. Improvements can also be seen for low values of $\mathcal{B}$, where the stop-0ℓ analysis and the $t\nu b\ell$ analysis contribute, showcasing the advantages of combining the statistical models of complementary analyses.
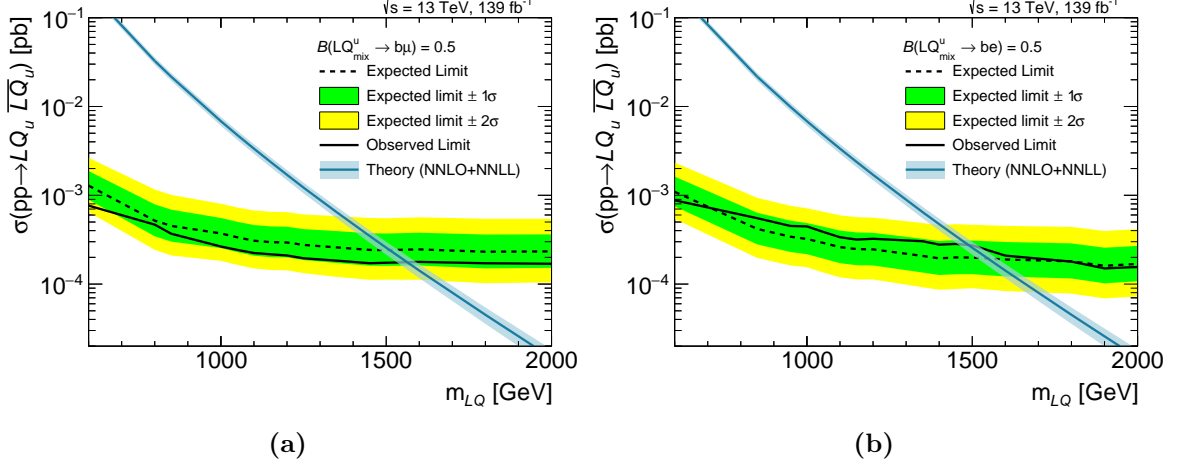
Figure 11.5: Expected and observed 95 % CL exclusion limits on the cross-section for (a) up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$ and for (b) up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.5$ as a function of the LQ mass, m_{LQ}.
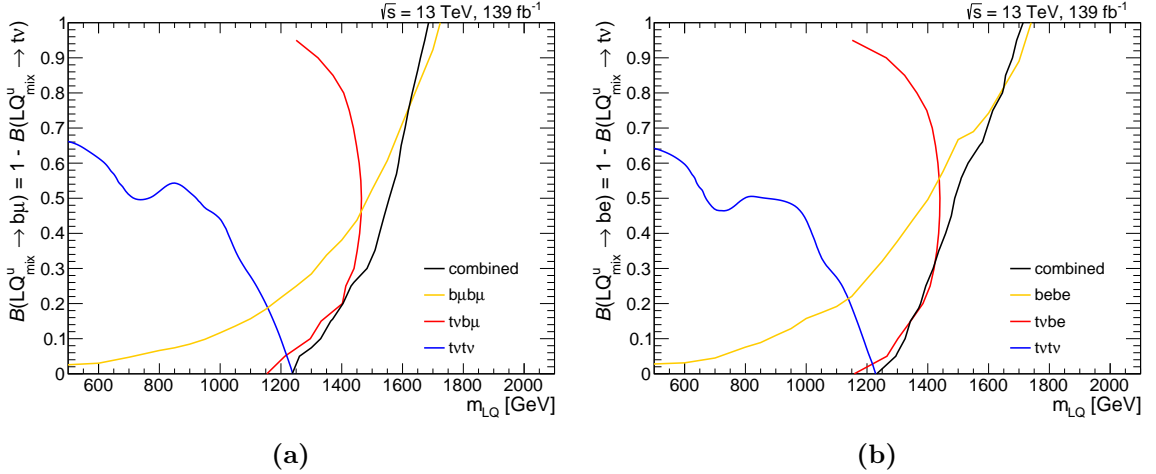

Figure 11.6: Observed 95 % CL exclusion limits on the LQ mass, m_{LQ}, as a function of the branching ratio into charged leptons for (a) up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons and for (b) up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and electrons.

Conclusion and Outlook

In this dissertation, a search for pair-produced leptoquarks in proton-proton collisions at a centre-of-mass energy of 13 TeV is presented. The analysis utilises the full Run 2 dataset collected with the ATLAS detector in the years 2015 to 2018, corresponding to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. In particular, leptoquarks coupling to quarks of the third generation and to leptons of the first or the second generation are targeted, motivated by hints at lepton flavour universality violation.

This analysis represents the first search for such leptoquarks in events containing exactly one charged electron or muon. It is most sensitive to events with one leptoquark decaying into a charged lepton and the other leptoquark decaying into a neutrino, but also covers events with two up-type leptoquarks decaying into a top quark and a neutrino each. The analysis therefore complements previous searches in the zero-lepton and the two-lepton channel.

In addition to a single charged lepton, events of interest contain large amounts of missing transverse momentum, and at least four jets, at least one of which is identified as originating from a bottom quark. A kinematic reweighting procedure is applied on events containing singly- and pair-produced top quarks as a function of the effective mass and the jet multiplicity in order to compensate for mismodelling at large top quark momenta due to higher-order corrections missing in the simulation. Monte Carlo simulations incorporating next-to-next-to-leading-order contributions could thus lead to improvements.

Neural networks are trained for several signal hypotheses to ensure good discriminating power against background processes. The analysis relies on a binned profile likelihood fit of the neural network output distribution and additionally includes control regions to constrain the normalisation for the major backgrounds, namely from top-quark–antitop-quark, single top-quark, and W+jets production. The compatibility of observed data with the Standard Model prediction is checked. In the absence of significant deviations, exclusion limits on the signal cross-section and the signal mass are derived. In total, eight signal models are considered. The limits depend on the branching ratio of the leptoquark into charged leptons, with observed limits on the leptoquark mass of up to 1460 GeV and 1440 GeV in the case of up-type leptoquarks coupling to muons and electrons, respectively. For scalar down-type leptoquarks, a maximum exclusion reach of 1370 GeV and 1390 GeV is achieved for leptoquarks coupling to muons and electrons.

For the first time, dedicated neural networks are trained also for vector leptoquarks. Corresponding exclusion limits reach as high as 1980 GeV and 1900 GeV in the Yang-Mills coupling

scenario with decays into muons and electrons, respectively. In the minimal coupling scenario, vector leptoquarks with masses of up to 1710 GeV are excluded in the muon channel and up to 1620 GeV in the electron channel.

While the analysis is primarily statistically limited, a better understanding of the systematic uncertainties presents a potential avenue for improvement. In particular, the treatment of the interference between top-quark–antitop-quark production and the production of a single top quark in association with a W boson and a bottom quark leads to large uncertainties. A reduction of these uncertainties could be achieved using a common simulation of $WbWb$ events, automatically taking the interference effects into account [201].

The sensitivity can be further enhanced through a statistical combination with complementary searches. Prospects for such a combined analysis are presented with a focus on scalar up-type leptoquarks. Exclusion limits on the leptoquark mass can potentially be improved by around 100 GeV, assuming equal decay rates of the leptoquarks into charged leptons and bottom quarks or neutrinos and top quarks, respectively. The analysis presented in this thesis is particularly suitable for the combination, as it provides a link between the zero-lepton and the two-lepton channel and drives the combined exclusion limits across a large range of low and intermediate values of the branching ratio of the leptoquark decaying into charged leptons. In the future, similar improvements are expected in the search for down-type leptoquarks, as a combination will be possible with analyses targeting down-type leptoquarks exclusively decaying into top quarks and charged leptons.

In light of the start of Run 3 in 2022, enhanced production cross-sections for the signal processes due to the increased centre-of-mass energies of 13.6 TeV promise great opportunities for future leptoquark searches. Building on the findings presented in this dissertation, future analysis efforts can target even higher mass exclusion limits and extend the leptoquark search programme to more specific signal models.

A | **Appendix**

In this section, additional figures related to the NNs are collected, complementing those shown in the main body of the thesis. Distributions of input variables to the training not part of the representative subset examined in Section 7.3.1 are included in Appendix A.1.

Correlations among all of the input variables are shown in Appendix A.2 for data, the combined background processes and a representative signal of scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$.

Distributions of NN_{out} in the training region prior to the profile likelihood fit can be found in Appendix A.3.

Distributions of NN input variables in the CRs are shown in Appendix A.4.

Lastly, distributions of NN_{out} in the signal region after the profile likelihood fit are shown in Appendix A.5.

A.1 Neural Network Input Variables in the Training Region

(a)

(b)

(c)

