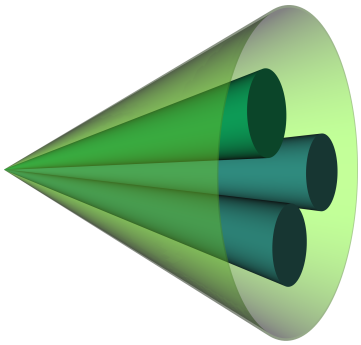
Searches for vector-like quarks with 13 TeV at the ATLAS experiment and development of a boosted-object tagger using a deep neural network



Dissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften

vorgelegt von
Elena M. Freundlich
geboren in Wiesbaden

Lehrstuhl für Experimentelle Physik IV
Fakultät Physik
Technische Universität Dortmund

2020

Diese Dissertation wurde der Fakultät Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften vorgelegt.

1. Gutachter:	Prof. Dr. Kevin Kröninger
2. Gutachter:	Prof. Dr. Dr. Wolfgang Rhode
Vorsitzender der Prüfungskommission:	Prof. Dr. Roland Böhmer
Vertreter der wiss. Mitarbeiter:	Dr. Davide Bossini

Datum des Einreichens der Arbeit: 03.06.2020
Datum der mündlichen Prüfung: 04.09.2020

Diese Doktorarbeit ist durch ein Promotionsstipendium der Konrad-Adenauer-Stiftung finanziert.

Kurzfassung

Diese Arbeit präsentiert zwei Suchen nach der Paarproduktion von vektorartigen Top-Quarks (VLT-Quarks) und vektorartigen Bottom-Quarks (VLB-Quarks) mit Datensätzen bestehend aus $36.1\,\text{fb}^{-1}$ und $139\,\text{fb}^{-1}$. Die Daten wurden in Proton-Proton-Kollisionen bei $\sqrt{s} = 13\,\text{TeV}$ mit dem ATLAS-Detektor genommen, welcher zu den Hauptexperimenten am Large Hadron Collider am CERN gehört. Die Zerfallssignatur dieser Suche nach $T\bar{T}$- und $B\bar{B}$-Produktion enthält ein Z-Boson, welches aus zwei Leptonen (e oder μ) mit gegensätzlicher Ladung und gleichem Flavor rekonstruiert wird und welches zusammen mit einem Quark der dritten Generation produziert wird. Zusätzlich verlangen beide Suchen Objekte mit hohem Transversalimpuls p_T im Endzustand.

In dieser Arbeit wird außerdem ein Multiklassen-Tagger für geboostete Objekte (MCBOT) entwickelt und präsentiert. Dieser verwendet ein tiefes neuronales Netz, um den Ursprung sogenannter *reclustered* Jets mit einem großen Radiusparameter eindeutig zu identifizieren. Dabei wird der Jet einer der vier Klassen V-Boson, H-Boson, Top-Quark oder Untergrund zugeordnet. In der Suche nach VLT- und VLB-Quarks mit $139\,\text{fb}^{-1}$ wird das Vorkommen einer Vielzahl an hadronischen Objekten mit hohem p_T ausgenutzt um mehrere Regionen zu definieren, die sich in der Anzahl der MCBOT- und b-tags unterscheiden.

Mit den zwei verwendeten Datensätzen von $36.1\,\text{fb}^{-1}$ und $139\,\text{fb}^{-1}$ wurden keine Abweichungen von der Standardmodellerwartung beobachtet, weshalb Ausschlussgrenzen auf die Produktionswirkungsquerschnitte der Paarproduktion und die Massen von VLT- und VLB-Quarks unter der Annahme verschiedener Verzweigungsverhältnisse gesetzt werden. Für $36.1\,\text{fb}^{-1}$ werden VLT-Quarkmassen unterhalb von $1102\,\text{GeV}$ unter der Annahme der Dublettmodelle (T, B) und (X, T) ausgeschlossen. Für VLB-Quarks werden untere Massenlimits von $1063\,\text{GeV}$ für das (B, Y) Dublettmodell gesetzt. Die Ergebnisse dieser Suche tragen signifikant zur Sensitivität zweier Kombinationssuchen nach $T\bar{T}$- und $B\bar{B}$-Produktion bei, die vom ATLAS-Experiment mit $36.1\,\text{fb}^{-1}$ durchgeführt wurden und einen größeren Phasenraum für die Existenz von vektorartigen Quarks ausschließen.

In der Suche nach VLT- und VLB-Quarks mit dem gesamten Run-2-Datensatz von $139\,\text{fb}^{-1}$ werden unter der Annahme der genannten Dublettmodelle untere Limits auf die VLT-Quarkmasse von $1350\,\text{GeV}$ gemessen und VLB-Quarkmassen unterhalb von $1316\,\text{GeV}$ ausgeschlossen. Die beobachteten Limits dieser Suche stellen die höchsten Ausschlusslimits für $T\bar{T}$- und $B\bar{B}$-Produktion unter der Annahme eines 100%-igen Verzweigungsverhältnisses in Endzustände mit Zt/b und – im Fall von VLB-Quarks – des (B, Y) Dublettmodells dar, welche bis zum heutigen Tage gesetzt werden konnten.

Abstract

In this thesis, two searches for the pair production of vector-like top (VLT) and vector-like bottom (VLB) quarks are presented using $36.1\,\text{fb}^{-1}$ and $139\,\text{fb}^{-1}$ of data. The data was taken at $\sqrt{s} = 13\,\text{TeV}$ in proton-proton collisions with the ATLAS detector, which is one of the major experiments at the Large Hadron Collider at CERN. The decay signature of this search for $T\bar{T}$ and $B\bar{B}$ production is characterized by the presence of a Z-boson, which is reconstructed from two opposite-charge same-flavor leptons (e or μ) and is produced alongside a third-generation quark. In addition, the analyses target final-state objects with high transverse momentum, p_{T}.

Within the scope of this thesis, the Multi-Class Boosted-Object Tagger (MCBOT) is developed and characterized. It uses a deep neural network, which identifies reclustered large-radius jets originating from V-bosons, H-bosons, top-quarks and background unambiguously. For the search with $139\,\text{fb}^{-1}$, the multiplicity of hadronic high-p_{T} objects is exploited by defining multiple regions according to the number of MCBOT- and b-tags.

No deviation from the Standard Model expectation is found in the observed data for $36.1\,\text{fb}^{-1}$ and $139\,\text{fb}^{-1}$, and exclusion limits on the pair production cross-section and the mass of VLT and VLB quarks are set depending on various branching-ratio assumptions. With $36.1\,\text{fb}^{-1}$, VLT quark masses below $1102\,\text{GeV}$ are excluded with respect to the (T, B) and (X, T) doublet models. For VLB quarks, lower mass limits of $1063\,\text{GeV}$ are set for the (B, Y) doublet. The results of this search significantly contribute to the sensitivity of two combination searches for $T\bar{T}$ and $B\bar{B}$ production performed at the ATLAS experiment with $36.1\,\text{fb}^{-1}$, which constrain the phase-space for vector-like quark existence further.

For the full Run-2 search with $139\,\text{fb}^{-1}$ of data, lower limits on the VLT quark mass of $1350\,\text{GeV}$ are observed and VLB quark masses below $1316\,\text{GeV}$ are excluded assuming the mentioned doublet models. The observed limits of this search for $T\bar{T}$ and $B\bar{B}$ production represent the highest exclusion limits to date assuming the 100% branching ratio into Zt/b final states and – in case of VLB quarks – for the (B, Y) doublet model.

Contents

List of contributions

The research performed in the scope of this thesis was embedded into the research of the ATLAS collaboration. Various resources like the event simulation, object reconstruction, many systematic uncertainties and data taking are provided by the collaboration. In addition, different frameworks for the event selection, the hypothesis testing and a multi-class tagger were used as a starting point for adjustments, advancements and new developments in order to derive the results presented in this thesis. In particular, I collaborated with colleagues from the *Heavy Quarks, Top and Composite Higgs* subgroup within the analysis group of *Exotics*, as well as with the Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ combined performance group of the ATLAS collaboration. My contributions are described in the following:

- For the Multi-Class Boosted-Object Tagger (MCBOT) in Chapter 6, the idea of a multi-class tagger was adapted from Ref. [1] and preexisting code was revised to create the inputs for MCBOT with reclustered jets on flat-in-p_{T} samples. This includes a less memory-dependent approach for the processing of large amounts of Monte Carlo (MC) simulations and the introduction of a two-dimensional reweighting. MCBOT was developed, tested and characterized in this thesis and also presented in novel ways. I performed all studies in Chapter 6 in this thesis. Additionally, I provided a tool to the collaboration in order to apply the tagging decisions of the standard and the b-tag enriched taggers for three b-tagging working points to the ntuples of an analysis.

- For the vector-like top (VLT) and vector-like bottom (VLB) quark search with 2015+2016 data, presented in Chapter 8, I developed the analysis strategy, where first studies were performed in my closely related master thesis [2]. The master thesis presents studies about the boosted topology of the searched channel $Q\bar{Q} \to Z(\ell\ell)q + X$, an optimization of the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts, which were reoptimized and changed in this thesis, and a combination of a resolved and boosted channel in order to derive expected exclusion limits with $3.2\,\mathrm{fb}^{-1}$ of Asimov data using statistical uncertainties only. In this thesis, I performed the main optimization with the larger dataset from 2015+2016 including the creation of control regions and the inclusion of data. In addition, I built the model of systematic uncertainties and used it in a profile likelihood fit in order to perform hypothesis tests. I also derived the results for the $2\ell \geq 2J$ channel with different interpretations about the vector-like quark exclusion assuming various branching ratios. My results were then used by the collaboration and published in a combination of all channels of the $Z(\ell\ell)t/b + X$ analysis [3] and in a collaboration-wide combination of all pair production searches [4] in order to exclude a larger phase-space.

- For the full Run-2 search for VLT and VLB quarks, I developed a new and advanced analysis strategy compared to the previous analysis. Based on MCBOT, a complex categorization was built depending on the number of V-, H- and top-tags. I performed all studies and derived all results presented in Chapter 9. A publication of these results and MCBOT is planned in an ATLAS paper, which is currently being prepared.

List of public contributions:

- Analyzer of the $2\ell \geq 2J$ channel in the $Z(\ell\ell)t/b + X$ search for pair production of VLT and VLB quarks with 2015+2016 data in Ref. [3].
- Poster contribution at LHCP, 2018. Placed first for the poster prize. Title: *Search for pair- and single-production of vector-like quarks in final states with at least one Z boson decaying into a pair of electrons or muons in pp collision data collected with the ATLAS detector at $\sqrt{s} = 13$ TeV*.
- Proceedings [5] for the poster presented at LHCP, 2018.
- Talk at DPG 2017 with the title *Search for Pair Production of Vector-like Quarks in the Boosted Dilepton Channels $Q\bar{Q} \to Zq + X$ with $\sqrt{s} = 13$ TeV at the ATLAS Experiment*.
- Talk at DPG 2018 with the title *Search for Pair Production of Vector-like Quarks in the Boosted Dilepton Channels $Q\bar{Q} \to Zq + X$ with $\sqrt{s} = 13$ TeV at the ATLAS Experiment*.
- Talk at DPG 2019 with the title *Multi-Class Boosted Object Tagger for Reclustered Jets at the ATLAS Experiment*.

Chapter 1
Introduction

Humankind's curiosity has always driven the urge to understand the basic principles of nature in order to gain a better understanding of the past, presence and possibly the future of our universe. At its heart is the pursuit of knowledge of the fundamental building blocks of matter and their interactions. A key discipline aiming for the exploration of the subatomic world is elementary particle physics. The theoretical model which represents our best knowledge to date is called the Standard Model (SM) of particle physics [6–9]. Elementary particles and their interactions at subatomic scales are described and the SM's validity has been tested in a multiplicity of experiments with high precision [10].

However, the SM is unable to answer several open questions in particle physics. The unified description of all four fundamental interactions including gravity has not been achieved yet. Furthermore, the origin of dark matter [11] and dark energy [12, 13], representing the largest part of the matter and energy in the universe, is unknown and one of the biggest puzzles of modern physics. In addition, the origin of neutrino masses [14, 15] and of the baryon-antibaryon asymmetry [16] has not yet been explained. A major milestone in the validation of the SM was reached in 2012, when the Higgs boson (H-boson) was discovered by the ATLAS and CMS experiments [17, 18] and the theoretical description giving mass to most SM particles was verified. However, the size of the H-boson mass, measured to be $(125.38 \pm 0.14)\,\text{GeV}$[1] [19], opens another theoretical question, the *naturalness problem* [20]. In the theoretical formulation of the H-boson mass, enormous fine-tuning of the parameters is required, which is assumed to be improbable. A possible solution is proposed by a variety of models, which introduce supersymmetry, extra-dimensions, an additional strong dynamic, a new global symmetry or combinations of these. Several models predict the existence of very massive fermions with novel properties. These new fermions have the same transformation properties of the left- and right-handed representation under the symmetry group of the weak interaction $SU(2)_L$ and are thus called *vector-like*. In this thesis, vector-like top (VLT) and vector-like bottom (VLB) quarks[2] are of special interest, which carry a fractional elementary charge of $\frac{2}{3}e$ and $-\frac{1}{3}e$, respectively. The hypothesized decay modes for VLT and VLB quarks are $Q \to H/Z + q$ and $Q \to W^{\pm} + q'$, where Q denotes the vector-like quark (VLQ) and q, q' are third-generation quarks. VLQs could be produced and observed via their decay products in high energy collisions at particle accelerators.

The world's most powerful particle accelerator, the Large Hadron Collider (LHC) at CERN, provided proton-proton collisions at a center-of-mass energy of up to $\sqrt{s} = 8\,\text{TeV}$ in Run-1[3] and $\sqrt{s} = 13\,\text{TeV}$ in Run-2[4]. With the taken data, important measurements of, e.g. top-quark and H-boson properties, are performed, but also tests of physics beyond the SM are made. Searches for

[1] Natural units with $\hbar = c = 1$ are used throughout this thesis for energies and masses, given in GeV.

[2] No differentiation between particles and antiparticles is made. The signs of electric charges arise from the context.

[3] Data taking period from 2010–2012 with $\sqrt{s} = 7\,\text{TeV}$ from 2010–2011 and $\sqrt{s} = 8\,\text{TeV}$ in 2012 at the LHC.

[4] Data taking period from 2015–2018 with $\sqrt{s} = 13\,\text{TeV}$ at the LHC.

VLQs are conducted by the ATLAS and CMS experiments, which are two of the four major experiments at the LHC. With the data from Run-1, both experiments were unable to observe VLQs and thus exclusion limits were set. The ATLAS collaboration excluded the existence of VLT and VLB quarks with masses below 715–950 GeV [21] and 575–813 GeV [21] with a dataset of $20.3\,\text{fb}^{-1}$ at 95% confidence level, respectively. The CMS collaboration was able to exclude a similar mass range for a dataset with $19.7\,\text{fb}^{-1}$: VLT quarks with $m_T < 720$–$920\,\text{GeV}$ [22] are excluded and VLB quarks with masses lower than 740 to $900\,\text{GeV}$ [23] are not compatible with the SM at 95% confidence level.

The increase of the LHC's center-of-mass energy from $8\,\text{TeV}$ to $13\,\text{TeV}$ together with a rise of instantaneous luminosity and the increased length of data taking led to a significant boost in sensitivity. Thus, the search for VLQs with Run-2 data is especially interesting and is expected to probe a larger phase-space which is plausible for VLQ existence at $m = 1\,\text{TeV}$ or slightly above [24].

In this thesis, two searches for VLT and VLB quarks are performed with $36.1\,\text{fb}^{-1}$ and $139\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$. The targeted topology requires one Z-boson which is reconstructed from a lepton pair e^+e^- or $\mu^+\mu^-$. With no additional leptons selected, the analyses in this thesis are denoted as the dilepton channel of the $Z(\ell\ell)t/b + X$ search. Both analysis strategies are optimized for final states with high momentum, a so-called *boosted* topology, since the decaying, extremely massive VLQs pass the surplus of mass in form of momentum to their SM decay products.

The first VLQ search [3] using $36.1\,\text{fb}^{-1}$ is a cut-based analysis, which requires at least two large-R jets. A variable with good separation between signal and background is used in a profile likelihood fit to perform hypothesis tests between the SM expectation and data. No excess above the SM expectation is observed and thus exclusion limits are set. In addition, a combination of all channels of the $Z(\ell\ell)t/b + X$ analysis [3] is presented which improves the exclusion limits further. Subsequently, a complex search consisting of a combination of all ATLAS pair production searches [4] with $36.1\,\text{fb}^{-1}$, including this thesis' results, is summarized and again enhances the exclusion limits.

This thesis also presents the development and characterization of a novel tool for the unambiguous identification of a specific type of large radius jet, the reclustered jet, which originates from V-bosons[5], H-bosons, top-quarks and background. A deep neural network is used to build the Multi-Class Boosted-Object Tagger (MCBOT).

Compared to the previously used dataset, the full Run-2 dataset with $139\,\text{fb}^{-1}$ is almost four times larger. Thus, it is promising to renew the $Z(\ell\ell)t/b + X$ search in order to test the SM. The second analysis strategy developed in this thesis is advanced by the use of MCBOT, which increases the sensitivity significantly. An event categorization based on b-, V-, H- and top-tags is simultaneously used in the profile likelihood fit in order to test the SM.

This thesis is structured as follows: An introduction to the SM, a motivation for VLQs and an overview of their properties is given in Chapter 2. It is followed by the experimental setup in Chapter 3, in which the LHC and the ATLAS detector are presented. Subsequently, the analyzed datasets are described and the basic principles of event simulation are given in Chapter 4. The reconstruction of physics objects in the ATLAS detector is discussed in Chapter 5, while MCBOT is presented in Chapter 6. In Chapter 7, statistical methods for hypothesis testing are introduced. The search for VLT and VLB quarks using $36.1\,\text{fb}^{-1}$ is presented in Chapter 8 and followed by the advanced, full Run-2 search in Chapter 9. In Chapter 10, a summary and conclusions are given.

[5] V-boson is a collective term for $W^{\pm}$- and Z-bosons.

Chapter 2
At a glance: Standard Model and beyond

The theoretical model which describes three of the four fundamental forces and the elementary particles interacting accordingly is called the Standard Model (SM) of particle physics. It describes many observations to high precision and is tested exhaustively in low and high energy particle physics. However, many open questions remain and the constant pursuit of a more comprehensive understanding drives this fundamental research. The SM, but also new theories are tested at the frontier of particle physics.

An introduction to the SM and to theories predicting physics beyond the scope of the SM is given in the following. In Section 2.1, the interactions and particles of the SM are described, while in Section 2.2 insufficiencies of the SM are discussed. As a possible solution to some of these shortcomings, theories predicting vector-like quarks (VLQs) are introduced in Section 2.3. The properties of VLQs and their interplay with the SM are presented in order to lay the foundation for experimental searches for these new particles in the later scope of this thesis. Furthermore, an overview of Run-1 searches for pair production of VLQs is given and Run-2 searches at the Large Hadron Collider (LHC) are motivated.

2.1 An introduction to the Standard Model

The SM [6–9] is a relativistic and renormalizable field theory, which describes the elementary particles and three of the four known fundamental interactions. The electromagnetic force acts on particles with an electric charge. With its overarching influence on every-day life, it is the most commonly known interaction besides the gravitational force, which in contrast is not part of the SM. There is also the strong interaction which is the force that holds the constituents of, e.g. protons and neutrons together. The nuclear force is a remnant of it, keeping protons inside atomic nuclei. The third interaction described in the SM is the weak interaction, which for example induces radioactive decays. The symmetry group of the SM is $SU(3)_C \times SU(2)_L \times U(1)_Y$ with one group representation for each interaction. The interactions are expressed as operators, while interacting particles are described by spinor fields.

Fermions

An overview of all elementary particles in the SM and their properties is given in Figure 2.1. The two classes of matter particles are called *fermions*, which are groups of six particles each with spin $s = \frac{1}{2}$. The first group of fermions is called *quarks*, which carry a fractional elementary charge of $\frac{2}{3}e$ for up-type quarks and $-\frac{1}{3}e$ for down-type quarks, where e describes the elementary charge. The up-

and down-type denomination originates from the third component of the *weak isospin* T_3, which is $\frac{1}{2}$ or $-\frac{1}{2}$ for weakly interacting fermions, respectively.

A further subdivision into three generations is done by ordering the quarks according to their mass: The up- and down-quark (u, d) are the lightest quarks, which are the most characteristic constituents of protons and neutrons. They are followed by the charm- and strange-quark (c, s) in the second generation and the top- and bottom-quark (t, b) in the third generation. The top-quark is the most massive elementary particle with a mass of $(173.34 \pm 0.27\,(\text{stat}) \pm 0.71\,(\text{syst}))$ GeV [25].

The other group of elementary matter particles is composed of six *leptons*. As for quarks, an up and a down type with respect to T_3 exist and a subdivision into three generations is present. The two leptons within one generation also differ by their electric charge. While the charged leptons (electron, muon, tau) carry a charge of $-1e$, their partners, the *neutrinos*, are electrically neutral. The mass of the charged leptons increases significantly from the first to the third generation, where the electron is the lightest and the tau the most massive lepton.

All matter consists of fermions. In particular, u- and d-quarks are the building blocks of protons and neutrons, while electrons build the atomic shell.

A counterpart for each fermion exists and is called *antiparticle*. The properties of matter and anti-matter are the same except for opposite additive quantum numbers like the electric charge. Throughout this thesis, particles and physics processes are formulated with particles only; however, the notation always includes the corresponding antiparticles.

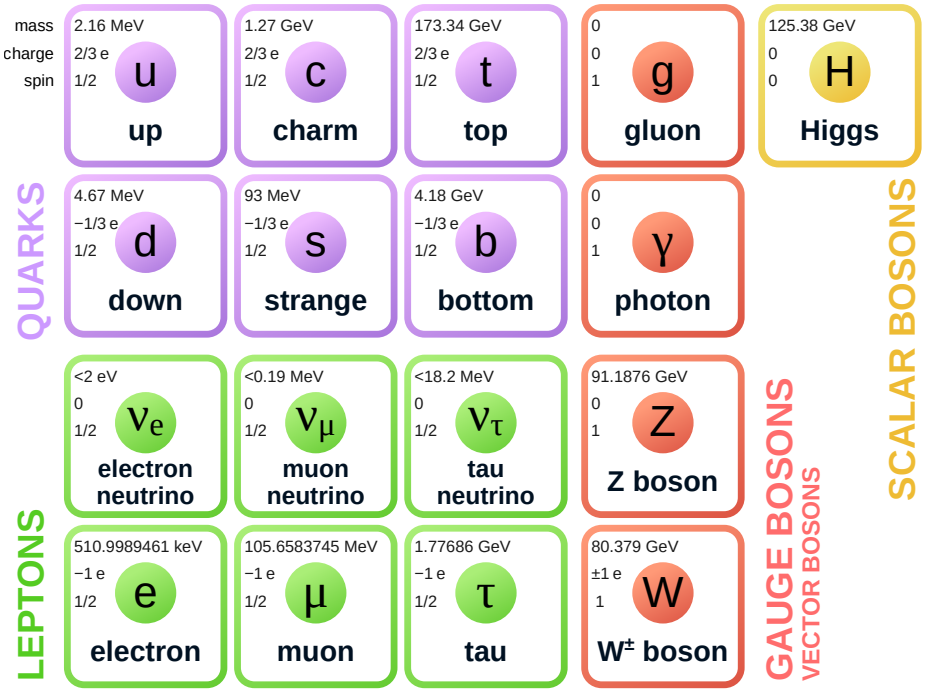


Figure 2.1: Illustration of the elementary particles of the SM and their properties. The graphic is adapted from Ref. [26]. For each particle the mass, charge and spin are shown, as well as their placement in the group of leptons, quarks, vector or scalar bosons. The particle's masses are mainly taken from Ref. [10], while the top-quark and H-boson masses are taken from Refs. [25] and [19], respectively.

Fundamental interactions and their gauge bosons

In the following, a closer look at the interactions between fermions is taken. The electromagnetic (EM) interaction is described by *quantum electrodynamics* (QED) in an abelian gauge theory with the symmetry $U(1)$. The *photon* γ is its mediator, which is an integer spin particle and thus denoted as *boson*. The photon carries spin $s = 1$, is massless and couples to the quantum number *electric charge* Q. Therefore, the photon couples to all quarks and the charged leptons (as well as to the $W^{\pm}$ boson, described in the next paragraph).

The weak interaction is the only interaction which is mediated by massive force carriers. The Z- and $W^{\pm}$-bosons with $s = 1$ interact on very short ranges ($\approx 10^{-16}\,\mathrm{m}$) due to their large masses of $m_W = (80.379 \pm 0.012)\,\mathrm{GeV}$ [10] and $m_Z = (91.1876 \pm 0.0021)\,\mathrm{GeV}$ [10]. The theoretical model of the weak interaction is a non-abelian gauge theory $SU(2)_L$ with a vector$-$axialvector structure. As a result, the weak mediators only couple to left-handed[1] particles and right-handed[2] antiparticles. The non-abelian theory leads to self-interactions of the mediators Z and $W^{\pm}$. The $W^{\pm}$- and Z-boson couple to quarks and leptons, which carry the quantum number T_3 of the weak interaction.

The strong interaction is formulated in a theory called *quantum chromodynamics* (QCD). The non-abelian group $SU(3)_C$ of QCD is a generalization of QED with a force carrier called *gluon* with $s = 1$. The gluons couple to themselves with the quantum number *color charge* C, since these non-abelian gauge bosons carry color charge but are also mediators of the QCD. This quantum number can take the values *red*, *green* and *blue*, or the corresponding anti-colors. Quarks are the only strongly interacting elementary fermions, and they carry one of these colors or anti-colors. The color is changed by gluons, which carry one color and one anti-color and occur in eight configurations. By gluon exchange the color of a strongly interacting particle is changed with the requirement of color conservation. On a macroscopic scale, only color-neutral states are present, which is realized by bound states of three quarks carrying all three (anti-)colors or two quarks carrying a color and the corresponding anti-color. These color-neutral states are denoted as *hadrons* and carry an integer electric charge. Compounds of three quarks are denoted as *baryons*, which includes protons and neutrons. Furthermore, *mesons* exist, which are composed of two quarks.

The coupling constants of the EM, strong and weak interactions are not constant as the name suggests. Instead, the coupling strength is dependent on the energy scale and, thus, called *running coupling*. Due to charge screening in QED, the effective charge decreases with larger distance until it reaches a plateau at the well-known value of the EM coupling constant $\alpha_{\mathrm{em}} \approx \frac{1}{137}$. In contrast, the QCD coupling strength decreases for smaller distances, corresponding to high energies, with values of $\alpha_{\mathrm{s}} \approx 0.1$; however, it increases for larger distances. At observable scales, particles interacting according to QCD experience a strong coupling and thus are bound by so-called *confinement*. On the other hand, they are *asymptotically free* at high energies, allowing a mathematical formulation in perturbative QCD. The strength of the weak interaction is small at low energy scales, but for energies corresponding to the W- and Z-boson mass scale the strength is approximately $1/30$, which exceeds the EM coupling strength. With a coupling of 10^{-38}, gravity is significantly weaker than the other interactions. Thus, the contribution of gravity is negligible in particle physics experiments. However, its description in the same theory with the three interactions of the SM is a very desirable goal.

[1] If the vector of spin and propagation direction of a particle point in opposed directions, the particle is left-handed.
[2] Opposite of left-handed: Spin and direction of motion align.

CKM matrix

The observation of the symmetry violation of the charge conjugation in combination with parity, denoted as CP, in neutral kaon decays in 1964 [27] led to the postulation of the third quark generation by Kobayashi and Maskawa [28]. The introduction of a unitary 3×3 matrix with a complex phase was able to explain the CP violation in the presence of three quark generations. This matrix is denoted as the *CKM matrix* and relates the electroweak eigenstates of the quarks (d', s', b') to the mass eigenstates (d, s, b). While all weak interactions of the quark sector are described in the basis of the weak eigenstates, the mass eigenstates are mixtures of those and represent the observables in experiments. The CKM matrix is defined as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} = V_{\text{CKM}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.1}$$

with the matrix elements V_{ij}, where i describes the up-type quarks u, c, t and j denotes the down-type quarks d, s, b. The CKM matrix is the only mechanism enabling transitions of the quark flavor in the SM. A summary of the current knowledge of the CKM matrix is given in Ref. [10].

Flavor changing neutral currents

The weak interaction is unique in its property to change the flavor of fermions. A process mediated by the $W^{\pm}$-bosons is denoted as *charged current*, while a process via the Z-boson is called a *neutral current*. A change of flavor, like in the β^- decay $d \rightarrow u e^- \bar{\nu}_e$, is allowed on tree level via the charged current. Flavor changing processes with the same charge in the initial and final state are forbidden at tree level and suppressed at loop level by the GIM mechanism [29]. This mechanism supports experimental observations of kaon decays, where the charged version, e.g. $K^+ \rightarrow \mu^+ \nu_\mu$, was observed, while the neutral decay $K_L^0 \rightarrow \mu^+ \mu^-$ was not. The neutral decay or transition is explained with a loop process involving a charged current, which leads to the postulated suppression due to higher order diagrams.

Electroweak unification

The EM and weak interactions are unified in the electroweak interaction with the symmetry groups $SU(2)_L \times U(1)_Y$. The unification was postulated by Glashow, Salam and Weinberg in the 1950's and 60's [7, 8, 30]. The theory is based on the four massless gauge bosons of the electroweak interaction W^1, W^2, W^3 and B^0. The associated quantum numbers are the weak isospin T_3 and the hypercharge Y, which connects Q and T_3 via $Y = 2(Q - T_3)$. The observable gauge bosons $W^{\pm}$, Z and γ are mixtures of these massless bosons and are defined as follows:

$$W_\mu^{\pm} = \frac{1}{\sqrt{2}}(W_\mu^1 \pm i W_\mu^2) \tag{2.2}$$

$$\gamma_\mu = \cos(\theta_W) B_\mu^0 + \sin(\theta_W) W_\mu^3 \tag{2.3}$$

$$Z_\mu = -\sin(\theta_W) B_\mu^0 + \cos(\theta_W) W_\mu^3 \tag{2.4}$$

The weak mixing angle θ_W is directly connected to the mass difference of the Z- and $W^{\pm}$-bosons by $\cos\theta_W = \frac{m_W}{m_Z}$ and is measured to be $\sin^2\theta_W = 0.23122(4)$ [10]. However, the emerging particles Z, $W^{\pm}$ and γ are massless in theory, as all other particles of the SM. The mechanism which enables mass terms in a renormalizable gauge theory is the Higgs mechanism and is described in the next subsection.

A direct consequence of the electroweak unification is the explanation for the difference in Z-boson couplings to quarks and leptons. Due to a mixture of the electroweak gauge bosons B^0 and W^3 for the Z-boson and the photon, the Z-boson does not only couple to T_3, but also to the electric charge. Thus, the different couplings of the Z-boson to quarks and leptons, which carry the same magnitude of T_3, but different electric charges, is explained.

Higgs mechanism

During the exploration of the SM particles and their interactions in the 20$^{\text{th}}$ century, the masses of many elementary particles were measured. However, the theoretical description of the SM did not provide an explanation for the creation of masses. In 1964, Higgs, Brout and Englert suggested the famous *Higgs mechanism* [31–36], where the masses of gauge bosons emerge from the spontaneous symmetry breaking (SSB) of the SM symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. This mechanism was appended to the theory of electroweak unification.

The Higgs field is introduced as a complex doublet of a scalar field $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ with spin zero and with the potential

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2 \, . \tag{2.5}$$

If the parameter μ^2 is negative, the shape of the potential shows a minimum, which is defined as the vacuum expectation value $v \approx 246\,\text{GeV}$ and is responsible for the electroweak symmetry breaking. The EM interaction remains unbroken by the SSB and thus the symmetry group of the SM becomes $SU(3)_C \times U(1)_{\text{em}}$. The unbroken symmetries of QED and QCD lead to massless gauge bosons, while the gauge bosons of the weak interaction acquire their masses. Furthermore, a new particle is created, which is associated with the H-boson with a measured mass of $(125.38 \pm 0.14)\,\text{GeV}$ [19].

The masses of fermions m_f are introduced by the Yukawa coupling, Y_f, of the Higgs field to the fermions f, where the coupling is defined as

$$Y_f = \sqrt{2}\frac{m_f}{v} \, . \tag{2.6}$$

It can be seen that the largest Yukawa coupling in the SM is assigned to the top-quark due to its large mass. A measurement of this coupling [37, 38] provides a test of the SM. The properties of the top-quark, e.g. its large mass show the importance of top-quark physics and might open a window to physics beyond the SM.

2.2 Open questions of the Standard Model

The SM has shown to be an accurate theory with respect to the description of the EM, weak and strong interactions. Many observations up to high precision have confirmed the SM's validity. However, there are inconsistencies within the SM. Furthermore, there is a variety of fundamental questions, which cannot be explained with the SM. An overview of open questions is given in the following section.

An obvious inability of the SM is the missing incorporation of the fourth fundamental interaction, the gravity. A distant goal is the formulation of a *theory of everything* (TOE), where all four interactions are unified close to the Planck scale at $\approx 10^{19}$ GeV [39]. However, the unification of the three interactions described in the SM is an intermediate step, where the SM might be describable in a *grand unified theory* (GUT) at an extremely high energy scale, estimated to be approximately 10^{15}–10^{16} GeV [40].

Furthermore, cosmological observations have shown that the amount of observable matter is significantly too low in order to describe, e.g. rotation curves of galaxies. Thus, the presence of an additional source of matter, denoted as *dark matter* [11], is indirectly measured and its amount is estimated to be of the size of $\approx 84.4\%$ [10] of the total matter density in the universe. The remainder is composed of the commonly known matter described according to the SM. The description of a dark matter particle within the SM is not feasible and experiments searching directly for these particles did not show significant deviations from the SM expectation, which translates to constraints on dark matter models [41–43].

Beside dark matter, there is strong evidence for the existence of so-called *dark energy*, which is associated with the accelerated expansion of the universe [12, 13]. Only the existence of a significantly larger amount of energy than the currently known forms of energy or the modification of general relativity, both summarized under the term dark energy, can explain the rate of the expansion. More details are given in Ref. [44]. Dark energy is assumed to incorporate approximately three quarters of the total energy density of the universe, of which only $\approx 5\%$ are composed of energy known in the SM.

Another open question of the SM is the abundance of matter over antimatter [16]. There is no mechanism in the SM which favors the production of baryonic matter, although no anti-baryonic matter is observed beside the production in particle accelerators or the occurrence in cosmic rays.

More subtle but also significant insufficiencies of the SM are related to a missing description of neutrino masses. In the SM, the form of a mass term for neutrinos is not known; however, the observation of neutrino oscillations [14, 15] indicates that their masses are non-zero and thus need to be described. Furthermore, the SM has a *strong CP problem* [45], which describes the lack of CP violation in QCD. Naturally, the SM Lagrangian contains a term, which allows for a CP violation; however, it is not observed [46].

A question of the SM which motivates the subsequent searches of new physics is the so-called *naturalness problem* [20]. In quantum field theory, the observable H-boson mass m_H receives radiative corrections from the self-energy of the H-boson field [47]:

$$m_H^2 = m_{\text{bare}}^2 + \frac{Y_t}{16\pi^2}\Lambda^2 + \delta\mathcal{O}(m_{\text{weak}}^2) \tag{2.7}$$

The mass term, m_{bare}, is defined as the H-boson mass of the Lagrangian before renormalization, which is a technique controlling divergences in the calculations. The parameter Λ is the cutoff value of the momentum in the top-quark loop, which is the dominant contribution to the self-energy of the H-boson. All massive particles also contribute; however, the Yukawa coupling of the top-quark, Y_t, is of the order of one and the largest Yukawa coupling of any elementary particle. Contributions from the scale around $100\,\text{GeV}$, summarized in the term $\delta\mathcal{O}(m^2_{\text{weak}})$, are negligible for the formulation of the naturalness problem. The validity of the SM is expected to reach up to the Planck scale, which defines the cutoff value Λ and leads to an enormous size of the second term in Eq. 2.7 proportional to Λ^2. In order to obtain the measured H-boson mass of $125.38\,\text{GeV}$, an enormous amount of fine-tuning of the bare mass term m_{bare} at a scale of $10^{35}\,\text{GeV}^2$ is required in order to cancel the contributions of radiative corrections proportional to Λ^2. This approach is considered to be unnatural, and it can be solved by the introduction of heavy fermion partners, which cancel the divergences. However, a fourth chiral family is excluded [48, 49] by electroweak precision tests and the cross-section measurements times branching ratio (BR) of the H-boson [50, 51] in combination with lower mass limits from direct searches for a fourth generation by the ATLAS and CMS collaborations [52, 53].

Instead, theories predicting physics beyond the SM provide a solution to the hierarchy problem and also provide answers about the origin of the SSB. Models of supersymmetry [54] are considered in this regard, which predict an additional particle for each SM particle with a half-integer spin difference. However, there is a variety of models beside supersymmetry. One important class consists of models which predict the H-boson as a pseudo-Nambu-Goldstone boson; thus, they are denoted as PNGB theories. An additional global symmetry is introduced, which is spontaneously broken and allows for the SM H-boson to emerge. The H-boson mass is stabilized by contributions from new heavy fermions, gauge bosons or scalars and thus fine-tuning is avoided. Groups of PNGB theories are the *Little Higgs* [55, 56] and *Composite Higgs* [57, 58] models. A new strong dynamic is predicted, which mixes with the SM, and might lead to a non-elementary H-boson as a compound of new, interacting particles. New resonances like relatively light top-quark partners could emerge at or slightly above the TeV scale [24], which could make them accessible at the LHC. A summary of models beside supersymmetry to address the naturalness problem is given in Ref. [59].

2.3 Vector-like quarks as a window to new physics

New physics models, like Little Higgs or Composite Higgs, are able to solve the naturalness problem of the SM. These theories predict the existence of new massive fermions, which represent a new particle type called *vector-like quarks* (VLQs). A direct search for VLQs is the main subject of the research presented in this thesis. Therefore, their properties, production and decay modes, as well as a motivation for pair production searches at the LHC with Run-2 data are presented.

2.3.1 Properties

An overview of the properties [60] of VLQs is necessary to construct a successful search strategy. The denotation *vector-like* describes the property that the right- and left-handed representation of the new vector-like fermions have the same transformation properties with regard to the symmetry group $SU(2)_L$. Thus, the constraint for the weak interaction to only couple to left-handed particles and

right-handed antiparticles is lifted. The interaction according to the strong and EM force are not altered, and follow the laws of the colored and electromagnetically charged fermions of the SM. Up to four flavors of VLQs are predicted and are denoted as T, B, X and Y. While the vector-like top (VLT) and vector-like bottom (VLB) quarks carry the typical SM up- and down-type electric charges ($Q_T = \frac{2}{3}e$, $Q_B = -\frac{1}{3}e$), the VLX and VLY quarks are predicted with exotic charges of $Q_X = \frac{5}{3}e$ and $Q_Y = -\frac{4}{3}e$. The VLQs can be arranged in seven multiplets, which are illustrated in Table 2.1. VLT and VLB quarks could exist as singlets and together in doublets, while VLX and VLY would only occur in doublets or triplets. In the notation below, the zero superscript describes weak eigenstates, while no superscript denotes mass eigenstates. For VLX and VLY quarks, both eigenstates are identical, when only one of each is present in the multiplet [61]. Only VLB and VLT quarks are considered in the following. Depending on the multiplet assumption, the BRs of the VLQ decays vary, which is described in Subsection 2.3.2.

Table 2.1: The seven gauge-covariant multiplets [61] for VLQs with renormalizable couplings to the scalar doublets of $SU(2)_L$. A zero superscript denotes weak eigenstates. No superscript is assigned for concurring mass and weak eigenstates, which is the case for VLX and VLY quarks.

	Possible multiplets
Singlets	$T^0_{L,R}$, $B^0_{L,R}$
Doublets	$(X\,T^0)_{L,R}$, $(T^0\,B^0)_{L,R}$, $(B^0\,Y)_{L,R}$
Triplets	$(X\,T^0\,B^0)_{L,R}$, $(T^0\,B^0\,Y)_{L,R}$

A mixing of VLQs with the SM quarks is essential to explain their decay topology into H-, $W^{\pm}$- and Z-bosons alongside a chiral quark. Only the third generation quarks, t and b, are considered in the following, motivated by the discussion of the Lagrangian below. Analogously to the mechanism of the CKM matrix, the weak and mass eigenstates of a VLQ and the corresponding third generation quark mix according to a mixing angle $\theta^q_{L,R}$, which is different for the left- and right-handed parts L, R and the up- or down-sector $q = u, d$. The mixing for vector-like top partners is described as

$$\begin{pmatrix} t_{L,R} \\ T_{L,R} \end{pmatrix} = U^u_{L,R} \begin{pmatrix} t^0_{L,R} \\ T^0_{L,R} \end{pmatrix} = \begin{pmatrix} \cos\theta^u_{L,R} & -\sin\theta^u_{L,R} e^{i\phi_u} \\ \sin\theta^u_{L,R} e^{-i\phi_u} & \cos\theta^u_{L,R} \end{pmatrix} \begin{pmatrix} t^0_{L,R} \\ T^0_{L,R} \end{pmatrix} \tag{2.8}$$

with the unitary matrix $U^u_{L,R}$ and the phase ϕ_u. The mixing matrix for the down sector is defined likewise. The mixing of chiral and vector-like quarks modifies the couplings of the SM quarks to the gauge bosons.

The constraints on the Yukawa coupling by the measurements of the H-boson cross-section times BR suggests that any additional massive fermions do not acquire their mass by the Higgs mechanism and the related Yukawa coupling. Thus, the masses of VLQs are defined as direct mass terms in the Lagrangian

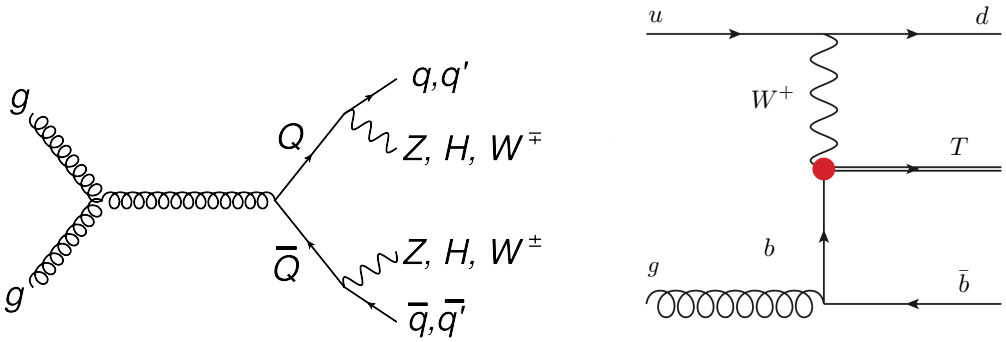
$$\mathcal{L}_{\text{mass}} = -\begin{pmatrix} \bar{t}^0_L & \bar{T}^0_L \end{pmatrix} \begin{pmatrix} Y^u_{33}\frac{v}{\sqrt{2}} & Y^u_{34}\frac{v}{\sqrt{2}} \\ Y^u_{43}\frac{v}{\sqrt{2}} & M^0 \end{pmatrix} \begin{pmatrix} t^0_R \\ T^0_R \end{pmatrix} + B \text{ mixing} + H.c.\,, \tag{2.9}$$

which contains the Yukawa coupling Y^q_{ij} for the the up- and down-type sector $q = u, d$ and the generation index i and j. Together with the vacuum expectation value v the usual chiral fermion mass

of the third generation is defined, while VLQs are assigned a bare mass term M^0, which ensures gauge invariance. Analogous to the mixing of the top-quark and the VLT quark, the mixing in the down-sector is defined with VLB and b-quarks in the Lagrangian in Eq. 2.9. The Lagrangian also motivates the assumption to only consider VLQ decays into third generation quarks: The hierarchy in the mass matrix with larger Yukawa couplings for increasing generations suggests a larger mixing of VLQs with the most massive, third quark generation. Therefore, only the mixing with the third quark generation is studied in the scope of this thesis.

2.3.2 Production and decay

The production of VLQs is possible through the strong interaction, since VLQs transform as triplets of the symmetry group $SU(3)_C$, as any other SM quark. A possible production mechanism of pair produced VLQs, $Q\bar{Q}$, via gluon fusion is illustrated in Figure 2.2(a). Furthermore, the three possible decay modes are depicted, where the notation of W-boson abbreviates both gauge bosons $W^{\pm}$ and is kept throughout this thesis. Decays into a H-, W- and Z-boson are allowed due to the weak mixing with SM quarks and the mass term in the Lagrangian (see Eq. 2.9). These bosons are produced alongside a third generation quark q, q', which is either a top- or b-quark depending on the conservation of electric charge.



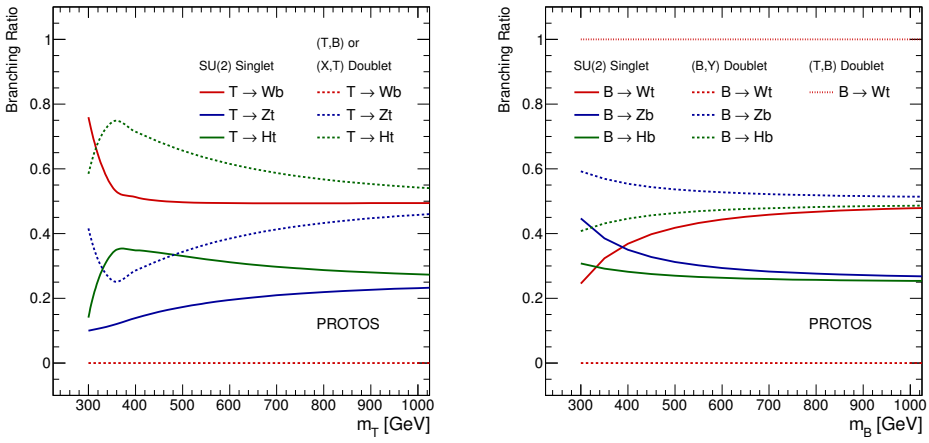
(a) (b)

Figure 2.2: Feynman diagrams [62] (a) for the pair production of VLQs $Q\bar{Q}$ via gluon fusion and the subsequent decay into a W-, Z- or H-boson and a third generation quark q or q'. In (b), single production of the VLT quark in the t-channel is shown.

Additionally, VLQs can be produced in single production via the weak interaction, as shown in Figure 2.2(b) for VLT quark production in the t-channel as an example. Via the exchange of a W- or Z-boson, VLT or VLB quarks are created and decay according to the modes mentioned above. The single production via a H-boson is significantly suppressed.

Another feature of VLT and VLB quark decays should be noted. The decay via a neutral current Z shows that these vector-like fermions introduce flavor changing neutral currents [63, 64] on tree-level due to their mixing with SM quarks.

The BRs of the VLT and VLB quark decay modes depend on their arrangement in a specific multiplet, which is shown in Figures 2.3(a) and 2.3(b), respectively. For VLT quarks, a $SU(2)$ singlet, but also the (T, B) and (X, T) doublets are illustrated as a function of the VLT quark mass hypothesis. It



(a) (b)

Figure 2.3: BRs for a) VLT and b) VLB quark decays for the assumption of different multiplets as a function of the VLQ mass hypothesis, calculated with PROTOS [65, 66]. For VLT quarks, the singlet and the (T, B) and (X, T) doublets are shown. For VLB quarks, the singlet and the (B, Y) and (T, B) doublets are illustrated [62].

can be seen that the BR for the decay $T \to Wb$ is the largest channel in the singlet configuration with approximately 50%. The decays $T \to Z/Ht$ converge towards a BR of $\approx 25\%$ at masses approaching $1\,\text{TeV}$. For the two doublets (T, B) and (X, T), the BRs are equal and the charged current decay is forbidden. The BRs of the final states Z/Ht become more similar at higher masses, and the BRs tend to the high mass limit of 50% for both decays.

In Figure 2.3(b), the VLB quark shows a similar behavior for the singlet BRs, where the charged current decay is dominant. The (B, Y) and (T, B) doublet differ for VLB quarks: For the (T, B) doublet, only the decay $B \to Wt$ is allowed. For the (B, Y) doublet, the charged current is forbidden and the decays into a b-quark and a neutral boson approach a BR of 50%.

It should be noted that an assumption is made on the sum of BRs into the Z/Hq and Wq' final states, which is set to one. This only allows the coupling of VLQs to SM quarks, and no additional new physics is assumed which could alter the sum of BRs into SM final states to a value smaller than one [67].

Furthermore, the choice of BRs into the three decay modes Z/Hq and Wq' with the requirement to sum up to one can be visualized with an isosceles right-angled triangle: The two shorter sides give the BR of two VLQ decay modes and the third BR results from the difference to unity. With this approach, any choice of BRs, which adds up to one, can be visualized as a point in the triangle and allows for a model independent interpretation of search results. The three corners of the triangle are denoted as search *corners* with specific sensitivity into a final state, since they contain the 100% BRs into one decay mode, and the triangle is called *BR plane*. The analyses presented in Chapters 8 and 9 are especially sensitive to the Z-boson corner, which gives maximum sensitivity for $\text{BR}(Q \to Zq) = 100\%$.

2.3.3 Motivation for Run-2 pair production searches

Searches for vector-like quarks have been performed at the LHC since Run-1, which refers to the first years of operation of the LHC [68] from 2010 until 2012 with center-of-mass energies of 7 TeV at the beginning and 8 TeV in 2012.

At the ATLAS [69] experiment, the highest mass limits on the pair production of VLT quarks at 8 TeV with a dataset of $20.3\,\text{fb}^{-1}$ were set at a lower threshold of 715–950 GeV [21] at 95% confidence level (CL). The mass range corresponds to results from a scan of all possible BR combinations $\text{BR}(T \to Ht) + \text{BR}(T \to Wb) + \text{BR}(T \to Zt) = 1$. Two different final state topologies $T\bar{T} \to Wb + X$ and $T\bar{T} \to Ht + X$ were used. The ATLAS collaboration also provided exclusion limits for pair production searches for VLB quarks at 8 TeV. From all VLB quark pair production searches [21, 62, 70, 71], the most sensitive exclusion limits were taken and VLB quark masses lower than 575–813 GeV [21] were excluded at 95% CL depending on possible combinations of BRs into Hb, Zb and Wt final states. A combined search at 8 TeV from the CMS collaboration set limits on the VLT mass of m_T >720–920 GeV [22] at 95% CL. For VLB quarks, the lower mass limits range from 740 to 900 GeV [23] at 95% CL using a dataset of 19.7fb^{-1}.

Since the pair production of VLQs depends solely on the coupling of the strong interaction, the number of parameters to model and test in a search is lower and thus more general than for single production, which cannot be formulated in a model-independent way. The parameter phase-space for single production is larger, since the couplings, c_W and c_Z, of the VLQs to W- and Z-bosons are unknown. These couplings influence the width of the VLQ mass and also the mixing angle with the third generation quarks, which also alters the BRs of the VLQ decay modes. Therefore, multiple parameters under specific assumptions are constrained in single production searches. Under certain assumptions of the SM coupling, often production cross-section times BR limits are set. A few searches for single production have been performed by the ATLAS collaboration [62, 72, 73] in Run-1. Limits on the production cross-sections and an interpretation on the coupling c_W were derived.

With more data taken by the LHC experiments in Run-2, which describes the data taking period from 2015–2018, it is reasonable to repeat and intensify the search program of the ATLAS and CMS experiments. However, not only the larger amount of recorded data motivates a continuation of the VLQ searches, but also the increased center-of-mass energy to create new particles and the increased parton luminosity at 13 TeV compared to 8 TeV suggest Run-2 searches. In order to assess the reach of a collider, it is more meaningful to investigate the parton luminosity than the pp luminosity of the LHC, since it directly describes the interacting constituents of colliding protons and thus provides information about the distribution of the additional center-of-mass energy among the constituents. The parton luminosity ratio of 13 TeV and 8 TeV is illustrated in Figure 2.4(a) and shows a comparison of different parton interactions, which are gluon fusion gg, quark-gluon qg and all quark-antiquark $\sum q\bar{q}$ interactions. The available energy to create new particles is denoted as M_X, which corresponds to $2m_{\text{VLQ}}$ in this context. It can be seen that the ratio increases steep as a function of M_X for all parton interactions. At a hypothetical VLQ mass of 1 TeV, the parton luminosity for VLQ pair production is 15–20 times larger at 13 TeV than for a center-of-mass energy of 8 TeV, assuming gluon fusion as the production mechanism. Therefore, searches for new physics with Run-2 data have an increased chance of opening a window to new physics.

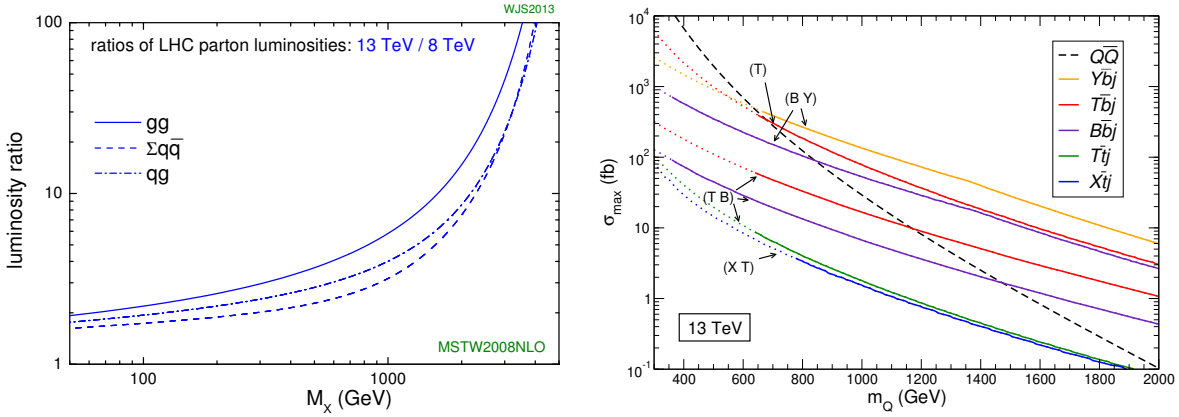


Figure 2.4: In a), the parton luminosity ratios [74] for 13 TeV and 8 TeV are illustrated. The energy transfer M_x for the production mechanisms gluon fusion gg, quark-gluon qg and all quark-antiquark $\sum q\bar{q}$ interactions is shown. In b), a depiction of the maximum cross-section limits [60] on VLQ single production is given, derived from indirect constraints in Ref. [60]. Dotted lines show excluded cross-sections by direct searches. Furthermore, the pair production limit is depicted with the dashed line.

A comparison for pair and single production cross-sections according to the multiplets with largest allowed weak SM coupling is given in Figure 2.4(b) for a center-of-mass energy of 13 TeV. The maximum production cross-section from indirect constraints on single production is shown and excluded parts from direct searches are illustrated as dotted lines. The illustration shows that the production cross-section for single production can be larger than for pair production and it varies depending on the multiplet assumption. However, these are the most optimistic cases of multiplets and the maximum production cross-sections from indirect constraints. If the couplings of VLQs to the gauge bosons W and Z are very small, the single production cross-sections are altered and might not be large enough for detection at an experiment. On the other hand, for single production the center-of-mass energy in the collision only needs to be large enough to create a single VLQ in contrast to pair production. Thus, it is not evident whether pair or single production are more promising and complementary search programs are pursued by the ATLAS and CMS collaborations.

The searches that are described in this thesis focus on pair production, since VLQs with a mass hypothesis around the TeV scale could mainly be produced as pairs and thus discovered or excluded with the best limits. The analyses performed in this thesis are embedded in a larger search program of the ATLAS collaboration. The results of the first search using $36.1\,\text{fb}^{-1}$ of the Run-2 data are compared with the current state of research in Section 8.9 to 8.11. Up to date, no published VLQ searches using the full Run-2 dataset are accessible. Thus, the search presented in Chapter 9 presents first results of VLQ pair production searches for the full Run-2 dataset with $139\,\text{fb}^{-1}$.

Chapter 3
Experimental setup

In order to test the Standard Model (SM), the energy for the production of its particles needs to be available. In nature, extremely high energy particles are of astronomical origin and can be studied with astrophysics experiments, using the earth's atmosphere as a particle detector. An alternative is the construction of a powerful particle accelerator, which initiates particle collisions at unprecedented man-made energies. The center-of-mass energy of the collisions is converted into the creation of new particles, and their decay products are detected in order to gain insight into the production and decay mechanisms of the SM. Not only precision measurements of the SM are done, but also searches for new physics are performed. In the following chapter, the Large Hadron Collider (LHC) at the European Laboratory for Particle Physics, abbreviated as CERN, is introduced in Section 3.1. Subsequently, the ATLAS detector, located at the Insertion Point 1 of the LHC, is presented in Section 3.2. It was used to collect the data, which are analyzed in this thesis.

3.1 The Large Hadron Collider at CERN

The LHC [68] at CERN is the largest particle accelerator world-wide with a circumference of $26.7\,\mathrm{km}$. With a center-of-mass energy of 13 TeV, reached in proton-proton collisions during Run-2 in the years 2015–2018, it is also the most powerful accelerator ever built. In beam pipes with ultra high vacuum, two proton beams circle the LHC ring in opposite directions. They are focused and crossed at four interaction points. At the interaction point in the ATLAS detector, bunches of up to 10^{11} protons collide with a frequency of $40\,\mathrm{MHz}$ with a peak luminosity of $19 \cdot 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [75]. Inelastic interactions occur 10^9 times per second, assuming 25 proton-proton collisions per bunch crossing. All of these simultaneous interactions except the triggered event are denoted as pile-up and lead to the creation of around 1000 particles every $25\,\mathrm{ns}$. The highest means of interactions per crossing, $<\mu>$, were achieved in the years 2017 and 2018 with $<\mu>=37.8$ and $<\mu>=36.1$ [76], respectively, leading to more activity in the detector and a higher recorded luminosity compared to Run-1.

The LHC is not only operated with proton bunches, but also performs heavy ion runs at the end of a year. Lead or xenon nuclei are accelerated and brought to collision at a center-of-mass energy of up to $5.44\,\mathrm{TeV}$ per nucleon, and the creation of extremely dense states, formed in collisions, allows the study of the quark-gluon-plasma.

In order to reach these unprecedented energies at a particle accelerator, a system of preaccelerators is used, before the beams are injected into the LHC. First, protons are obtained from hydrogen atoms, which are dismantled using an electromagnetic (EM) field. Then, the protons are accelerated in the Linear Accelerator 2 (LINAC 2) up to an energy of $50\,\mathrm{MeV}$ and then passed on to the Proton Synchrotron Booster (PSB). When the proton beams have reached an energy of $1.4\,\mathrm{GeV}$, they are

transferred to the Proton Synchrotron (PS), which increases the particle energy to 25 GeV. The last of the four preaccelerators is the Super Proton Synchrotron (SPS), where the energy is again increased by an order of magnitude up to 450 GeV. Then, the protons reach the LHC and are brought to their maximum energy of 6.5 TeV per proton beam. Using a system of superconducting dipole electromagnets for deflection and quadrupole electromagnets for focusing, the protons are guided through the ring.

The LHC provides the four major experiments, ATLAS [69], ALICE [77], CMS [78] and LHCb [79], with particle interactions, which they need in order to perform their measurements. ALICE is a heavy-ion detector and studies the quark-gluon plasma. LHCb searches for differences between matter and antimatter and utilizes b-hadrons for high precision measurements. ATLAS and CMS are built with coinciding objectives and their independent research provides high confidence in matching results.

3.2 Design of the ATLAS detector

The ATLAS (A Toroidal LHC ApparatuS) detector [69] is built as a general purpose detector. The observation of the H-boson was one of the main goals of the ATLAS experiment, and was achieved in July, 2012 [17]. This break-through is supported by the same observation [18] made by the CMS experiment. The ATLAS detector is also designed to measure different decay channels and other quantities of the H-boson. Furthermore, the top quark, which is the most massive elementary particle, is produced in high numbers at the LHC and could provide a window to new physics. SM properties like the H-boson and top-quark masses are measured with high precision. Another goal is the search for new physics, in which the existence of, e.g. supersymmetric, dark matter or vector-like particles is tested.

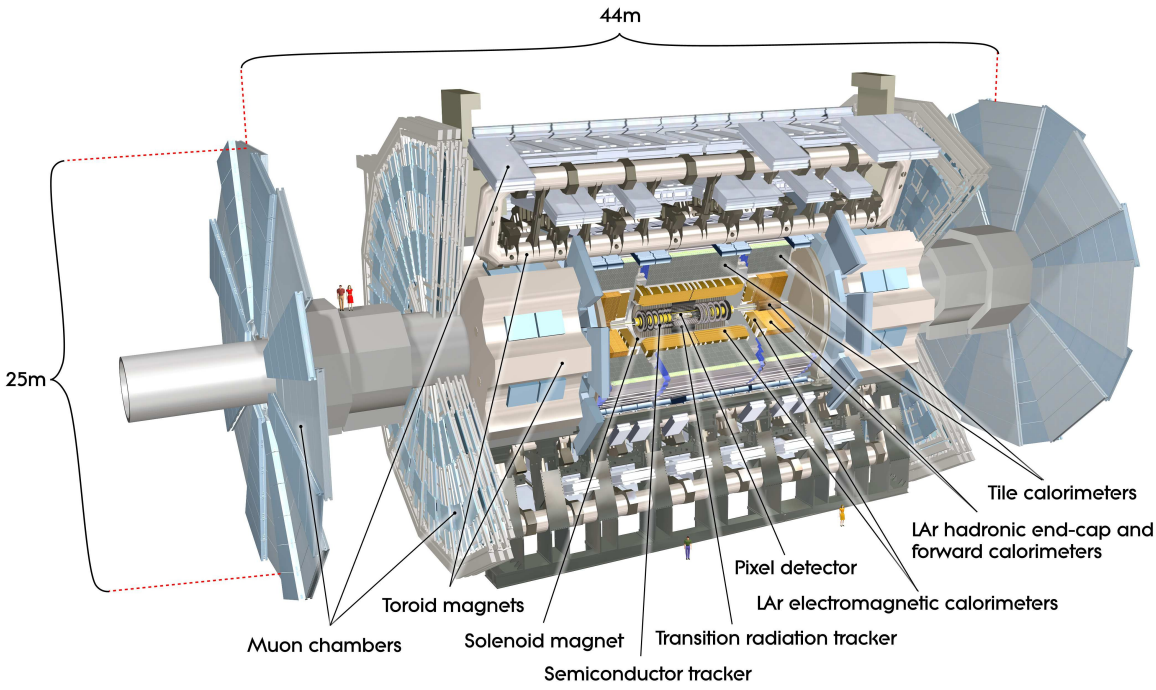


Figure 3.1: Illustration of the different detector layers of the ATLAS detector [69].

In Figure 3.1, a cut-away view of the ATLAS detector is shown. With a height of $25\,\mathrm{m}$, a length of $44\,\mathrm{m}$ and a mass of 7000 tons, it is the detector with the largest volume constructed for a particle collider up to the present day. Four subsystems work together for the detection of particles and their decays. A cylindrical design is chosen for maximum hermeticity, which ensures that the particles, emerging from the collision point and spreading in all directions, are detected. Detector parts, which are arranged cylindrically around the beam axis are denoted as *barrel* parts, while arrangements perpendicular to the beam axis are called *end caps*. Around the interaction point, a particle tracker, called the *Inner Detector* (ID), is located, enclosed by a strong magnetic field. It is surrounded by an *electromagnetic and hadronic calorimeter system*. The outermost part consists of *muon chambers*, again embedded into a magnetic field, together denoted as the Muon Spectrometer (MS). All elementary particles, except neutrinos, can be detected with these systems directly or via their decay products. An indirect technique to measure neutrino contributions is discussed in Chapter 5. The four subsystems and a subsequent trigger system, to filter events of interest, are presented in the following subsections.

The coordinate system of the ATLAS detector is set as follows: The interaction point is defined as the origin of the coordinate system, and the z direction points along the beam axis. Perpendicular to z, the x-axis is directed towards the LHC's center and the y-axis points upwards. This x-y-plane can also be expressed in spherical coordinates as the R-ϕ-plane, with the radius R as distance from the beam axis and the azimuthal angle ϕ. The polar angle θ specifies the inclination, which is expressed as the pseudorapidity $\eta = -\ln(\tan\frac{\theta}{2})$. The plane, which is perpendicular to the beam axis and intersects with the origin, corresponds to $\eta = 0$. The detector's dimensions extend up to $|\eta| = 4.9$. Distances between objects in the detector are calculated with the measure $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.2.1 Inner Detector

The ID is a tracking detector, which is covered in a $2\,\mathrm{T}$ solenoid field. Charged particles leave hits in multiple, position-sensitive detector layers, which are reconstructed to tracks. The track curvature due to the magnetic field provides a measurement of the momentum and charge of the traversing particles. Furthermore, good reconstruction of the primary and secondary vertex is accomplished. These are crucial elements for efficient, robust and accurate object reconstruction.

The layout of the ID is depicted in Figure 3.2 and shows the Pixel Detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). Both the Pixel Detector and SCT are semiconductor trackers, and consist of arrays of reversely biased diodes, implemented in silicon. By applying a high operating voltage, the depletion zone in the diodes is enlarged and charged particles, passing the detector layers, cause strong signals. The Pixel Detector has the highest granularity of the sub-detectors, with pixels mostly of the size of $50 \times 400\,\mu\mathrm{m}^2$, which allow high spatial resolution. During the long shutdown after Run-1, a new detector layer, the Insertable B-Layer (IBL) [80] was added to the Pixel Detector. It reduces the distance of the innermost layer to the beam pipe from $5\,\mathrm{cm}$ to $3.3\,\mathrm{cm}$, and provides even smaller granularity. The IBL is followed by three pixel layers, which cover the region $|\eta| < 2.5$. A charged particle passes four layers yielding four space points to build a track. Altogether, the Pixel Detector provides around 92 million readout channels.

The SCT consists of silicon microstrips, which are $12\,\mathrm{cm}$ long in direction of the beam axis and $80\,\mu\mathrm{m}$ wide. Each module is built in a way that one layer of microstrip sensors is attached on the front and rear of the module. One layer is rotated by $40\,\mathrm{mrad}$ against the other, which allows a significantly

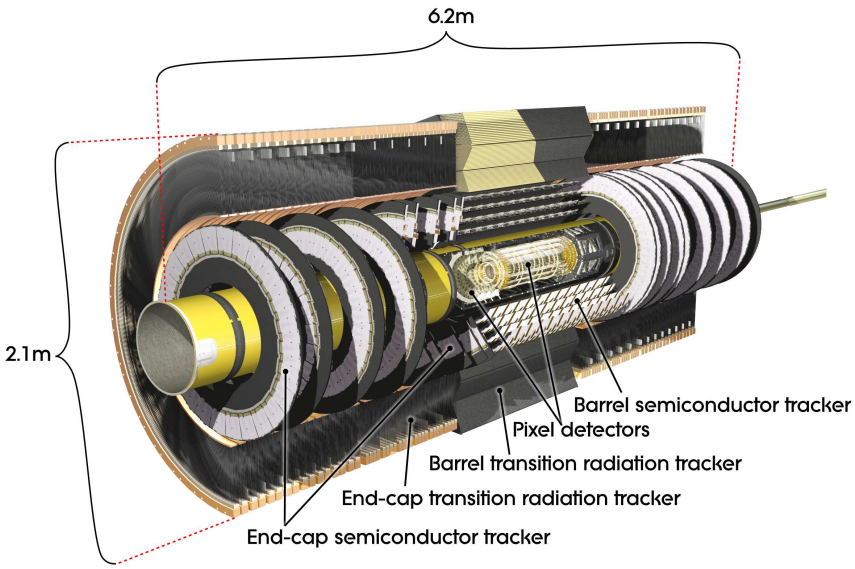


Figure 3.2: Illustration of the ID layers of the ATLAS detector [69].

improved resolution of $0.56\,\mathrm{mm}$ in z direction, compared to the length of $12\,\mathrm{cm}$. A total of eight layers is present to reconstruct a charged particle's track, but altogether, only four space points can be extracted. The SCT covers both the barrel and the disk regions up to $|\eta| < 2.5$ with a total of 6.3 million readout channels.

In the TRT, charged particles provide information for tracking by gas ionization. The TRT also uses the property of highly relativistic particles to emit transition radiation when passing different materials with different permittivities. The detector consists of straw tubes with a diameter of $4\,\mathrm{mm}$, which are filled with a xenon based gas mixture. In the middle of the straw, a tungsten wire, plated with gold, is used as anode. The intermediate gaps between straws are filled with polypropylene fibers, where transition radiation occurs, which is then absorbed by the xenon gas. As the probability for the emission of photons is a function of the γ factor, electrons can be discriminated from heavier particles. In the barrel region, the $144\,\mathrm{cm}$ long straws are read out from each end, while in the disks the straws with $37\,\mathrm{cm}$ length correspond to one channel. A total of 351000 readout channels is provided, which supports the tracking of the Pixel Detector and the SCT up to a region with $|\eta| < 2.0$. The TRT improves the momentum resolution significantly, due to a large number (≈ 36 in the barrel region) of hits per track.

3.2.2 Calorimeters

The calorimeter system is divided into the electromagnetic and hadronic calorimeters and covers a detector region up to $|\eta| < 4.9$. The calorimeters are used for an energy measurement together with the determination of the location of the energy deposition. For the energy measurement, the containment of the electromagnetic and hadronic showers inside the calorimeters plays a crucial role. The radiation length of the calorimeters is optimized for full coverage, so that punch-through effects, describing the leakage of showers beyond the calorimeters, are minimized. All calorimeter systems

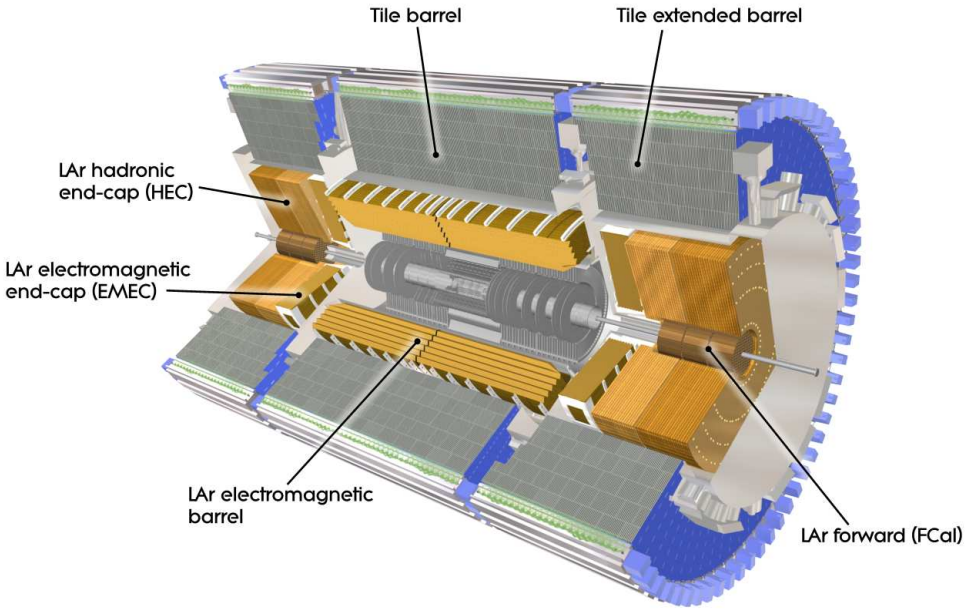


Figure 3.3: Illustration of the calorimeter system of the ATLAS detector [69].

are sampling calorimeters with alternating layers of active and absorber material, reducing the size and costs in comparison to a homogeneous calorimeter, which solely consists of active material.

The electromagnetic calorimeter stops any purely electromagnetically interacting particles, which are electrons and photons. A finer granularity of the calorimeter is used in the central detector region, while for higher $|\eta|$ the granularity is reduced. The barrel and two end-cap calorimeters are illustrated in Figure 3.3. For all electromagnetic calorimeters liquid argon is used as active material, permeated with lead as absorber.