(d)
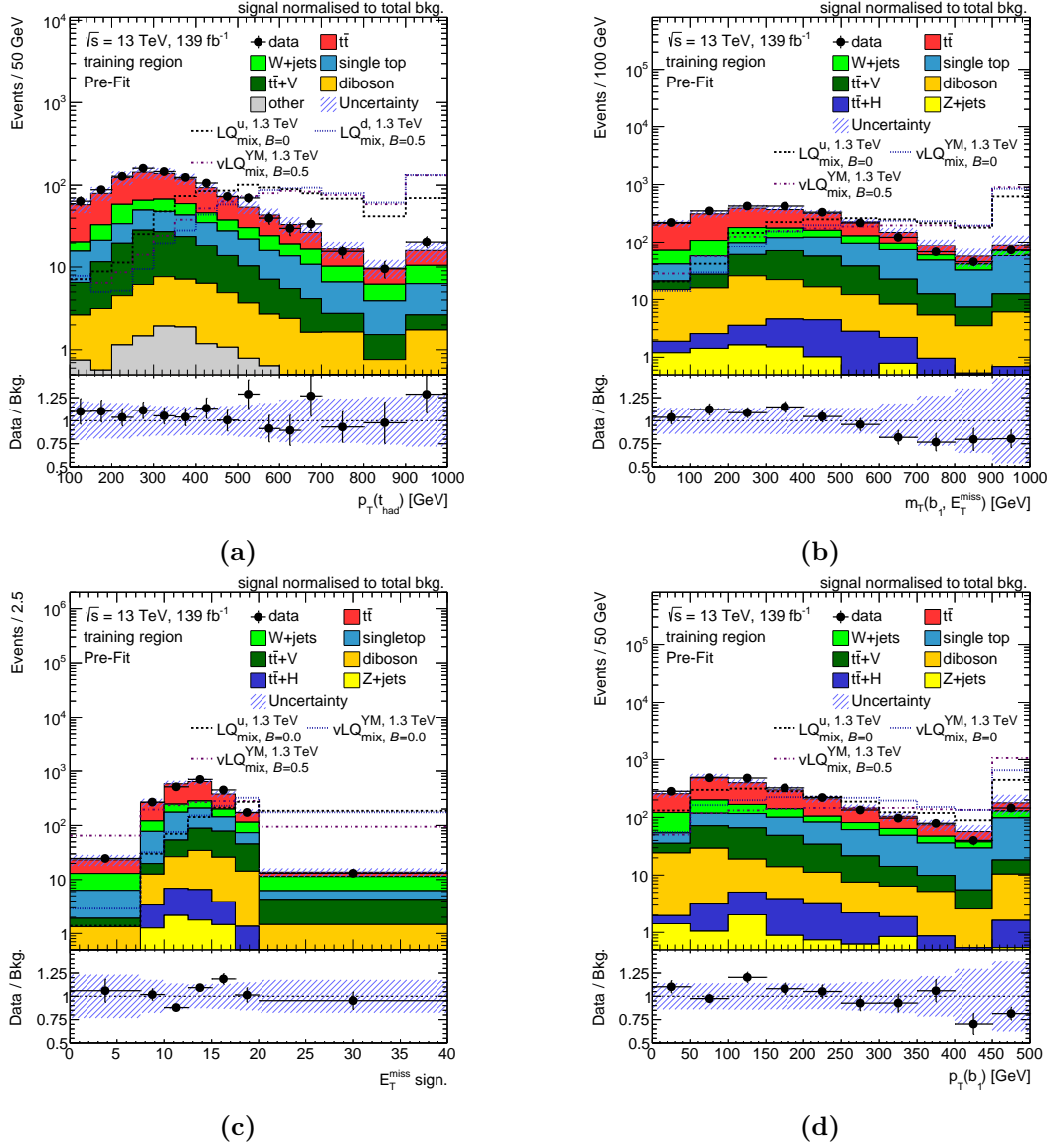
Figure A.1: Distributions of exemplary discriminating variables, namely (a) the transverse momentum of the reclustered hadronically decaying top quark, $p_{\mathrm{T}}(t_{\mathrm{had}})$, (b) the transverse mass of the leading-p_{T} b-jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, $m_{\mathrm{T}}(b_1, E_{\mathrm{T}}^{\mathrm{miss}})$, (c) the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, and (d) the transverse momentum of the leading-p_{T} b-jet, $p_{\mathrm{T}}(b_1)$. The figures show the SM background expectation in comparison with the observed data. The hatched bands include statistical and systematic uncertainties. The ratios of observed and expected background events are shown in the bottom panels of the figures. The last bin contains the overflow. Three exemplary signal processes, normalised to the background expectation, are overlaid as dashed lines.
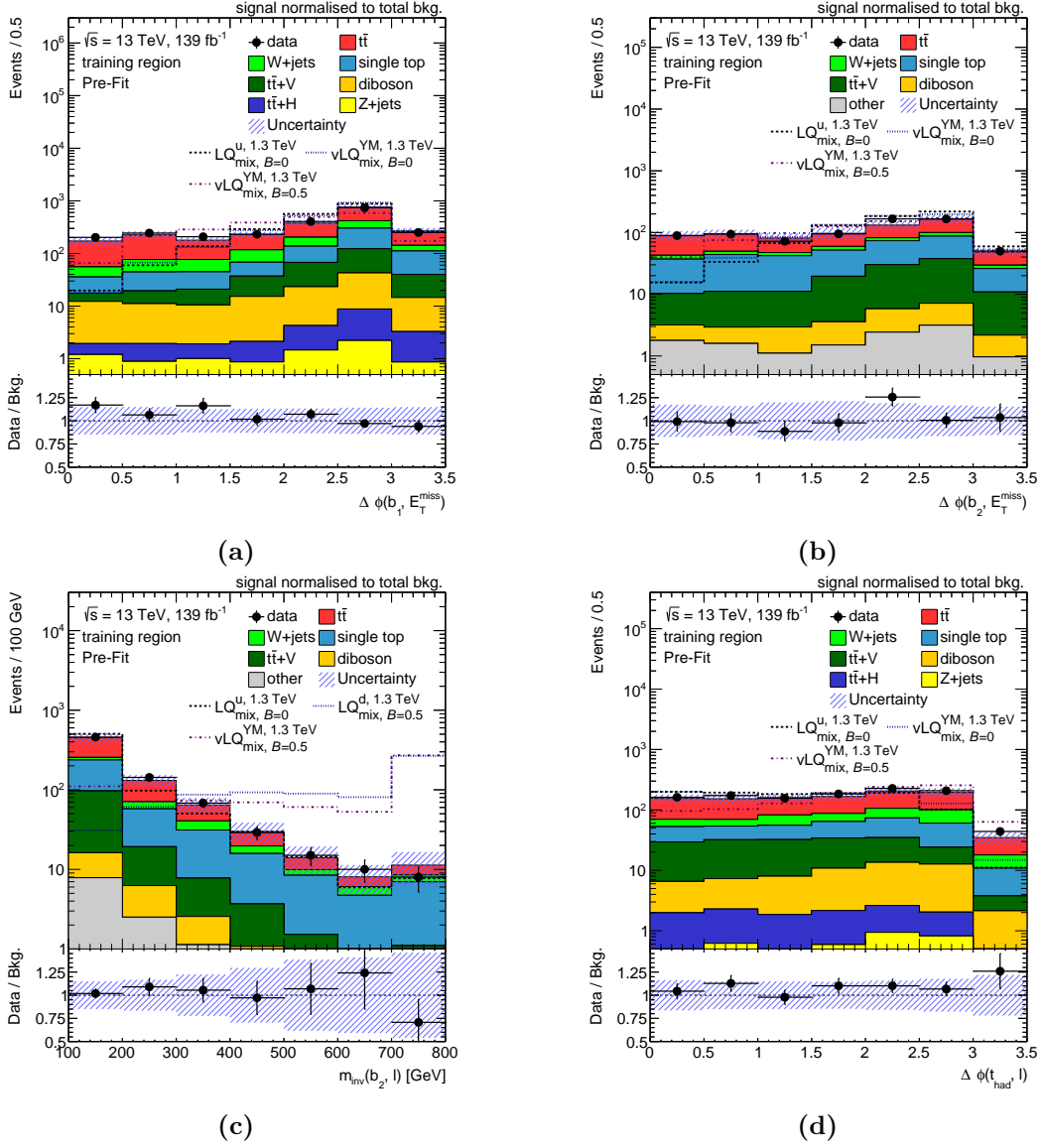
Figure A.2: Distributions of exemplary discriminating variables, namely (a) the azimuthal separation between the leading-p_{T} b-jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(b_1, E_{\mathrm{T}}^{\mathrm{miss}})$, (b) the azimuthal separation between the subleading-p_{T} b-jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(b_2, E_{\mathrm{T}}^{\mathrm{miss}})$, (c) the invariant mass of the subleading-p_{T} b-jet and the lepton, $m_{\mathrm{inv}}(b_2, \ell)$ and (d) the azimuthal separation between the reclustered hadronically decaying top quark and the lepton, $\Delta\phi(t_{\mathrm{had}}, \ell)$. The figures show the SM background expectation in comparison with the observed data. The hatched bands include statistical and systematic uncertainties. The ratios of observed and expected background events are shown in the bottom panels of the figures. The last bin contains the overflow. Three exemplary signal processes, normalised to the background expectation, are overlaid as dashed lines.

135

A.2 Correlations between Input Variables
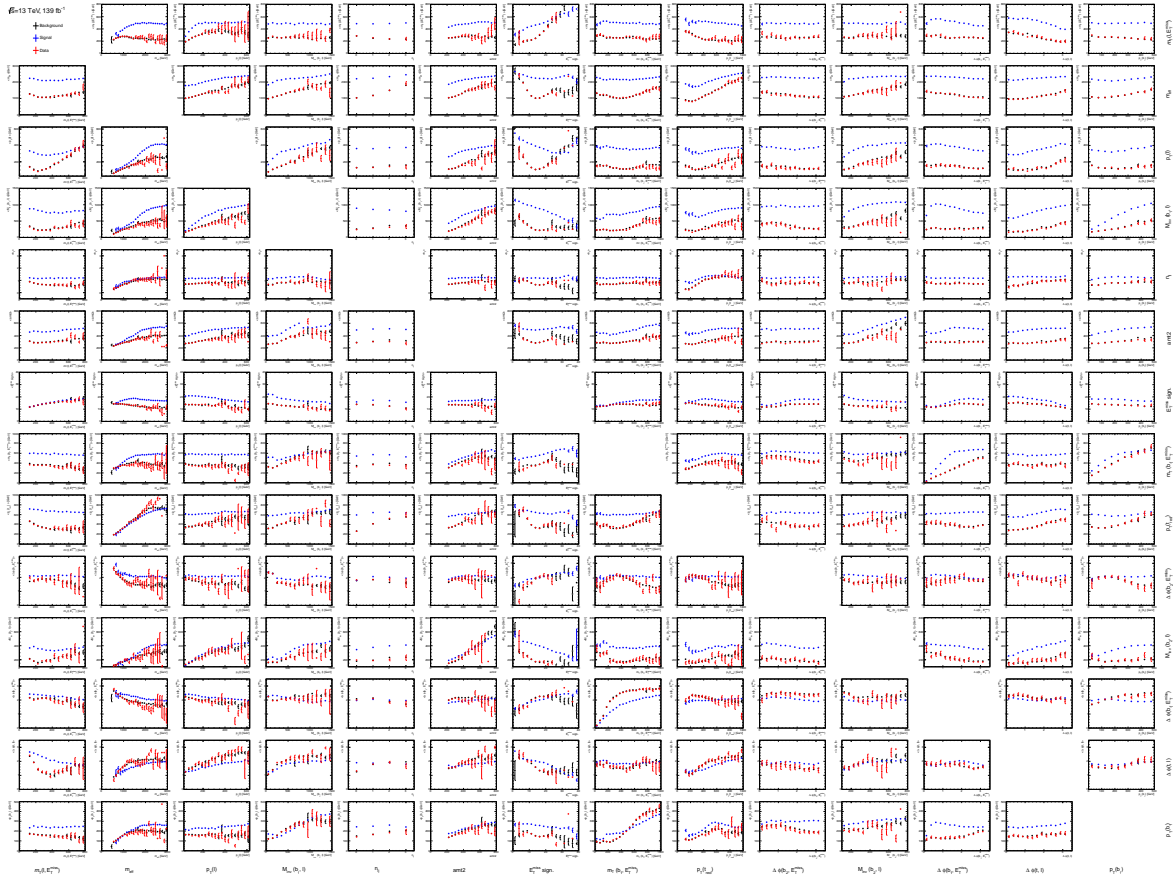
Figure A.3: Profile plots visualising the correlations between the NN input variables. MC simulated SM background events are shown in black, signal events in blue, and data in red. Here, scalar up-type LQs with a mass of $m_{\mathrm{LQ}} = 1300\,\mathrm{GeV}$ and $\mathcal{B}(\mathrm{LQ}^{\mathrm{u}}_{\mathrm{mix}} \to b\mu) = 0.5$ are shown as signal.

A.3 Pre-Fit Neural Network Output Distributions in the Training Region
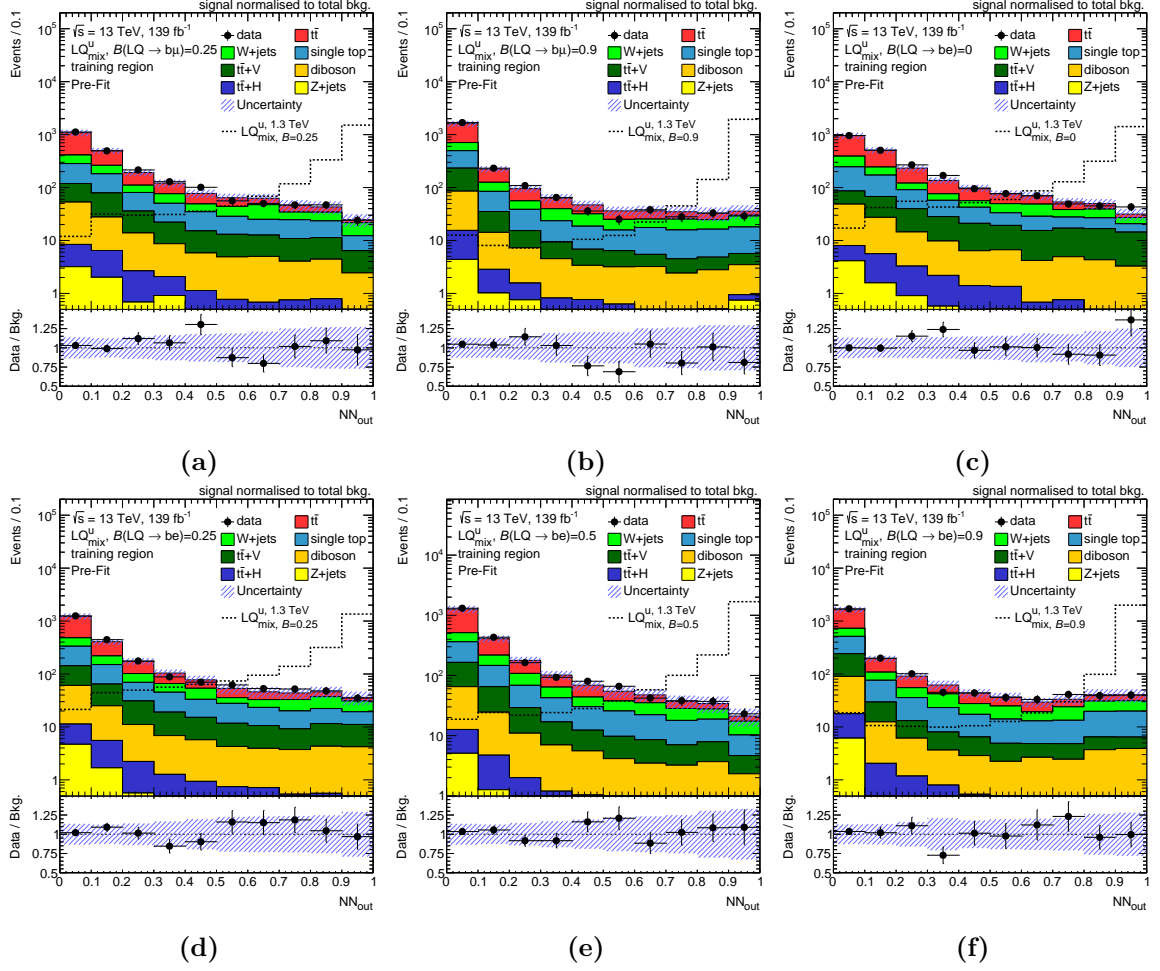
Figure A.4: Distributions of NN_{out} obtained from NNs trained (a) on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.25$, (b) on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.9$, (c) on scalar up-type LQs decaying into top quarks and neutrinos with $\mathcal{B} = 0$, (d) on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.25$, (e) on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.5$, and (f) on scalar up-type LQs decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.9$.
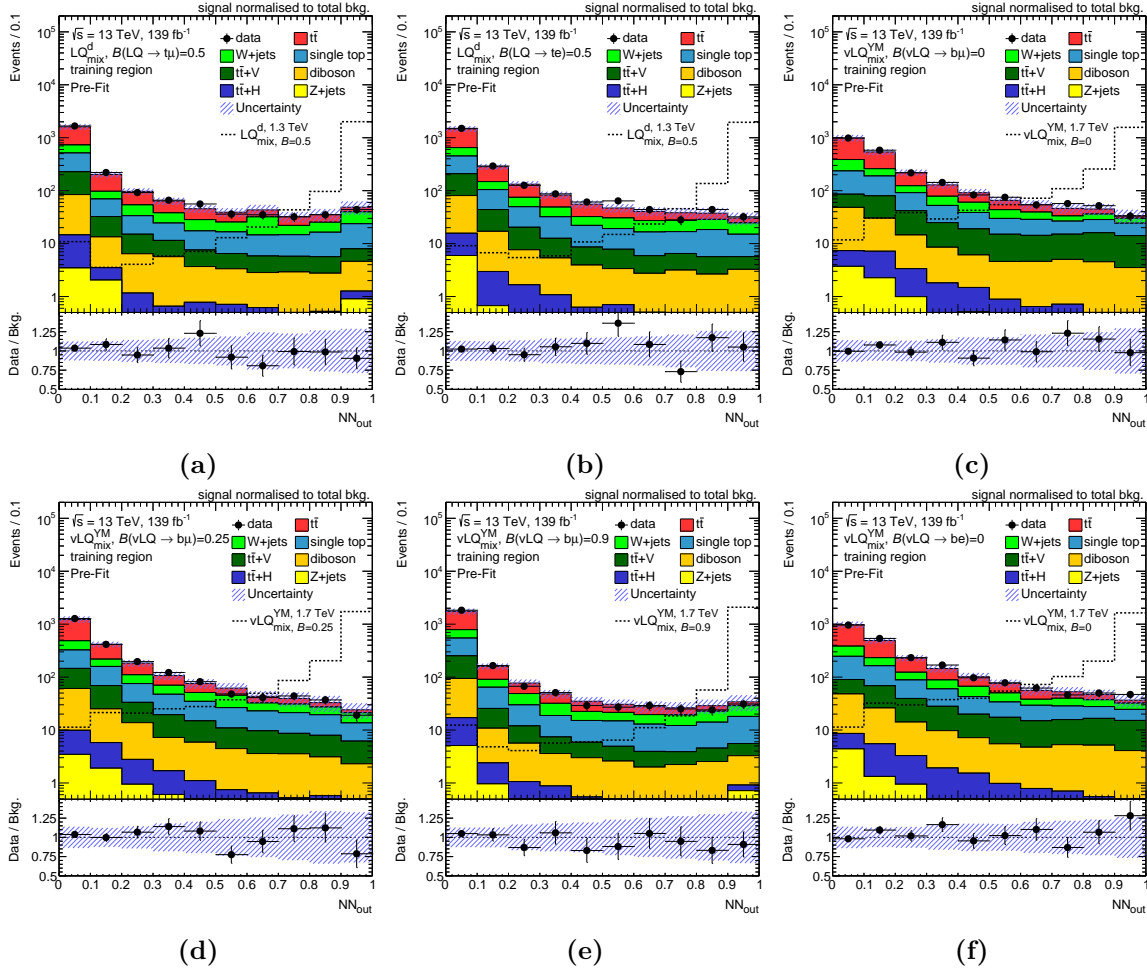
Figure A.5: Distributions of NN_{out} obtained from NNs trained (a) on scalar down-type LQs decaying into top quarks and muons or bottom quarks and neutrinos with $\mathcal{B} = 0.5$, (b) on scalar down-type LQs decaying into top quarks and muons or bottom quarks and neutrinos with $\mathcal{B} = 0.5$, (c) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos with $\mathcal{B} = 0$, (d) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.25$, (e) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.9$, and (f) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos with $\mathcal{B} = 0$.
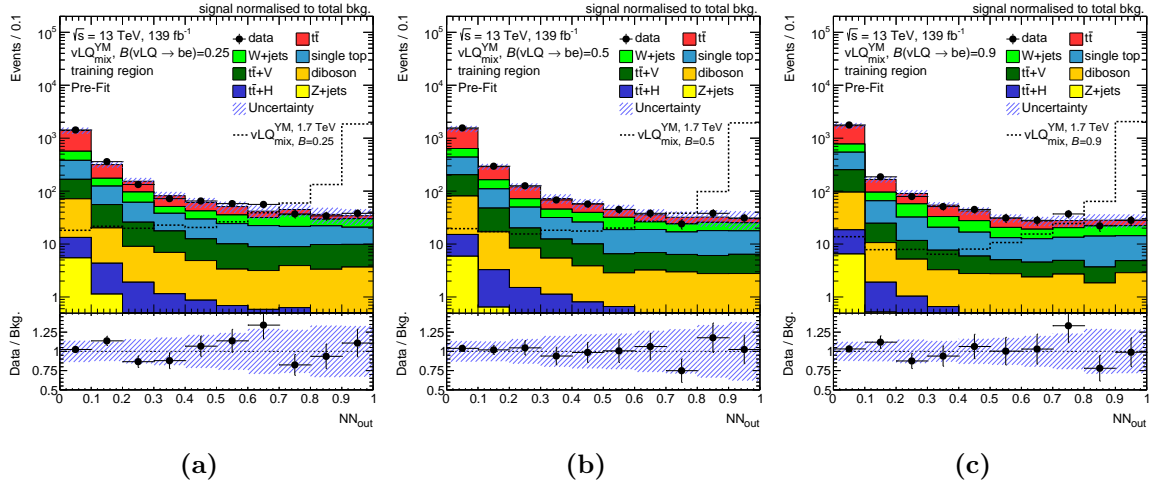
Figure A.6: Distributions of NN_{out} obtained from NNs trained (a) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.25$, (b) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.5$, and (c) on vector LQs in the Yang-Mills coupling scenario decaying into top quarks and neutrinos or bottom quarks and electrons with $\mathcal{B} = 0.9$.

A.4 Neural Network Input Variables in Control Regions

(a)

(b)

(c)