The hadronic calorimeters are constructed for the energy measurement of hadronic cascades, which are longer and wider than electromagnetic showers. The accurate measurement is important for the reconstruction of jets and the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, which describes the momentum imbalance per event and is further explained in Chapter 5. In the central detector region, the hadronic tile calorimeter is located, which uses steel as an absorber and scintillating tiles to measure energy deposits. The same holds for the extended tile barrel calorimeters, enlarging the reach of the barrel up to $|\eta| < 1.7$. Furthermore, there are two end-caps, which both utilize alternating layers of liquid argon and copper plates.

The forward calorimeters provide energy measurements up to $|\eta| < 4.9$, reaching beyond $|\eta| = 3.2$, which is the boundary for the end-cap calorimeters. Therefore, the coverage for jet reconstruction is increased significantly. As active material, again, liquid argon is chosen. The first layer of the forward calorimeter consists of copper as absorber, the second and third layers are composed of tungsten. The use of a higher density material than copper ensures that hadronic showers are stopped within the shorter calorimeter extension compared to the end-caps.

3.2.3 Muon spectrometer

The only detectable particles which traverse the calorimeter system due to their minimal ionizing property are muons. A three layered gas detector system in the barrel and end-cap regions up to $|\eta| < 2.7$ is used to measure the muon momentum and charge with precision tracking chambers. Furthermore, trigger chambers collect information about interesting events in the region $|\eta| < 2.4$, e.g. by providing a quick momentum measurement. Additionally, they complement the tracking chambers with a second coordinate measurement.

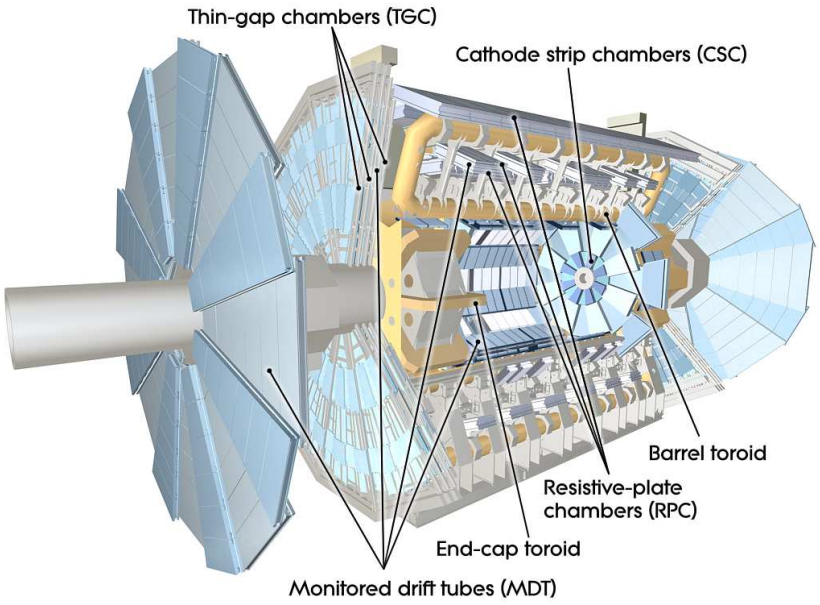


Figure 3.4: Illustration of the muon system of the ATLAS detector [69].

As for the other detector parts, the barrel system is arranged cylindrically around the beam axis, while the end-caps consist of large planes perpendicular to it. The barrel muon system is enclosed in a magnetic field, provided by three toroidal superconducting magnets with eight coils each. As illustrated in Figure 3.4, one large barrel toroid with a length of $25.3\,\mathrm{m}$ and a diameter of $20.1\,\mathrm{m}$ encloses the central detector part. Together with smaller end-cap magnets, starting at $|\eta| > 1.6$, a magnetic field transverse to the incoming muons is created with a maximum strength of $4\,\mathrm{T}$ in a region up to $|\eta| < 2.7$.

The tracking chambers mostly consist of Monitored Drift Tubes (MDTs) with a diameter of $3\,\mathrm{cm}$, operated with an Ar/CO_2 gas mixture at $3\,\mathrm{bar}$. Traversing muons ionize the gas and the liberated electrons are collected at a central wire at an operating voltage of $3000\,\mathrm{V}$. One coordinate in the bending plane is measured, and no measurement along the tubes is possible. In the first layer of the barrel for high $|\eta|$, the drift tubes are not suitable due to high activity. Instead, Cathode-Strip Chambers (CSCs) are used beyond $|\eta| = 2.0$, arranged in two disks, as depicted in Figure 3.4. These multiwire proportional chambers are filled with an Ar, CO_2 and CF_4 gas mixture and function well in high particle densities. An alignment system of 12000 optical sensors is used in combination with track-based alignment algorithms to provide the required resolution in the tracking chambers of a few tens of $\mu\mathrm{m}$.

The trigger chambers consist of Resistive Plate Chambers (RPCs) for $|\eta| < 1.05$, which are sets of parallel electrodes filled with gas and operated with an electric field of $4.9\,\mathrm{kV/mm}$. Following the RPCs, Thin Gap Chambers (TGCs) are used in the detector region up to $|\eta| < 2.4$, which are multiwire proportional chambers. Fast signals are provided by both detector systems, which are fundamental for triggering. As mentioned earlier, the trigger chambers complement the MDT coordinate measurement with a second one in the non-bending direction, which is the ϕ coordinate.

3.2.4 Trigger system

A large amount of data for analysis is necessary in order to measure or detect signals with low production cross-sections or small branching ratios. The high interaction rate with a frequency of $40\,\mathrm{MHz}$ creates enormous amounts of data to be recorded and stored. However, it is not feasible from a technical point of view to record the signals from over 100 million readout-channels at such a high rate. Furthermore, it is not necessary to record every interaction, for example interactions that do not provide benefit for searched signals. Therefore, a trigger system is used in order to filter data and select only events which contain processes of interest. Especially, high momentum particles or final states with high (missing) energy are targeted.

At first, the Level-1 (L1) trigger uses basic, less granular information from calorimeter triggers and the trigger chambers of the muon system. With this selection, regions of interest (RoI) are built, and the reduced data rate of $100\,\mathrm{kHz}$ is passed to the High-Level Trigger (HLT) [81]. Subsequently, the information from the RoIs is processed with high precision information from all detector systems. This leads to a further decrease in data, and the total rate reduction to $1\,\mathrm{kHz}$ by L1 and HLT is four orders of magnitude. In the end, recorded events are transferred to mass storage.

Chapter 4
Run-2 dataset and event simulation

The proton-proton collisions, provided by the Large Hadron Collider (LHC), are used by the ATLAS detector to record as much data for physics analyses as possible. Except during shutdowns for technical interventions, the ATLAS detector is operated continuously and records data during so-called *runs*, which are periods of high intensity stable beams in the LHC. The recorded dataset, utilized in this thesis, is presented in Section 4.1. Subsequently, the principle of Monte Carlo (MC) simulation in high energy physics (HEP) and its major steps are introduced in Section 4.2. At last, the concepts of *signal* and *background* are explained in Section 4.3.

4.1 Analyzed datasets

The dataset analyzed for the two analyses presented in Chapters 8 and 9 are recorded in the years 2015+2016 and 2015 to 2018 of Run-2, respectively. The amount of recorded data is specified with the integrated instantaneous luminosity $\mathcal{L}$, which gives a measure for the number of events N produced for a cross-section σ:

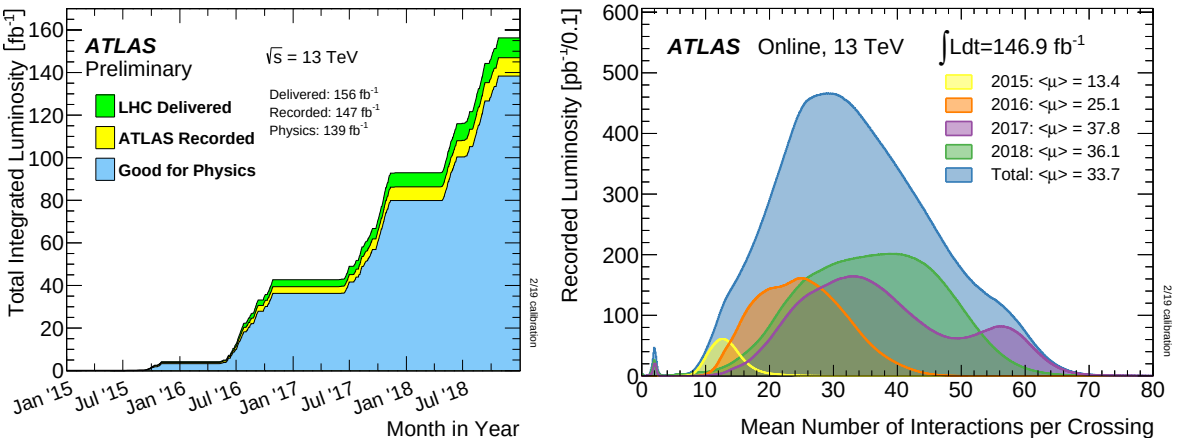
$$\int \mathcal{L}\,\mathrm{d}t = \frac{N}{\sigma} \tag{4.1}$$

Two forward detectors [82, 83] are installed in the ATLAS environment for the luminosity measurement. While LUCID-2[1] [84] measures the luminosity during data taking, ALFA[2] performs a calibration of LUCID in order to derive the absolute luminosity. The calibration is done in dedicated run conditions.

The integrated luminosity is illustrated in Figure 4.1(a) as a function of the years of data taking during Run-2. A total of $\int \mathcal{L}\mathrm{d}t = 156\,\mathrm{fb}^{-1}$ was delivered by the LHC in stable beams, of which $147\,\mathrm{fb}^{-1}$ were recorded by the ATLAS detector in Run-2. A small time delay between the moment of the stable-beam flag, set by the LHC operators, and the full effective operation of the ATLAS detector is caused by the so-called *warm start*. The detector is set into a ready state by turning on the high voltage of the Inner Detector. Afterwards, data-quality requirements are enforced and the final dataset, to be used by analysts, is listed as "Good for Physics". The total integrated luminosity matching these criteria in Run-2 yields $139\,\mathrm{fb}^{-1}$. It can also be observed that the integrated luminosity in the years 2017 and 2018 increased significantly compared to 2015+2016. Due to a larger average collision rate of bunches, as well as stronger focusing of the beams at the interaction point, the instantaneous luminosity was driven up.

These changed beam properties also lead to an increased activity in the detector. The average number of pile-up events $<\mu>$ is depicted in Figure 4.1(b). It can be seen that the mean increased

[1] **L**uminosity measurements **U**sing **C**herenkov **I**ntegrating **D**etector
[2] **A**bsolute **L**uminosity **F**or **A**TLAS



(a) (b)

Figure 4.1: Illustration of a) the integrated luminosity recorded during Run-2 and b) the average number of pile-up events $<\mu>$ per bunch-crossing for all years of data taking during Run-2 of the ATLAS detector [76].

significantly and that bunch-crossings with up to approximately 70 simultaneous inelastic proton-proton collisions occur. This environment makes great demands on the detector systems, presented in Chapter 3, and also the reconstruction of physics objects, to be discussed in Chapter 5.

4.2 Monte Carlo simulation in HEP

In order to compare the observed dataset to the predictions made by the Standard Model (SM), MC simulations are used in HEP. These simulations try to mimic nature as accurately as possible from the incoming particles of the hard scattering to the response of the detector with all its randomness. Repeated random sampling according to the probability density function (pdf) of a specific SM process is applied. Whenever new physics beyond the SM is tested, MC simulations can also be generated for a modified probability model with respect to the SM including new particles or interactions. The multiplicity of steps taken to obtain an MC simulation is described in the following.

At first, the hard scattering is modeled, which describes the inelastic scattering of two incoming protons. Protons are not only composed of valence quarks, but also of sea quarks and gluons, as shown by lepton-nucleon scattering experiments [85, 86]. These hadron constituents are called *partons*. The probability density functions of partons within, e.g. a proton, can be measured and are described as *parton distribution functions* (PDFs). They describe the probability of finding a parton a with momentum fraction x in a larger hadronic compound, like the proton, whose PDF is denoted as $f_a(x_a, \mu_F)$ in the following. When an inelastic scattering at the LHC occurs, any of the partons from both protons can interact with momenta following the underlying PDFs. These are some of the unknown properties of a collision, which make hadron collisions complex systems. Scatterings of high momentum partons are the desired processes to be studied at the LHC, since only they can provide the required energies for particle production in top-quark, H-boson or beyond-SM physics. Therefore, MC simulations are performed in the phase-spaces of interest.

The cross-section of the process, which creates n final state particles from the interaction of partons a and b from two protons, is described with the *factorization* theorem:

$$\begin{aligned}\sigma &= \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b \int f_a(x_a, \mu_F) f_b(x_b, \mu_F) \mathrm{d}\sigma_{ab\to n}(\mu_F, \mu_R) \\ &= \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b \int \mathrm{d}\Phi_n f_a(x_a, \mu_F) f_b(x_b, \mu_F) \cdot \frac{1}{2\hat{s}} |\mathcal{M}_{ab\to n}|^2(\Phi_n; \mu_F, \mu_R) \end{aligned} \tag{4.2}$$

A sum over all possible parton collisions, possibly leading to the final states n is done, as well as all possible momentum fractions of these partons are summed. The PDFs are integrated over the parton-level cross-section $\mathrm{d}\sigma_{ab\to n}(\mu_F, \mu_R)$. This can also be expressed with the squared matrix element $|\mathcal{M}_{ab\to n}|^2$, which is averaged over spin and color charge of partons a and b, and the parton flux $\frac{1}{2\hat{s}} = \frac{1}{2x_a x_b s}$ with the squared center-of-mass energy s. The integral is then calculated over the differential phase-space element $\mathrm{d}\Phi_n$ for all final state particles n. The factorization theorem allows the splitting of the perturbative and non-perturbative regime: The PDFs describe the non-perturbative hadron content of a proton and are a proton property independent of the kinematics of the process. They are separated from the cross-section of the hard scattering at the factorization scale μ_F and need to be determined with data. The hard scattering happens at a high energy scale, where the strong coupling is small, the partons are asymptotically free and the process can be described with perturbative quantum chromodynamics (QCD).

The simulation of the hard scattering starts with the generation of all possible Feynman diagrams up to a specified order in QCD (typically next-to-leading order (NLO)). The corresponding matrix elements are then calculated numerically and the cross-section is obtained ($\sigma \propto |\mathcal{M}|^2$). As a result, an accurate simulation of the momenta of the outgoing particles is achieved.

In order to incorporate higher order corrections to the simulation of the hard scattering, parton showers are used. Especially soft and collinear radiation from any of the incoming or outgoing particles is poorly described by fixed order matrix elements, while parton showers perform well in this regime. The color charged states emerging from the hard interaction emit gluons, which can in turn radiate quark pairs or gluons themselves. The radiation can either be emitted from the initial or final state particles, which is called ISR or FSR, respectively. As for the hard scattering, this cascade of strongly interacting particles occurs at high energy scales, describable with perturbative QCD. The parton shower reduces a small number of high energy particles to a multiplicity of low momentum final states, where the non-perturbative regime sets in at low momentum transfers at the QCD scale $\Lambda_{\mathrm{QCD}} \approx 1\,\mathrm{GeV}$ and the partons are confined in hadrons. In order to exploit the advantages of the matrix element calculation and the parton shower algorithm, matching and merging [87] of both is performed. Processes of quantum electrodynamics (QED) are simulated with the same procedure, by using the electromagnetic charge instead of the color charge.

For the modelling of any non-perturbative contributions, phenomenological models are utilized. The transition from partons to quarks and gluons, which are confined in hadrons, is described by the process of *hadronization*. Different models are implemented in the following multi-purpose event generators: The string model [88, 89] is used in PYTHIA [90], while the cluster hadronization model [91, 92] is incorporated in HERWIG [93] and SHERPA [94].

Furthermore, contributions from low momentum gluons, which are denoted as *soft* radiation, are considered. As described earlier, there is a multiplicity of partons present inside a proton. Therefore, multiple interactions between the partons of two colliding protons can occur, which can lead to additional hard scatterings, but mostly contribute soft radiation. All soft contributions emerging outside of the hard-scattering, as well as outside of ISR and FSR, are called the *underlying event*. These contributions to the MC simulation are determined with tuned parameters using data. As an additional step of the MC simulation, the modelling of unstable hadron decays is performed with, e.g. EVTGEN [95].

Subsequently, the detector response is simulated in order to get a full picture of the expected signal from a hard scattering. The ATLAS detector simulation [96] uses the software package GEANT4 [97].

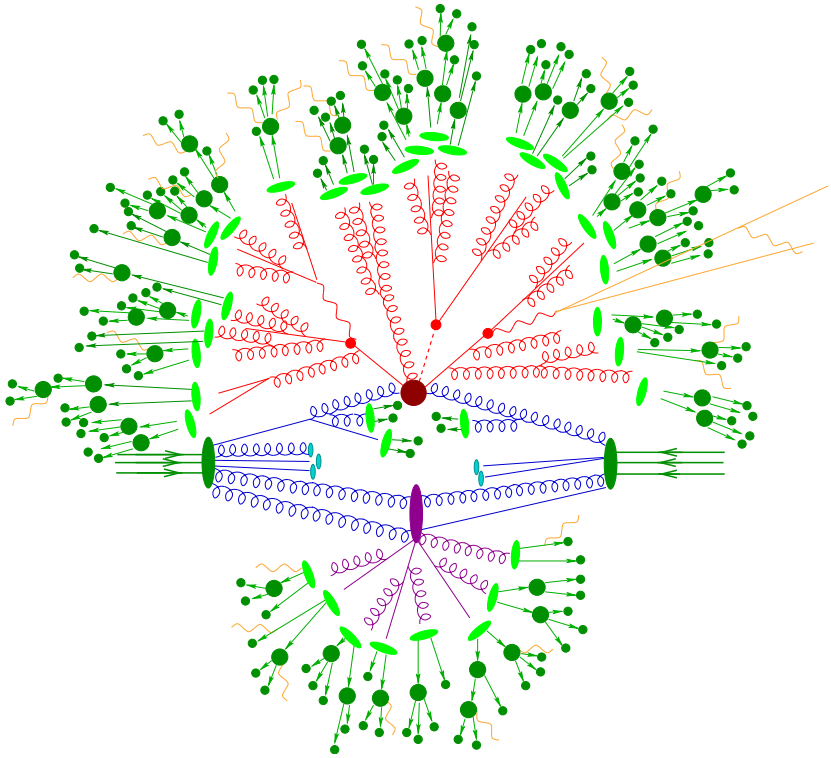


Figure 4.2: Adapted depiction of the MC simulation of a $t\bar{t}H$ event [98]. The high momentum parton-parton interaction is shown as a red dot with subsequent decays with smaller red dots. The hard scattering and parton shower are illustrated with red lines. A second parton-parton interaction is depicted in purple. The non-perturbative effects of hadronization and unstable hadron decays are shown with light and dark green dots and lines, respectively. QED radiation (orange) can be emitted along the way.

The illustration of the steps performed by an MC event generator for $t\bar{t}H$ production is depicted in Figure 4.2. The hard scattering is shown as a large red dot with a top-quark pair and a H-boson emerging, and the small red dots show their decays. All red lines correspond to the perturbative regime, namely the matrix element calculation and the parton shower. The purple dot illustrates additional parton-parton interactions with the subsequent parton shower. The light green dots mark the start of the non-perturbative regime with the initiation of hadronization. The dark green structures illustrate the unstable hadron decays. In orange, QED radiation is depicted, which can originate from all charged partons or particles during the whole process.

The different intermediate particles and states are often saved by the MC generators for the user. The use of information of intermediate steps of the simulation lacks physical meaning, since only the observables of the simulation, like the energy, momentum and the spatial distribution of final state particles, are able to model nature accurately. However, a reasonable approach is to get insight into the origin of a specific decay. Particle information, accessed from the simulation of the hard-scattering and parton shower, is called *truth*-information.

4.3 The concept of signal and background

The dataset measured by the ATLAS experiment and the concept of MC simulations in HEP were introduced in this chapter. When a search for new physics or a measurement of SM properties is performed, the results often rely on a comparison of data with MC simulation. An analysis looks for a desired physics process, which can be part of the SM or contain new physics, and is called *signal*. The aim is to isolate signal events, while suppressing the "rest" of the SM events, which diminish the sensitivity. These are referred to as *background*. Thus, signal and background are terms, which describe different processes depending on the use case. In the following analyses, described in Chapters 8 and 9, signal refers to any MC simulation of vector-like quarks with varying parameters like the branching ratio or the mass hypothesis. Background includes all SM processes, especially those, which lead to signal-like signatures in the detector. A detailed description of the MC simulations used in the searches for vector-like quarks are given in Sections 8.4 and 9.2.

The MC simulations are scaled to the integrated luminosity of the measured dataset, described in Section 4.1. The MC-sample luminosity $\int \mathcal{L}_{\text{MC}} \mathrm{d}t$, which usually exceeds the integrated luminosity $\int \mathcal{L}\, \mathrm{d}t$ of data by far, is scaled according to

$$w = \frac{\int \mathcal{L}\, \mathrm{d}t}{\int \mathcal{L}_{\text{MC}} \mathrm{d}t} = \frac{\int \mathcal{L}\, \mathrm{d}t}{\frac{N_{\text{MC}}}{\sigma_{\text{NLO}}}} \tag{4.3}$$

with the weight w. The number of unweighted MC events N_{MC} in combination with the sample cross-section σ_{NLO} in QCD, usually derived at NLO, is used to determine the luminosity of the MC sample.

Chapter 5
Object reconstruction and identification

The reconstruction of physics objects in the detector is a crucial step in an analysis. In the high pile-up environment in the ATLAS detector, an enormous number of activity is present in each event, such as charged-particle trajectories (tracks), energy deposits in the calorimeters, as well as signatures in the muon chambers. Therefore, the reconstruction and unambiguous identification of the decay chain and its physics objects become challenging. Different particle types have different signatures in the detector, but there is still ambiguity between them and thus the possibility of misidentification. In the following, an overview of the basic reconstruction and identification techniques within the ATLAS experiment is given. Objects of importance for the contents of this thesis are electrons, muons, small-R jets, large-R and reclustered jets, as well as the missing transverse momentum, discussed in Section 5.1 to 5.6. Also, the identification of small-R jets from b-quarks is crucial, which is abbreviated as b-tagging and described in Section 5.7.

5.1 Electrons

Electrons of interest originate from the hard-scattering process or from the subsequent decay chain as *prompt* electrons. Hadrons and non-prompt electrons can be misidentified as these signal electrons. Non-prompt electrons are, for example, produced from photon conversion or the decay of heavy-flavor hadrons and need to be separated from prompt electrons. Reconstructed electrons [99, 100] have to pass certain trigger, reconstruction, identification and isolation criteria. Information from tracks in the Inner Detector (ID), as well as energy deposits in the calorimeters is used.

The electron reconstruction is available within $|\eta| < 2.47$ (excluding the transition region between $1.37 < |\eta| < 1.52$) and a minimum p_{T} of 28 GeV is required for all electron candidates. The energy clusters in the calorimeter are reconstructed and matched to charged-particle tracks in the ID using quality criteria. This is an important step, since electrons tend to emit bremsstrahlung in the detector material, which can create e^+e^- pairs, leading to multiple tracks pointing towards the same conversion vertex.

Afterwards, certain identification criteria have to be fulfilled, which are based on a likelihood approach. Variables like the hadronic leakage, the information from each electromagnetic-calorimeter layer or track conditions are used to get more insight into the origin of the object. For the following analyses the *tight* likelihood was used, which has an approximate efficiency of 95%.

Furthermore, isolation criteria are applied in order to separate prompt electrons from background. This is based on the knowledge that prompt electrons are often separated spatially from other detector activity. However, boosted decays with e^+e^- final-state products or the $t\bar{t}$ decays create dense and particle-rich environments. In order to identify isolated electrons, a variable-radius cone of size

$\Delta R = \min(10\,\text{GeV}/E_T, 0.2)$ [101] is chosen. In this cone, the transverse energy in the electromagnetic calorimeter or the transverse momentum, p_T, in the ID is summed. The sum of the transverse momentum or energy, excluding the electron's contribution, are required to be below 6% within this cone. In the following, the working point *FixedCutTightTrackOnly* is used, which only uses an isolation based on the ID. The electron candidates are matched to the primary vertex, using the transverse impact parameter d_0 with its significance fulfilling $|d_0/\sigma(d_0)| < 5$. The longitudinal impact parameter z_0 has to fulfill $|\Delta z_0 \sin(\theta)| < 0.5\,\text{mm}$.

In order to make sure that the reconstruction of electrons in Monte Carlo (MC) simulations emulates the behavior in data, scale factors are derived for the trigger, reconstruction, identification and isolation using tag-and-probe methods with $Z \to ee$ and $J/\psi \to ee$. Furthermore, the electron energy scale is calibrated [100, 102], accounting for differences in simulation and data.

5.2 Muons

For the reconstruction of muons [103], the ID and the Muon Spectrometer (MS) are used. Individual reconstructions in both detector parts are done and then combined in a fit to match tracks in the ID to ones in different layers of the MS. Muons from the hard scattering or from processes of interest have to be separated from pion or kaon decays.

Important quality criteria for muon identification are based on variables sensitive to differences between measurements in the ID and the MS, like the charge or p_T or a combination of them. *Medium* identification criteria are used in the following analyses, constructed for the minimization of systematic uncertainties of reconstruction and calibration.

The isolation of muons is constructed in a similar fashion as for electrons. The same working point *FixedCutTightTrackOnly* is chosen and uses the $p_\text{T}^{\text{varcone30}}$ method, which ensures that the sum of p_T in a variable-radius cone size $\Delta R = \min(10\,\text{GeV}/p_\text{T}^{\mu}, 0.3)$ is less than 6%. As for electron candidates, the muon candidate is matched to the primary vertex with the requirements of $|d_0/\sigma(d_0)| < 3$ and $|\Delta z_0 \sin(\theta)| < 0.5\,\text{mm}$.

Furthermore, the muon p_T is calibrated [103] to match the data distribution, inducing smaller differences for the momentum scale and resolution between data and MC simulation. $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ processes are used for this procedure. For muons, a minimum requirement of $p_\text{T} > 28\,\text{GeV}$ is made in the following. A loose definition of muons with identification based on calorimeter information exists as well and is used in Section 5.5.

5.3 Small radius jets

Due to the nature of the strong interaction, quarks are asymptotically free at high energies, but can only occur as bound states at low energies, where they are constrained by confinement. The radiation of gluons and quarks during the (perturbative) parton shower, as well as the subsequent (non-perturbative) processes of hadronization, create hadronic showers along the axis of the initial flight path of a hadronically decaying particle. At the end of the decay chain, color-neutral bound states are formed. The hadronic showers deposit their energy in the electromagnetic and hadronic calorimeters and are reconstructed as jets. The jet reconstruction, following the recommendations of the ATLAS

collaboration, is based on the use of topological clusters [104, 105] built from calorimeter cells, which are called *topoclusters* in the following. The energy measured in these cells is calibrated to the electromagnetic scale. In order to correct the measured jets to the desired hadronic scale at particle level, a jet energy scale (JES) correction [106] is applied: The different jet responses between data and MC simulation are corrected for jets in data. Furthermore, detector- and jet-related effects are compensated by this procedure. Differences between the jet energy measurement and the initial particle energy can arise due to losses in inactive detector material. Additionally, the simultaneous occurrence of multiple pp interactions leads to energy contributions in the calorimeters from other events and should be corrected for. Furthermore, during jet fragmentation, a particle could be bent out of the jet cone due to the magnetic field and therefore could not be reconstructed within the jet. These effects and others are compensated by a step-wise correction procedure until the particle-level energy is reconstructed as accurately as possible. Additionally, jets need to fulfill special quality criteria [107], and the pile-up contribution is reduced by applying the jet vertex tagger (JVT) [108]. Jets with a transverse momentum of up to 60 GeV within the region of $|\eta| < 2.4$ or 2.5 have to fulfill certain JVT requirements in the 2015+2016 and the full Run-2 analyses, respectively.

In order to reconstruct a jet, topoclusters are used as an input to the anti-k_t algorithm [109] using FastJet [110]. This algorithm creates conical jet shapes and is resilient to soft radiation. It uses the distance measure d_{ij} between two objects i and j and searches for the minimum between all inputs. If the minimum is found for object j, it is merged to object i, which becomes the reconstructed jet in the end. If, in contrary, the distance from object i to the beam axis is smaller than to any other of the objects, the clustering is stopped. The distance measure is defined as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \qquad \text{with} \qquad d_{iB} = k_{ti}^{2p} \,. \tag{5.1}$$

The $\Delta R_{ij}^2 = \Delta\eta^2 + \Delta\phi^2$ is the spatial distance of object i and j and k_{tx} is the transverse momentum of object x. For the anti-k_t algorithm, the exponent p is chosen to be -1. In Figure 5.1, the conical jet shapes reconstructed with the anti-k_t algorithm are shown. As a standard radius parameter 0.4 is chosen within the ATLAS collaboration. These jets are called small-R jets and a minimum requirement of 25 GeV is set on the p_{T} for all subsequent studies. Small-R jets within the detector region of $|\eta| < 2.5$ are denoted as central jets.

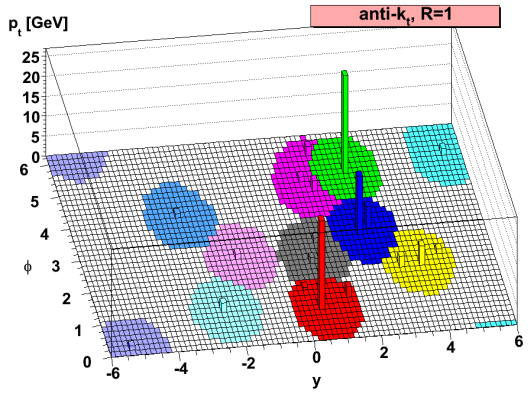


Figure 5.1: Jet reconstruction with the anti-k_t algorithm with a radius parameter of $R = 1.0$ [109].

5.4 Large radius jets

The use of small-R jets is sufficient for the reconstruction of hadronic decays in a low p_{T} range. When an analysis in a boosted regime is conducted, the use of additional jet collections becomes beneficial. In Figure 5.2, a top-quark decay at low and high p_{T} is depicted, which also illustrates the jets. It can be seen that for low p_{T} the decay products are separated spatially, but become collimated for a top-quark decay at high p_{T}. The top-quark constituents can be captured within one jet with a large radius parameter R, which is called a *boosted object* in the following. According to the rule of thumb for the cone radius R of a jet

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

the radius parameter of a jet can be calculated, which would be needed for capturing all of a particle's decay products within one jet. For a top-quark p_{T} of around $350\,\mathrm{GeV}$, all decay products are likely to be found in one jet with a radius of 1.0, which is the cone size used for large radius jets in this thesis.

In the following, two different approaches for the reconstruction of jets with large radius parameter are given. On the one hand, a jet collection called *large-R jets* is presented, which is reconstructed from topological clusters and is then calibrated to match the hadronic scale. On the other hand, so-called *reclustered* (RC) jets are introduced, which are reconstructed from calibrated small-R jets and do not need an additional calibration.

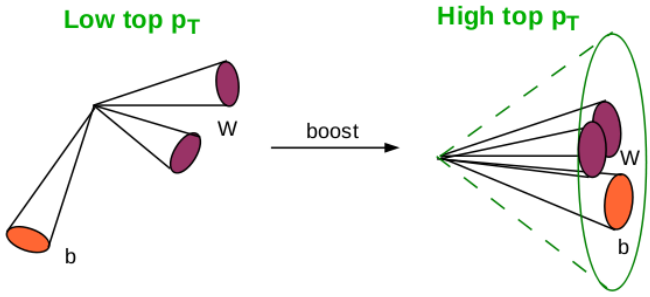


Figure 5.2: Illustration of a top-quark decay at low and high p_{T} [111]. For increasing p_{T} the decay products are collimated and might be reconstructed within a large radius jet.

5.4.1 Large-R jets

Analogous to small-R jets, large-R jets are reconstructed from topological clusters built from calorimeter cells using the anti-k_t algorithm. At this stage, their energy scale is calibrated at the electromagnetic scale. In contrast to the procedure for small-R jets, their energy scale is then locally corrected for each topocluster, using the *local hadronic cell weighting* (LCW) [104]. The method aims to compensate energy losses due to smaller calorimeter responses for hadrons (non-compensating calorimeter with $e/\pi > 1$) and signal losses due to inactive detector material or the noise suppression of the topoclustering. The algorithm evaluates for each topocluster which proportion of it is likely to originate from an electromagnetic or hadronic shower. The topocluster is weighted accordingly and different sources of inefficiency are compensated for each type, yielding a larger correction for hadronic parts of topoclusters.

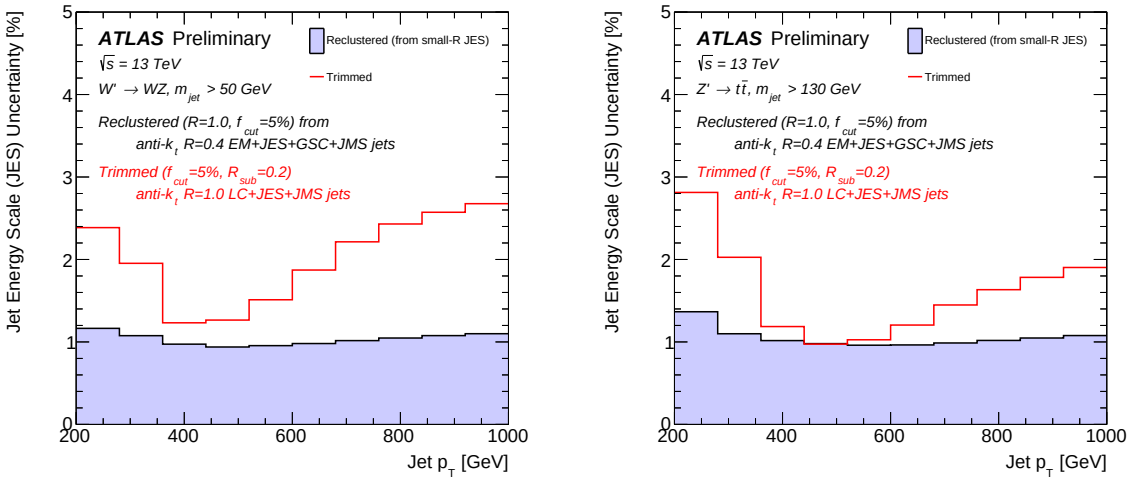
As a next step, *grooming* is applied, which keeps the substructure of a jet, but removes soft radiation in form of pile-up or the underlying event from the large-R jet. The method used here is called *trimming* [112] and applies a reclustering algorithm on the large-R jet using the k_t algorithm [113–115] with a cluster radius of $R_{\text{sub}} = 0.2$. Subsequently, any of the 0.2-jets with energy contributions smaller than 5% of the total jet energy is removed from the large-R jet. Thus, the jet energy and mass scale move closer to the true particle-level jet scale. Next, a calibration procedure for the jet mass [116] and jet energy [117] follows, in order to further correct for different scales in data and simulation. The jet mass scale (JMS) is brought to within 3% [116] of the true particle-level jet mass with this procedure. The calibration induces minimum thresholds on the p_{T}, mass and η distributions of a large-R jet of $p_{\text{T}} > 200\,\text{GeV}$, $m > 50\,\text{GeV}$ and $|\eta| < 2.0$. The mass cut is used to reduce the contribution from light jets, for example gluons, but also jets from b-quarks, since large-R jets target the reconstruction of heavier particles like W-, Z-, H-bosons or top-quarks.

5.4.2 Reclustered jets

Another option for the use of large radius jets are reclustered (RC) jets [118, 119]. Instead of using topological clusters as input for the anti-k_t algorithm, calibrated small-R jets are used. The RC jets are clustered with a radius of $R = 1.0$ and then trimmed: All small-R jets with $p_{\text{T}} < 5\%$ of the RC jet p_{T} are discarded from the RC jet. The minimum p_{T} for small-R jets is $25\,\text{GeV}$, which is effectively also a trimming cut.

Since the reclustering procedure is a rather new technique used in the ATLAS collaboration, the performance of RC jets is compared to large-R jets, reconstructed from topological clusters, in detailed studies in Ref. [119]. Studies based on MC simulations and small-R jets with calibrated JES and JMS show that reclustering of small-R jets recovers the mass and p_{T} distribution of RC jets to their values at particle-level. In Figure 5.3, the comparison of the JES uncertainty is illustrated for jets originating from W-bosons and top-quarks. The uncertainty is found to be significantly smaller for RC jets than for large-R jets in the range under study up to $1\,\text{TeV}$. Furthermore, over a large p_{T} range (200–$1000\,\text{GeV}$) the jet mass resolution (JMR) shows a better performance for RC jets, ranging from 22% for high-p_{T} W-bosons to 29% for top-quarks, compared to large-R jets. Another important topic of study is the impact of close-by-effects for RC jets. The original calibrations of small-R jets are derived in rather isolated environments using dijet, multijet, γ+jets and Z+jets processes [106]. So it is not clear that the transition from isolated small-R jets to an environment with small-R jets that are very close-by or even overlapping creates good jet modelling. The study of dense environments has shown that no additional uncertainties are needed [119].

It can be concluded that the use of RC jets can be very beneficial for an analysis, since the systematic uncertainties can be propagated from small-R jets. Additionally, no additional calibration is needed, which leads to an easy application. The following studies in this thesis use small-R jets with a JES calibration only, since an estimation of the uncertainties for a JMS calibration were unavailable at the time of the generation of MC simulations.



(a) (b)

Figure 5.3: Comparison of the JES uncertainties [119] for RC jets and large-R jets truth-matched to W-bosons produced via a) $W' \to WZ$ and top-quarks from b) $Z' \to t\bar{t}$ decays. Over the whole p_{T} spectrum the JES uncertainty of RC jets is smaller.

5.5 Overlap removal

In order to resolve ambiguities between electrons, muons, small-R jets and large-R jets, an overlap removal procedure is applied with the following sequential steps. Muon candidates, which are tagged to be minimally ionizing in the calorimeter, are rejected against electron candidates, if both objects share a track in the ID. Subsequently, all electrons with the same track in the ID as any other muon candidate are removed. The ambiguity between jets and electrons is resolved by removing small-R jets which are within a cone of $\Delta R < 0.2$ of an electron. On the other hand, electrons are removed, if they remain in the $\Delta R < 0.4$ cone of a small-R jet. A small-R jet with less than three tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ which is within $\Delta R = 0.04 + 10\,\mathrm{GeV}/p_{\mathrm{T}}(\mu)$ of a muon candidate is rejected. On the other hand, a muon is rejected in favor of a small-R jet, if the small-R jet has at least three tracks within this ΔR cone. No overlap removal is applied between small-R jets and large-R jets, since no combination of properties from both objects is used. The only overlap removal for large-R jets is performed for electrons, where the large-R jet is rejected, if any electron is within its radius of 1.0.

5.6 $E_{\mathrm{T}}^{\mathrm{miss}}$

Not all of the Standard Model (SM) particles leave traces or energy deposits in the detector. Neutrinos pass all detector layers without interaction and thus cannot be detected directly. Besides, theories predicting physics beyond the SM can also contain particles, which do not interact with the detector. In order to get a handle on the detection of such signals, the missing transverse momentum with the magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ [120] is used. From all hard objects reconstructed and used in an event, like leptons and jets, the momentum missing from the momentum balance is calculated. Additionally, a soft-event contribution is added, which contains ID tracks matched to the hard-scatter event. Among others,

this can contain contributions from the underlying event or objects not fulfilling the quality criteria to be identified as jet or lepton. Both the hard- and soft-event contributions are rather insensitive to pile-up: Hard objects are reconstructed and calibrated such that they are as independent of pile-up as possible. Additionally, soft-event contributions suppress pile-up due to only being based on matched tracks. Furthermore, a signal-ambiguity resolution is performed, since an object can pass more than one of the reconstruction algorithms for jets and leptons, yielding a wrong magnitude for the missing transverse momentum.

5.7 b-tagging

The identification of jets originating from b-quarks is an important technique used at the ATLAS experiment. For example, the flavor information helps to identify top-quark decays or has enabled the observation of the process $H \rightarrow bb$ [121]. As b-tagging algorithm, a multivariate technique, called MV2 [122] is used. It is based on a boosted decision tree (BDT), which takes the p_{T} and η of the small-R jet and the output of other b-tagging [123] algorithms as input. Three kinds of algorithms are used to improve the b-tagging performance by using different information. First of all, impact-parameter-based algorithms (IP2D, IP3D) are used in order to exploit the different lifetimes of b-hadrons and other hadrons, treated as background. Depending on the lifetime, distances between the primary vertex and a displaced vertex vary and are longer for b-tagged jets. Furthermore, a secondary vertex finding algorithm SV [124] is used, which looks for two-track pairs and uses them in a χ^2 fit in order to find a secondary vertex. The third algorithm is the decay chain multi-vertex algorithm JETFITTER [125], which is based on information about the decay chain from the primary vertex to a b- and then a c-hadron.

The output of the MV2 classifier is shown in Figure 5.4. It can be seen that a good discrimination of jets originating from b-quarks and other quarks is achieved. Additionally, c-jets are more likely

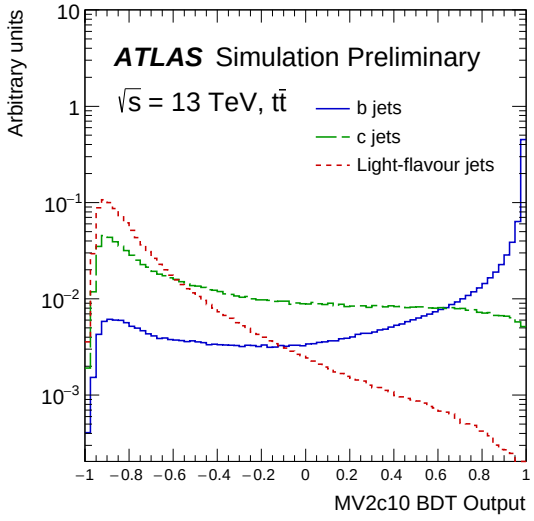


Figure 5.4: Output of the b-tagging classifier MV2c10 [122] applied on a sample of $t\bar{t}$ events. The signal class of b-jets is shown, as well as the background classes of c-jets and light jets.

to be misidentified as b-jets. The rejection for light jets is very good at high classifier values. The classifier denotation MV2c10 specifies the composition of the training sample. It always consists of 2.5 million jets, of which 40% are b-jets (signal). The rest of the sample is split into light (u-, d-, s- and gluon-jets) and c-jets. For MV2c10 the c-jet fraction is enriched to 7%, which increases the rejection of c-jets compared to no c-fraction in the training.

Chapter 6
Multi-Class Boosted-Object Tagger using a deep neural network

A novel tool for the identification of boosted objects is presented in this chapter. The multi-class tagger uses a deep neural network (DNN) to separate four different classes of boosted large radius jets from each other. As input to the DNN, reclustered (RC) jets and their subjet properties are used. In the end, ambiguous and unambiguous tagging of jets from V-bosons[1], H-bosons, top-quarks and background summarizing lighter quarks and gluons is possible. A short motivation is given in Section 6.1, followed by an introduction to the general concepts of neural networks (NNs) in Section 6.2. The DNN structure used for the Multi-Class Boosted-Object Tagger (MCBOT) is described in Section 6.3. The simulations, used for the training of the *standard* MCBOT, and the input variables, as well as data preprocessing are described in Sections 6.4 and 6.5, respectively. Subsequently, two additional variations of MCBOT, trained on different samples, are presented in Sections 6.6 and 6.7. A comparison of all three tagger versions is done in Sections 6.8 and 6.9. The DNN output distributions, their specific features for all four classes and their performance are discussed. At last, a recommendation for assigning unambiguous tags for V-boson, H-boson and top-quark jets is presented in Section 6.10, followed by a short summary in Section 6.11.

6.1 Motivation

As discussed in Chapter 5, the reconstruction of physics objects is crucial for gaining knowledge about the hard-scattering process and its subsequent decay products. The identification of the particle type initiating a jet is called *tagging* and is an important tool for the classification of events. In Section 5.7, b-tagging was introduced, which is performed on small-R jets. Tagging of jets with a larger radius parameter has become more important in Run-2 searches at the ATLAS experiment, e.g. see Refs. [126] and [127]. Due to the increased center of mass energy in Run-2, more boosted objects occur and can be captured in one large radius jet, as discussed in Section 5.4. These jets have distinct features like the number of small-R jets inside the radius, depending on the particle they originate from. There is a variety of top-quark and boson tagging algorithms [128, 129]. However, these taggers are only available for the identification of one or two classes of boosted objects. They are based on large-R jets, which use topological clusters from the calorimeters as input and come with their own set of systematic uncertainties. This increases the total systematic uncertainty in the subsequent search or measurement. Furthermore, combined usage of multiple taggers is not advisable due to unclear handling of correlated systematic uncertainties and ambiguous tags.

[1] V-boson describes either a W- or a Z-boson.

In order to avoid these drawbacks, a multi-class tagger for the simultaneous identification of V-boson, H-boson, top-quark and background jets is developed. Hence, tagging of all these objects can be done in a consistent way. No distinction between W- and Z-bosons is made here, since W- and Z-boson masses are relatively close. Thus, a discrimination would be more difficult and is not attempted in the scope of this thesis. The background output class is a collective category for jets initiated by u-, d-, s-, c-, b-quarks or gluons. Furthermore, any jets with mixtures of these quarks and gluons are assigned to the background class. Background jets occur at any stage of the decay and hadronization chain, but they are not part of the interesting physics an analysis searches for, which usually contains V-bosons, H-bosons or top-quarks, denoted as signal in the following. However, the background jets are often irreducible and therefore the rejection of them is the main priority.

This novel approach for a multi-class tagger was first applied in a search for vector-like quarks (VLQs) in all-hadronic final states [1] and was based on variable-R reclustered (vRC) jets. The idea was adapted and developed further in the scope of this thesis for a wider applicability within the ATLAS collaboration. The following tagger is based on RC jets with a fixed radius of $R = 1.0$, which are reconstructed from small-R jets with the anti-k_t algorithm. Small-R jets are calibrated and therefore, all inputs for the multi-class tagger are calibrated as well, making any additional calibration unnecessary. Furthermore, small-R jets come with a full set of systematics, which can be propagated through the tagger. No additional systematic uncertainties are needed.

6.2 Introduction to neural networks

Machine learning (ML) techniques have grown to be a substantial element in addressing problems of high complexity in high-energy physics, but also in every day life. A computer is used to solve complex tasks by using a large amount of data and learning its patterns. NNs are a class of ML techniques, which can go beyond linear dependencies of input and output data. At first, they were proposed for the simulation of biological systems [130–133], which dates back to the 1940s. The connected structures, called *nodes*, are inspired by the human brain with its neurons.

ML algorithms can be sorted into two classes, *supervised* and *unsupervised*. In supervised learning the target vectors for the solution of the problem are given to the NN during the training process. Classification and regression problems are the subcategories of supervised learning. In unsupervised learning, no information about the truth is given. In the following introduction to ML, a supervised ML approach is taken. In order to gain better insight into the NN's learning mechanism, an NN with one hidden layer is chosen as an example. It is a fully-connected feed-forward NN, which means that all nodes of one layer are connected to all nodes of the adjacent layers.

The great potential of NNs to learn complex data relations is based on the adaptive creation of a mathematical model. In Figure 6.1, an example of an NN with one hidden layer is depicted. At first, input variables x_1 to x_D are selected, which can also be written as elements of a vector $\vec{x}$ of length D. With the input variables $\vec{x}$, a linear combination is built with coefficients, called weight elements w_{ji} from a weight vector $\vec{w}_j$, in order to derive the so-called *activations* a_j:

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)} = \sum_{i=0}^{D} w_{ji}^{(1)} x_i \tag{6.1}$$

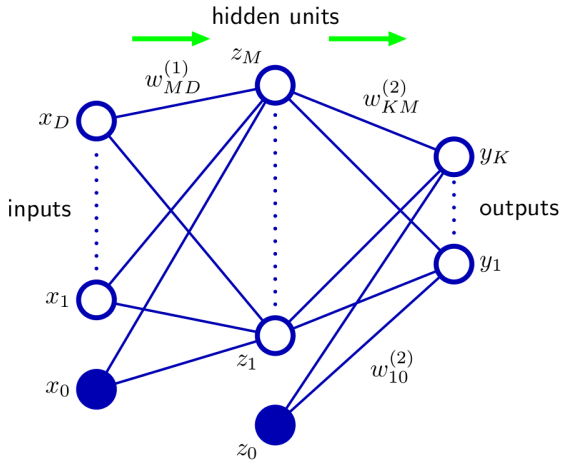


Figure 6.1: Example of a fully-connected feed-forward NN with one hidden layer [134].

The index j belongs to node z_j of the hidden layer and ranges from $j = 1, ..., M$. The weight element w_{j0} is called bias and can also be absorbed into the sum assuming $x_0 = 1$. The bias accounts for differences in the mean of the target values and the mean of the linear combination $\sum_{i=1}^{D} w_{ji}^{(1)} x_i$. The superscript (1) indicates that the weights correspond to the derivation of the activations for the first layer, calculated with formula 6.1. One weight vector $\vec{w}_j$ is defined for each activation a_j.

The activations a_j are then fed into an activation function $h(a_j) = z_j$. As a second step, all newly obtained values in the vector $\vec{z}$ are again subject to a linear combination with a new weight vector and another transformation f, yielding the output vector $\vec{y}$. Thus, the NN output values y_k are obtained by

$$y_k = f(a_k) = f\left(\sum_{j=0}^{M} w_{kj}^{(2)} z_j\right) \tag{6.2}$$

Combining both transformations performed in Eqs. 6.1 and 6.2, the compact solution for the output y_k of an NN with one hidden layer is

$$y_k(\vec{x}, \vec{w}) = f\left(\sum_{j=0}^{M} w_{kj}^{(2)} \cdot h\left(\sum_{i=0}^{D} w_{ji}^{(1)} x_i\right)\right) \ . \tag{6.3}$$

In this notation, all weight vector elements are arranged in one weight vector $\vec{w}$. As activation functions, non-linear functions like $\tanh(a)$ or the rectified linear unit (ReLU) with

$$f_{\text{ReLU}}(a) = a^+ = \max(0, a) \tag{6.4}$$

are often used. The $\max$ function sets all negative a to zero, while yielding the value a for all positive input values. The ReLU function is illustrated in Figure 6.2(a). For the output layer of binary classification problems, the sigmoid function with

$$f_{\text{sigmoid}}(a) = \frac{1}{1 + \exp(-a)} \tag{6.5}$$

is a popular choice, while for multi-class classification problems often the softmax function is used:

$$f_{\text{softmax},k}(a_k) = \frac{\exp(a_k)}{\sum_j \exp(a_j)} \ . \tag{6.6}$$

The softmax activation is the *normalized exponential function* and the sum of the outputs of all j classes, whose individual output range is zero to one, is unity. The softmax function is depicted in Figure 6.2(b). The non-linear activation functions in combination with increasing depth of the NN enable the network to learn more complex features. If a linear activation function was chosen, it would be possible to replace the subsequent application of transformations for multiple layers with one linear transformation, making more than one hidden layer unnecessary.

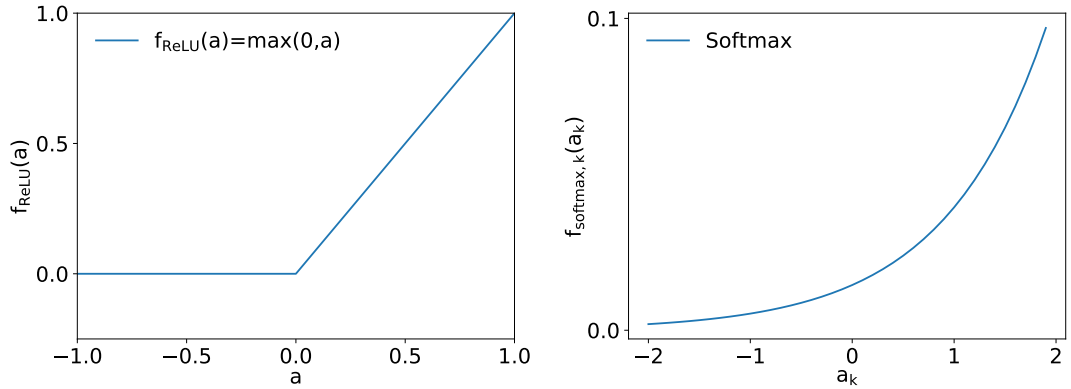


Figure 6.2: Illustration of the a) ReLU and b) softmax activation functions.

In order to find a generalized solution to a given problem with an NN, usually a set of non-linear basis-functions, here the hidden units z, are used. In order to learn problems of high dimensionality due to many parameters and their correlations, these basis functions need to be adjusted to suit the problem. In case of NNs, this is achieved with adaptive weights. At first, the flow of information through an NN occurs from the input nodes x_k to the output nodes y_k and is called *forward propagation*. In contrast, the subsequent process for deriving adaptive weights is a backward propagation: Once the output values y_k are derived, the so-called *error backpropagation* is initiated. The aim is to minimize the *loss function* of the NN's output values y_k and the target values t_k in order to obtain a good generalization of the problem. For a multi-class classification problem the cross-entropy E is often used as loss function. In the NN example with one hidden layer, it can be written as

$$E = -\sum_{n=1}^{N}\sum_{k=1}^{K} t_{kn} \ln y_k(\vec{x}_n, \vec{w}) \,, \tag{6.7}$$

where K denotes the number of output classes. The target value t_k can either be zero or one, depending on the correct classification for this output node. Here, the full set of input vectors $\vec{x}_n$ with $n = 1, ..., N$ enters the calculation of the loss function.

The minimum of the multi-dimensional minimization of E can be addressed with different approaches. The simplest method is the use of *gradient descent* [135], which starts at the initial position of the weight vector $\vec{w}^{(\tau)}$ at iteration τ and moves in the direction of the negative gradient of E with a given learning rate δ:

$$\vec{w}^{(\tau+1)} = \vec{w}^{(\tau)} - \delta \nabla E(\vec{w}^{(\tau)}) \tag{6.8}$$

There are more advanced methods with an improvement of training time available, like the optimizer Adam [136] used for MCBOT. Adam is efficient for problems of high dimensionality.

If the training of an NN is performed too long, the NN memorizes the training data and the mathematical model is not suitable for generalization. This state is called *over-fitting* and should be avoided. An important method to prevent over-fitting is *regularization*: Penalty terms enter the calculation of the loss function E. In the following, a layer-wise *LASSO* [137] regularizer, also called L1, is used, which adds a linear penalty term to all weights in each layer: All of the elements of one layer's weights w_{lq} are summed according to $\sum_{q,l} |w_{lq}|$. The sum is assigned with a global scale factor λ (to be optimized for each ML task) and the penalty is propagated to the loss function of the NN.

In order to investigate the performance of an NN and check for over-fitting, the total available dataset is split in different samples with the same data composition. The *training* sample is used for training the NN and deriving the mathematical model. A *validation* sample is important during training time in order to check the loss function and decide, if over-fitting occurs. At last, a *test* sample is created in order to evaluate the performance of the training.

Important parameters for the construction of an NN were discussed in this section and can be summarized under the term *hyperparameters*. An NN with more than one hidden layer is denoted as DNN.

6.3 DNN structure

MCBOT is a feed-forward DNN and its architecture consists of four hidden fully-connected layers. It was trained with Keras [138], and Tensorflow [139] is used as a backend. The input layer has 18 input nodes. The second layer has an increased size of 32 nodes, which drops to 27, 14 and 12 for the second, third and fourth hidden layer. The output layer is four-dimensional with one node each for V-boson, H-boson, top-quark and background jets. In Figure 6.3, a schematic of the structure is shown.

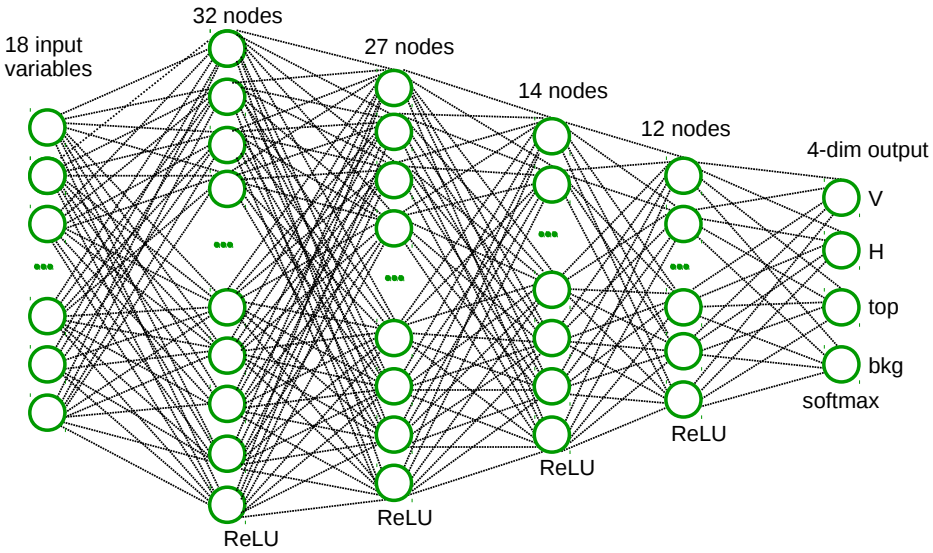


Figure 6.3: Architecture of the DNN for MCBOT with four hidden layers with 32, 27, 14 and 12 nodes. ReLU is used as activation function of the hidden layers, while the softmax function is chosen for the output layer with four output nodes for V-boson, H-boson, top-quark and background jets.

As activation the ReLU function is used for the hidden layers, while softmax is used for the output layer. The softmax activation function of the last layer ensures that the output values of all four classes add up to one. Together with the categorical cross-entropy as loss function these values can be interpreted as probabilities. A summary of all used hyperparameters is listed in Table 6.1. The architecture as well as most of the hyperparameters are taken from a multi-class boosted-object tagger based on vRC jets used in a search for VLQ with all-hadronic final states [1]. Most of the hyperparameters have been explained in Section 6.2, except for the batch size, number of epochs and batch normalization, which are discussed in the following: One epoch lasts until the NN has processed the whole training dataset once. The batch size determines into how many slices one epoch is subdivided. For example, a training dataset could consist of 10^6 input vectors $\vec{x}$. If the batch size is 400 (like for MCBOT), 400 of these input vectors $\vec{x}$ are drawn randomly from the training dataset. Then, the NN calculates the forward and backward propagation, as described in Section 6.2, in order to derive the adaptive weights. Afterwards, a new batch is loaded into memory and the procedure is repeated until the whole dataset is processed, which would take 2500 iterations in this example. As discussed in Section 6.5, the dataset is preprocessed in a specific way for enhanced fit stability, speed and performance. Batch normalization [140] contributes to this improvement. The input distributions for each layer are normalized for each batch at training time in order to train on batches with data having a mean of zero and a standard deviation of one.

Table 6.1: Hyperparameters for MCBOT.

Hyperparameter	Specification
Loss function	Categorical cross-entropy
Learning rate	10^{-5}
Batch size	400
L1 regularizer	0.001
Optimizer	Adam [136]
Batch normalization	Yes
Number of epochs	50

Figure 6.4 provides information about the training and validation loss over the 50 epochs of training. It can be seen that both the training and validation loss decline quickly, approaching a less steep line at the end of the training, ideally being a horizontal line. The training length was optimized for a training on a dataset, which only includes the RC jet p_{T} reweighting in contrast to a two-dimensional reweighting. The difference between both reweightings is described in Section 6.5. In future studies, the training should be performed slightly longer than 50 epochs, e.g. 60 to 70 epochs, for the MCBOT tagger including the two-dimensional reweighting. However, the main optimization of MCBOT is finished after 50 epochs, since the distribution in Figure 6.4 is expected to reach a plateau after some epochs. Furthermore, it might seem surprising that the validation loss is lower than the training loss in this plot. In general, the loss calculated on the training data is expected to be optimal, since it is the dataset the NN learns from. Therefore, it should be lower than the loss calculated on the validation data. The reason for a lower validation loss in Figure 6.4 is the way it is calculated. For the training

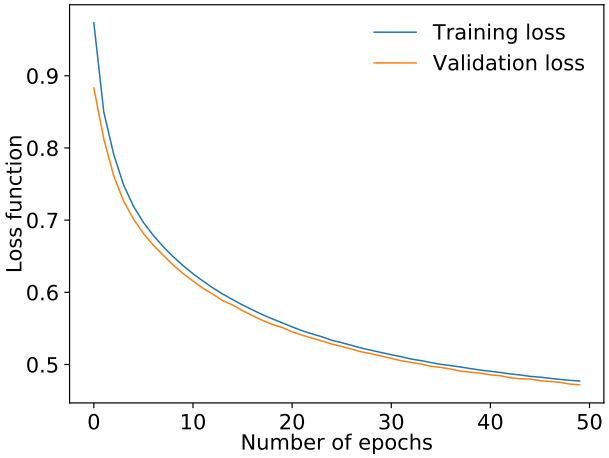


Figure 6.4: Training and validation loss for MCBOT over 50 epochs.

loss the categorical cross-entropy is derived after each batch and then averaged over the whole epoch. For the validation loss it is only calculated at the end of an epoch once. Therefore, the higher values of losses, corresponding to a worse performance, at the beginning of one epoch are not considered for the validation set and lead to a lower validation loss. This effect should flatten out at the end of training, when the weight updates are marginal. The validation loss is still lower than the training loss due to the L1 regularizer, which is applied during training, but switched off during validation.

6.4 Input Monte Carlo simulations

In order to train a DNN for the discrimination of V-boson, H-boson, top-quark and background jets, input data is generated with Monte Carlo (MC) simulations. For boosted top-quark jets, $Z' \to t\bar{t}$ events are generated with PYTHIA 8 [90] with the A14 set [141] of tuned parameters as well as the NNPDF23LO parton distribution function (PDF). The decay of b- and c-hadron decays was simulated with EVTGEN [95]. W-boson and Z-boson jets are obtained from $W' \to WZ$ events, generated with the same generator, tune and PDF. For the production of boosted H-bosons, graviton events with $G \to HH$ are generated with a BR$(H \to bb) = 100\%$, using PYTHIA 8 with the previously mentioned tune and PDF. The same generator, tune and PDF are used for the generation of background jets, where events with the production of two jets, denoted as *dijet*, are simulated in multiple p_{T} ranges. All generated MC processes are passed to the ATLAS detector simulation.

The p_{T} spectrum of a physics process is modified in the generation of all samples. In general, the p_{T} spectrum of the decay products and thus also the RC jets decreases exponentially due to radiation of particles in the hadronic decay chain during the parton shower and hadronization. Hence, the phase-space for production of objects with high p_{T} is much smaller. All of the MC samples, used for MCBOT, are enhanced with events at high p_{T} in order to give the NN more input data to learn features in the high RC jet p_{T} range above 1 TeV.

Furthermore, information about the RC jet origin is needed in order to be able to obtain training data with their target values, denoted as *labels* for this chapter's classification problem. The MC simulations contain information at generator level and reconstructed objects at detector level. The

reconstructed RC jets are obtained after the matrix element and shower generation, as well as the detector simulation. No information about the origin of an RC jet is present at this point. Combining the MC generator information with the reconstructed objects is called truth matching. The ΔR distance of a "true" particle and each RC jet is calculated. If it is smaller than $\Delta R = 0.75$, the RC jet is assigned one of four labels: V-boson jet, H-boson jet, top-quark jet or background jet.

Three datasets are constructed for training, testing and validating the performance of MCBOT. They are all constructed in the same way and contain the same proportions of V-boson, H-boson, top-quark and background jets. Half of the dataset is composed of background jets from the dijet sample. The other half consists of one third of truth-matched V-boson, H-boson and top-quark jets each. 25 million RC jets are used in the training, while the test set contains 5.5 million jets. A small validation sample with 3.8 million RC jets is used for monitoring the performance while training.

6.5 Input variables and preprocessing

MCBOT takes properties of an RC jet and its subjet properties as input. The properties of the three subjets with largest p_{T} are used as input features: The four-momentum vectors (p_{T}, E, η, ϕ) give insight into the kinematics. Additionally, b-tagging information for each small-R jet is used to gain knowledge whether the jet origin was a b-quark or not. This procedure ensures consistency in the analysis, presented in Chapter 9, which uses b-tagging and MCBOT at the same time. The 77% signal efficiency working point is chosen. If a subjet is b-tagged, the b-tagging decision is set to one. If no b-tag is present, a zero is assigned. Furthermore, combined properties from the small-R jets are given to the DNN: The mass and p_{T} of the RC jet, as well as its number of constituents are used. This makes the study of correlations and feature learning more convenient from a technical point of view. However, the RC jet quantities can be reconstructed from the subjet properties.

MCBOT is trained for an RC jet p_{T} range from $150\,\mathrm{GeV}$ up to $3\,\mathrm{TeV}$, as motivated below. As already introduced, the tagger is aimed at the differentiation of V-boson, H-boson and top-quark jets against background jets. The V-boson class contains the lightest object, which is the W-boson with a mass of $(80.379 \pm 0.012)\,\mathrm{GeV}$ [10]. The radius, according to Eq. 5.2, of a W-boson jet is almost covered in a radius of $R = 1.0$ at an RC jet p_{T} of $150\,\mathrm{GeV}$. Thus, an RC jet p_{T} of $150\,\mathrm{GeV}$ is a reasonable lower bound with the lightest boosted object being fully-contained. Furthermore, the lower bound helps to remove a lot of background jets before starting the training. Therefore, the tagger is trained on signal and background RC jets, which are created in the same kinematic regime. Their separation is especially interesting, since an analysis often selects jets based on minimum criteria on the p_{T}.

For increasing RC jet p_{T} it becomes more difficult to distinguish between different objects, since the subjets are merged within $R = 0.4$ and b-tagging performance drops [142]. Thus, an upper bound of $3\,\mathrm{TeV}$ is chosen for the RC jet p_{T}.

Furthermore, a minimum requirement on the RC jet mass of $40\,\mathrm{GeV}$ is made. In Figure 6.5, the invariant RC jet mass is shown including a minimum requirement on the RC jet p_{T} of $150\,\mathrm{GeV}$ for all four MCBOT classes. It can be seen that V-boson, H-boson and top-quark jets show distinct peaks at their expected masses. For top-quark jets, there is a peak at the W-boson mass as well, which corresponds to top-quark jets, where the b-jet is not captured in the RC jet reconstruction with $R = 1.0$. These jets are called partially-contained in contrast to fully-contained jets, which include all

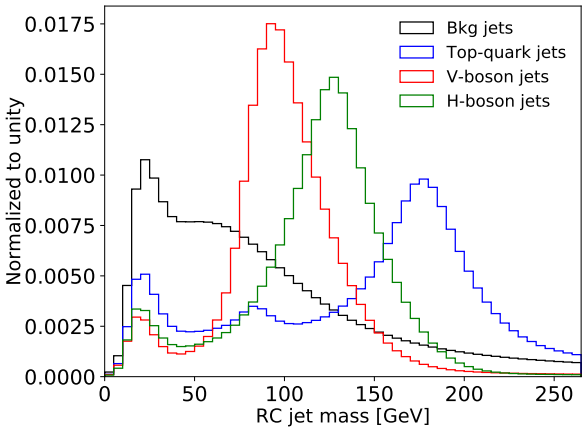


Figure 6.5: The RC jet mass before the minimum requirement of 40 GeV on its mass for V-boson, H-boson, top-quark and background jets, shown in different colors. Only RC jets with a p_{T} between 150 GeV and 3 TeV are illustrated.

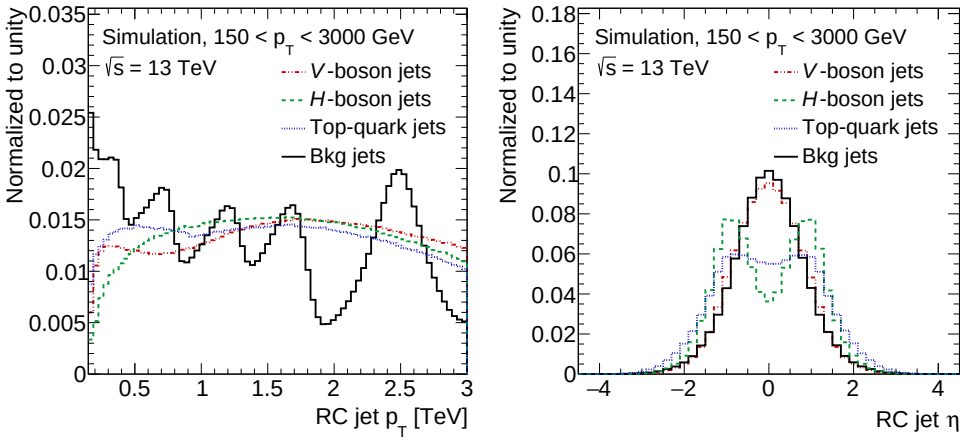
decay products from the top-quark. For background jets, a falling spectrum can be seen, which shows that the RC jet mass does not correspond to a resonance and that most background RC jets carry lower masses than the signal jets. A particular high peak for masses below 40 GeV is observed for signal and background jets, which corresponds to RC jets with light constituents. In case of signal, these low masses can also occur if not all decay products are captured within one RC jet. Below 40 GeV there is no clear distinction possible for all four classes and therefore these jets are removed.

In a fully-connected DNN the number of input features is constant for all events. Thus, the number of subjets given to MCBOT is fixed as well and is chosen to be three. Since the number of subjets within an RC jets can be smaller than three subjets, the empty subjet values are filled with unphysical default values, listed in Table 6.2. Hence, the DNN can recognize existing subjets, which do not carry zero energy and p_{T} or a b-tagging decision of -1.

Table 6.2: Default values, if a second or third RC jet constituent is not present. The four momentum vector as well as the b-tagging decision are set to the values below.

Subjet variable	η	ϕ	E	p_{T}	b-tagging decision
Default value	RC jet η	RC jet ϕ	0	0	-1

The aim of MCBOT is to train an NN to learn from differences in properties of RC jets and its subjets from V-boson, H-boson, top-quark or background jets. These properties are, e.g. the RC jet mass or features of the subjets. What should not be learned from are differences in input distributions, which originate from differences in the simulated physics processes. In Figure 6.6(a), the RC jet p_{T} distributions for each class are shown. As mentioned before, the distributions are enhanced with events at high RC jet p_{T}. The background sample is generated in p_{T} slices, making sure that the high p_{T} regions are enhanced with events. Using these p_{T} distributions of the four classes in the training would introduce a dependency on this kinematic. The same statements hold for the η-distributions, as shown in Figure 6.6(b). The RC jet η between the four classes follows different shapes due to



(a) (b)

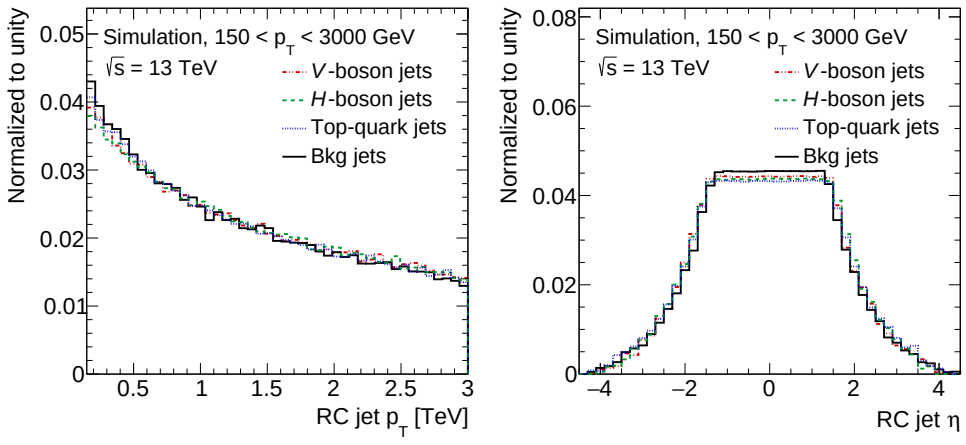
Figure 6.6: RC jet a) p_{T} and b) η distributions without a reweighting. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

the underlying physics process. Although the RC jet η is not used as an input variable, it is highly correlated with the subjet η, which is used in the training for up to three subjets. In order to be independent of the kinematics of the event, which the RC jet was produced in, and train equally on events from all RC jet p_{T} and η regions, a two-dimensional reweighting in RC jet p_{T} and η is applied during the data preprocessing. A fine binning in p_{T} and η is chosen with 150 and 90 bins, respectively, and each bin is normalized to $\frac{1}{\text{number of bins}}$ in order to normalize the reweighted histogram to one. The weights obtained from this histogram are used as weight factors during the training and scale the corresponding phase-space to flat distributions, which gives each p_{T}-η-bin the same importance. In Figure 6.7, the RC jet p_{T} and η distributions after the application of the two-dimensional reweighting are shown. While the two-dimensional distribution of p_{T} and η is flat, the projected distributions do not necessarily have to be flat. For high p_{T}, not all η values occur, which leads to a smaller fraction of p_{T} and η values filled in the tails. This leads to the falling spectrum of the RC jet p_{T} in Figure 6.7(a) and the falling edges for RC jet η in Figure 6.7(b).