(d)
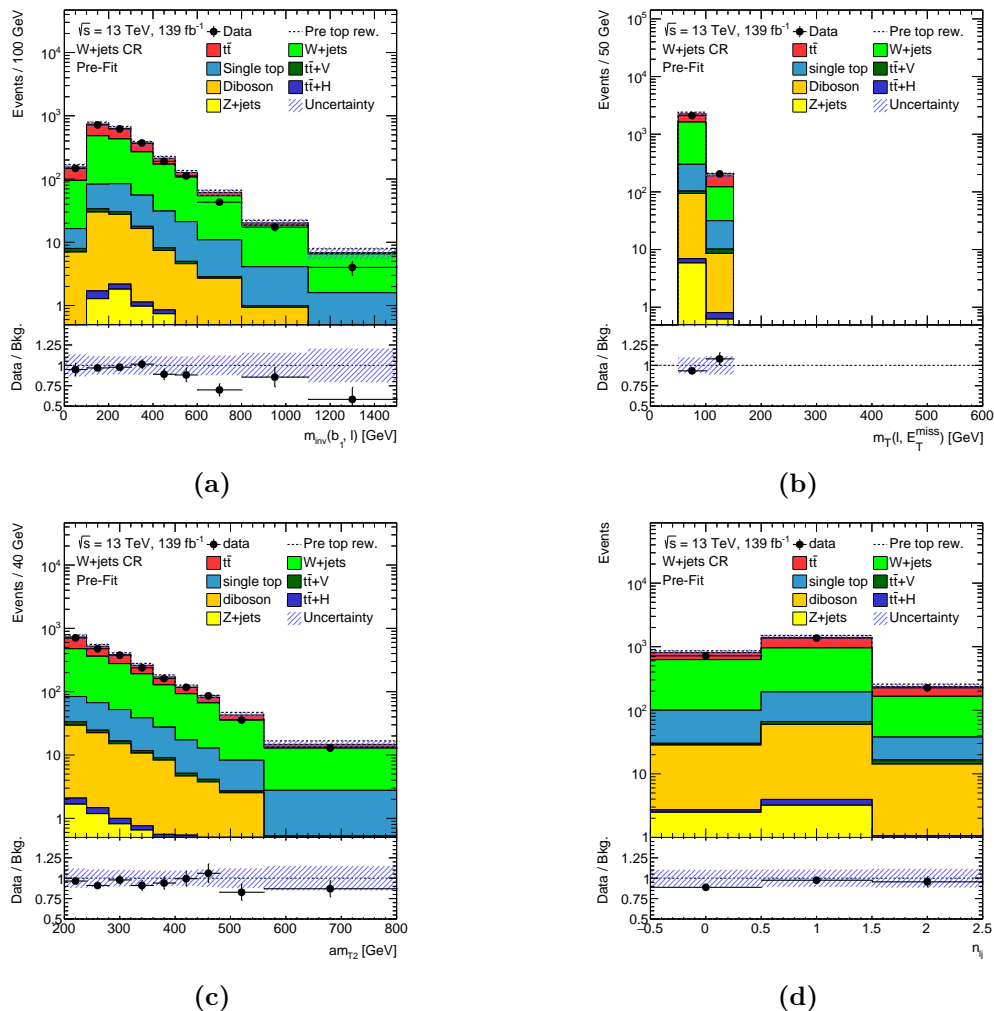
Figure A.7: Distributions (a) of $m_{\mathrm{inv}}(b_1, \ell)$, (b) of $m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$, (c) of am_{T2}, and (d) of n_{lj} in the W+jets CR after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.
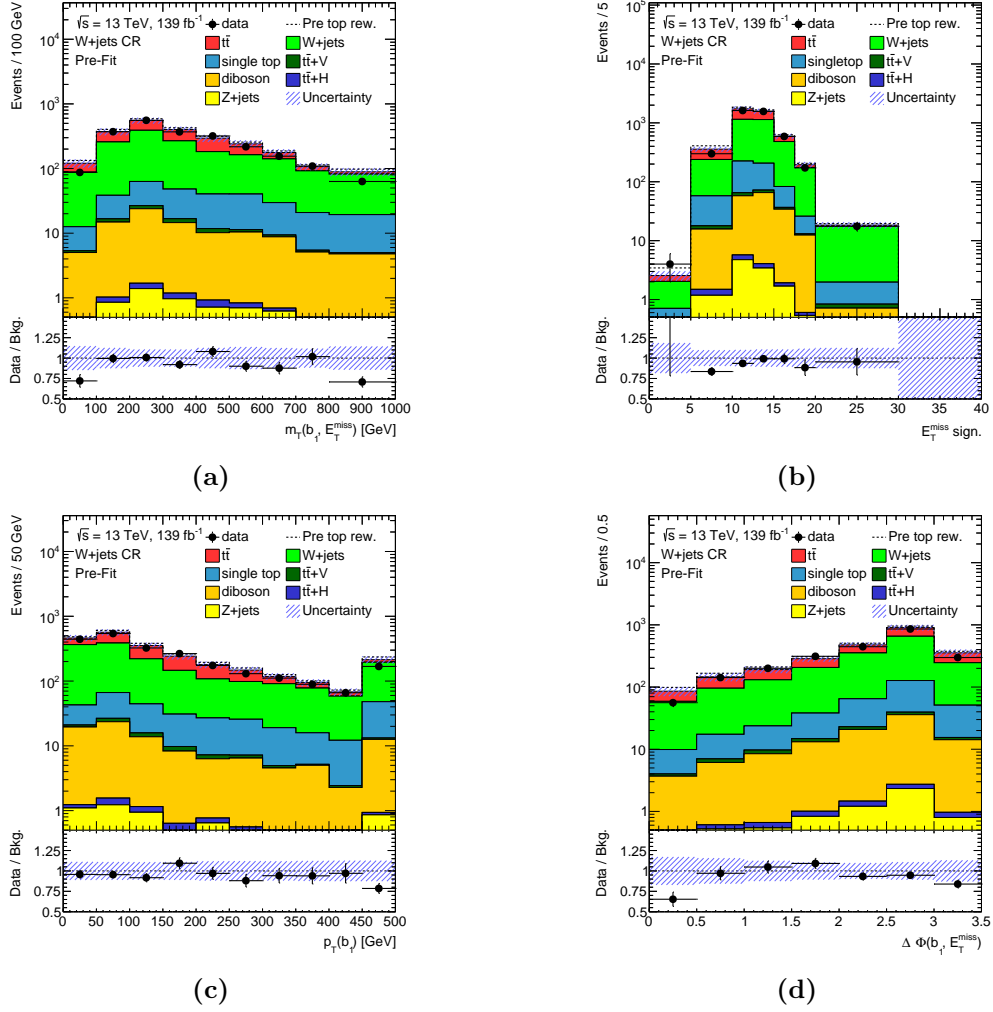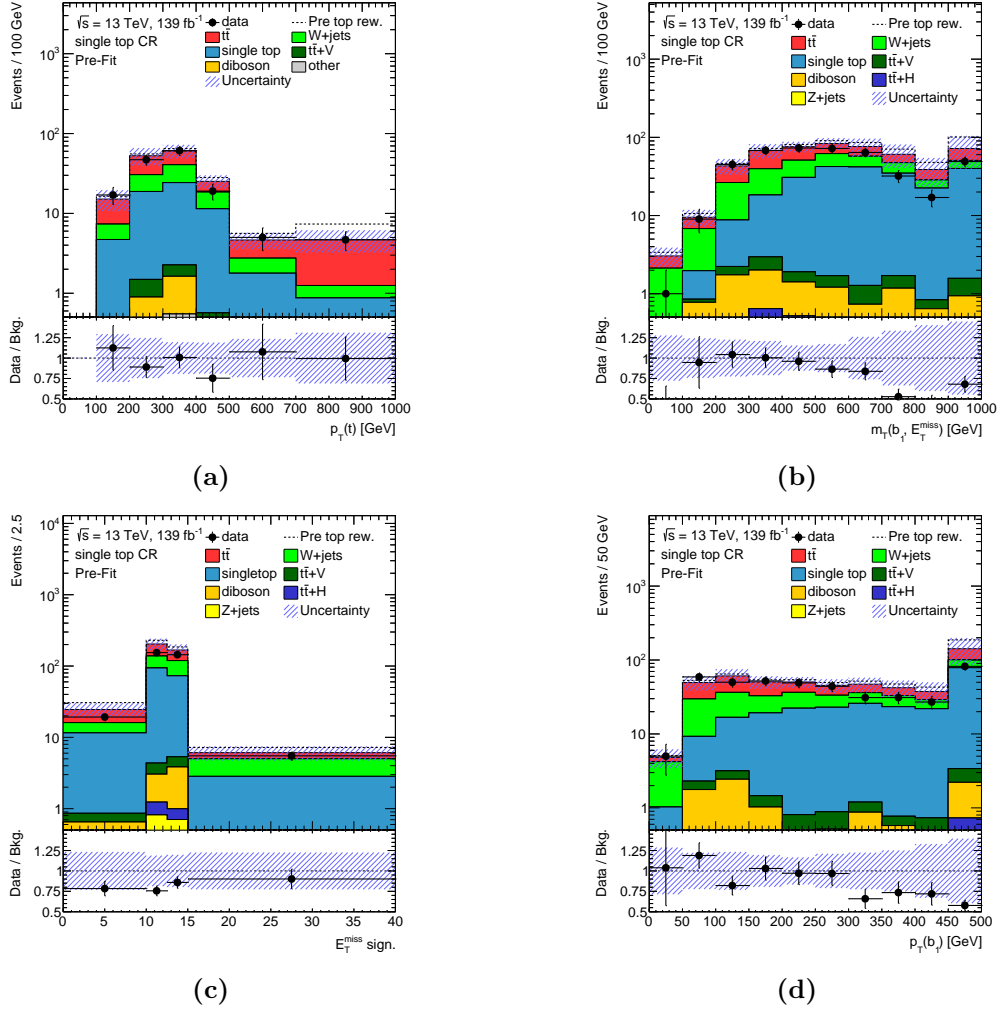
Figure A.8: Distributions (a) of $m_T(b_1, E_T^{miss})$, (b) of E_T^{miss} significance, (c) of $p_T(b_1)$, and (d) of $\Delta\phi(E_T^{miss}, b_1)$ in the W+jets CR after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.

(a)

(b)

(c)

(d)
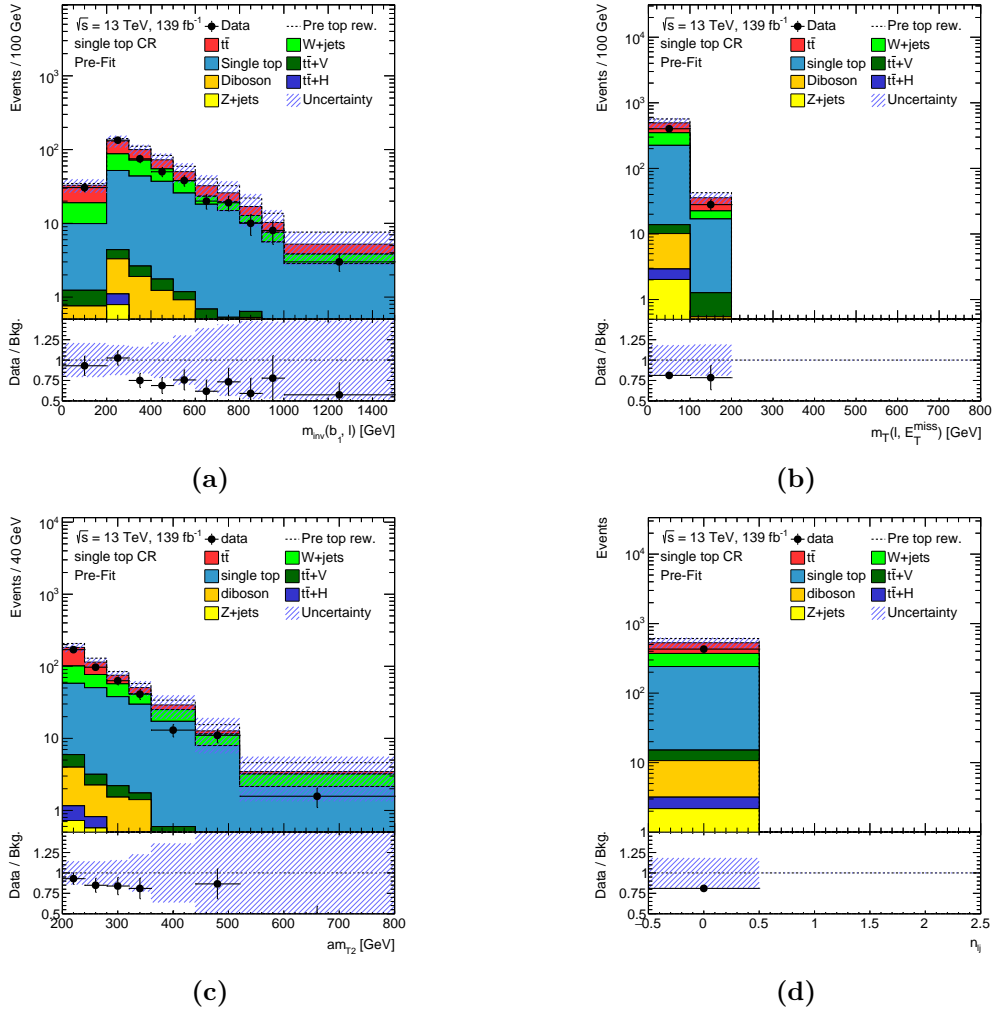
Figure A.9: Distributions (a) of $m_{\text{inv}}(b_1, \ell)$, (b) of $m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}})$, (c) of am_{T2}, and (d) of n_{lj} in the single top CR after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.