A subset of input variables after all preprocessing steps is shown in Figure 6.8 for the four classes. In Figure 6.8(a), the RC jet mass is illustrated, which includes the cut on $m > 40\,\mathrm{GeV}$ and shows the peak structure as described for Figure 6.5.

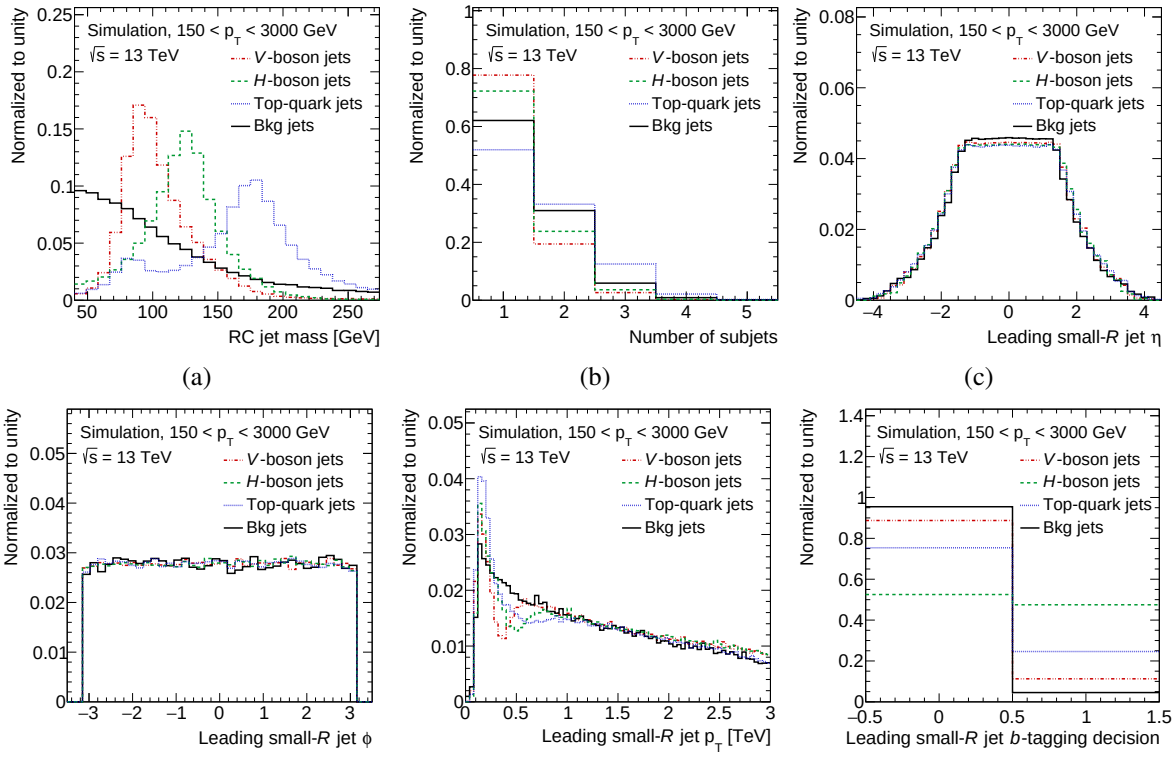
The number of subjets in Figure 6.8(b) shows that for all classes one subjet inside an RC jet occurs most frequently. It should be noted that the p_{T} range extends from $150\,\mathrm{GeV}$ up to $3\,\mathrm{TeV}$, where the subjets merge at higher RC jet p_{T}. According to Eq. 5.2, the extension of a top-quark decay in R becomes smaller than the small-R jet radius of $R = 0.4$ for an RC jet p_{T} of around $900\,\mathrm{GeV}$. This means that RC jets with $p_{\mathrm{T}} > 900\,\mathrm{GeV}$ mainly contribute to the bin with one subjet. In general, the top-quark jets show the largest subjet multiplicities, where three subjets correspond to a full hadronic top-quark decay, while for V- and H-bosons two subjets are expected.

The η distribution of the leading subjet (in p_{T}) is shown in Figure 6.8(c). The distribution looks very similar to the RC jet η distribution after the two-dimensional reweighting in Figure 6.7, since both are highly correlated. In Figure 6.8(d), the ϕ distribution of the leading subjet is flat across the whole range for all classes, which is expected due to isotropic decays around the beam axis.



(a) (b)

Figure 6.7: Reweighted RC jet a) p_T and b) η distributions according to the two-dimensional reweighting in RC jet p_T and η. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.



(a) (b) (c)

(d) (e) (f)

Figure 6.8: Most important input distributions for MCBOT, weighted according to the p_T-η-reweighting. In a) and b), the RC jet mass and the number of subjets are shown, respectively. Furthermore, properties of the subjet with largest p_T ("leading") are illustrated: The η and ϕ distributions are depicted in c) and d), while the p_T and b-tagging decision are shown in e) and f). The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

An interesting feature occurs in the p_{T} distribution of the leading subjet, illustrated in Figure 6.8(e): The p_{T} spectrum is falling as expected, but shows a shallow peak for V-boson and H-boson jets starting at $400\,\mathrm{GeV}$ or $500\,\mathrm{GeV}$, respectively. This peak is caused by the merging of subjets with $R = 0.4$ above a p_{T} threshold of $400 - 450\,\mathrm{GeV}$ for V-boson jets and $650\,\mathrm{GeV}$ for H-boson jets, according to Eq. 5.2. At this threshold the leading subjet is likely to have the same content as two subjets at lower RC jet p_{T}. This causes a rise in the occurrence of higher p_{T} subjets. The same feature is also visible for top-quark jets close to $1\,\mathrm{TeV}$.

At last, the b-tagging decision is depicted in Figure 6.8(f) for the leading subjet. For H-boson and top-quark jets, a value of one appears more often, which corresponds to a b-tagged jet at 77% signal efficiency. This is very plausible when considering the large branching ratios of top-quark and H-boson decays into final states with at least one b-quark. While $\mathrm{BR}(t \to Wb) \approx 100\%$ is naturally high, the H-boson decays used to obtain H-boson jets are simulated according to an artificially increased branching ratio with $\mathrm{BR}(H \to bb) = 100\%$, as described in Section 6.4.

After applying the two dimensional reweighting, the data are almost prepared for the training. The last step is the scaling of the input data. The aim is to obtain variables with a mean of zero and a standard deviation of one. This ensures that the variables used for training are numerically stable, because the NN's initialization with random weights and their adjustments during the training process are of the same magnitude. The scaling also speeds up the training process.

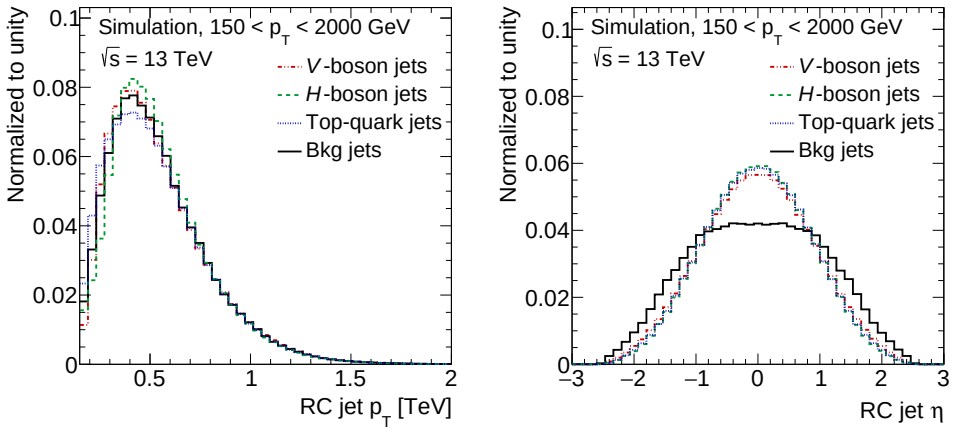
6.6 Alternative tagger enriched with b-tagged jets

In this thesis, a search for vector-like bottom (VLB) and vector-like top (VLT) quarks is performed. A characteristic feature of their final states is a high multiplicity of b-tagged jets, which are produced in the decays of $T \to Z/Ht$, $T \to Wb$, $B \to Z/Hb$ and $B \to Wt$. Therefore, an alternative tagger is trained, targeting a better discrimination of V-boson, H-boson, top-quark and b-tagged jets. The DNN structure as well as the input variables remain the same, but the input sample for background jets is adjusted. Instead of using the natural quark and gluon composition from this sample, it is enriched with b-tagged jets using the 70% signal efficiency working point. The composition is chosen such that 50% of the background contribution consists of RC jets with b-tagged jets. The 70% b-tagging working point is chosen in order to keep as many RC jets as possible, while applying a working point as tight as possible for high b-hadron purity. A truth-matching analogous to the one described in Section 6.5 is not possible, since no generator information of the initial particles is available for the background jet sample.

The training sample is composed of 20 million RC jets, while the testing and validation sample contain three and one million RC jets, respectively. The alternative tagger is called *b-tag enriched tagger* in the following, while the setup described previously with the normal background jet composition is denoted as *standard tagger*. The subsequent studies are done for the standard and the b-tag enriched tagger.

6.7 Non-flat-in-p_T tagger

Before MC simulations with a p_T-spectrum enhanced at high p_T were available, a tagger with the same input features and the same hyperparameters was trained. In fact, the input samples corresponded to the ones used for the multi-class tagger based on vRC jets [1], mentioned before. Several VLB and VLT quark samples for the singlet and doublet models at different masses from 500 to 1400 GeV were combined for RC jets from V-bosons, H-bosons and top-quarks. PROTOS v2.2 [65, 66] with the NNPDF2.3 LO PDF was used for the generation of the hard-scattering process, while PYTHIA 8.1 [90] was used for the parton shower. The background jets were simulated with PYTHIA 8.1 with the same PDF as for the PROTOS samples. For all samples the A14 [141] set of tuned parameters was used. This tagger is abbreviated as NFLIP (**N**on-**FL**at-**I**n-$\mathbf{p_T}$) tagger in the following, while the standard and b-tag enriched taggers described in the previous sections are shortened to FLIP (**FL**at-**I**n-$\mathbf{p_T}$) taggers.



(a) (b)

Figure 6.9: RC jet a) p_T and b) η distributions for the NFLIP tagger after a p_T reweighting. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

The p_T-range of this tagger reaches from 150 GeV up to 2 TeV. In contrast to the FLIP taggers, the MC simulation is not enhanced with events in the high p_T region. Analogous to the multi-class tagger using vRC jets, only an RC jet p_T reweighting is applied. The background RC jet p_T distribution is reweighted to the sum of the V-boson, H-boson and top-quark jet distributions. Figure 6.9(a) shows the RC jet p_T and it can be seen that the majority of RC jets is located between 150 GeV and 1 TeV. The p_T reweighting is applied in this depiction and the differences between all four classes are small. Thus, the NFLIP tagger is mainly trained on RC jets with p_T below 1 TeV, which decreases the performance for higher p_T. This fact should be kept in mind for the following comparisons between all three tagger versions, which investigate different kinematic regions more closely. Another feature should be noted: Figure 6.9(b) shows the RC jet η distribution after applying the RC jet p_T reweighting. This previous version of the tagger was constructed analogous to the vRC jet tagger, omitting a shape difference in η, which is induced by the p_T reweighting. The different shape of background jets compared to the signal jets could in principle be an unphysical feature learned by the NN.

6.8 DNN output distributions and feature investigation

First insight into the discrimination power, performance and feature learning of the MCBOT can be gained by looking at the DNN output distributions. The output of MCBOT is four-dimensional and each dimension of the output vector corresponds to one class. Each dimension can also be considered as a tagger, since it has the ability to differentiate the corresponding RC jet origin from the other classes. Therefore, the four-dimensional output of MCBOT gives a V-tagger, H-tagger, top-tagger and background-tagger.

Construction of test samples

An important ingredient for a consistent comparison of different taggers is the choice of a test sample. Depending on the composition of the four classes and the kinematic phase-space in the test sample, e.g. low or high p_{T}, the performance of a tagger can vary significantly. The same statement holds for the training of a tagger: The NFLIP tagger trained on the VLQ samples has mostly seen RC jets with $p_{\mathrm{T}} < 1\,\mathrm{TeV}$, while the FLIP taggers trained on the flat-in-p_{T} samples extend up to a higher p_{T} range. Therefore, the DNN output distributions are compared in three RC jet p_{T} slices from $150\,\mathrm{GeV}$ to $1\,\mathrm{TeV}$, 1 to $2\,\mathrm{TeV}$ and 2 to $3\,\mathrm{TeV}$. Also the composition of the background jet sample is important. While the standard FLIP tagger is expected to perform best on the standard background jet fraction, the b-tag enriched FLIP tagger is not. The following comparisons are made in three p_{T} slices on the standard test sample and only the FLIP taggers are compared in more detail, since they are used in the analysis presented in Chapter 9. The distributions for tests on the b-tag enriched sample are shown in Appendix A.2. Observing by eye, there is no large difference between the DNN output evaluated on the standard and b-tag enriched FLIP test sets. A closer investigation of the performance is done in Section 6.9.

6.8.1 V-tagger

In Figure 6.10, the DNN output distributions for the V-boson tagger are depicted. The left column shows the NFLIP tagger, while the middle and right column illustrate the standard and b-tag enriched FLIP taggers, respectively. The upper row gives information about the first p_{T} slice from $150\,\mathrm{GeV}$ to $1\,\mathrm{TeV}$ and the lower row for 1–2 TeV. The DNN output distributions for the V-, H-, top- and background-taggers for an RC jet p_{T} from 2–3 TeV are depicted in Appendix A.1. Overall, the distributions in Figure 6.10 show that the NFLIP and FLIP taggers are able to separate the V-boson class from the three other classes. The separation is better for the taggers trained on flat-in-p_{T} samples, which can be concluded from the area under the histogram being larger at higher DNN output values. This corresponds to a higher probability for the FLIP taggers to identify a V-boson jet correctly.

For the p_{T} range from 1–2 TeV in Figures 6.10(e) and 6.10(f), a larger proportion of H-boson and background jets is seen for higher DNN output values compared to the p_{T} range from $150\,\mathrm{GeV}$ to $1\,\mathrm{TeV}$. These are explained by RC jets with an RC jet mass similar to $m_V \approx 85\,\mathrm{GeV}$. This relation is obtained from Figure 6.11, which shows the two-dimensional histograms with the RC jet mass plotted against the DNN output of the V-tagger for RC jets from V-bosons, H-bosons and background. H-boson and background jets typically have a higher and lower mass, respectively, but also V-boson-like

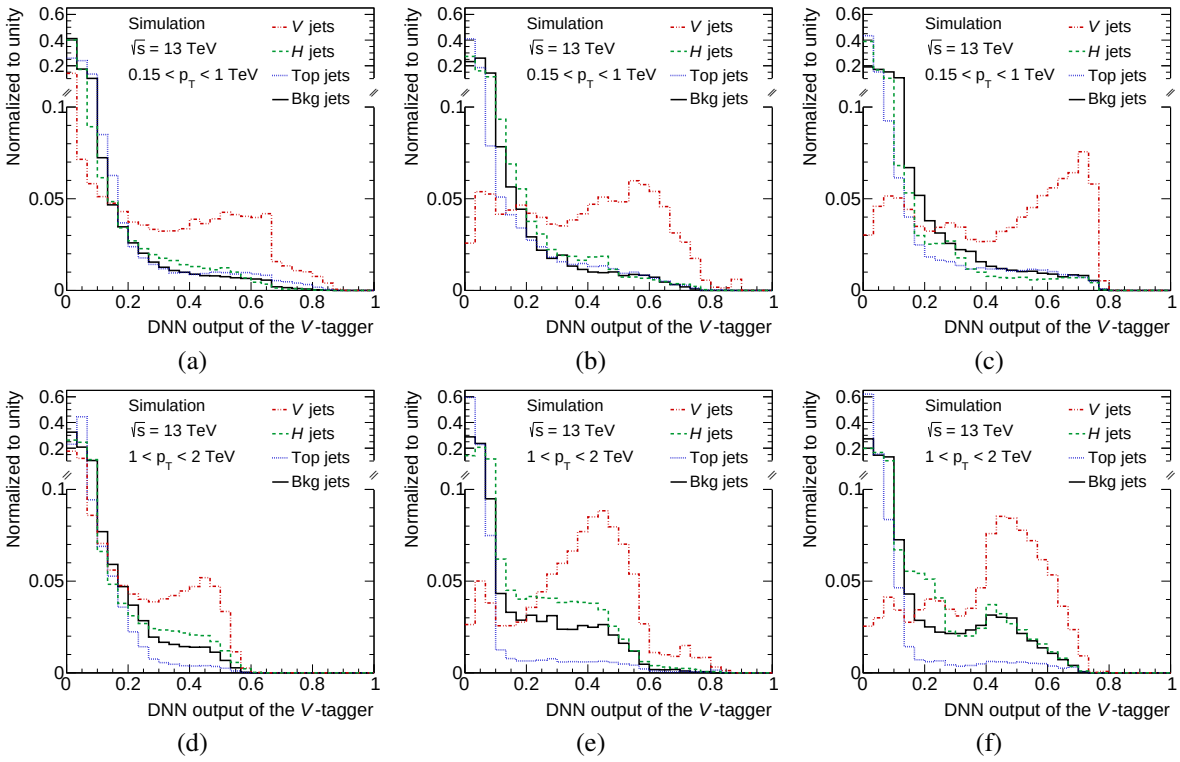


Figure 6.10: DNN output distributions of the V-tagger for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top and 1–2 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

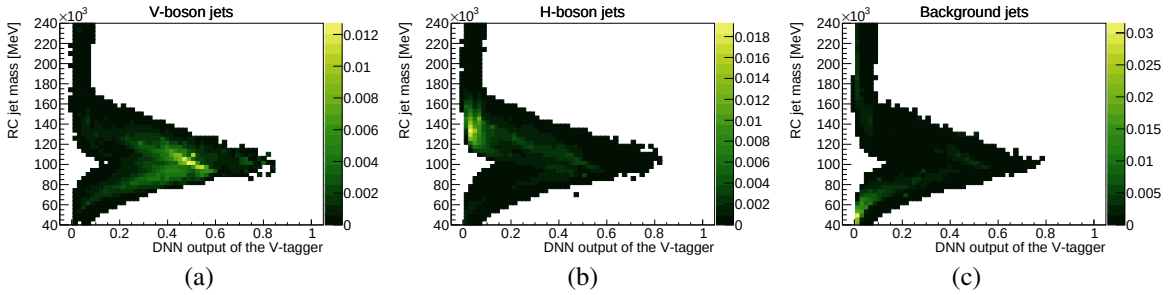


Figure 6.11: RC jet mass plotted against the DNN output of the V-tagger for a) V-boson, b) H-boson and c) background jets for the standard FLIP tagger evaluated on the standard FLIP test samples from 1–2 TeV. All histograms are normalized to unity.

masses occur, as illustrated in Figure 6.8(a). For increasing RC jet p_{T}, the subjets of the V-boson jets merge and thus the mass becomes an important variable for discrimination.

Furthermore, there are some V-boson jets, which occur for output values larger than 0.6 for the standard tagger in Figure 6.10(e). This feature is explained with the η distributions of the leading subjet in Figure 6.12. The V-tagger as well as the background-tagger are shown for V-boson jets in Figures 6.12(a) and 6.12(b), respectively. From the V-tagger it can be seen that the DNN output

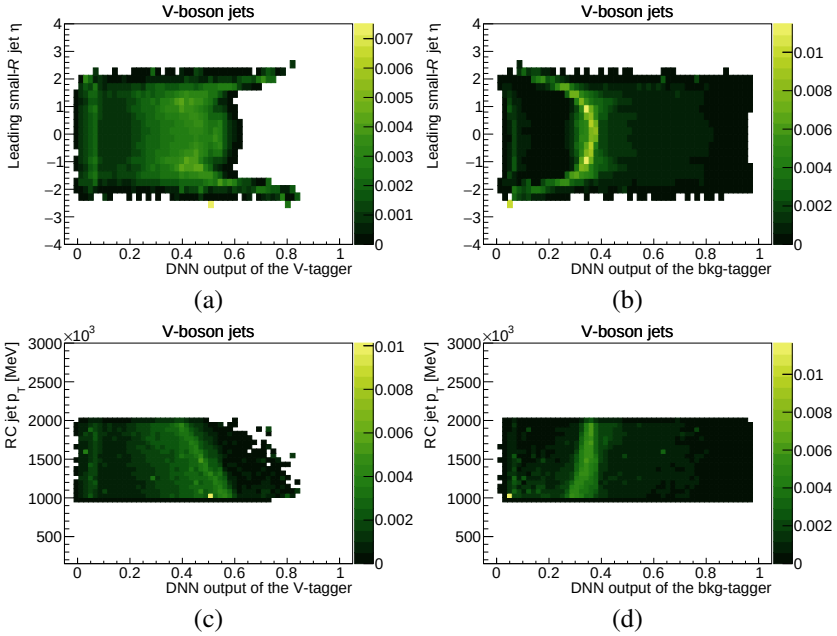


Figure 6.12: η distribution of the leading subjet in a) and b), and the RC jet p_{T} in c) and d) plotted against the DNN output of the V-tagger and background-tagger for V-boson jets for the standard FLIP tagger evaluated on the standard FLIP test samples from 1–2 TeV. All histograms are normalized to unity.

values above 0.6 are obtained for V-boson jets with $|\eta|$ at the largest possible values. Furthermore, DNN output values around 0.55 are reached for very central V-boson jets. A slightly lower DNN output of around 0.4 is more likely for $|\eta| \approx 1$. These observations are inverted when looking at the background tagger for V-boson jets in Figure 6.12(b). The highest V-tagger output values of some V-boson jets above 0.6 are introduced by the two-dimensional reweighting. The η distributions for signal jets are slightly wider than for background jets across the whole MCBOT p_{T} range from $150\,\mathrm{GeV}$ to $3\,\mathrm{TeV}$, as illustrated in Figure 6.6(b). While for this inclusive p_{T} range from $150\,\mathrm{GeV}$ to $3\,\mathrm{TeV}$ all η values occur for signal and background jets, this changes for higher p_{T}: Signal jets occur at high η values, while no background jets are present in the corresponding η region. Thus, the RC jets at these highest $|\eta|$ values are solely V-boson jets. These differences are picked up by the DNN and high η regions are identified with V-boson jets - the same is valid for H-boson and top-quark jets. This means that a small dependence of MCBOT on the kinematics is introduced. However, the fraction of events in these higher η regions is small, as illustrated in Figure 6.7(b), where the differences between the reweighted distributions are very small.

The η variable is also correlated with the p_{T} of an RC jet, as illustrated in Figures 6.12(c) and 6.12(d). A V-boson jet is more likely to be identified correctly, if the RC jet p_{T} is at the lower bound of the p_{T} range ($\approx 1\,\mathrm{TeV}$), while a larger p_{T} leads to a larger value for the background-tagger for V-boson jets.

6.8.2 H-tagger

A depiction of the DNN output of the H-tagger is shown in Figure 6.13 with the same allocation of p_{T} slices and taggers as for the V-tagger. Again, the separation power of the FLIP taggers improves compared to the NFLIP tagger. It is also interesting to see that for very high DNN output values above 0.8 there are a lot of H-boson jets, which are identified correctly and which was not the case for the V-tagger. For the standard tagger in Figure 6.13(b), there are three peaks at 0.4, 0.8 and 0.9 in the distribution for H-boson jets for 150 GeV to 1 TeV. For the b-tag enriched tagger in Figure 6.13(c), there are only two peaks at slightly different, but also high values. In order to investigate these features, important input variables are plotted against the DNN output for a specific output class. Looking at H-boson jets, the number of constituents, the b-tagging decision for the leading subjet and the RC jet mass help to understand the peak structure.

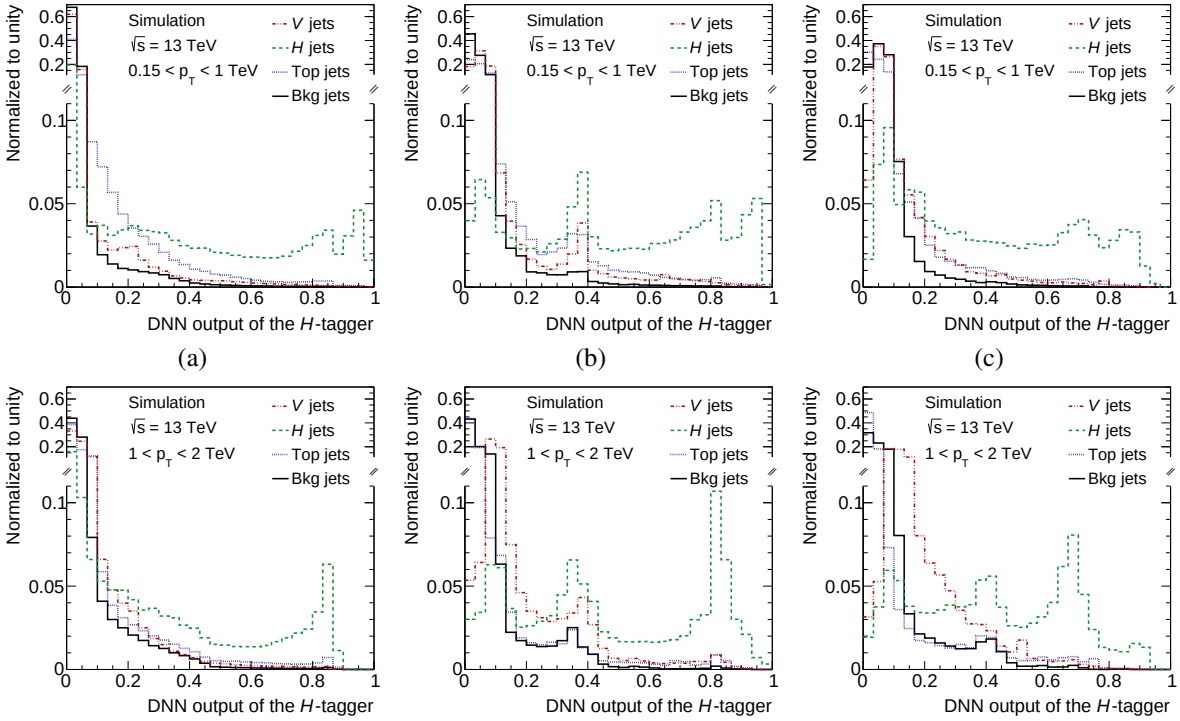


Figure 6.13: DNN output distributions of the H-tagger for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top and 1–2 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

In Figure 6.14(a), the number of subjets of the standard tagger is shown for 150 GeV to 1 TeV. Together with Figure 6.16(a), which shows the b-tagging decision in the same p_{T} range, it can be concluded that the peak at 0.4 comes from RC jets with one or two subjets, which are often b-tagged. The two highest peaks for the standard tagger at 0.8 and 0.9 have similar features: RC jets originating from a H-boson are likely to be tagged correctly, if they consist of two subjets, which are b-tagged. This is how the highest peak at 0.9 is created. The slightly lower peak at 0.8 is built from an RC jet

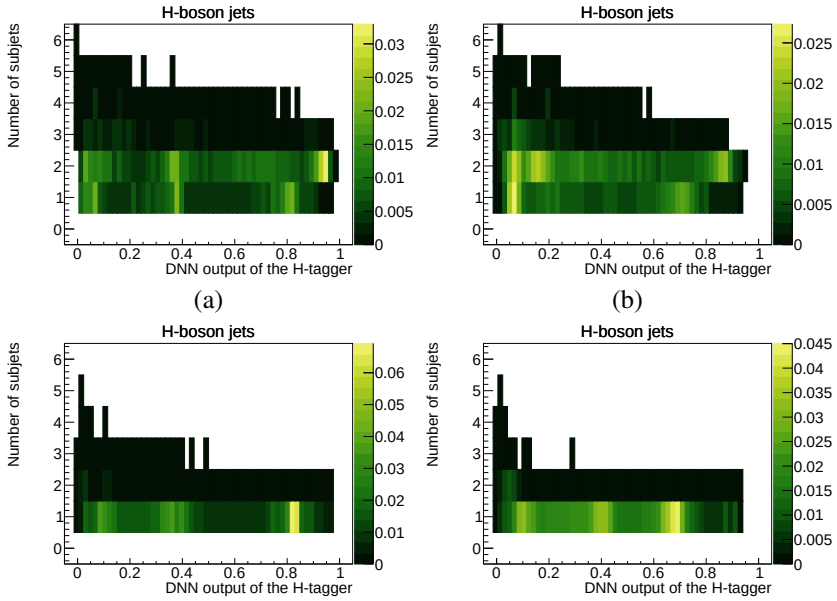


Figure 6.14: Number of constituents plotted against the DNN output of the H-tagger for H-boson jets for the standard FLIP tagger on the left and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from $150\,\mathrm{GeV}$ to $1\,\mathrm{TeV}$ is shown on the top and 1–2 TeV on the bottom. All histograms are normalized to unity.

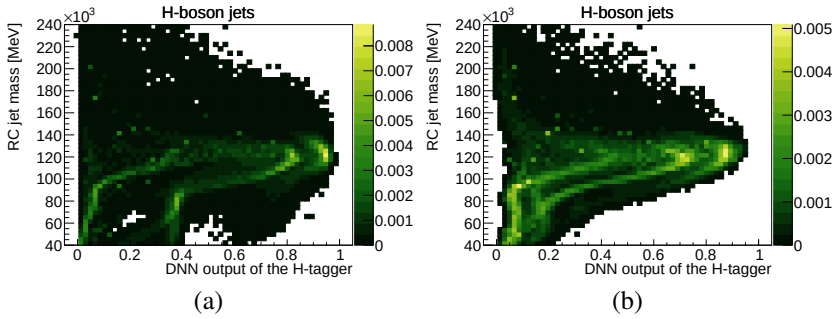


Figure 6.15: RC jet mass plotted against the DNN output of the H-tagger for H-boson jets for a) the standard tagger and b) the b-tag enriched tagger. Both are evaluated on the standard test sample from $150\,\mathrm{GeV}$ to $1\,\mathrm{TeV}$. All histograms are normalized to unity.

with one subjet, which is also b-tagged. This opens the question why the tagger can separate some RC jets with one or two subjets, which are b-tagged, better than others. The feature adding discrimination power is the RC jet mass, which can be seen in Figure 6.15(a). The peak at 0.4 comes mainly from RC jets with smaller masses than the H-boson mass, which could come from partially contained decay products within the jet radius $R = 1$. V-boson and background jets, but also partially contained top-quark jets show a peak at 0.4 as well, since their masses tend to be smaller and therefore are hard to be separated from H-boson jets with smaller masses than $100\,\mathrm{GeV}$.

For the b-tag enriched tagger the same statements hold for the two-peak-structure at the highest output values (of 0.7 and 0.85 in this case) in the p_{T} range from $150\,\mathrm{GeV}$ to $1\,\mathrm{TeV}$. This is supported by Figures 6.14(b) and 6.16(b), which show the number of subjets and the b-tagging decision for the b-tag enriched FLIP tagger, respectively. In contrast, the peak at 0.4 is not present and is shifted to

smaller DNN values at approximately 0.2 with a smaller size compared to the 0.4 peak for the standard tagger. Analogous to the standard tagger, it also corresponds to H-boson jets with lower masses, as illustrated in Figure 6.15(b).

For 1–2 TeV, only one large peak for H-boson jets at high DNN values remains for both FLIP taggers, as can be seen in Figures 6.13(e) and 6.13(f). Figures 6.14(c) and 6.14(d) show that the high peak originates from RC jets with one subjet. As depicted in Figures 6.16(c) and 6.16(d), this peak comes from a subjet, which is b-tagged. For an RC jet with higher p_{T}, the subjets are more likely to merge and therefore, the two-peak-structure mentioned previously disappears. The peak at 0.4 for H-boson jets is caused by RC jets with one subjet which is mostly not b-tagged. V-boson, top-quark and background jets with these features and a mass similar to the H-boson mass build a small peak underneath the H-boson jet peak around 0.4. It is present for the standard and the b-tag enriched taggers.

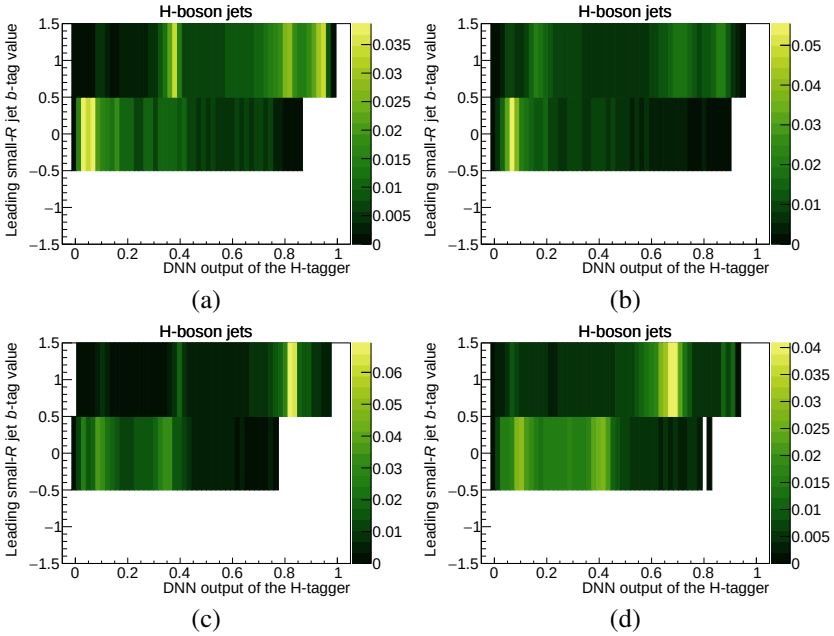


Figure 6.16: b-tagging decision für the leading subjet plotted against the DNN output of the H-tagger for H-boson jets for the standard FLIP tagger on the left and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top and 1–2 TeV on the bottom. All histograms are normalized to unity.

6.8.3 Top-tagger

The DNN output distributions for the three different versions of the top-tagger are depicted in Figure 6.17. Starting from the test sample from 150 GeV to 1 TeV in the top row, it can be seen that the standard and b-tag enriched FLIP taggers have a significantly larger area under the histogram for high DNN output values compared to the NFLIP tagger, corresponding to a better separation power of the FLIP taggers for top-quark jets.

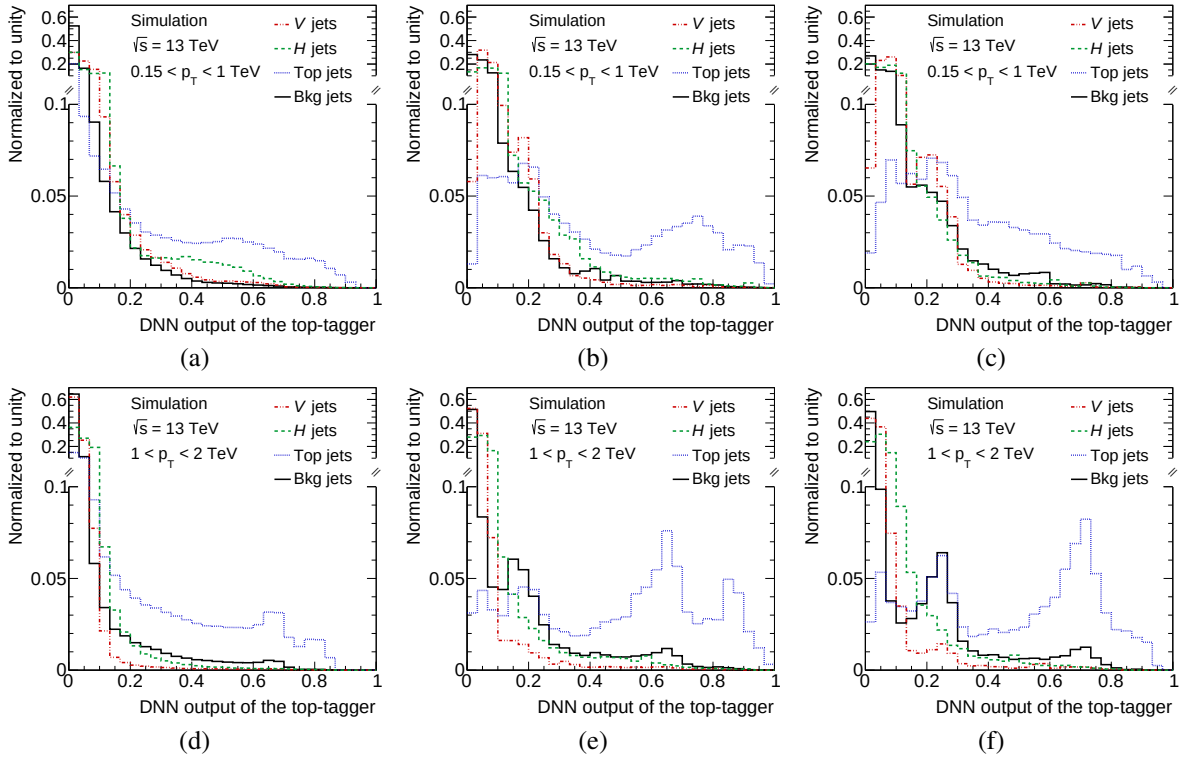
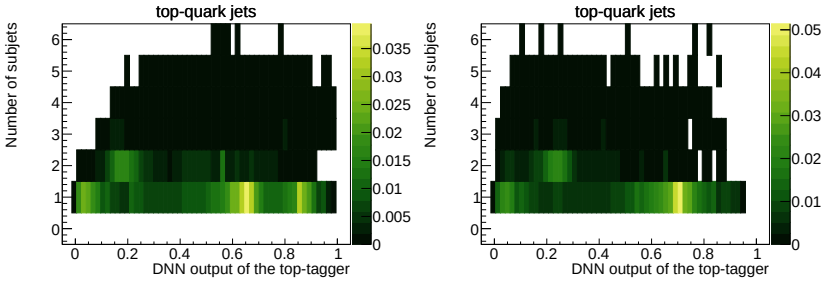


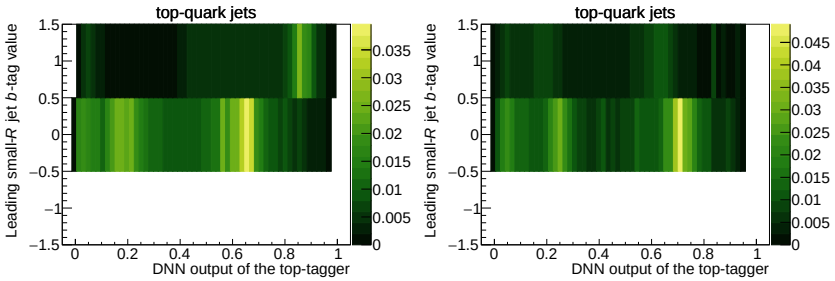
Figure 6.17: DNN output distributions of the top-tagger for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top and 1–2 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

For 1–2 TeV, the same statement holds. Furthermore, a different peak structure for top-quark jets is present in the DNN output in Figures 6.17(e) and 6.17(f) with two peaks at high DNN values for the standard tagger and one peak for the b-tag enriched tagger. In Figures 6.18(a) and 6.18(b), information for both FLIP taggers is provided about the number of constituents in this p_{T} region. It can be seen that the majority of RC jets consists of one constituent between 1–2 TeV. Combined with the illustration of the b-tagging decision of the leading subjet, presented in Figures 6.19(a) and 6.19(b), the structures can be explained. For the standard tagger, the two peaks originate from RC jets, which are either b-tagged or not b-tagged, while consisting of one subjet. For the b-tag enriched tagger, top-quark jets are more likely to be identified correctly, if this RC jet with one constituent is not b-tagged. Combining this information with the b-tag enriched H-tagger, it can be concluded that



(a) (b)

Figure 6.18: Number of constituents plotted against the DNN output of the top-tagger for top-quark jets for a) the standard FLIP tagger and b) the b-tag enriched FLIP tagger. The p_{T} range is 1–2 TeV. All histograms are normalized to unity.



(a) (b)

Figure 6.19: b-tagging decision für the leading subjet plotted against the DNN output of the top-tagger for top-quark jets for the a) standard FLIP tagger and the b) b-tag enriched FLIP tagger. The p_{T} range is 1–2 TeV. All histograms are normalized to unity.

an RC jet with one b-tagged subjet is more likely to be tagged as a H-boson jet than a top-quark jet. Once again, it should be noted that the H-boson is assumed to decay exclusively into $b\bar{b}$ final states.

Furthermore, there are a few less pronounced features for 1–2 TeV. The peak for top-quark and background jets at DNN values around 0.2 originates from RC jets, which have two subjets without a b-tag. It is present for both FLIP taggers. Furthermore, there is a small peak for background jets at DNN values of around 0.7 for both taggers for 1-2 TeV. The structure can be found underneath the peak in the top-quark jet spectrum that corresponds to the non-b-tagged peak. To further investigate these features, the RC jet mass is plotted against the DNN output of the top-tagger in Figure 6.20 for top-quark and background jets for both taggers. Again, the mass is a very discriminating feature with the majority of the real top-quark jets located around the top-quark mass, depicted in Figures 6.20(a) and 6.20(b) for the standard and b-tag enriched FLIP taggers respectively. As seen in Figures 6.20(c) and 6.20(d), for background jets there are some jets, which have a very large mass, comparable to the top-quark mass. These jets are likely to be mistagged as top-quark jets, especially, if they are not b-tagged. It also means that the taggers shape the background mass distribution, resulting in mistagged background jets with a similar mass as the signal.

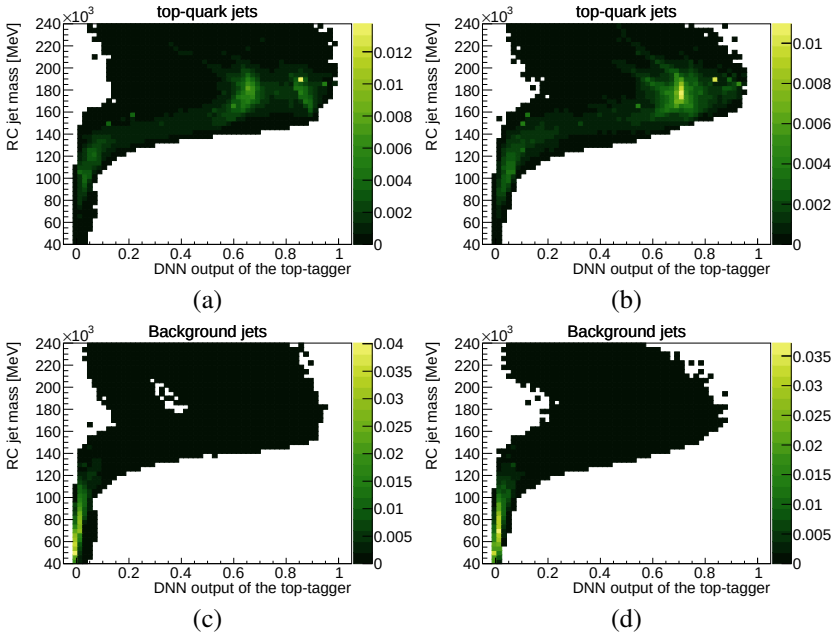


Figure 6.20: RC jet mass plotted against the DNN output of the top-tagger for top-quark jets in the top row and background jets in the bottom row for the standard tagger in a) and c) and the b-tag enriched tagger in b) and d). The standard test sample from 1–2 TeV is used. All histograms are normalized to unity.

6.8.4 Background-tagger

The DNN output distributions for the background-tagger are depicted in Figure 6.21. For the standard and b-tag enriched FLIP taggers in the p_{T} range between 150 GeV to 1 TeV, there is good discrimination for background jets in comparison to V-boson, H-boson and top-quark jets. Especially, the high peak of V-boson, H-boson and top-quark jets close to one disappears, which is quite pronounced for the NFLIP tagger in Figure 6.21(a). This indicates that the FLIP taggers perform better at rejecting background jets.

The peak structures for background jets between 150 GeV and 1 TeV, seen in Figures 6.21(b) and 6.21(c), are investigated with the number of constituents and the b-tagging decision for the leading subjet. The corresponding distributions in this p_{T} range can be seen in Figures 6.22(a) and 6.23(a) for the standard tagger and in Figures 6.22(b) and 6.23(b) for the b-tag enriched tagger. The number of constituents shows that the largest peak comes from RC jets with one subjet, while the adjacent peak consists of RC jets with two constituents. In combination with the b-tagging decision, it can be concluded that the background-tagger is likely to tag an RC jet correctly, if it is not b-tagged. The same statement holds for the RC jet p_{T} range from 1–2 TeV.

Between 150 GeV and 1 TeV, the standard tagger also shows a quite pronounced peak for V-boson jets around 0.25, as depicted in Figure 6.21(b). In Figures 6.24(a) and 6.24(c), the b-tagging decision for the leading subjets and the number of constituents are shown for V-boson jets. It follows that the large peak comes from RC jets with one subjet without a b-tag. Furthermore, the mass of V-boson jets is typically lower than for H-boson and top-quark jets, which makes them more likely to be confused with background.

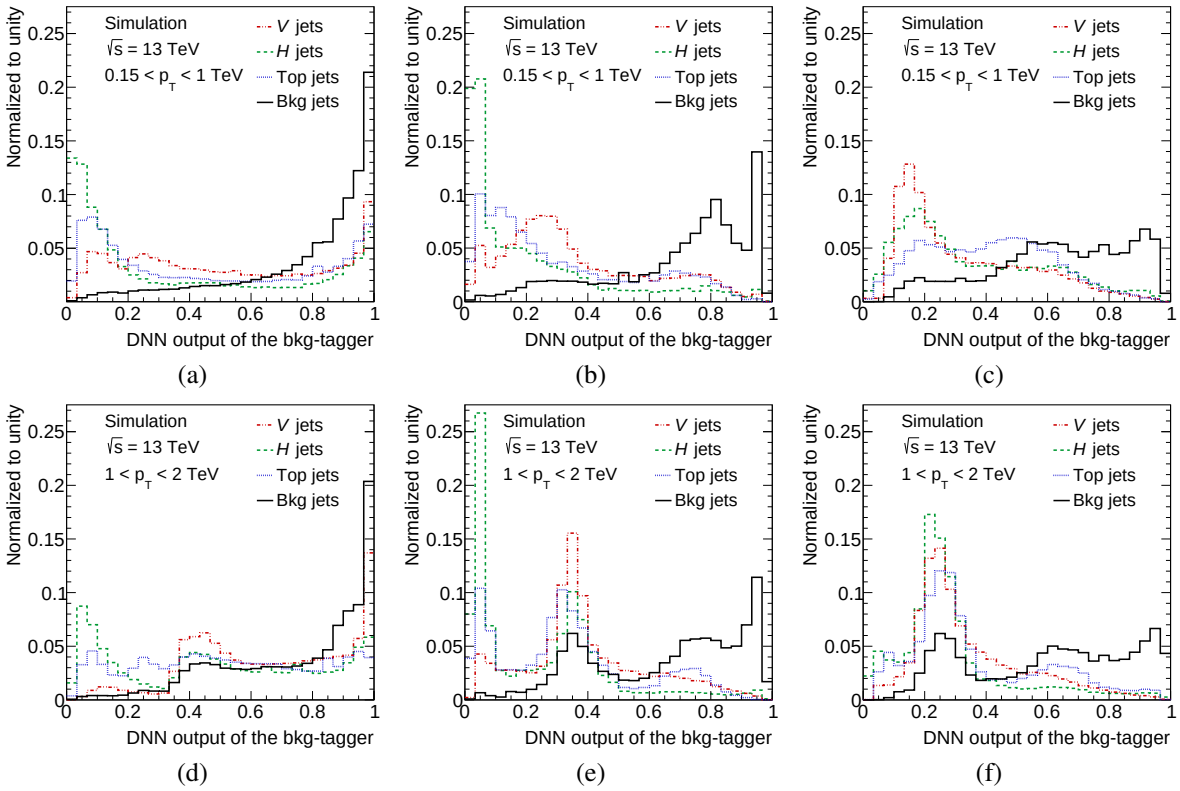
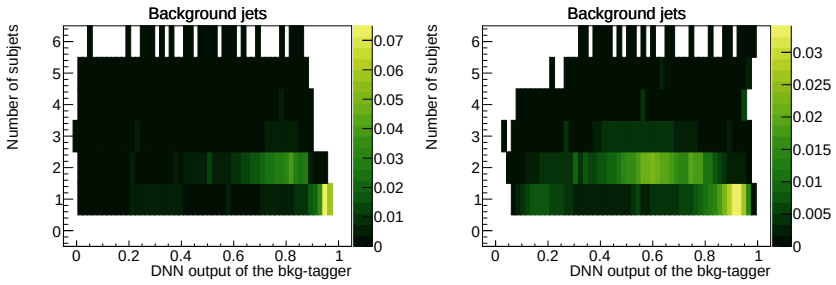


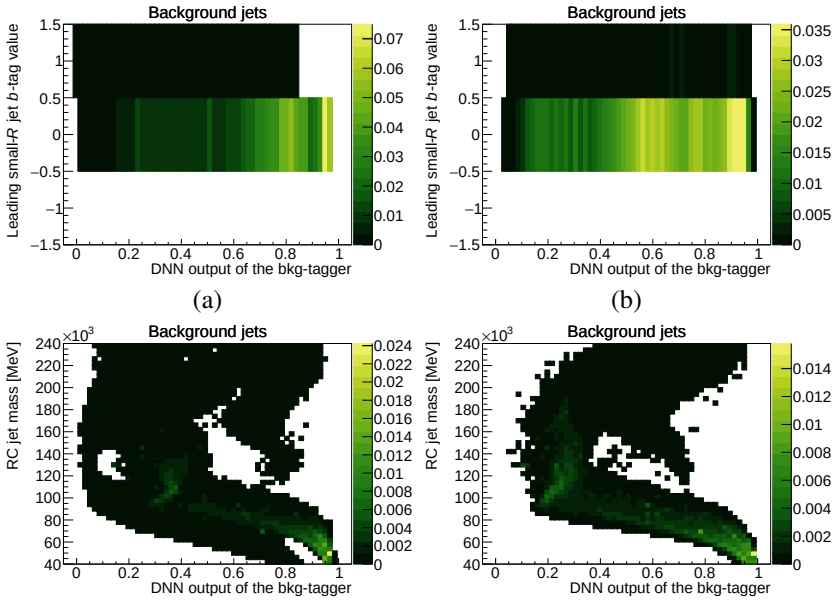
Figure 6.21: DNN output distributions of the background-tagger for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top and 1–2 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

For the b-tag enriched tagger in Figure 6.21(c), there are eye-catching features for top-quark jets. Compared to the standard tagger, a second peak associated with top-quark jets emerges at DNN values of approximately 0.5. This can again be explained with the b-tagging information from the leading subjet and the number of subjets, depicted in Figures 6.24(b) and 6.24(d). The background-tagger is more likely to tag a top-quark jet incorrectly, if it consists of two subjets, where the leading subjet is often not b-tagged, but can be b-tagged as well. The difficulty of identifying RC jets with exactly two subjets can also be seen in the other DNN distributions. All the structures from V-boson and top-quark jets underneath the second largest peak in the background spectrum in Figures 6.21(b) and 6.21(c), as well as Figures 6.21(e) and 6.21(f) originate from this feature. Looking at the p_{T} range from 1–2 TeV and at Figures 6.25(c) and 6.25(d), a small peak is present for two subjets at higher values. From these two figures also another interesting feature can be deducted: In combination with the b-tagging decision of the leading subjet in Figures 6.25(a) and 6.25(b), it explains the two versus one peak-structure of V-boson, H-boson and top-quark jets in the DNN output for both FLIP taggers. There are two peaks for these three classes for the standard tagger at values close to zero and around 0.35. For the b-tag enriched tagger the peak at zero almost disappears entirely, while a large peak at 0.25 is present. The standard tagger uses information about b-tagging of exactly one subjet to achieve a better discrimination. Since for the b-tag enriched tagger the background class has a lot of b-tagged



(a) (b)

Figure 6.22: Number of constituents plotted against the DNN output of the background-tagger for background jets for a) the standard FLIP tagger and b) the b-tag enriched FLIP tagger for a p_{T} range from 150 GeV to 1 TeV. All histograms are normalized to unity.



(a) (b)

(c) (d)

Figure 6.23: b-tagging decision for the leading subjet in a) and b) and RC jet mass in c) and d) plotted against the DNN output of the background-tagger for background jets for the standard FLIP tagger on the left and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top and 1–2 TeV on the bottom. All histograms are normalized to unity.

subjets, it focuses more on other features and thus no peak related to a b-tagged subjet is present at zero. On its own test sample, it performs best, as shown in Section 6.9.

For the p_{T} range from 1–2 TeV, an additional peak at approximately 0.25 or 0.35 occurs in the background jet spectrum for the standard and the b-tag enriched tagger, respectively. Figures 6.23(c) and 6.23(d) show the RC jet mass for the background class. It can be concluded that for the standard and b-tag enriched taggers this peak comes from background jets with high masses. It is the counter part of the feature exhibited for the top-tagger, and shows that similar RC jet masses from different origins are difficult to distinguish.

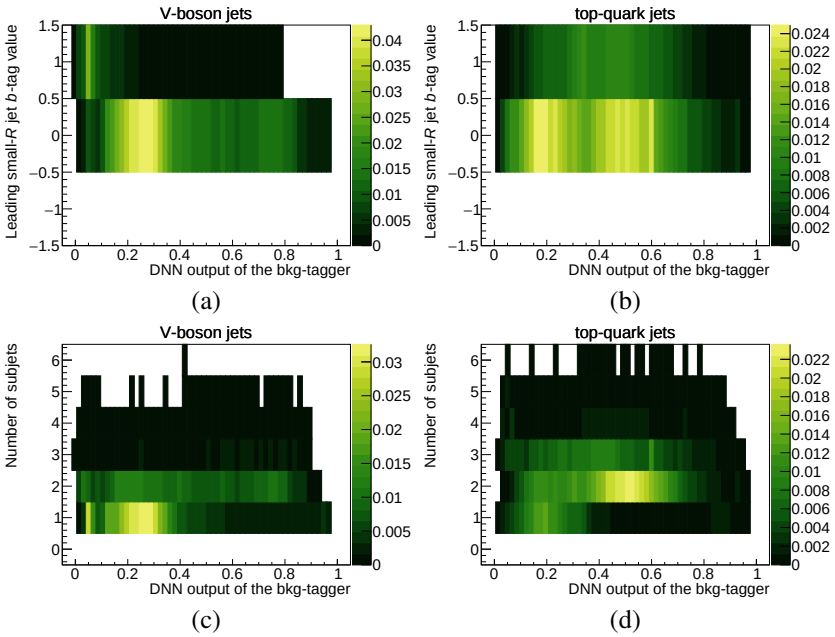


Figure 6.24: Input features for the background tagger from 150 GeV to 1 TeV for V-boson jets and the standard tagger on the left. The b-tag enriched tagger and top-quark jets are shown on the right. In a) and b), the b-tagging decision for the leading subjet is plotted. In c) and d), the number of constituents is shown. All histograms are normalized to unity.

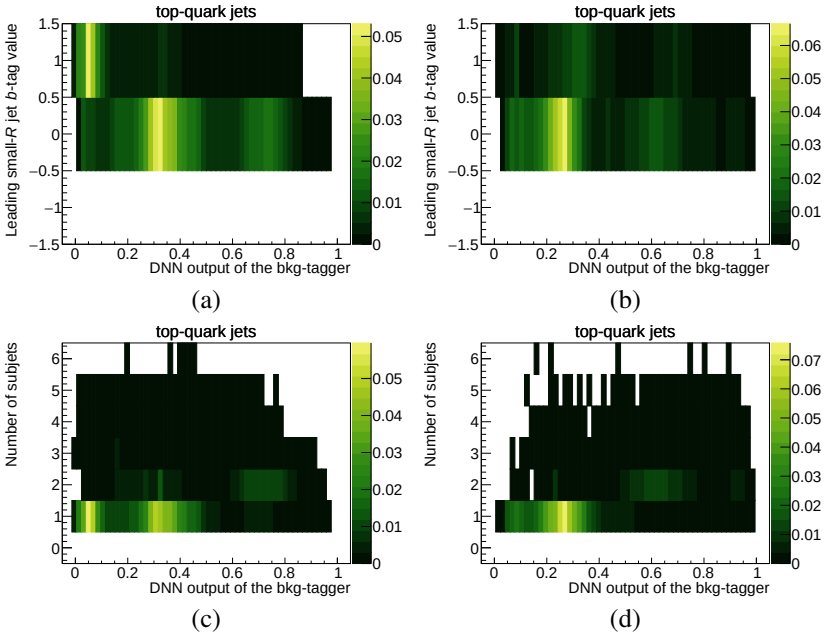


Figure 6.25: Input features for the background tagger from 1–2 TeV. In a) and b), the b-tagging decision for the leading subjet is shown for top-quark jets. In c) and d), the number of constituents are depicted. The standard and b-tag enriched FLIP taggers are shown on the left and right, respectively. All histograms are normalized to unity.

6.8.5 Summary of feature learning

A short summary of the knowledge gained in this section is given. The FLIP taggers showed improved performance in comparison to the NFLIP tagger for the p_{T} range from 150 GeV to 2 TeV. Appendix A.1 supports this statement for RC jets with a p_{T} from 2–3 TeV. The following statements are valid for both FLIP taggers if not stated otherwise.

The most important features used by the DNN to separate the four classes are the RC jet mass, the number of constituents of an RC jet and the b-tagging decision from the leading subjet. Below an RC jet p_{T} of 1 TeV, the DNN obtains more information from the subjet multiplicity than for $p_{\mathrm{T}} > 1$ TeV, where the constituents of the V-boson, H-boson and top-quark jets are mostly merged into one subjet. It was also shown that the RC jet mass highly contributes to the discrimination between the four classes. V-boson jets are more likely to be tagged as background jets, since their RC jet masses are often more similar to background jet masses than H-boson or top-quark jet masses. The RC jets with a high DNN output have shown to carry an RC jet mass similar to that tagger's signal class. For example, background jets with similar masses as the top-quark mass have high DNN outputs of the top-tagger.

In general, the DNN learned that a b-tagged subjet is an indication for the presence of a H-boson or top-quark jet, which is expected due to the high branching ratios into final states with b-quarks. From the b-tag enriched H- and top-tagger for 1–2 GeV, it was seen that RC jets with exactly one subjet which is b-tagged are rather tagged as H-boson jets than as top-quark jets. This is a contrary to the standard FLIP H- and top-tagger, where a b-tagged subjet is not only an indication for a H-boson jet, but also for a top-quark jet.

Both FLIP taggers were able to learn features enabling a good discrimination between signal and background jets. The property of background jets to contain one constituent that is mostly not b-tagged was learned by the DNNs. It was also shown that signal RC jets consisting of two subjets without a b-tag are more likely to be mistagged as background.

In summary, the features learned by MCBOT are more complex than simple cuts on, e.g. the RC jet mass or the b-tagged jet multiplicity, which can be applied directly in an analysis. In addition, different features from different p_{T} ranges of the input RC jets were learned by MCBOT and are summarized in a powerful tagger for simultaneous discrimination of V-boson, H-boson, top-quark and background jets.

6.9 Performance

In the previous section, the features of the DNN output distributions for all four classes were studied. First insight into the discrimination power between the classes for each tagger was gained and a qualitative test was obtained. Nevertheless, a quantitative comparison is still needed. In order to investigate the performance more closely, receiver operating characteristics (ROCs) are used. With these curves, the efficiency of one tagger to identify its corresponding truth-matched jet class correctly is plotted against the rejection of another class. The rejection is defined as $\frac{1}{\epsilon}$, where ϵ is the efficiency of the class to be rejected at a specific cut value on the DNN output.

In the following, the separation power of MCBOT's three V-boson, H-boson and top-quark classes are investigated compared to the background class. No comparison of the signal rejection of one class

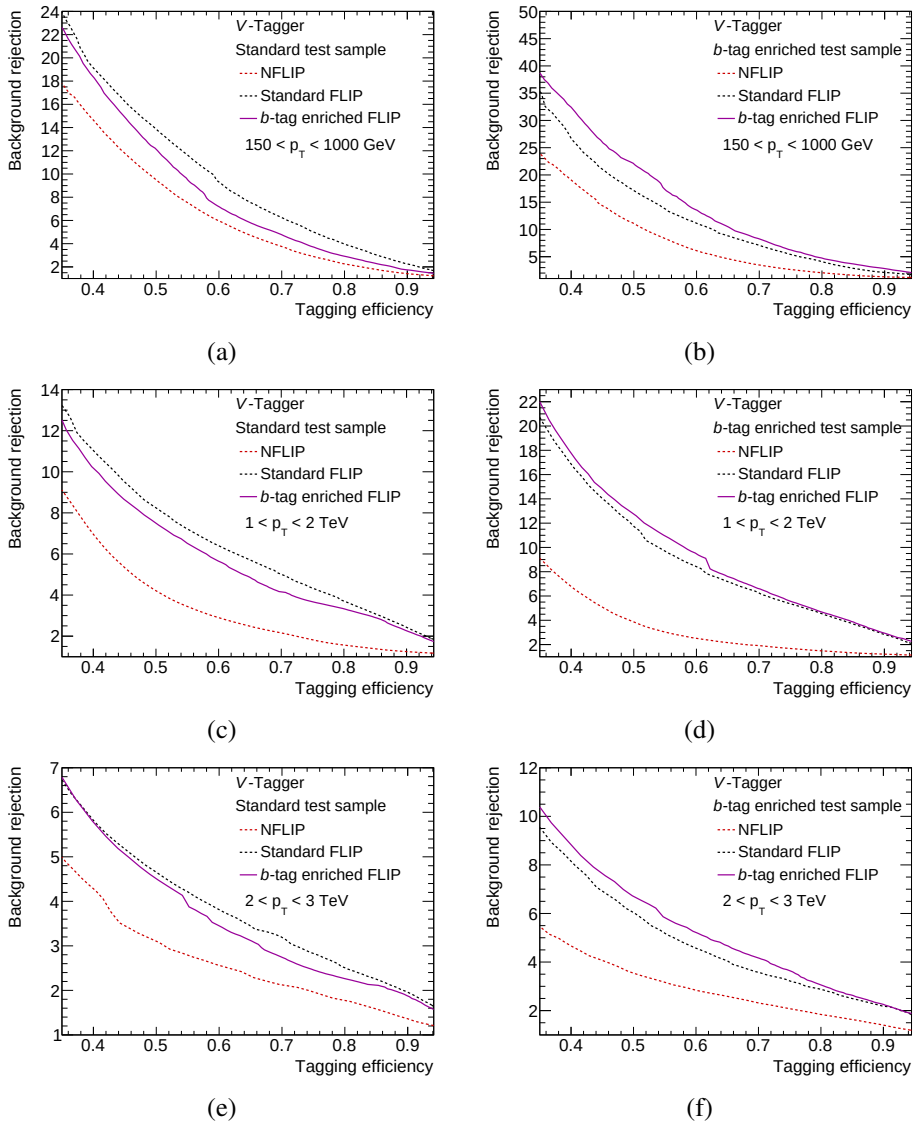


Figure 6.26: ROCs for the V-tagger on the standard and b-tag enriched FLIP test samples on the left and right side, respectively. The NFLIP, the standard FLIP and the b-tag enriched FLIP taggers are shown. All three RC jet p_{T} regions are depicted: On the top, 150 GeV to 1 TeV is shown, while in the middle and on the bottom 1–2 TeV and 2–3 TeV are illustrated, respectively.

against another is done, since the main focus is the rejection of background jets. The comparison is made for all three p_{T} slices of the test sample, as introduced before: 150 GeV to 1 TeV, 1 to 2 TeV and 2 to 3 TeV. The studies in Section 6.8 were done on the test sample with the composition of the training sample of the standard FLIP tagger. For the ROC curves in this section, also the b-tag enriched test samples are depicted.

In Figure 6.26, the signal efficiency for the V-tagger is plotted against the background rejection. The top row with Figures 6.26(a) and 6.26(b) shows the first p_{T} slice from 150 GeV to 1 TeV. In the left column the standard test sample is used, while in the right column the performance comparison is made with the b-tag enriched test sample. The middle row with Figures 6.26(c) and 6.26(d) depicts the

second p_{T} slice for 1–2 TeV. The last row shows the highest p_{T} slice of 2–3 TeV in Figures 6.26(e) and 6.26(f). All of the distributions show that the FLIP taggers perform better than the NFLIP tagger. It can also be seen that the rejection power decreases for higher RC jet p_{T}, indicating that the tagger is more suitable for a phase-space where the individual subjets are not merged and carry more information about e.g. the 2- or 3-prong structure. Furthermore, it can be concluded that a tagger performs best on its corresponding test sample, as one would expect: The standard tagger always shows a higher rejection on the standard composition test sample in all three p_{T} regions. Similarly, the b-tag enriched tagger performs better on the b-tag enriched test sample.

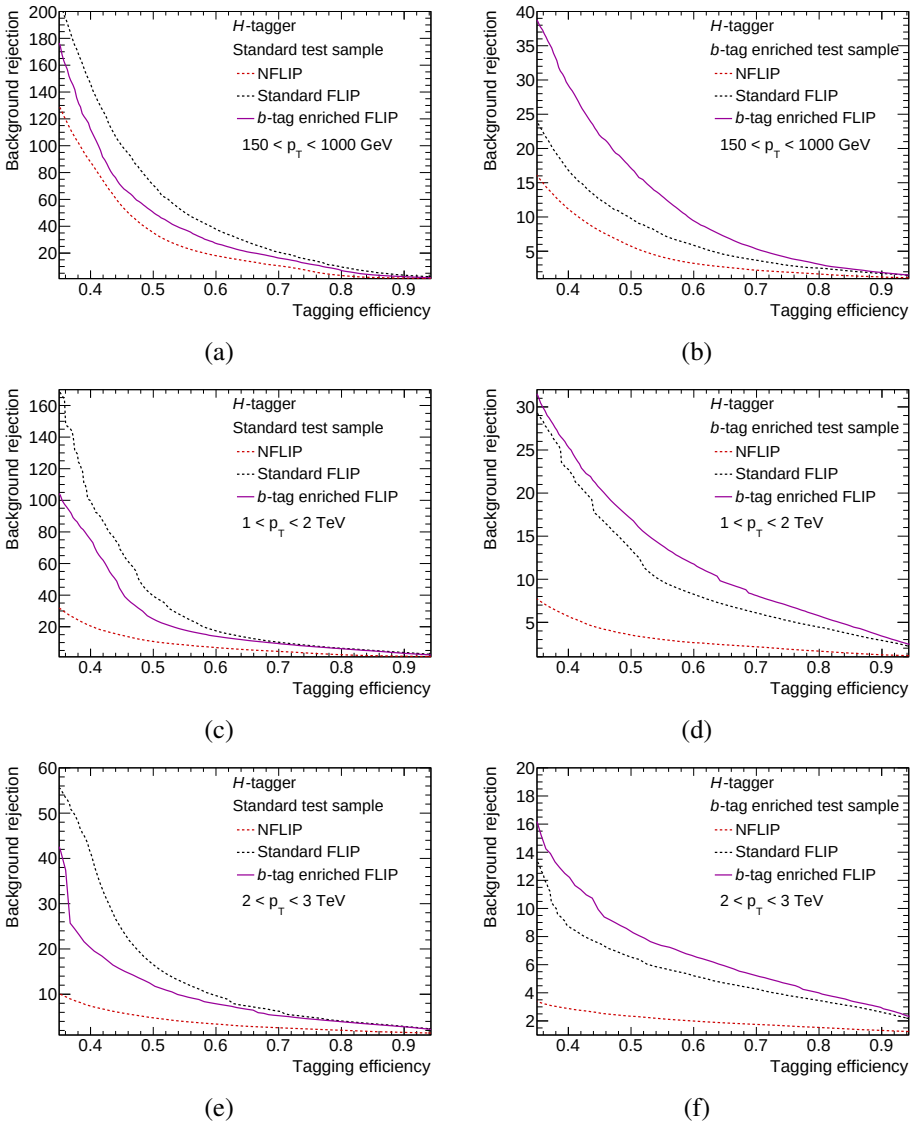


Figure 6.27: ROC curves for the H-tagger on the standard and b-tag enriched FLIP test samples on the left and right side, respectively. The NFLIP, the standard FLIP and the b-tag enriched FLIP taggers are shown. All three RC jet p_{T} regions are depicted: On the top, 150 GeV to 1 TeV is shown, while in the middle and on the bottom 1–2 TeV and 2–3 TeV are shown, respectively.

For the H-tagger, the ROC curves can be found in Figure 6.27. The same order of plots is kept as in Figure 6.26. In all of the distributions the FLIP taggers perform better than the NFLIP tagger, even in the p_{T} range from 150 GeV to 1 TeV, where most of the NFLIP tagger's training p_{T} range is located. Again, the FLIP taggers perform best on their own test samples. An additional feature can be seen for 150 GeV up to 2 TeV: The background rejection is significantly higher for the standard tagger on the standard test sample than for the b-tag enriched tagger on its own test sample. The background rejection is higher by approximately a factor of five in these two p_{T} regions. The larger rejection can be explained by the composition of the test sample. RC jets, which originate from H-bosons, are likely to have one or two b-tagged subjets. In the standard test sample most of the background

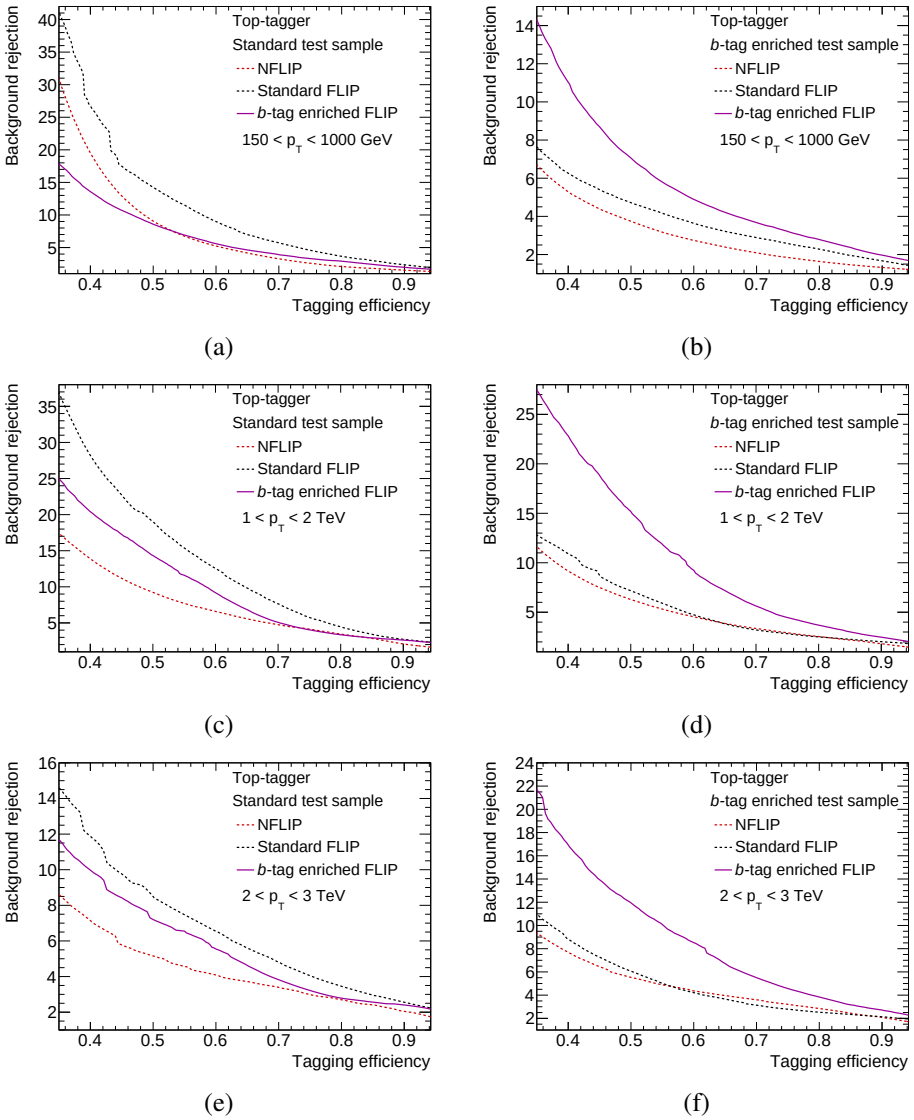


Figure 6.28: ROCs for the top-tagger on the standard and b-tag enriched FLIP test samples on the left and right side, respectively. The NFLIP, the standard FLIP and the b-tag enriched FLIP taggers are shown. All three RC jet p_{T} regions are depicted: On the top, 150 GeV to 1 TeV is shown, while in the middle and on the bottom 1–2 TeV and 2–3 TeV are shown, respectively.