Figure A.10: Distributions (a) of $p_T(t_{had})$, (b) of $m_T(b_1, E_T^{miss})$, (c) of E_T^{miss} significance, and (d) of $p_T(b_1)$ in the single top CR after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.
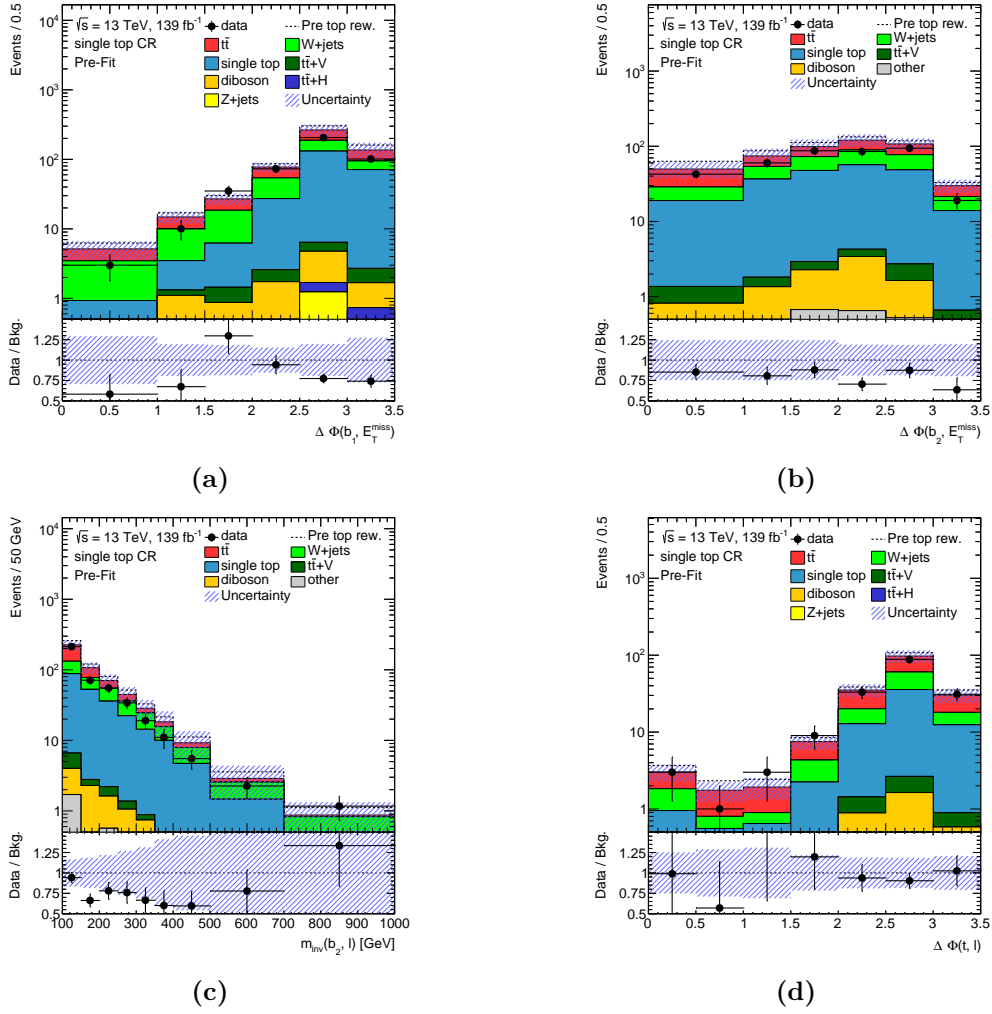
Figure A.11: Distributions (a) of $\Delta\phi(E_\mathrm{T}^\mathrm{miss}, b_1)$, (b) of $\Delta\phi(E_\mathrm{T}^\mathrm{miss}, b_2)$, (c) of $m_\mathrm{inv}(b_2, \ell)$, and (d) of $\Delta\phi(t_\mathrm{had}, \ell)$ in the single top CR after applying the top reweighting, before the fit to data in CRs and SR. The hatched bands include statistical and systematic uncertainties. The total background expectation before applying the top reweighting is shown as a dashed line. The ratios of observed and expected background events are shown in the bottom panels of the plots. The last bin contains the overflow.

A.5 Post-Fit Neural Network Output Distributions in the Signal Region

(a)   **(b)**   **(c)**

(d)   **(e)**   **(f)**
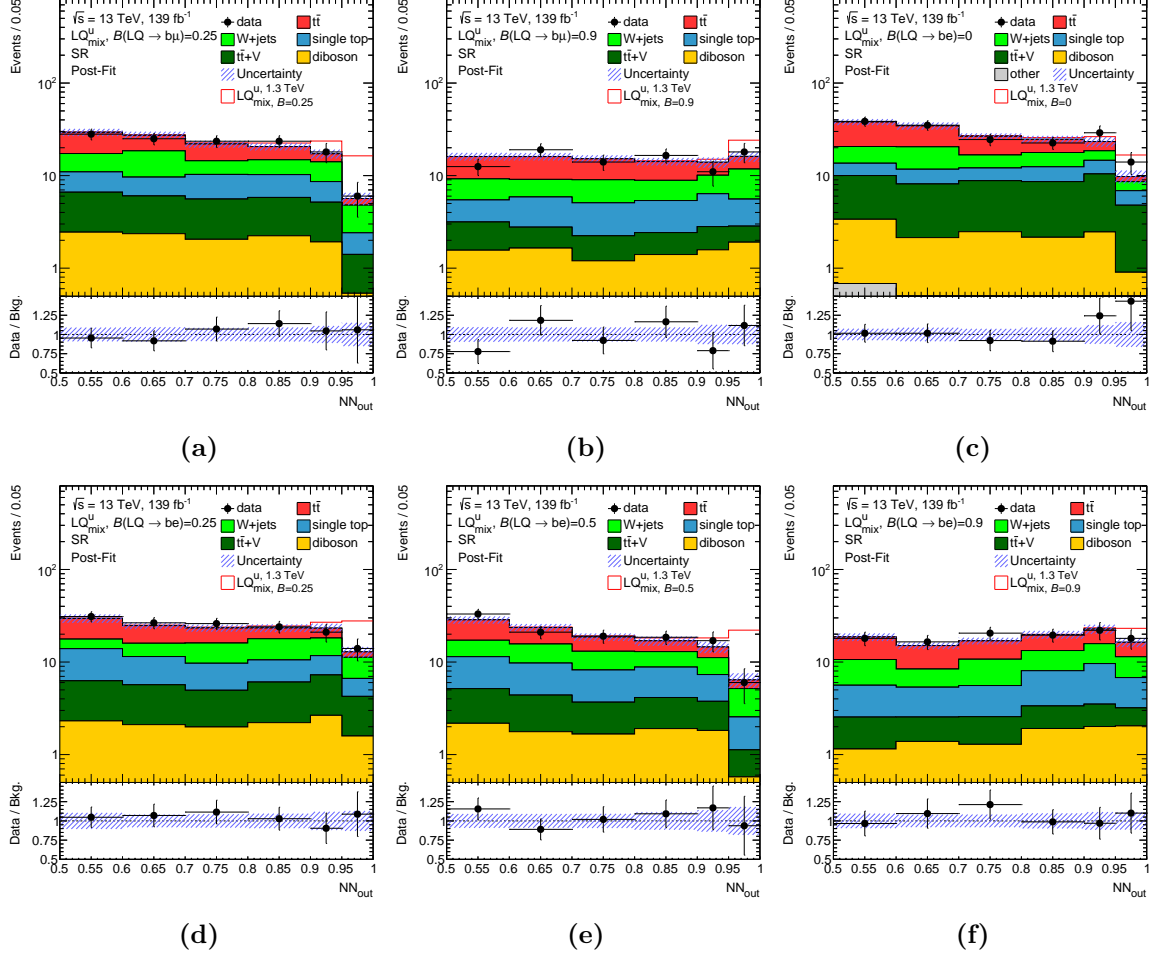
Figure A.12: Post-fit distributions of NN_{out} in the signal region (a) for an NN trained on up-type scalar LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B} = 0.0$, (b) for an NN trained on up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, (c) for an NN trained on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, (d) for an NN trained on up-type scalar LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B} = 0.0$, (e) for an NN trained on up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, and (f) for an NN trained on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. The pre-fit signal expectation at $m_{\mathrm{LQ}} = 1.3\,\mathrm{TeV}$ for scalar LQs and at $m_{\mathrm{LQ}} = 1.7\,\mathrm{TeV}$ is added on top of the post-fit background yields.
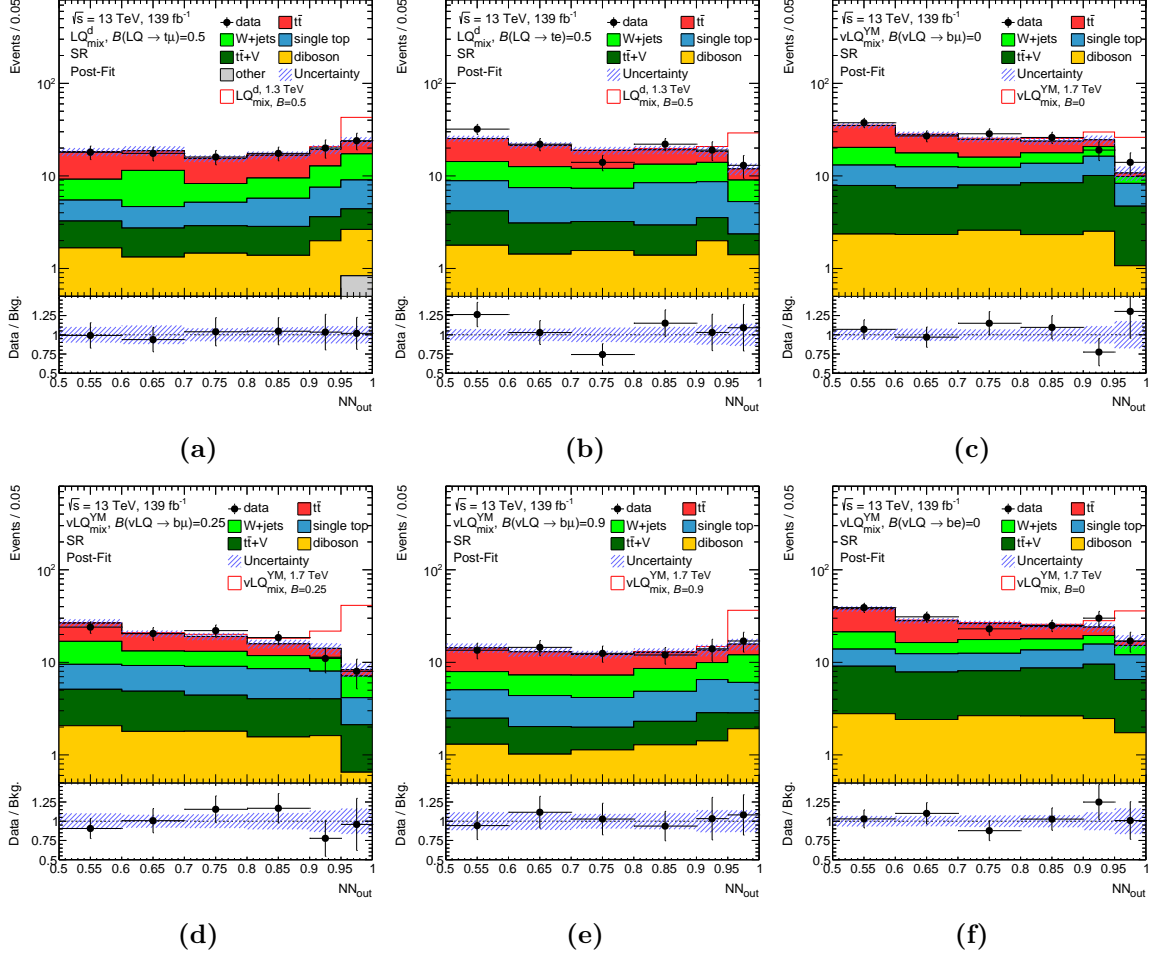
Figure A.13: Post-fit distributions of NN_{out} in the signal region (a) for an NN trained on up-type scalar LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B} = 0.0$, (b) for an NN trained on up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, (c) for an NN trained on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, (d) for an NN trained on up-type scalar LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B} = 0.0$, (e) for an NN trained on up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, and (f) for an NN trained on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. The pre-fit signal expectation at $m_{LQ} = 1.3\,\text{TeV}$ for scalar LQs and at $m_{LQ} = 1.7\,\text{TeV}$ is added on top of the post-fit background yields.
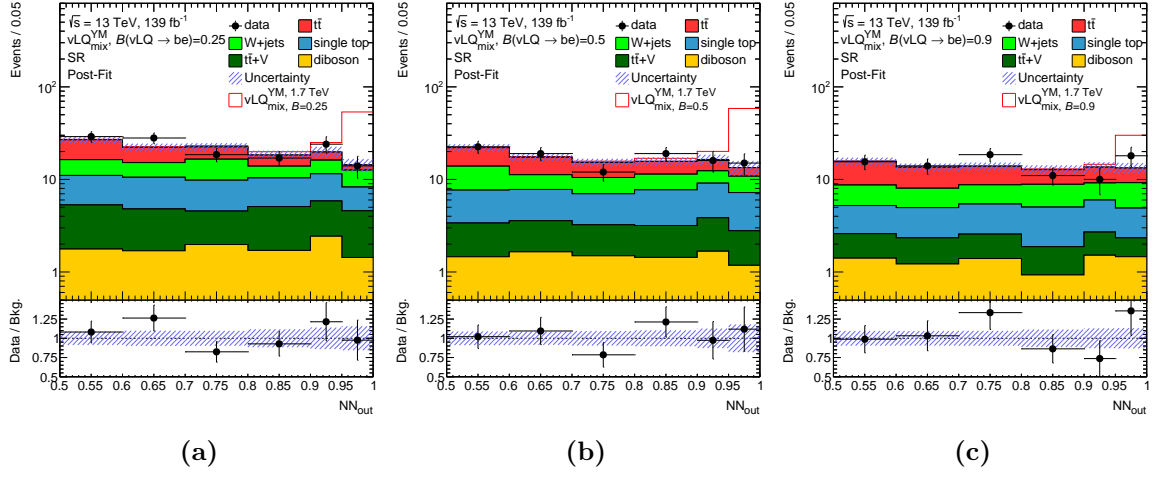
Figure A.14: Post-fit distributions of NN_{out} in the signal region (a) for an NN trained on up-type scalar LQs decaying exclusively into top quarks and neutrinos, i.e. $\mathcal{B} = 0.0$, (b) for an NN trained on up-type scalar LQs decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$, and (c) for an NN trained on vector LQs in the YM coupling scenario decaying into top quarks and neutrinos or bottom quarks and muons with $\mathcal{B} = 0.5$. The pre-fit signal expectation at $m_{\mathrm{LQ}} = 1.3\,\mathrm{TeV}$ for scalar LQs and at $m_{\mathrm{LQ}} = 1.7\,\mathrm{TeV}$ is added on top of the post-fit background yields.