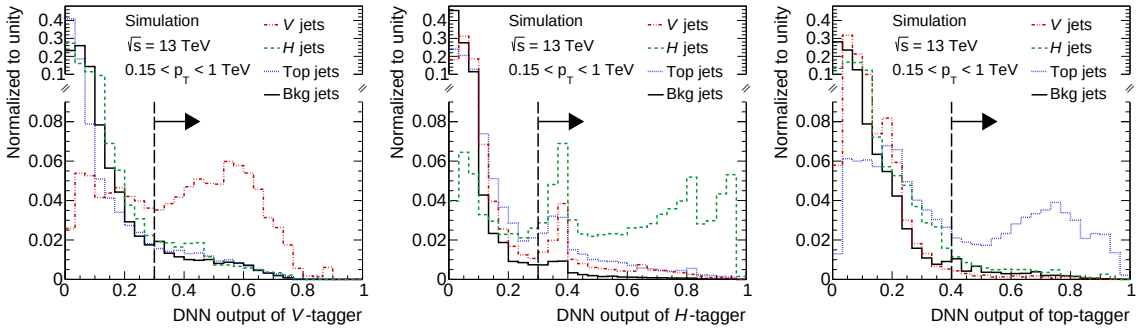
jets originate from light quarks or gluons, while for the b-tag enriched test sample, their fraction is reduced in favor of the enrichment of b-tagged jets. Thus, the separation of H-boson and background jets is more difficult. On the other hand, it can be seen that the enrichment with b-tagged jets leads to a H-tagger, which is better at separating H-boson jets and b-tagged jets compared to the standard tagger. This statement is supported by Figure 6.27(b), 6.27(d) and 6.27(f), where the b-tag enriched tagger is constantly better at rejecting the background class.

The ROC curves for MCBOT's top-tagger are depicted in Figure 6.28. In Figure 6.28(a), it can be seen that in the lowest p_{T} range the b-tag enriched tagger performs worse on the standard test sample compared to the NFLIP tagger. Nevertheless, the standard tagger always performs better than the NFLIP tagger and the b-tag enriched tagger on the standard test sample. As shown in Figure 6.28(b), on its own test sample the b-tag enriched tagger performs best again. For the two higher p_{T} ranges, the same statement holds. Especially for the test on the b-tag enriched sample the b-tag enriched tagger outperforms the two other taggers significantly. Again, the b-tag enriched tagger is more suitable for an environment with a higher number of b-tagged jets in the background composition. Thus, the background objects present in an analysis provide guidance about whether the standard or b-tag enriched tagger should be used.

6.10 Unambiguous identification of V-boson, H-boson and top-quark jets

The output of the multi-class tagger is constructed in a way that all four classes add up to one for each RC jet. As mentioned before, the softmax activation function in combination with the categorical cross-entropy allows for a probability interpretation. The four numbers associated to each RC jet give information about which class the jet is most likely to belong to or not. When optimizing an analysis, there are specific requirements on the final-state objects like the presence of one or multiple boosted objects. The signal efficiency of a tagger is important in order to optimize the analysis selection. Depending on the strategy, a high signal efficiency or a high background rejection is necessary. The DNN output value corresponding to a specific signal efficiency is called working point (WP).

For MCBOT, the WP can be obtained by cutting on the DNN output distributions of V-boson, H-boson and top-quark jets. In Figure 6.29, three example distributions with three WPs are illustrated. All RC jets above the thresholds, regardless of their origin obtained by truth-matching, are tagged. This leads to ambiguities in the MCBOT tags: Some jets may be tagged as two or even three of the V-boson, H-boson and top-quark classes. The ambiguities have to be resolved in favor of a unique tag and the most likely choice to tag a jet correctly. In principle, always assigning the tag with the highest probability could be used, since the unambiguous tags are assigned according to the best possible decision of the DNN. This option, denoted as "natural WP", was tested and found to be less useful due to its inflexibility in optimizing the analysis sensitivity. Choosing the highest DNN output as tag means that only one WP for each class can be chosen. Assuming that an analysis is very rich in one of the boosted objects, but not in the others, this might induce a worse performance. Therefore, the choice of a WP is crucial in the analysis optimization. Instead of the natural WP, another procedure is followed. Three desired WPs are chosen, one for each signal tagger, and if a double tag occurs, the RC jet is assigned to the class with the highest DNN output and a tag. If three tags are present, also the



(a) (b) (c)

Figure 6.29: DNN output for the standard FLIP V-, H- and top-tagger for the standard FLIP test sample in the RC jet p_{T} range from 150 GeV to 1 TeV. The arrows show three independent choices of a WP for the V-, H- and top-tagger, possibly leading to ambiguous tags. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

highest DNN output for this RC jet among the three signal taggers is used to assign an unambiguous tag. This approach leaves full flexibility to the user and tends towards the tagging decisions with the same signal efficiencies as for the natural WP.

Analogous to the argumentation with different test samples influencing the performance, the signal efficiency in an analysis is always dependent on the signal process under study. Therefore, a truth matching can be done on MC generator level in order to quote the signal efficiency. On the other hand, the knowledge of the signal efficiency is not needed for the optimization of an analysis. Instead, different WPs can be tested, corresponding to an unknown signal efficiency, in order to find the best combination of cuts for the V-tagger, H-tagger and top-tagger. The MCBOT working point optimization of the VLQ analysis presented in Chapter 9 is performed in Section 9.3.2.

6.11 Summary

In this chapter, the Multi-Class Boosted-Object Tagger (MCBOT) on RC jets was presented and characterized. The four classes to be separated are V-boson, H-boson, top-quark and background jets. The tagger uses RC jet properties like the mass, p_{T} and the number of subjets. Furthermore, the four-momentum vector of the three subjets with highest p_{T}, as well as their b-tagging decision are used. The WPs for the V-tagger, H-tagger and top-tagger can be chosen independently of each other. If multiple tags occur, the tag is chosen which belongs to the highest DNN output, corresponding to the highest probability to be the correct tag.

Two taggers were trained on flat-in-p_{T} samples: One contains the standard background jet composition as provided by the MC simulation with a high fraction of jets originating from light jets and gluons, denoted as standard FLIP tagger. The other tagger's training sample is enriched in b-tagged jets and provides a better discrimination between signal and background for environments with a lot of b-tagged jets. This version is called b-tag enriched FLIP tagger.

Corresponding to the development of two taggers, two test samples were used for the standard and b-tag enriched compositions of the background jet class to evaluate the performance of MCBOT.

Each of these types is split in three p_{T} slices in order to be able to compare the performance in phase-spaces with different kinematics. A performance gain of the taggers trained on flat-in-p_{T} samples was observed in comparison to the NFLIP tagger. A tagger similar to the NFLIP tagger and based on variable-R reclustered (vRC) jets was used in the all-hadronic VLQ search of the ATLAS collaboration with $36.1\,\mathrm{fb}^{-1}$ [1]. It was also found that each FLIP tagger performs best on its corresponding test sample. Therefore, the choice between the standard or b-tag enriched tagger should be based on the background composition in an analysis. It should also be noted that the flat-in-p_{T} MCBOT is a novel tool provided to the ATLAS collaboration. It extends the idea and implementation of a variable-R RC jet based tagger introduced in Ref. [1] to the more widely used RC jets, while improving the performance due to the flat-p_{T} spectrum and a two-dimensional reweighting used in the training.

Chapter 7
Statistical concepts for hypothesis testing

In order to interpret the observed data at an experiment with respect to the hypothesis that vector-like quarks (VLQs) exist, statistical methods are utilized. The aim is to obtain a result which offers a statistical significance that enables to claim a discovery or exclude a specific parameter space, where the existence of VLQs can be negated with the data. The sensitivity to VLQs depends on parameters like their production cross-section and their branching ratios into searched final states. Both quantities can be very low and therefore, the collection of more data might enable more stringent limits or even a discovery in the future.

Whether a discovery can be claimed or exclusion limits are set, is decided by performing hypothesis tests. The concepts of background and signal were discussed in Chapter 4 and are important for the following hypothesis definitions. If a hypothesized signal is not represented in nature, a test of the hypothesis S+B assuming the presence of signal and Standard Model (SM) background is rejected at a specified confidence level (CL). Then, the background-only (B-only) hypothesis is assumed to be true. In contrast, the rejection of the B-only hypothesis corresponds to a discovery claim.

In the following, basic quantities needed for hypothesis testing are introduced in Section 7.1. Moreover, the use of likelihood ratios, especially the *profile likelihood ratio* (PLR), as test statistic is motivated in Section 7.2. Subsequently, methods to determine the probability density function of the test statistic are described in Section 7.3. At last, the CL_s method is described in Section 7.4, which is utilized for deriving exclusion limits.

7.1 p-value and significance Z

The decision to reject a hypothesis depends on the size of the so-called p-value of a hypothesis H_μ for an observation in data, where the subscript μ defines which hypothesis is tested. A distribution t_μ is defined, which can separate between different hypotheses H_μ and in which each of the possible outcomes of a measurement is mapped to a single value. The smaller the p-value, the larger is the incompatibility between measured data and a hypothesis H_μ. The one-sided p-value is defined as

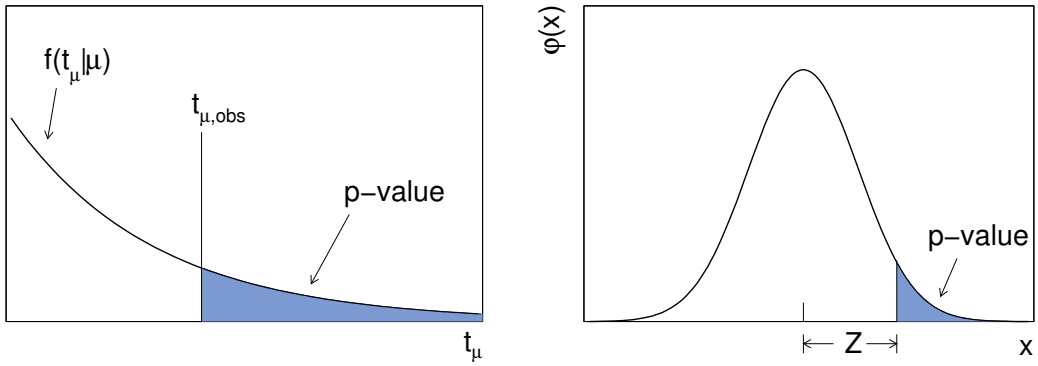
$$p_\mu = \int_{t_{\mu,\text{obs}}}^{\infty} f(t_\mu|\mu)\mathrm{d}t_\mu \ . \tag{7.1}$$

It is evaluated for the observed value $t_{\mu,\text{obs}}$. The function $f(t_\mu|\mu)$ with the signal strength μ to be tested in data and the property $\int f(t_\mu|\mu)\mathrm{d}t_\mu = 1$ describes the probability density function (pdf) of the distribution t_μ. In Figure 7.1(a), the p-value for an hypothesis H_μ at the observation $t_{\mu,\text{obs}}$ is visualized in addition to the depiction of the pdf $f(t_\mu|\mu)$.

The p-value is closely related to the significance Z by

$$Z = \phi^{-1}(1 - p) \tag{7.2}$$

with the quantile function of the standard normal distribution ϕ^{-1}. In Figure 7.1(b), a standard Gaussian distribution $\varphi(x)$ is illustrated as pdf. The significance Z then states the relative location of this observation with respect to the mean of the distribution, also giving insight into the likeliness of a measurement. In particle physics, a significance of 5σ is the convention for claiming a discovery, which corresponds to the rejection of the B-only hypothesis. In a one-sided p-value calculation, this corresponds to $p = 2.87 \cdot 10^{-7}$. This means: If the B-only hypothesis was true, the probability to measure data that are at least as incompatible with this hypothesis as observed is $p = 2.87 \cdot 10^{-7}$, which is extremely unlikely.



(a) (b)

Figure 7.1: Illustration of a) the p-value calculation for hypothesis H_μ at the observation $t_{\mu,\text{obs}}$ using the test statistic t_μ with the pdf $f(t_\mu|\mu)$. In b), the relation between the significance Z and the p-value for a standard Gaussian pdf $\varphi(x)$ [143] is shown.

7.2 Likelihood ratios for hypothesis tests

During the construction of an analysis, variables are searched for which discriminate well between signal and background in order to identify regions with an increased ratio of signal and background events. At the end, a final discriminant is chosen according to the highest discrimination power. There is an abundance of possibilities, like kinematic variables, invariant masses or output values of multivariate techniques. The final discriminant is used to construct a *test statistic*, which preserves discrimination power between the S+B and B-only hypothesis and reduces the dimensionality from a one-dimensional distribution to a single number. According to the Neyman-Pearson lemma [144], the most powerful test statistic t_μ at the *level of significance* α is the likelihood ratio

$$t_\mu(\boldsymbol{x}) = \frac{L_\mu(\boldsymbol{x})}{L_{\mu'}(\boldsymbol{x})}\,, \tag{7.3}$$

with the likelihood functions L for the hypotheses H_μ and $H_{\mu'}$. The likelihood functions are evaluated at the observation vector $\boldsymbol{x}$, which contains all data points of a measurement. The level of

significance α describes the size of the error, assigned to the rejection of H_μ, although it is true. Furthermore, the power of the test β, defined as accepting H_μ although $H_{\mu'}$ is true, is minimized for a given value of α due to the use of the likelihood ratio.

The simplified definition of the likelihood ratio in Eq. 7.3 is only dependent on one variable, which is the parameter of interest μ. Since an analysis incorporates a large number of systematic uncertainties, related to experimental and theoretical sources, a so-called *profile likelihood ratio* (PLR) is utilized. It includes all systematic uncertainties as *nuisance parameters* (NPs) and evaluates those during the fitting procedure, where the numerator and denominator are maximized separately. In order to obtain the best guess for each NP, additional information in form of estimated constraints or auxiliary measurements is provided to the fit. Sources of systematic uncertainties are described in Sections 8.5 and 9.4. The PLR $\lambda(\mu)$ is defined as

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}\,. \tag{7.4}$$

The parameter of interest is the signal strength μ, which specifies the amount of signal S in addition to the SM background expectation B. The total number of expected events n including a signal hypothesis is defined as $n = \mu S + B$. The denominator $L(\hat{\mu}, \hat{\boldsymbol{\theta}})$ is called the *unconditional maximum likelihood estimator*. The parameter space is adjusted to maximize the likelihood, and the best fit values $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ for the vector of NPs $\boldsymbol{\theta}$, as well as for μ are derived. For the *conditional maximum likelihood estimator* $L(\mu, \hat{\hat{\boldsymbol{\theta}}})$, the signal strength μ according to the size of a hypothetical signal is given as an input to the fit, and the NPs are maximized depending on the choice of μ.

With this PLR, the B-only hypothesis is defined as $\mu = 0$, while any choice of $\mu > 0$ is denoted as an S+B hypothesis.

A comprehensive likelihood function not only includes the systematic uncertainties as NPs, but also provides a handle on the limited number of simulated events. As discussed in Chapter 4, the full simulation of a Monte Carlo (MC) process is computationally intensive, and probes of the SM at an experiment's sensitivity boundary to new physics can also lead to a reduced number of MC events. The likelihood function utilized in the analyses presented in Chapters 8 and 9 is taken from Ref. [145] and is defined as follows for a distribution with N bins:

$$L = \prod_{i=1}^{N} \text{Pois}(n_i|\mu s_i(\boldsymbol{\theta}) + \gamma_i b_i(\boldsymbol{\theta})) \cdot \text{Pois}(m_i|\gamma_i\tau_i) \cdot \prod_{k=1}^{M} G(\theta_k|\theta_k^0, \Delta\theta_k) \tag{7.5}$$

The first Poisson term is the heart of the likelihood and evaluates the probability to observe n_i events in bin i for an expectation of $\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})$ events. Additionally, a so-called γ-factor of the order of one is introduced in each bin, which treats the statistical uncertainty of the limited number of MC events and is multiplied to the background expectation $b_i(\boldsymbol{\theta})$. The γ-factors in each bin scale the background expectation and are derived from the auxiliary measurement $\text{Pois}(m_i|\gamma_i\tau_i)$. This auxiliary measurement is a Poisson constraint term, where the number of unweighted MC events m_i of each bin, calculated according to $m_i = (b_i/\delta_i)^2$, follows a Poisson distribution with variance $\gamma_i\tau_i$. Here, δ_i is the total statistical uncertainty in bin i. For the following purposes, τ_i is always fixed to m_i, since no fluctuation of the mean is included and thus the standard deviation of the mean is zero. More information is given in the HISTFACTORY documentation in Ref. [145].

Furthermore, Gaussian constraint terms for each of the M systematic uncertainties are included in Eq. 7.5, which carry information about the systematic uncertainties. For experimental sources, these NPs are obtained by an auxiliary measurement, and the measured uncertainty is then added as the ± 1 standard deviation σ of the Gaussian constraint. In case of systematic uncertainties from theoretical sources, additional information from alternative simulation approaches are added in the same fashion. The mean value is described with θ_k^0, derived from an auxiliary measurement, and the standard deviation of the Gaussian distribution is fixed to $\Delta\theta_k$. There is one NP for each systematic uncertainty which is adjusted by the fit, and the modified mean and standard deviation are extrapolated to all N bins.

The likelihood function in Eq. 7.5 has been defined in one so-called *region* or *category* with N bins. However, the setup of an analysis may be more complex. Multiple phase-space regions may be defined and analyzed for different purposes. For example, the use of control regions is an important concept. By including regions of phase-space, which are rich in events, orthogonal to the search region and still close so the signal kinematics, auxiliary measurements of NPs and γ-factors are performed. A combined likelihood in all regions is created by multiplication of the individual likelihoods for each region. Thus, a combined fit is performed across all regions. Furthermore, a correlation scheme of NPs is chosen, which correlates similar effects across regions and decorrelates phase-space and background-process specific NPs. The correlation of NPs may also lead to further constraints on systematic uncertainties.

A test statistic q_μ is built from the PLR:

$$q_\mu = -2 \ln \lambda(\mu) \tag{7.6}$$

The test statistic only depends on the parameter of interest μ. Only models with $\mu \geq 0$ are considered in this thesis, since the contribution from a VLQ signal cannot diminish the number of background events. Therefore, the test statistic is modified to exclude the parameter space with $\mu < 0$. When testing for a discovery, the test statistic q_0 for positive signal from Ref. [143] is utilized:

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \text{for } \hat{\mu} \geq 0 \\ 0, & \text{for } \hat{\mu} < 0 \end{cases} \tag{7.7}$$

For the calculation of exclusion limits, the test statistic $\tilde{q}_\mu$, also taken from Ref. [143], is used with a modified PLR $\tilde{\lambda}$, incorporating positive signal contributions only:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(0), & \text{for } \hat{\mu} \leq \mu \\ 0, & \text{for } \hat{\mu} > \mu \end{cases} = \begin{cases} -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0, \hat{\hat{\boldsymbol{\theta}}}(0))}, & \text{for } \hat{\mu} < 0 \\ -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}, & \text{for } 0 \leq \hat{\mu} \leq \mu \\ 0, & \text{for } \hat{\mu} > \mu \end{cases} \tag{7.8}$$

It can be seen that for the case $\hat{\mu} < 0$, derived from the unconditional maximum likelihood estimator, $\hat{\mu}$ is replaced with zero, corresponding to background contributions only. This assumption is useful, since the maximum agreement of data and any hypothesis with non-negative μ is reached for $\mu = 0$, if there is no positive fitted signal yield.

In the course of this thesis, the so-called *Asimov* dataset is often used for certain analysis checks. The Asimov dataset corresponds to the B-only expectation with statistical uncertainties matching the Poisson uncertainties of the background expectation in each bin.

After the definition of different test statistics, hypothesis tests for discovery or exclusion limits can be performed. However, the derivation of exclusion limits, using the test statistic $\tilde{q}_\mu$, is the focus of the rest of this chapter, since exclusion limits are utilized to a larger extent in Chapters 8 and 9.

Exclusion limits can be calculated with the Asimov dataset, but also with the observed dataset, and their meaning is quite different. In order to specify the sensitivity of an experiment, it is useful to derive an expected sensitivity based on all possible outcomes of a measurement according to the background expectation, and thus only MC simulation in form of the Asimov dataset is used. At first, the pdf of the test statistic $\tilde{q}_\mu$ from Eq. 7.8 with $\mu = 0$ is derived, which is described in more detail in Section 7.3. Furthermore, the pdf of the S+B hypothesis for a choice of μ is calculated. The value of $\tilde{q}_\mu$ corresponding to the median of the pdf $f(\tilde{q}_\mu|0)$ is used to calculate the p-value for the S+B hypothesis, described by $f(\tilde{q}_\mu|\mu)$. The signal strength μ is varied until the p-value of the S+B hypothesis is equal to the chosen level of significance α. Then, the expected limit is reached, which is quoted in the optimization process of an analysis.

After the observed data is included instead of the MC expectation, the observed limit on μ is calculated following Eq. 7.1. Instead of the median of the distribution $f(\tilde{q}_\mu|0)$, the observed value $\tilde{q}_{\mu,\text{obs}}$ is used and the μ of the S+B hypothesis is varied until the level of significance α is reached. This is the final result of the analysis and constrains the phase-space for new physics.

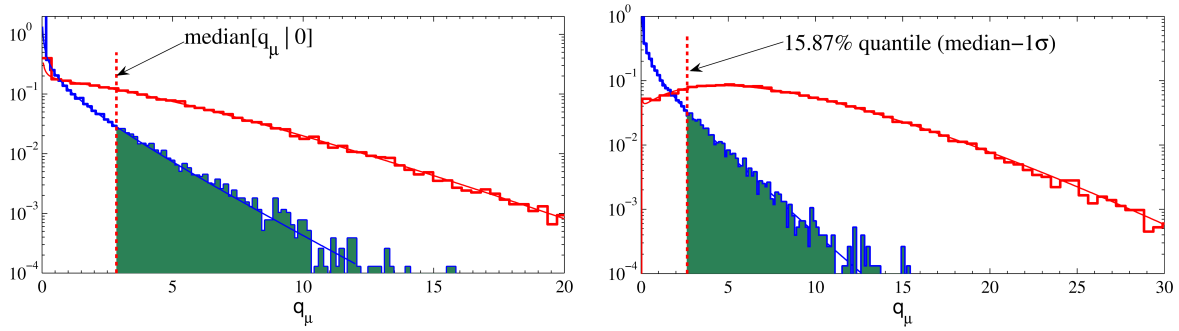


Figure 7.2: Probability density functions (pdfs) $f(q_\mu|0)$ (red) and $f(q_\mu|\mu)$ (blue) for the B-only and S+B hypotheses, respectively [143]. The p-value calculation of the S+B hypothesis for a) the median (expected) 95% CL limit and b) the -1σ fluctuation of this limit are shown. Furthermore, the estimation of the pdfs with the asymptotic approximation is shown as a solid line, which is overlaid with a histogram obtained from pseudo-experiments.

The derivation of the expected limit is visualized in Figure 7.2(a). The pdfs $f(q_\mu|0)$ and $f(q_\mu|\mu)$ are depicted in red and blue, respectively. Instead of $\tilde{q}_\mu$, a different test statistic q_μ from Ref. [143] is used, which is equal to $\tilde{q}_\mu$ in the approximation to be discussed in Section 7.3. The pdf $f(q_\mu|\mu)$ corresponds to the S+B hypothesis with μ larger than zero. In order to obtain the expected sensitivity, the median of $f(q_\mu|0)$ is calculated and the p-value of the S+B hypothesis is derived at this median, illustrated as green area in Figure 7.2. In this example, the pdf $f(q_\mu|\mu)$ is chosen such that the p-value is 0.05, which corresponds to a significance of $\alpha = 5\%$ and a $1 - \alpha = 95\%$ CL.

Insight about the effect of statistical fluctuations in data can be gained by not only quoting the expected (median) significance, but also a $\pm 1\sigma$ fluctuation of $\hat{\mu}$ for $f(q_\mu|0)$. In Figure 7.2(b), a -1σ fluctuation is illustrated, which corresponds to the 15.87% quantile of $f(q_\mu|0)$. The same procedure can also be applied for a $+1\sigma$ fluctuation, when the 84.13% quantile is taken instead. The p-value is derived for the S+B hypothesis with the pdf $f(q_\mu|\mu)$, and the expected sensitivity can now be quoted with statistical uncertainty bands. In the following, the analysis results always show ± 1 and ± 2 uncertainty bands around the expected limit at 95% CL, which then include statistical and systematic uncertainties.

7.3 Pseudo-experiments and asymptotic approximation

In order to derive a p-value for a specific hypothesis, as introduced in Eq. 7.1, the pdf of the test statistic is required. For the calculation of expected and observed limits, the pdfs $f(\tilde{q}_\mu|\mu)$ and $f(\tilde{q}_\mu|0)$ of the test statistic in Eq. 7.8 need to be derived for various hypotheses of μ. These pdfs can be obtained by a large number of pseudo-experiments, which randomize the inputs of the likelihood function. For each pseudo-experiment, values from the pdf of the final discriminant are drawn in each bin and one value of a test statistic is obtained. A large number of these pseudo-experiments is performed for the B-only hypothesis with $\mu = 0$, but also for all hypothesized values of μ of the S+B hypothesis. Furthermore, the mean values of the Gaussian constraint terms in the likelihood, presented in Eq. 7.5, are varied during this procedure. By generating a multitude of pseudo-experiments (e.g. 10^5), the pdfs $f(\tilde{q}_\mu|\mu)$ and $f(\tilde{q}_\mu|0)$ of both hypotheses are sampled. Subsequently, it is possible to derive the median of $f(\tilde{q}_\mu|0)$ due to the availability of the pdf of the B-only hypothesis, as already described in Section 7.2. An interpolation between different hypotheses for μ allows an identification of the upper limit on μ, which is found for the equality of the p-value and a level of significance α. Following the same procedure for the calculation of the p-value and the upper limit on μ, the observed limit can be calculated for the observation $\tilde{q}_{\mu,\text{obs}}$.

Instead of generating pseudo-experiments, the pdfs can be calculated with *asymptotic approximations* [143], based on Refs. [146] and [147]. These asymptotic formulae for different test statistics reflect the exact pdfs in the large sample limit and give good approximations for small sample sizes down to a few events. In order to model the tails of the test statistic's pdf with good precision up to 2–3 σ a minimum of 2–5 events is sufficient [146]. The test statistic $\tilde{q}_\mu$ for exclusion limits is approximated to

$$\tilde{q}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2}, & \hat{\mu} < 0 \\ \frac{(\mu-\hat{\mu})^2}{\sigma^2}, & 0 \leq \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu\,, \end{cases} \tag{7.9}$$

where σ is the standard deviation of a Gaussian distribution $\hat{\mu}$ with mean μ'. The corresponding approximated pdf $f(\tilde{q}_\mu|\mu)$ can be found in Ref. [143]. All important quantities like the significance Z and the exclusion limit for μ can be calculated using $\tilde{q}_\mu$ and $f(\tilde{q}_\mu|\mu)$. A comparison of the pdfs $f(q_\mu|\mu)$ and $f(q_\mu|0)$ derived with pseudo-experiments and the asymptotic approximation is illustrated in Figure 7.2. Pseudo-experiments are shown as histograms, while the approximation is drawn as a continuous function, and no significant differences between them can be seen. As already mentioned

in Section 7.2, the asymptotic approximations for the pdfs of the test statistic q_μ, shown in Figure 7.2, and $\tilde{q}_\mu$, used in the following limit calculations, yield identical exclusion limits in the large sample limit. Thus, this conclusion drawn from Figure 7.2 also holds for $\tilde{q}_\mu$.

7.4 CL_s method

As discussed in the previous sections, the criterion for rejecting a specific hypothesis is the p-value in combination with the level of significance α. Therefore, a simple way to test for discovery or set exclusion limits has been shown. Nevertheless, the calculation of exclusion limits at a specific CL can be improved compared to a simple p-value calculation for the S+B hypothesis according to Eq. 7.1. Thus, the following section focuses on an improved recipe for the derivation of one-sided limits.

The main focus of a search for new physics, which did not find a significant signal contribution, is the exclusion of the S+B hypothesis. Since background can never be fully suppressed, a signal-only hypothesis cannot be tested. However, this is exactly what is desired: A statement about the presence of signal and no convoluted conclusion about an S+B hypothesis. If the signal contribution compared to background is very small, the pdfs of the B-only and S+B hypotheses can overlap. In Figure 7.3, an example of two largely overlapping pdfs for a test statistic q is illustrated with the corresponding p-values for the S+B and B-only hypotheses, calculated as follows:

$$p_{s+b} = P(q \geq q_{\text{obs}}|s+b) = \int_{q_{\text{obs}}}^{\infty} f(q|s+b)\mathrm{d}q \tag{7.10}$$

$$p_b = P(q \leq q_{\text{obs}}|b) = \int_{-\infty}^{q_{\text{obs}}} f(q|b)\mathrm{d}q \tag{7.11}$$

For the derivation of exclusion limits in this thesis, the pdfs $f(q|s+b)$ and $f(q|b)$ are replaced with the pdfs $f(\tilde{q}_\mu|\mu)$ and $f(\tilde{q}_\mu|0)$, corresponding to the S+B and B-only hypothesis, respectively. With these p-value definitions, a small value for p_{s+b} corresponds to small compatibility between the observed

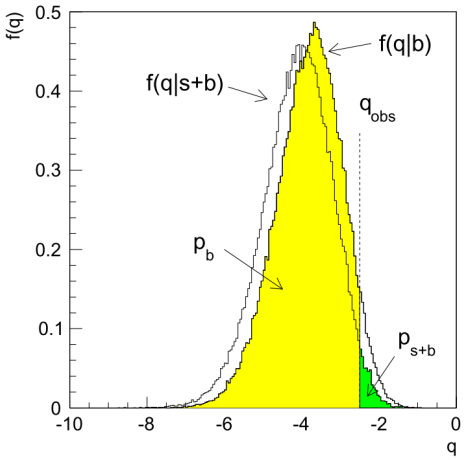


Figure 7.3: Largely overlapping probability density functions (pdfs) $f(q|\text{s+b})$ and $f(q|\text{b})$ for the S+B and B-only hypotheses, respectively [148]. The corresponding p-value calculation $p_{\text{s+b}}$ and p_{b} are shown for q_{obs}.

data and the S+B hypothesis. The same statement about incompatibility of data and B-only hypothesis holds for small values of $1 - p_b$. In case of negligible signal contribution with largely overlapping pdfs, it becomes obvious that the exclusion of the S+B hypothesis, which is very similar to the B-only hypothesis, rather makes a statement about the background than the signal. A downward fluctuation of q_{obs} in a search with very small signals could cause the scenario in Figure 7.3. In this case, the validity of the background model could be questioned, but this should not influence the conclusions about the presence of a new signal. To summarize, a hypothesis test should not exclude a signal contribution, which it is not sensitive to and should only make statements about the signal hypothesis.

The following technique is chosen to circumvent this problem and is denoted as the CL_s method. It is defined as

$$CL_s = \frac{p_{s+b}}{1 - p_b} < \alpha \, . \tag{7.12}$$

The upper limit for the parameter of interest μ is obtained, when the value of CL_s is smaller than the level of significance α, typically chosen to be 5%, which is then quoted as a $1 - \alpha = 95\%$ CL limit. In contrast to the calculation of limits using $p_{s+b} < \alpha$, as introduced in Section 7.2, exact coverage of the CL derived with the CL_s method is not given. The CL_s method is always more conservative: In case of well separated pdfs for S+B and B-only hypotheses, the observed p_b is zero and thus the denominator equals unity. For increasing overlap of the pdfs, p_b rises and $1 - p_b$ becomes smaller. Subsequently, the CL_s value rises and $\alpha = 5\%$ is excluded for a larger signal strength as it would be for the calculation of p_{s+b} only, which corresponds to a weaker limit. In the following analyses, the test statistic $\tilde{q}_\mu$ together with the asymptotic approximation (where applicable) is used for the limit derivation with the CL_s method. The software package TRExFitter is used for this purpose. It is based on ROOT [149] and RooFit [150, 151] and widely used within the ATLAS collaboration.

Chapter 8
Search for vector-like bottom and top quarks with 36.1 fb^{-1}

In this chapter, a search for vector-like bottom (VLB) and vector-like top (VLT) quarks is presented. Data taken in the years 2015 and 2016 during Run-2 at a center-of-mass energy of 13 TeV with an integrated luminosity of 36.1 fb^{-1} are used in a cut-based analysis, targeting boosted final states in the dilepton channel. The analysis strategy is presented in Section 8.1 and is followed by a signal region (SR) optimization in Section 8.2. Then, two control regions (CRs) for the main backgrounds are defined in Section 8.3. The signal and background modelling is discussed in Section 8.4. Subsequently, the systematic uncertainties that are relevant for this analysis are introduced in Section 8.5, and the data and Monte Carlo (MC) simulation agreement is investigated in Section 8.6. The model of a B-only fit is studied in Section 8.7. As a next step, hypothesis tests are performed for VLB and VLT quarks for discovery and exclusion limits in Section 8.8. These results are published in Ref. [3] together with a combination of two other channels targeting the Z-boson corner in the $Z(\ell\ell)t/b + X$ search. This search is presented in Section 8.9, and a more inclusive combination of all pair production vector-like quark (VLQ) searches by the ATLAS collaboration is presented in Section 8.10. An overview of other VLQ searches, e.g. performed by the CMS collaboration or targeting single production with the same dataset are given in Section 8.11.

Previous studies of the dilepton channel of the $Z(\ell\ell)t/b + X$ search including the derivation of exclusion limits with statistical uncertainties only have been performed in a closely related master thesis [2].

8.1 Analysis strategy for the $2\ell \geq 2J$ channel in the Z-boson corner

As presented in Section 2.3, vector-like quarks, if existent, decay into a Z-, W- or H-boson, and their corresponding branching ratios (BRs) are not known. The following analysis targets the pair production of VLB and VLT quarks. The search region is composed of final states with one reconstructed Z-boson, which decays into two leptons (electrons or muons) and is labeled as Z-boson candidate in the following. The charge of the leptons has to be of opposite sign and the leptons carry the same flavor (OSSF), which means that only events with e^+e^- and $\mu^+\mu^-$ final states are accepted. An inclusive decay of the second VLQ is allowed, and thus three boson combinations in the final state are possible: ZZ, ZW and ZH. This topology is associated with the Z-boson corner in the BR plane, since the sensitivity for Z-boson final states is especially high in comparison to final states with W- or H-bosons.

The topology is illustrated in Figure 8.1. VLQs are expected to be very massive, reaching and - later to be shown - exceeding the TeV scale [4]. This means that they are produced at low transverse momentum, p_{T}, while their decay products are highly boosted due to the large difference in mass of the VLQ and its decay products. The maximum sum of masses from decay products can yield $m_t + m_H \approx 300\,\mathrm{GeV}$. A benchmark mass of $900\,\mathrm{GeV}$ is chosen for VLB and VLT quarks in this search, since this mass is close to the exclusion limits [21, 23] at the time of the analysis development. This means that there is a surplus of $600\,\mathrm{GeV}$ in momentum, which is evenly distributed among the decay products. Thus, a Z-boson candidate would carry high p_{T}, denoted as $p_{\mathrm{T}}\,(Z)$ in the following. Furthermore, the creation of high-p_{T} jets leads to a high sum of all scalar transverse jet momenta, denoted as $H_{\mathrm{T}}(\mathrm{jets})$ in the following. Signal events are also expected to have at least one or two b-tagged jets. Another property of a VLQ decay is a non-zero number of large-R jets due to the boosted decay products.