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[4] 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. B **716** (2012) 1, arXiv: `1207.7214` `[hep-ex]`.

[5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: `1207.7235` `[hep-ex]`.

[6] P. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020), 083C01.

[7] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*,
Phys. Lett. **12** (1964) 132.

[8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*,
Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor.

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble,
Global Conservation Laws and Massless Particles,
Phys. Rev. Lett. **13** (1964) 585, ed. by J. C. Taylor.

[10] J. Goldstone, *Field Theories with Superconductor Solutions*,
Nuovo Cim. **19** (1961) 154.

[11] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*,
Phys. Rev. Lett. **10** (1963) 531.

[12] M. Kobayashi and T. Maskawa,
CP Violation in the Renormalizable Theory of Weak Interaction,
Prog. Theor. Phys. **49** (1973) 652.

[13] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*,
Phys. Rev. Lett. **30** (1973) 1343, ed. by J. C. Taylor.

[14] K. G. Wilson, *Confinement of Quarks*,
Phys. Rev. D **10** (1974) 2445, ed. by J. C. Taylor.

[15] CDS Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*,
Phys. Rev. Lett. **74** (1995) 2626, arXiv: `hep-ex/9503002`.

[16] D0 Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632, arXiv: `hep-ex/9503003`.

[17] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

[18] J. E. Augustin et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.

[19] J. J. Aubert et al., *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[20] Y.-S. Tsai, *Decay Correlations of Heavy Leptons in $e^+e^- \to \ell^+\ell^-$*, Phys. Rev. D **4** (1971) 2821, [Erratum: Phys.Rev.D 13, 771 (1976)].

[21] M. L. Perl et al., *Evidence for Anomalous Lepton Production in e^+–e^- Annihilation*, Phys. Rev. Lett. **35** (1975) 1489.

[22] S. W. Herb et al., *Observation of a Dimuon Resonance at $9.5\,\mathrm{GeV}$ in $400\,\mathrm{GeV}$ Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.

[23] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540\,\mathrm{GeV}$*, Phys. Lett. B **122** (1983) 103.

[24] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider*, Phys. Lett. B **122** (1983) 476.

[25] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around $95\,\mathrm{GeV}/c^2$ at the CERN SPS Collider*, Phys. Lett. B **126** (1983) 398.

[26] UA2 Collaboration, *Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider*, Phys. Lett. B **129** (1983) 130.

[27] ATLAS Collaboration, *Standard Model Summary Plots February 2022*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2022-009: CERN, 2022, URL: `http://cds.cern.ch/record/2804061`.

[28] ATLAS Collaboration, *Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, Phys. Lett. B **784** (2018) 345, arXiv: `1806.00242 [hep-ex]`.

[29] J. Haller et al., *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, Eur. Phys. J. C **78** (2018) 675, arXiv: `1803.01853 [hep-ph]`.

[30] CDF Collaboration,
High-precision measurement of the W boson mass with the CDF II detector,
Science **376** (2022) 170.

[31] R. J. Crewther, P. Di Vecchia, G. Veneziano, and E. Witten, *Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics*,
Phys. Lett. B **88** (1979) 123, [Erratum: Phys.Lett.B 91, 487 (1980)].

[32] C. Abel et al.,
Measurement of the Permanent Electric Dipole Moment of the Neutron,
Phys. Rev. Lett. **124** (2020) 081803, arXiv: 2001.11966 [hep-ex].

[33] Planck Collaboration, *Planck 2018 results. VI. Cosmological parameters*,
Astron. Astrophys. **641** (2020) A6, [Erratum: Astron.Astrophys. 652, C4 (2021)],
arXiv: 1807.06209 [astro-ph.CO].

[34] P. F. De Salas, S. Gariazzo, O. Mena, C. A. Ternes, and M. Tórtola, *Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects*,
Front. Astron. Space Sci. **5** (2018) 36, arXiv: 1806.11051 [hep-ph].

[35] D. Clowe et al., *A direct empirical proof of the existence of dark matter*,
Astrophys. J. Lett. **648** (2006) L109, arXiv: astro-ph/0608407.

[36] V. C. Rubin, N. Thonnard, and W. K. Ford Jr.,
Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/,
Astrophys. J. **238** (1980) 471.

[37] A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, Astron. J. **116** (1998) 1009,
arXiv: astro-ph/9805201.

[38] S. Perlmutter et al., *Measurements of Ω and Λ from 42 high redshift supernovae*,
Astrophys. J. **517** (1999) 565, arXiv: astro-ph/9812133.

[39] R. H. Cyburt, B. D. Fields, and K. A. Olive,
Primordial nucleosynthesis in light of WMAP, Phys. Lett. B **567** (2003) 227,
arXiv: astro-ph/0302431.

[40] A. D. Sakharov,
Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe,
Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[41] LHCb Collaboration, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*,
JHEP **08** (2017) 055, arXiv: 1705.05802 [hep-ex].

[42] LHCb Collaboration, *Test of lepton universality in beauty-quark decays*,
Nature Phys. **18** (2022) 277, arXiv: 2103.11769 [hep-ex].

[43] LHCb Collaboration,
      Tests of lepton universality using $B^0 \to K_S^0 \ell^+ \ell^-$ and $B^+ \to K^{*+} \ell^+ \ell^-$ decays,
      Phys. Rev. Lett. **128** (2022) 191802, arXiv: `2110.09501 [hep-ex]`.

[44] Belle Collaboration,
      Test of Lepton-Flavor Universality in $B \to K^* \ell^+ \ell^-$ Decays at Belle,
      Phys. Rev. Lett. **126** (2021) 161801, arXiv: `1904.02440 [hep-ex]`.

[45] Belle Collaboration, *Test of lepton flavor universality and search for lepton flavor
      violation in $B \to K\ell\ell$ decays*, JHEP **03** (2021) 105, arXiv: `1908.01848 [hep-ex]`.

[46] Y. S. Amhis et al.,
      Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018,
      Eur. Phys. J. **C81** (2021) 226, updated results and plots available at
      `https://hflav.web.cern.ch/`, arXiv: `1909.12524 [hep-ex]`.

[47] LHCb Collaboration, *Branching Fraction Measurements of the Rare $B_s^0 \to \phi\mu^+\mu^-$
      and $B_s^0 \to f_2'(1525)\mu^+\mu^-$- Decays*, Phys. Rev. Lett. **127** (2021) 151801,
      arXiv: `2105.14007 [hep-ex]`.

[48] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean
      observables for the binned analysis of $B- > K^*\mu^+\mu^-$ at large recoil*,
      JHEP **01** (2013) 048, arXiv: `1207.2753 [hep-ph]`.

[49] LHCb Collaboration,
      Measurement of CP-Averaged Observables in the $B^0 \to K^{*0}\mu^+\mu^-$ Decay,
      Phys. Rev. Lett. **125** (2020) 011802, arXiv: `2003.04831 [hep-ex]`.

[50] D. Hanneke, S. F. Hoogerheide, and G. Gabrielse, *Cavity Control of a
      Single-Electron Quantum Cyclotron: Measuring the Electron Magnetic Moment*,
      Phys. Rev. A **83** (2011) 052122, arXiv: `1009.4831 [physics.atom-ph]`.

[51] B. Abi et al.,
      Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm,
      Phys. Rev. Lett. **126** (2021) 141801, arXiv: `2104.03281 [hep-ex]`.

[52] S. Borsanyi et al.,
      Leading hadronic contribution to the muon magnetic moment from lattice QCD,
      Nature **593** (2021) 51, arXiv: `2002.12347 [hep-lat]`.

[53] H. Georgi and S. L. Glashow, *Unity of All Elementary Particle Forces*,
      Phys. Rev. Lett. **32** (1974) 438.

[54] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*,
      Phys. Rev. D **10** (1974) 275, [Erratum: Phys.Rev.D 11, 703–703 (1975)].

[55] H. Fritzsch and P. Minkowski, *Unified Interactions of Leptons and Hadrons*,
      Annals Phys. **93** (1975) 193.

[56] H. Georgi, *The State of the Art - Gauge Theories*,
AIP Conf. Proc. **23** (1975) 575, ed. by H. C. C. E. W. Carlson.

[57] S. Dimopoulos and L. Susskind, *Mass Without Scalars*,
Nucl. Phys. B **155** (1979) 237, ed. by A. Zichichi.

[58] S. Dimopoulos, *Technicolored Signatures*,
Nucl. Phys. B **168** (1980) 69, ed. by A. Zichichi.

[59] E. Eichten and K. D. Lane, *Dynamical Breaking of Weak Interaction Symmetries*,
Phys. Lett. B **90** (1980) 125.

[60] L. F. Abbott and E. Farhi, *Are the Weak Interactions Strong?*,
Phys. Lett. B **101** (1981) 69.

[61] L. F. Abbott and E. Farhi, *A Confining Model of the Weak Interactions*,
Nucl. Phys. B **189** (1981) 547.

[62] W. Buchmuller, R. Ruckl, and D. Wyler, *Leptoquarks in Lepton - Quark Collisions*,
Phys. Lett. B **191** (1987) 442, [Erratum: Phys.Lett.B 448, 320–320 (1999)].

[63] 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. B **910** (2016) 1, arXiv: `1603.03568` [`hep-ph`].

[64] H. Nishino et al., *Search for Proton Decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in a Large Water Cherenkov Detector*, Phys. Rev. Lett. **102** (2009) 141801,
arXiv: `0903.0676` [`hep-ex`].

[65] G. Hiller and M. Schmaltz,
R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities,
Phys. Rev. D **90** (2014) 054014, arXiv: `1408.1627` [`hep-ph`].

[66] B. Gripaios, M. Nardecchia, and S. A. Renner,
Composite leptoquarks and anomalies in B-meson decays, JHEP **05** (2015) 006,
arXiv: `1412.1791` [`hep-ph`].

[67] M. Freytsis, Z. Ligeti, and J. T. Ruderman, *Flavor models for $\bar{B} \to D^{(*)}\tau\bar{\nu}$*,
Phys. Rev. D **92** (2015) 054018, arXiv: `1506.08896` [`hep-ph`].

[68] D. Buttazzo, A. Greljo, G. Isidori, and D. Marzocca,
B-physics anomalies: a guide to combined explanations, JHEP **11** (2017) 044,
arXiv: `1706.07808` [`hep-ph`].

[69] A. Crivellin, D. Müller, and T. Ota, *Simultaneous explanation of $R(D^{(\star)})$ and $b \to s\mu^+ \mu^-$: the last scalar leptoquarks standing*, JHEP **09** (2017) 040,
arXiv: `1703.09226` [`hep-ph`].

[70] J. E. Camargo-Molina, A. Celis, and D. A. Faroughy,
Anomalies in Bottom from new physics in Top, Phys. Lett. B **784** (2018) 284,
arXiv: `1805.04917` [`hep-ph`].

[71] D. Chakraverty, D. Choudhury, and A. Datta,
A Nonsupersymmetric resolution of the anomalous muon magnetic moment,
Phys. Lett. B **506** (2001) 103, arXiv: `hep-ph/0102180`.

[72] M. Bauer and M. Neubert,
Minimal Leptoquark Explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_\mu$ Anomalies,
Phys. Rev. Lett. **116** (2016) 141802, arXiv: `1511.01900 [hep-ph]`.

[73] C.-K. Chua, X.-G. He, and W.-Y. P. Hwang,
Neutrino mass induced radiatively by supersymmetric leptoquarks,
Phys. Lett. B **479** (2000) 224, arXiv: `hep-ph/9905340`.

[74] U. Mahanta, *Neutrino masses and mixing angles from leptoquark interactions*,
Phys. Rev. D **62** (2000) 073009, arXiv: `hep-ph/9909518`.

[75] S.-P. He, *A leptoquark and vector-like quark extended model for the simultaneous
explanation of the W boson mass and muon $g-2$ anomalies*, (2022),
arXiv: `2205.02088 [hep-ph]`.

[76] T. A. Chowdhury and S. Saad, *Leptoquark-vectorlike quark model for m_W (CDF),
$(g-2)_\mu$, $R_{K^{(*)}}$ anomalies and neutrino mass*, (2022), arXiv: `2205.03917 [hep-ph]`.

[77] L. Di Luzio, A. Greljo, and M. Nardecchia,
Gauge leptoquark as the origin of B-physics anomalies,
Phys. Rev. D **96** (2017) 115011, arXiv: `1708.08450 [hep-ph]`.

[78] N. Assad, B. Fornal, and B. Grinstein, *Baryon Number and Lepton Universality
Violation in Leptoquark and Diquark Models*, Phys. Lett. B **777** (2018) 324,
arXiv: `1708.06350 [hep-ph]`.

[79] L. Calibbi, A. Crivellin, and T. Li,
Model of vector leptoquarks in view of the B-physics anomalies,
Phys. Rev. D **98** (2018) 115002, arXiv: `1709.00692 [hep-ph]`.

[80] M. Bordone, C. Cornella, J. Fuentes-Martin, and G. Isidori,
A three-site gauge model for flavor hierarchies and flavor anomalies,
Phys. Lett. B **779** (2018) 317, arXiv: `1712.01368 [hep-ph]`.

[81] R. Barbieri and A. Tesi, *B-decay anomalies in Pati-Salam SU(4)*,
Eur. Phys. J. C **78** (2018) 193, arXiv: `1712.06844 [hep-ph]`.

[82] M. Blanke and A. Crivellin, *B Meson Anomalies in a Pati-Salam Model within the
Randall-Sundrum Background*, Phys. Rev. Lett. **121** (2018) 011801,
arXiv: `1801.07256 [hep-ph]`.

[83] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza, and E. Laenen,
*NNLL-fast: predictions for coloured supersymmetric particle production at the LHC
with threshold and Coulomb resummation*, JHEP **12** (2016) 133,
arXiv: `1607.07741 [hep-ph]`.

[84] W. Beenakker, M. Krämer, T. Plehn, M. Spira, and P. Zerwas,
Stop production at hadron colliders, Nucl. Phys. B **515** (1998) 3,
arXiv: `hep-ph/9710451`.

[85] W. Beenakker et al.,
Supersymmetric top and bottom squark production at hadron colliders,
JHEP **08** (2010) 098, arXiv: `1006.4771 [hep-ph]`.

[86] W. Beenakker et al., *NNLL resummation for stop pair-production at the LHC*,
JHEP **05** (2016) 153, arXiv: `1601.02954 [hep-ph]`.

[87] J. Alwall et al., *The automated computation of tree-level and next-to-leading order
differential cross sections, and their matching to parton shower simulations*,
JHEP **07** (2014) 079, arXiv: `1405.0301 [hep-ph]`.

[88] T. Mandal, S. Mitra, and S. Seth,
Pair production of scalar leptoquarks at the LHC to NLO parton shower accuracy,
Phys. Rev. D **93** (2016) 035018, arXiv: `1506.07369 [hep-ph]`.

[89] M. J. Baker, J. Fuentes-Martín, G. Isidori, and M. König,
High-p_T signatures in vector-leptoquark models, Eur. Phys. J. C **79** (2019) 334,
arXiv: `1901.10480 [hep-ph]`.

[90] ATLAS Collaboration, *Searches for third-generation scalar leptoquarks in $\sqrt{s} = 13$
TeV pp collisions with the ATLAS detector*, JHEP **06** (2019) 144,
arXiv: `1902.08103 [hep-ex]`.

[91] G. Aad et al., *Measurement of the W boson polarization in top quark decays with the
ATLAS detector*, JHEP **06** (2012) 088, arXiv: `1205.2484 [hep-ex]`.

[92] CMS Collaboration,
*Search for heavy neutrinos and third-generation leptoquarks in hadronic states of two
τ leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
JHEP **03** (2019) 170, arXiv: `1811.00806 [hep-ex]`.

[93] ATLAS Collaboration, *Search for pairs of scalar leptoquarks decaying into quarks
and electrons or muons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*,
JHEP **10** (2020) 112, arXiv: `2006.05872 [hep-ex]`.

[94] CMS Collaboration,
Inclusive nonresonant multilepton probes of new phenomena at $\sqrt{s} = 13$ TeV,
(2022), arXiv: `2202.08676 [hep-ex]`.

[95] CMS Collaboration,
Search for pair production of first-generation scalar leptoquarks at $\sqrt{s} = 13$ TeV,
Phys. Rev. D **99** (2019) 052002, arXiv: `1811.01197 [hep-ex]`.

[96]   CMS Collaboration, *Search for single production of scalar leptoquarks in proton-proton collisions at $\sqrt{s} = 8\ TeV$*,
Phys. Rev. D **93** (2016) 032005, [Erratum: Phys.Rev.D 95, 039906 (2017)],
arXiv: 1509.03750 [hep-ex].

[97]   ATLAS Collaboration, *Search for pair production of scalar leptoquarks decaying into first- or second-generation leptons and top quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 313,
arXiv: 2010.02098 [hep-ex].

[98]   ATLAS Collaboration, *Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector*,
Phys. Rev. D **104** (2021) 112005, arXiv: 2108.07665 [hep-ex].

[99]   ATLAS Collaboration,
Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a τ-lepton in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,
JHEP **06** (2021) 179, arXiv: 2101.11582 [hep-ex].

[100]   CMS Collaboration, *Search for singly and pair-produced leptoquarks coupling to third-generation fermions in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
Phys. Lett. B **819** (2021) 136446, arXiv: 2012.04178 [hep-ex].

[101]   CMS Collaboration, *Searches for physics beyond the standard model with the M_{T2} variable in hadronic final states with and without disappearing tracks in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **80** (2020) 3,
arXiv: 1909.03460 [hep-ex].

[102]   ATLAS Collaboration,
Search for new phenomena in final states with b-jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector,
JHEP **05** (2021) 093, arXiv: 2101.12527 [hep-ex].

[103]   ATLAS Collaboration,
Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector,
Eur. Phys. J. C **80** (2020) 737, arXiv: 2004.14060 [hep-ex].

[104]   *LHC Machine*, JINST **3** (2008) S08001, ed. by L. Evans and P. Bryant.

[105]   ALICE Collaboration, *The ALICE experiment at the CERN LHC*,
JINST **3** (2008) S08002.

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

[107] CMS Collaboration, *The CMS Experiment at the CERN LHC*,
JINST **3** (2008) S08004.

[108] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[109] C. De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*,
(2016), General Photo, URL: `https://cds.cern.ch/record/2119882`.

[110] B. Salvachua, "Overview of Proton-Proton Physics during Run 2",
9th LHC Operations Evian Workshop, 2019 7.

[111] ATLAS Collaboration, *Number of Interactions per Crossing*, 2018, URL: `https: //twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[112] J. Pequenao, "Computer generated image of the whole ATLAS detector", 2008,
URL: `https://cds.cern.ch/record/1095924`.

[113] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, Volume 1*,
ATLAS-TDR-4; CERN-LHCC-97-016, 1997,
URL: `https://cds.cern.ch/record/331063`.

[114] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, Volume 2*,
ATLAS-TDR-5, CERN-LHCC-97-017, 1997,
URL: `https://cds.cern.ch/record/331064`.

[115] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673,
arXiv: `1704.07983 [hep-ex]`.

[116] ATLAS Collaboration, *Production and Integration of the ATLAS Insertable B-Layer*,
JINST **13** (2018) T05008, arXiv: `1803.00844 [physics.ins-det]`.

[117] ATLAS Collaboration, *ATLAS Calorimeter Performance: Technical Design Report*,
ATLAS-TDR-1; CERN-LHCC-96-040, 1996,
URL: `https://cds.cern.ch/record/331059`.

[118] J. Pequenao, "Computer Generated image of the ATLAS calorimeter", 2008,
URL: `https://cds.cern.ch/record/1095927`.

[119] ATLAS Collaboration, *ATLAS Liquid Argon Calorimeter: Technical Design Report*,
ATLAS-TDR-2; CERN-LHCC-96-041, 1996,
URL: `https://cds.cern.ch/record/331061`.

[120] ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*,
ATLAS-TDR-10; CERN-LHCC-97-022, CERN, 1997,
URL: `https://cds.cern.ch/record/331068`.

[121] J. Pequenao, "Computer generated image of the ATLAS Muons subsystem", 2008,
URL: `https://cds.cern.ch/record/1095929`.

[122] G. Avoni et al.,
       The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS,
       JINST **13** (2018) P07017.

[123] V. Cindro et al., *The ATLAS beam conditions monitor*, JINST **3** (2008) P02004.

[124] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13\,TeV$
       using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019,
       URL: https://cds.cern.ch/record/2677054.

[125] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*,
       Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[126] ATLAS Collaboration, *Performance of the missing transverse momentum triggers
       for the ATLAS detector during Run-2 data taking*, JHEP **08** (2020) 080,
       arXiv: 2005.09554 [hep-ex].

[127] F. Siegert, *Monte-Carlo event generation for the LHC*,
       PhD thesis: Durham U., 2010.

[128] Y. L. Dokshitzer,
       *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^-
       Annihilation by Perturbation Theory in Quantum Chromodynamics.*,
       Sov. Phys. JETP **46** (1977) 641.

[129] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*,
       Sov. J. Nucl. Phys. **15** (1972) 438.

[130] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*,
       Nucl. Phys. B **126** (1977) 298.

[131] R. D. Ball et al., *Parton distributions from high-precision collider data*,
       Eur. Phys. J. C **77** (2017) 663, arXiv: 1706.00428 [hep-ph].

[132] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*,
       Eur. Phys. J. C **75** (2015) 132, arXiv: 1412.7420 [hep-ph].

[133] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand,
       Parton Fragmentation and String Dynamics, Phys. Rept. **97** (1983) 31.

[134] B. R. Webber,
       A QCD Model for Jet Fragmentation Including Soft Gluon Interference,
       Nucl. Phys. B **238** (1984) 492.

[135] GEANT4 Collaboration, Geant4 – *a simulation toolkit*,
       Nucl. Instrum. Meth. A **506** (2003) 250.

[136] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*,
       Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[137] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444.

[138] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: `1611.10235` [`physics.ins-det`].

[139] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: `1902.04655` [`physics.ins-det`].

[140] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data*, JINST **14** (2019) P03017, arXiv: `1812.03848` [`hep-ex`].

[141] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: `1603.05598` [`hep-ex`].

[142] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: `2012.00578` [`hep-ex`].

[143] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: `1703.10485` [`hep-ex`].

[144] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: `0802.1189` [`hep-ph`].

[145] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: `1111.6097` [`hep-ph`].

[146] M. Cacciari, "FastJet: A Code for fast k_t clustering, and more", *Proceedings, Deep inelastic scattering, 14th International Workshop, DIS 2006* (Tsukuba, Japan, Apr. 20–24, 2006) 487, arXiv: `hep-ph/0607071` [`hep-ph`].

[147] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 689, arXiv: `2007.02645` [`hep-ex`].

[148] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: `1510.03823` [`hep-ex`].

[149] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: `1907.05120` [`hep-ex`].

[150] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281.

[151] ATLAS Collaboration,
Measurement of the b-jet identification efficiency for high transverse momentum jets in $t\bar{t}$ events in the lepton + jets channel with the ATLAS detector using Run 2 data,
ATL-PHYS-PUB-2021-004, 2021, URL: https://cds.cern.ch/record/2753734.

[152] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex].

[153] ATLAS Collaboration,
E_T^{miss} *performance in the ATLAS detector using 2015–2016 LHC pp collisions*,
ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233.

[154] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*,
JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph].

[155] ATLAS Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using $36\,fb^{-1}$ of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, JHEP **06** (2018) 108,
arXiv: 1711.11520 [hep-ex].

[156] ATLAS Collaboration,
ATLAS data quality operations and performance for 2015–2018 data-taking,
JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[157] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*,
ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[158] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034,
arXiv: 1905.09127 [hep-ph].

[159] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. A **462** (2001) 152.

[160] S. Frixione, G. Ridolfi, and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126,
arXiv: 0707.3088 [hep-ph].

[161] P. Nason,
A new method for combining NLO QCD with shower Monte Carlo algorithms,
JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[162] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070,
arXiv: 0709.2092 [hep-ph].

[163] 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 **06** (2010) 043, arXiv: `1002.2581` [`hep-ph`].

[164] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: `1410.8849` [`hep-ph`].

[165] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: `1410.3012` [`hep-ph`].

[166] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: `1112.5675` [`hep-ph`].

[167] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: `0808.3674` [`hep-ph`].

[168] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: `1907.13071` [`hep-ph`].

[169] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: `1111.5206` [`hep-ph`].

[170] A. Denner, S. Dittmaier, and L. Hofer, COLLIER*: A fortran-based complex one-loop library in extended regularizations*, Comput. Phys. Commun. **212** (2017) 220, arXiv: `1604.06792` [`hep-ph`].

[171] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, Phys. Rev. D **69** (2004) 094008, arXiv: `hep-ph/0312266`.

[172] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: `1111.1220` [`hep-ph`].

[173] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: `1207.5030` [`hep-ph`].

[174] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: `hep-ph/0109231`.

[175] S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: `0903.1219` [`hep-ph`].

[176] M. Kramer, T. Plehn, M. Spira, and P. Zerwas,
       Pair production of scalar leptoquarks at the CERN LHC,
       Phys. Rev. D **71** (2005) 057503, arXiv: `hep-ph/0411038`.

[177] M. Kramer, T. Plehn, M. Spira, and P. Zerwas,
       Pair Production of Scalar Leptoquarks at the Fermilab Tevatron,
       Phys. Rev. Lett. **79** (1997) 341, arXiv: `hep-ph/9704322`.

[178] C. Borschensky, B. Fuks, A. Kulesza, and D. Schwartländer,
       Scalar leptoquark pair production at hadron colliders,
       Phys. Rev. D **101** (2020) 115017, arXiv: `2002.08971 [hep-ph]`.

[179] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*,
       Comput. Phys. Commun. **178** (2008) 852, arXiv: `0710.3820 [hep-ph]`.

[180] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias
       and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*,
       ATL-PHYS-PUB-2016-017, 2016, URL: `https://cds.cern.ch/record/2206965`.

[181] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*,
       (2004), arXiv: `physics/0402093`.

[182] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*,
       Nucl. Instrum. Meth. A **559** (2006) 190, ed. by J. Blumlein et al.

[183] CMS Collaboration, *Measurement of the Polarization of W Bosons with Large
       Transverse Momenta in W+Jets Events at the LHC*,
       Phys. Rev. Lett. **107** (2011) 021802, arXiv: `1104.3829 [hep-ex]`.

[184] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*,
       Eur. Phys. J. C **76** (2016) 196, arXiv: `1512.01178 [hep-ph]`.

[185] J. Bellm et al., *Herwig 7.1 Release Note*, (2017), arXiv: `1705.06919 [hep-ph]`.

[186] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*,
       J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]`.

[187] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White,
       Single-top hadroproduction in association with a W boson, JHEP **07** (2008) 029,
       arXiv: `0805.3067 [hep-ph]`.

[188] E. Re, *Single-top Wt-channel production matched with parton showers using the
       POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv: `1009.2450 [hep-ph]`.

[189] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the
       Nature of the Higgs Sector*, (2016), arXiv: `1610.07922 [hep-ph]`.

[190] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in
       proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
       Phys. Rev. D **99** (2019) 072009, arXiv: `1901.03584 [hep-ex]`.

[191] ATLAS Collaboration, *Multi-boson simulation for* 13 *TeV ATLAS analyses*,
ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986.

[192] ATLAS Collaboration, *Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Phys. Lett. B **759** (2016) 601, arXiv: 1603.09222 [hep-ex].

[193] J. Neyman and E. S. Pearson,
On the Problem of the Most Efficient Tests of Statistical Hypotheses,
Phil. Trans. Roy. Soc. Lond. A **231** (1933) 289.

[194] G. Cowan, K. Cranmer, E. Gross, and O. Vitells,
Asymptotic formulae for likelihood-based tests of new physics,
Eur. Phys. J. C **71** (2011) 1554, [Erratum: Eur.Phys.J.C 73, 2501 (2013)],
arXiv: 1007.1727 [physics.data-an].

[195] A. L. Read, *Presentation of search results: The CL(s) technique*,
J. Phys. G **28** (2002) 2693, ed. by M. R. Whalley and L. Lyons.

[196] ATLAS Collaboration, *Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **04** (2021) 174,
arXiv: 2012.03799 [hep-ex].

[197] L. Heinrich, M. Feickert, and G. Stark, *pyhf: v0.6.3*, version 0.6.3,
URL: https://github.com/scikit-hep/pyhf/releases/tag/v0.6.3.

[198] L. Heinrich, M. Feickert, G. Stark, and K. Cranmer,
pyhf: pure-Python implementation of HistFactory statistical models,
Journal of Open Source Software **6** (2021) 2823.

[199] K. Cranmer and A. Held, *Building and steering binned template fits with cabinetry*,
EPJ Web Conf. **251** (2021) 03067.

[200] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 089,
arXiv: 1805.01845 [hep-ex].

[201] ATLAS Collaboration, *Studies of $t\bar{t}/tW$ interference effects in $b\bar{b}\ell^+\ell^{-\prime}\nu\bar{\nu}^\prime$ final states with Powheg and MG5_aMC@NLO setups*, ATL-PHYS-PUB-2021-042, 2021,
URL: https://cds.cern.ch/record/2792254.

Acknowledgements