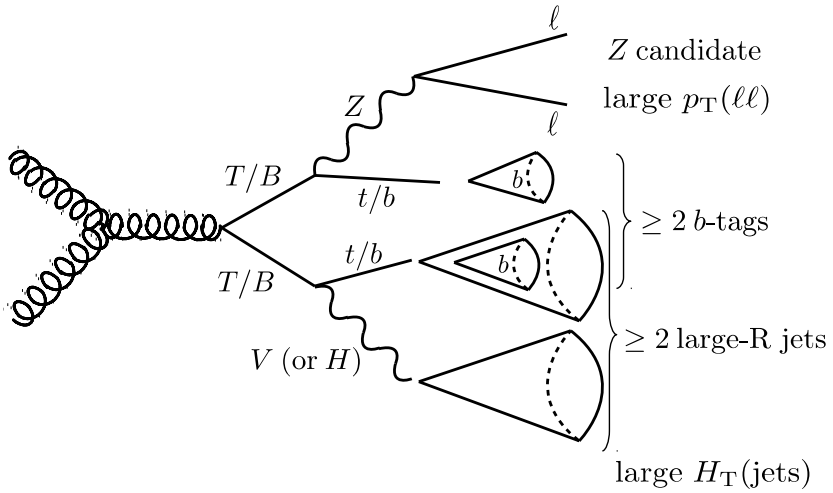


Figure 8.1: Illustration of the pair production of VLT and VLB quarks via gluon fusion. The requirement of one Z-boson candidate reconstructed from an OSSF lepton pair $\ell\ell$ is shown. Possible decays into W-, Z- or H-bosons are depicted. Also, the reconstructed physics objects are indicated with important features.

The main approach of an analysis searching for new physics is the separation of signal and background events, as explained in Chapter 4. The largest background is Z + jets, but significant contributions also come from $t\bar{t}$ and $t\bar{t} + X$ events. If the Z-boson of the Z + jets background decays leptonically and recoils against a number of jets, the final state might look like a VLQ decay. Especially, if the Z-boson is produced in association with a pair of b-tagged jets, it becomes harder to separate signal and background. Furthermore, $t\bar{t}$ events can be misinterpreted as signal if both top-quarks decay leptonically, meaning that the W-boson decays into $W \rightarrow \ell\nu$. This decay is almost always accompanied by the creation of a b-quark, bringing it closer to the topology of the search. A similar statement holds for $t\bar{t} + X$ events, where $t\bar{t}$ events are produced along with a V-boson. The cross-section of this process is a few hundred times smaller [152, 153] compared to $t\bar{t}$, but the final states are more similar to the VLQ decays: Two leptons emerge from a V-boson decay, and both top-quarks can decay hadronically, resulting in more hadronic activity in the detector. More details about what is considered as X is given in Section 8.4.

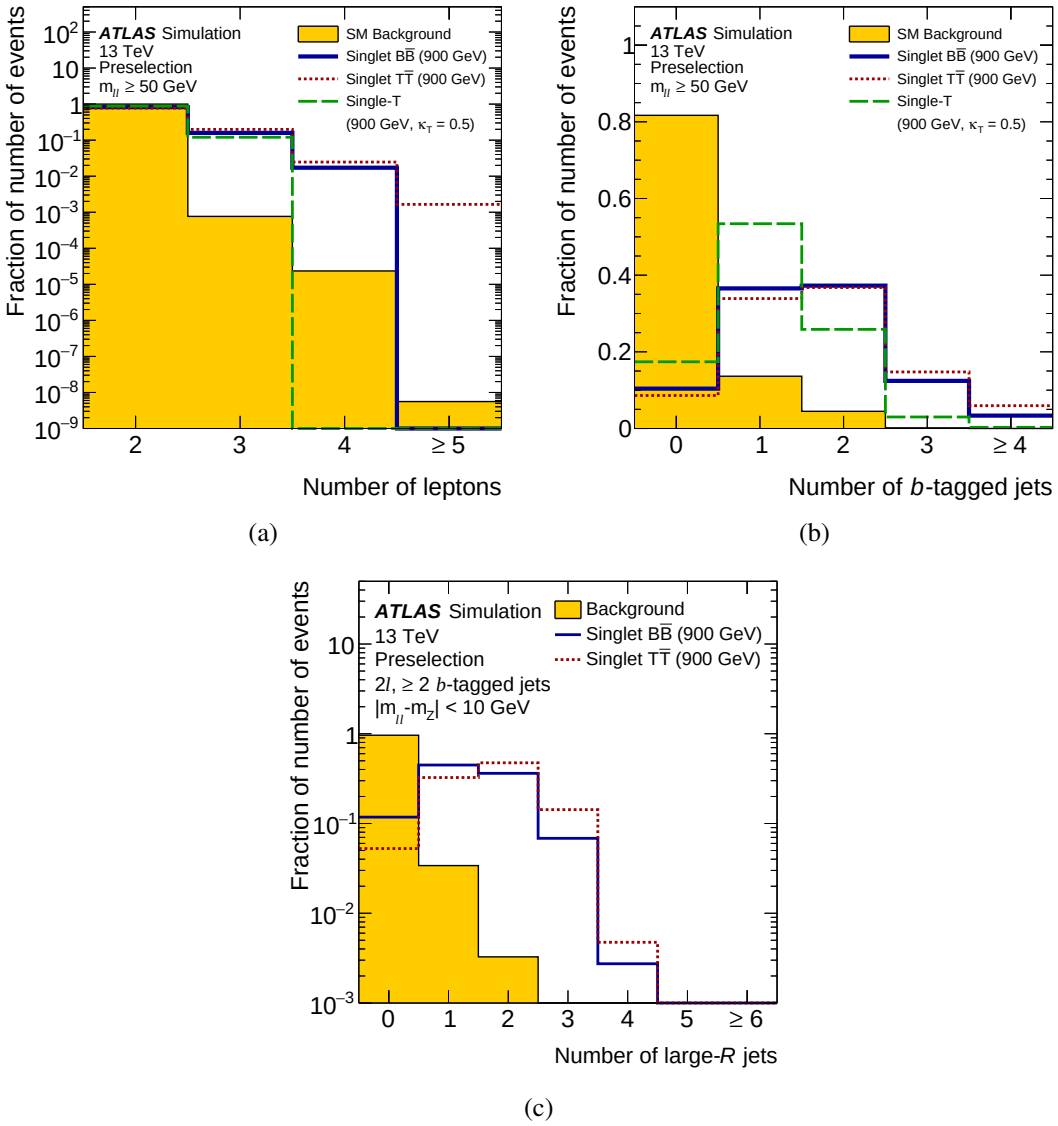


Figure 8.2: Shape comparison for signal and background events for the a) lepton and b) b-tagged jet multiplicity (for 77% signal efficiency) at preselection level and requiring $m_{\ell\ell} \geq 50\,\text{GeV}$ [3]. Furthermore, the c) large-R jet multiplicity is illustrated after the requirement of at least two b-tagged jets, exactly two OSSF leptons and a Z-boson candidate mass $m_{\ell\ell}$ with $|m_{\ell\ell} - m_Z| < 10\,\text{GeV}$ [154]. Both benchmark signal masses of the pair production at 900 GeV are shown, as well as one VLT quark single-production mass hypothesis for a) and b). Furthermore, the sum of all backgrounds is shown. Signal and background processes are normalized to unity. The last bin contains all overflow events.

A closer look at the distribution of object multiplicities is used to identify their discrimination power between signal and background events. In Figure 8.2, the lepton, b-tagged jet and large-R jet multiplicities are depicted. The histogram with background events contains the main backgrounds discussed previously, as well as the diboson (VV), triboson (VVV), W+jets, single top processes and 4-top production. Besides, both pair production signal mass points at 900 GeV are shown for the singlet BRs in addition to one signal hypothesis for single production of VLT quarks, which is not

investigated further in the scope of this thesis. All histograms are normalized to unity in order to only compare the shapes, because the VLQ proportion is very small at this inclusive stage of the selection. The preselection criteria included in all figures, are listed in Table 8.1. Figure 8.2(a) motivates the analysis strategy with a splitting of channels according to the lepton multiplicity, which are analyzed separately: There is a minimum of two leptons for the Z-boson candidate. Although most of the background events are located in the two-lepton bin, it has the advantage of high signal efficiency. For three or more leptons, a large fraction of background is suppressed, but also the amount of signal drops significantly. Therefore, the analysis strategy is split based on the lepton multiplicity into a dilepton and trilepton channel[1].

Table 8.1: Definitions of preselection and primary selection for the $2\ell \geq 2J$ channel. All cuts are applied in a sequence, so that the primary selection includes the preselection cuts.

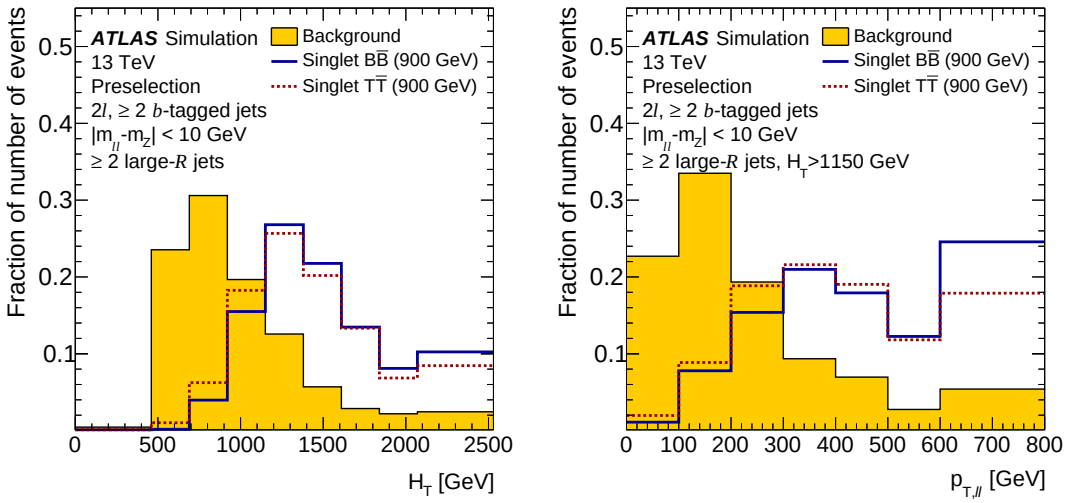
Preselection
≥ 2 central small-R jets
≥ 2 OSSF leptons (e/μ)
$m_{\ell\ell} < 400\,\text{GeV}$
Primary selection
≥ 2 b-tagged jets
exactly 2 leptons
≥ 2 large-R jets ($m > 50\,\text{GeV}$, $p_\text{T} > 200\,\text{GeV}$, $\lvert\eta\rvert < 2.0$)

Figure 8.2(b) provides information about the b-tagged jet multiplicity for the 77% signal efficiency working point. It can be seen that most background events show no b-tagged jets, while for signal the multiplicity is higher. For VLB and VLT quarks, over 50% of events have at least two b-tagged jets, reducing the amount of background events by around 95%. Therefore, a cut of at least two b-tagged jets is applied for the dilepton channel.

In Figure 8.2(c), the large-R jet multiplicity is illustrated, including cuts on at least two b-tagged jets and exactly two OSSF leptons. Furthermore, only events with a mass of the Z-boson candidate close to the Z-boson mass are selected with a requirement of $|m_{\ell\ell} - m_Z| < 10\,\text{GeV}$, since this VLQ search targets at least one on-shell Z-boson candidate. It can be seen that most of the background events have no large-R jets, which are only calibrated for a minimum p_T of $200\,\text{GeV}$ and a minimum mass of $50\,\text{GeV}$ in the detector region $|\eta| < 2.0$. For signal, events with at least one large-R jet represent the majority. Going to at least two large-R jets, the fraction of background events is highly suppressed, while preserving a significant amount of signal. Therefore, the dilepton channel is split in two more channels: The resolved channel[2] consists of events with zero or one large-R jet and is abbreviated as 2ℓ 0–1J channel, where J stands for large-R jet. The boosted channel is called $2\ell \geq 2J$ and is the focus of this chapter.

[1] The trilepton channel was analyzed within the HQT subgroup of the ATLAS collaboration.
[2] The dilepton resolved channel was analyzed within the HQT subgroup of the ATLAS collaboration.

After highlighting the benefit of some additional requirements in this section, a so-called primary selection is defined on top of the preselection. It is summarized in Table 8.1, and the requirements of at least two b-tagged jets, exactly two OSSF leptons and at least two large-R jets are added for the selection of events in the $2\ell \geq 2J$ channel. The Z-boson mass window cut $|m_{\ell\ell} - m_Z| < 10\,\text{GeV}$, also called *on-Z*, is not included in these selections, since it is later adjusted for a CR definition. Nevertheless, this cut is added for the following studies, targeting the construction of a SR selection. It reduces the contribution from $t\bar{t}$ events significantly, since the invariant mass of two leptons originating from leptonic top-quark decays is not correlated with the Z-boson mass.



(a) (b)

Figure 8.3: Shape comparison [154] for signal and background events for a) $H_{\mathrm{T}}(\mathrm{jets})$ at primary selection level and on-Z. Furthermore, the distribution of b) $p_{\mathrm{T}}\,(Z)$ is illustrated after an additional requirement of $H_{\mathrm{T}}(\mathrm{jets}) > 1150\,\text{GeV}$. Both benchmark signal masses of the pair production at $900\,\text{GeV}$ are shown. Furthermore, the sum of all backgrounds is shown. Signal and background processes are normalized to unity. The last bin contains all overflow events.

As mentioned earlier, the signal topology shows kinematic distinctiveness in form of large $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$. In Figure 8.3(a), $H_{\mathrm{T}}(\mathrm{jets})$ is illustrated in a normalized histogram for signal, and the sum of all background processes after the primary selection and the on-Z criterion are applied. It can be seen that both benchmark signal sample processes are characterized by very high $H_{\mathrm{T}}(\mathrm{jets})$, exceeding $H_{\mathrm{T}}(\mathrm{jets})$ of background processes significantly. Due to the good discrimination power between signal and background, a cut on $H_{\mathrm{T}}(\mathrm{jets})$ is beneficial.

In Figure 8.3(b), $p_{\mathrm{T}}\,(Z)$ is shown after adding a lower cut of $1150\,\text{GeV}$ on $H_{\mathrm{T}}(\mathrm{jets})$. The choice of this cut value is justified in Section 8.2. The variable $p_{\mathrm{T}}\,(Z)$ has also great separation power between signal and background and is chosen as selection requirement in this analysis. It should be noted that the p_{T} contributions of $H_{\mathrm{T}}(\mathrm{jets})$ and $p_{\mathrm{T}}\,(Z)$ are of different source and thus give complementary information: While $H_{\mathrm{T}}(\mathrm{jets})$ is sensitive to hadronic activity in the event, $p_{\mathrm{T}}\,(Z)$ represents the leptonic activity. All objects except the Z-boson candidate solely contribute to $H_{\mathrm{T}}(\mathrm{jets})$ in the dilepton channel.

8.2 Signal region optimization

In Section 8.1, it has been shown that cuts on $H_{\mathrm{T}}(\mathrm{jets})$ and $p_{\mathrm{T}}(Z)$ have the power to reduce the background contribution. Furthermore, a final variable with great separation power between signal and background is needed for hypothesis tests. As final discriminant, the invariant mass of the Z-boson candidate and the b-tagged jet with highest p_{T} is chosen, denoted as $m(Zb)$. Especially for VLB quarks, this variable is able to reconstruct the VLQ resonance, if the b-tagged jet with highest p_{T} belongs to the same VLB quark decay as the Z-boson candidate. In Figure 8.4(a), the distribution of $m(Zb)$ is shown at primary selection level and on-Z. It can be seen that the final discriminant separates background and signal well. For VLB quarks, a peak at the benchmark mass of 900 GeV is seen. For VLT quarks, $m(Zb)$ shows also good separation power, although the resonance is only partially reconstructed: The W-boson from the hadronic top-quark decay of $T \to Zt \to Z(\ell\ell)Wb$ is not captured in the invariant mass and thus yields smaller $m(Zb)$ values. Nevertheless, $m(Zb)$ has shown to perform well for VLB and VLT quarks. The use of the invariant mass, $m(Zt)$, with a top-tagged large-R jet has been tested and was found to perform worse [2].

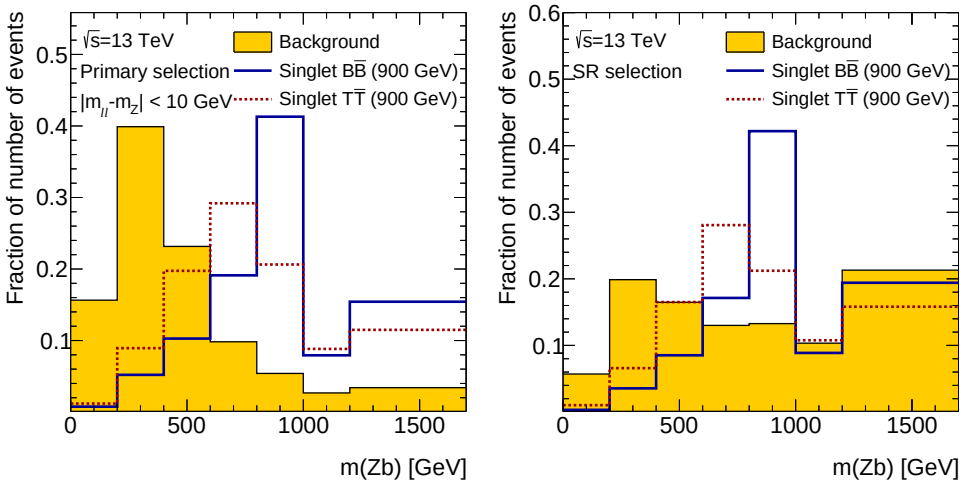


Figure 8.4: Final discriminant $m(Zb)$ at a) primary selection level and on-Z, and b) after the full SR selection. Both benchmark signal masses of the pair production at 900 GeV are shown. Furthermore, the sum of all backgrounds is shown. Signal and background processes are normalized to unity. The overflow bin is included.

In order to increase the sensitivity of the analysis, an optimization of the magnitude of the lower bounds on $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ is performed, based on the expected exclusion limits for VLB and VLT quarks. Both cuts are varied simultaneously and the highest mass limit, derived with statistical uncertainties only, is taken as the best cut combination. Tables 8.2 and 8.3 show the scan of different cuts for VLT and VLB quarks, respectively. The highest mass limits can be reached with a cut combination of $H_{\mathrm{T}}(\mathrm{jets}) > 1150$ GeV and $p_{\mathrm{T}}(Z) > 250$ GeV for both signals. These cuts in addition to the on-Z cut are added to the SR selection, which is denoted as the full SR selection.

In Figure 8.4(b), $m(Zb)$ is depicted after the full SR selection. Going to a more boosted regime including the $H_{\mathrm{T}}(\mathrm{jets})$ and $p_{\mathrm{T}}(Z)$ cuts, the remaining background events are more similar to signal events. The peak structure for VLB and VLT quarks from Figure 8.4(a) is still present and $m(Zb)$ shows good discrimination power to be used in a hypothesis test.

Table 8.2: Optimization of the $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts. The expected mass limits are shown for combinations of their cuts in units of GeV for the singlet benchmark model of the VLT quark at 900 GeV. Entries with "–" are not tested. Only statistical uncertainties are taken into account.

	$H_{\mathrm{T}}(\mathrm{jets}) >$ value [GeV]									
	700	800	850	900	950	1000	1050	1100	1150	1200
$p_{\mathrm{T}}(Z) > 200$ GeV	928	935	939	943	–	–	–	961	969	–
$p_{\mathrm{T}}(Z) > 250$ GeV	937	946	949	954	–	–	976	978	979	977
$p_{\mathrm{T}}(Z) > 300$ GeV	941	950	954	958	963	964	970	973	973	970

Table 8.3: Optimization of the $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts. The expected mass limits are shown for combinations of their cuts in units of GeV for the singlet benchmark model of the VLB quark at 900 GeV. Entries with "–" are not tested. Only statistical uncertainties are taken into account.

	$H_{\mathrm{T}}(\mathrm{jets}) >$ value [GeV]							
	700	800	850	900	1050	1100	1150	1200
$p_{\mathrm{T}}(Z) > 200$ GeV	938	946	948	951	–	966	966	–
$p_{\mathrm{T}}(Z) > 250$ GeV	953	960	962	965	980	983	984	983
$p_{\mathrm{T}}(Z) > 300$ GeV	955	963	966	969	978	982	983	–

8.3 Definition of control regions

The background modelling with respect to data has to be validated, since the results of a hypothesis test highly depend on the normalization and shape of the backgrounds entering the SR. Therefore, CRs for the main backgrounds are defined, which are orthogonal to the SR selection and thus have no overlap in events. The demand on a CR also includes a sufficient phase-space proximity to the SR in order to avoid different kinematic properties across regions. This is particularly important for a good extrapolation of the background normalization and shape to the SR in a binned profile likelihood fit. The largest backgrounds are Z + jets and $t\bar{t}$, and a CR for each of them is defined.

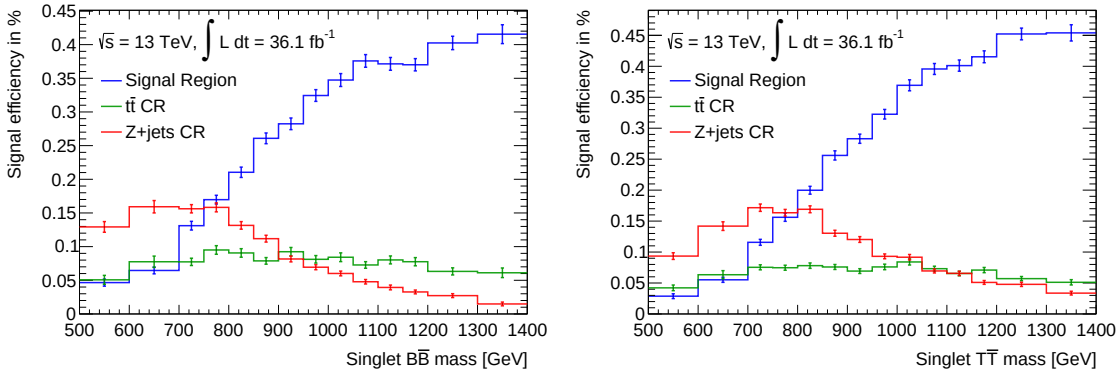
For the Z + jets CR, the Z-boson mass window cut is applied in order to select Z-bosons decaying into leptons, as done in the SR. Additionally, the $H_{\mathrm{T}}(\mathrm{jets})$ cut is inverted, which makes the region orthogonal to the SR. Due to the general requirement of at least two large-R jets, a boosted topology is given and the CR is close to the SR while keeping a high number of background events.

For a pure $t\bar{t}$ CR, the Z-boson mass resonance is excluded: Instead of selecting events with a Z-boson candidate close to m_Z, events outside of the region with $|m_{\ell\ell} - m_Z| < 10$ GeV are selected. The region $m_{\ell\ell} < 50$ GeV is excluded as well, due to large contributions from Drell-Yan processes with additional jet production.

Another important point to consider is the signal contribution in the CRs. Since CRs are used for background modelling, they are unblinded before the SR. Thus, the CRs should not be sensitive to the searched signal and a signal content of $\frac{S}{B} < 10\%$ is pursued for a VLT and VLB quark mass of 900 GeV, assuming singlet BRs, where S and B are the number of signal and background events, respectively.

Table 8.4: Definition of the SR and CRs in the $2\ell \geq 2J$ channel. The preselection and primary selection are defined in Table 8.1.

$t\bar{t}$ CR	Z + jets CR	SR
Preselection Primary selection		
$\lvert m_{\ell\ell} - m_Z \rvert > 10\,\text{GeV}$ and $m_{\ell\ell} > 50\,\text{GeV}$	$\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$	
$E_\text{T}^\text{miss} < 200\,\text{GeV}$	$H_\text{T}(\text{jets}) < 1150\,\text{GeV}$	$H_\text{T}(\text{jets}) > 1150\,\text{GeV}$
$p_\text{T}(Z) < 600\,\text{GeV}$		$p_\text{T}(Z) > 250\,\text{GeV}$
exclude $2.0 < \Delta R(Z, 1^\text{st}\text{ large-R jet}) < 2.8$		



(a) (b)

Figure 8.5: Signal efficiency for a) VLB and b) VLT quarks with the singlet BR assumption in the SR and CRs at all mass hypotheses. Statistical uncertainties on the number of MC events are shown.

Since the signal contribution in the $t\bar{t}$ CR is initially too high, requirements for a signal reduction are introduced. Cuts of $p_\text{T}\,(Z) < 600\,\text{GeV}$ and $E_\text{T}^\text{miss} < 200\,\text{GeV}$ are applied, which aim at differences in the boost of the final state particles for signal and background. Additionally, a requirement for excluding $2.0 < \Delta R(Z, 1^\text{st}\text{ large-R jet}) < 2.8$ is made for the Z-boson candidate and the large-R jet with the largest p_T. Since VLQs are produced at low p_T, their decay products, Z and t/b, are distributed rather back-to-back. Therefore, events in that region are removed and at the same time the amount of $t\bar{t}$ is kept high. The full selection of the SR and CRs is summarized in Table 8.4.

The benchmark VLQ mass point of 900 GeV was used for the optimization of the SR and CRs. Nevertheless, other mass points have other production cross-sections and might show a worse performance for this optimization. In Figure 8.5, the signal efficiencies in the SR and CRs are illustrated for all VLB and VLT quark mass hypotheses with BRs according to the singlet model. The signal efficiency is defined as the signal events passing the region's selection divided by the total number of signal events generated inclusively for all W-, Z- and H-boson decay modes of both pair produced VLQs at 36.1 fb^{-1}. In both cases, it can be seen that the efficiency for the SR continuously

rises from 500 to 1400 GeV, which means that the analysis performs well at the benchmark mass and above, where the VLQ signal might also be located. For the CRs, the signal efficiency decreases for increasing VLQ masses larger than 800 GeV. This fact is reassuring, since it means that for a higher mass hypothesis the CRs are less sensitive than for a VLQ mass of 900 GeV. Accidental unblinding is therefore avoided and the strategy is hence robust for an analysis in the boosted regime.

8.4 Signal and background modelling

The signal and background processes considered in this analysis are modeled with MC simulations. As discussed in Section 8.1, the main backgrounds are Z+jets and $t\bar{t}$, with $t\bar{t}+X$ also playing an important role. Minor backgrounds are VV, single top, W+jets, 4-top production and VVV.

For Z+jets production, SHERPA 2.2.1 [98, 155–157] with the PDF set NNPDF3.0 [158] NNLO is used for the simulation of the hard scattering and the parton shower. The samples are split into their heavy flavor composition.

For $t\bar{t}$ production, the matrix-element simulation is done with POWHEG-BOX v2 [159, 160] using the POWHEG method [161, 162] and the NNPDF3.0 NNLO PDF set. Events with at least one lepton are considered, since the signal topology requires at least two leptons. Because one jet faking an electron is more likely than two, the consideration of simulated events of fully hadronic $t\bar{t}$ events is not necessary. The subsequent parton shower and hadronization processes are simulated with PYTHIA 8 [90] with the A14 set of tunable parameters [141].

The $t\bar{t}+X$ process consists of multiple final states like the production of a $t\bar{t}$ pair and a W- or Z-boson simulated with MADGRAPH5_aMC@NLO [163] and $t\bar{t}+WW$ processes simulated with MADGRAPH 5 [164]. Decays of $Z \to \nu\bar{\nu}$, $q\bar{q}$ and $\ell\ell$ (with $\ell = (e, \mu, \tau)$) are considered. 4-top production ($t\bar{t}t\bar{t}$) is simulated at matrix-element level with MADGRAPH 5, and it is also considered as part of the $t\bar{t}+X$ process. The parton shower is produced using PYTHIA 8 with the A14 tunes.

Furthermore, the simulation of VV events is performed with the generator SHERPA 2.2.1 and SHERPA 2.2.2 using the NNPDF3.0 NNLO PDF set. For VVV, SHERPA is used, but with version 2.1 and the CT10 [165] PDF set. Additionally, the impact of W+jets events, simulated with POWHEG +PYTHIA 6 using the CTEQ6L1 [166] PDF set, was tested in the analysis. However, it was found to be negligible and is thus not mentioned in the following. At last, the hard scattering of the single top process is generated with POWHEG-BOX v1 interfaced to PYTHIA8 with Perugia 2012 [167] tunable parameters. The CT10 PDF set is used. The generator, shower program, PDF set and the order of the cross-section calculation are summarized in Table 8.5.

The signal processes are modeled with the generator PROTOS [66] and the parton shower algorithm PYTHIA 8 with the NNPDF2.3LO PDF. Different mass hypotheses are simulated in a range from 500 GeV to 1400 GeV with a step size of 50 GeV between 700 and 1200 GeV and 100 GeV outside of this range. The singlet VLQ models are taken as simulation basis with BRs of $1/3$ into W-, Z- and H-boson final states and are reweighted to the BRs for the desired multiplet and VLQ mass. The validity of the reweighting is shown in studies of the kinematic differences of the singlet and doublet model, which have shown [168] to be negligible. The VLQ production cross-sections are calculated with TOP++ [169] with NNLO precision with next-to-next-to-leading logarithm soft-gluon contributions and the MSTW 2008 NLO PDF set [170, 171].

Table 8.5: List of background MC samples [3]. Information about the MC generator, the parton shower and its set of tuned parameters is given. Furthermore, the utilized PDF sets in the matrix element (ME) and the order in QCD of the cross-section calculation are given.

	Generator	Shower program and tune	PDF set (ME)	Cross section
Z + jets	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF3.0 NNLO	NNLO
$t\bar{t}$	POWHEG-BOX v2	PYTHIA 8, A14	NNPDF3.0 NNLO	NNLO+NNLL
VV	SHERPA 2.2.1/2	SHERPA 2.2.1	NNPDF3.0 NNLO	NLO
$t\bar{t}+V$	MADGRAPH5_aMC@NLO	PYTHIA 8, A14	NNPDF3.0 NLO	NLO
$t\bar{t}+WW$	MADGRAPH 5	PYTHIA 8, A14	NNPDF2.3 LO	NLO
$t\bar{t}t\bar{t}$	MADGRAPH 5	PYTHIA 8, A14	NNPDF2.3 LO	NLO
Single top	POWHEG-BOX v1	PYTHIA8, Perugia 2012	CT10	NLO+NNLL
VVV	SHERPA 2.1	SHERPA 2.1	CT10	NLO

8.5 Systematic uncertainties

A major part of performing a search (or measurement) is the evaluation of systematic uncertainties. There are a lot of sources influencing the precision of a result. For example, the reconstruction of physics objects, introduced in Chapter 5, comes with a large set of systematic uncertainties. Furthermore, the application of flavor tagging or the precision of the cross-sections of the background processes have to be considered. Their impact on the analysis is evaluated by using each uncertainty as a nuisance parameter (NP) in a profile likelihood fit, as described in Chapter 7. The derivation of systematic uncertainties is a wide field. For the physics objects, the evaluation of the size of a systematic uncertainty is mostly done by the physics performance groups of the ATLAS collaboration. Other systematic uncertainties like the modelling uncertainties are estimated within the analysis. In the following, the treatment and estimation of systematic uncertainties are discussed.

Preparation steps

A few preparation steps are performed before the systematic uncertainties are used in a profile likelihood fit as Gaussian-shaped distributions. The standard deviations $\pm 1\sigma$ of these Gaussian distributions are derived as described in this section.

At first, a smoothing procedure is applied to all uncertainties. It aims at the reduction of seemingly large systematic uncertainties induced by statistical fluctuations due to low numbers of MC events. Each uncertainty can either have an up and a down variation, denoted as *two-sided*, or can be *one-sided*, which means that just one variation is present. For each uncertainty, a histogram with the nominal MC prediction and the one-sided or two-sided variations is built. In order to smooth the uncertainties, a rebinning is applied until the relative statistical uncertainty of merged bins is smaller than a predefined tolerance, starting at 8%. If the number of sign changes of the slope is higher than

a predefined value of four, the tolerance is halved and the procedure is repeated. Thus, the derived systematic uncertainties are less prone to statistical fluctuations, which enhances the fit stability.

Furthermore, the systematic uncertainties are symmetrized. From the two-sided absolute variations, Δup and Δdown, the standard variation σ_i^{sym} of the symmetrized distribution is derived by

$$\sigma_i^{\text{sym}} = \frac{\Delta\text{up} - \Delta\text{down}}{2} \,. \tag{8.1}$$

This standard deviation is then added or subtracted from the nominal yields in bin i, x_i^{nom}, in order to obtain the event yields for the variations. For one-sided variations, the systematic uncertainty is mirrored in order to derive a corresponding up or down variation. Having two-sided, symmetrized uncertainties increases the fit stability.

At last, a pruning procedure is applied, which removes the normalization and/or shape of systematic uncertainties from the fit, whose size is smaller than 1%. Therefore, a fit with less parameters is created, which improves the running time and the fit stability once more.

Physics objects

Uncertainties related to all steps of the object reconstruction are considered. For electrons and muons, this includes uncertainties related to the trigger decision, reconstruction, identification and isolation. The reference processes $Z \to \ell\ell$ and $J/\psi \to \ell\ell$ [99, 103] are used for this purpose. Furthermore, the uncertainties on the calibration of the electron energy and the muon momentum scale and resolution [102, 103] are evaluated.

For small-R jets, a large number of systematic uncertainties is taken into account. Most of them come from the JES calibration [106], which uses multijet, Z + jets and γ+jets processes. Each of these samples introduces its own systematic variations: For example, uncertainties are included from the reconstructed objects of these three calibration samples. Additionally, differences in the JES calibration in η are included for the small-R jet uncertainties. Furthermore, pile-up related uncertainties as well as punch-through effects, where a jet reaches beyond the depth of the calorimeters, are included. Moreover, there are uncertainties accounting for differences in data and simulation regarding the jet response of gluon-jets, a different composition of light quark- and gluon-jets, as well as a different composition of jets originating from b-quarks (BJES). All JES related uncertainties are listed in Table 1 of Ref. [106]. All of these systematic jet uncertainties are reduced to a set of uncertainties, which is smaller, preserves the most important correlations and is more suitable for the application in an analysis. The use of the jet vertex tagger (JVT) [108] in order to suppress pile-up, also introduces an uncertainty. At last, the accuracy of the jet energy resolution [172] is estimated.

For large-R jets, the R_{trk} method is used to evaluate the uncertainties [117, 173] related to JES and JMS calibrations: Calorimeter-to-track jet ratios are used to build double ratios and estimate the uncertainties on the modelling or differences between data and simulation. Double ratios are also used for uncertainties related to the tracking efficiency. Another source of uncertainties for large-R jets is related to their jet mass and p_{T} resolution. The uncertainties are evaluated by randomly smearing the resolution of the large-R jet mass and p_{T} distribution. In order to achieve this, the standard deviation of the large-R jet properties is derived from a fit to the response function of the large-R jet mass or p_{T} distribution. This fitting procedure is performed on $Z' \to t\bar{t}$ events and uses large-R jets with

$p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and bins of $\frac{m}{p_{\mathrm{T}}}$. In the analysis, the large-R jet mass and p_{T} are drawn randomly from a 20% smeared Gaussian distribution and are subsequently fed to the analysis selection. This leads to acceptance effects due to requirements on the large-R jet mass and p_{T} in the analysis and gives an estimation of the resolution uncertainty.

There are also systematic uncertainties related to $E_{\mathrm{T}}^{\mathrm{miss}}$ [120]. The hard-objects in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation consist of the reconstructed and calibrated objects used in the analysis, which come with their own set of systematic uncertainties, as described above. Therefore, no further calculation of systematic uncertainties is needed. For the soft part, uncertainties for the energy scale and resolution are assigned by varying/smearing these contributions and evaluating the impact on $E_{\mathrm{T}}^{\mathrm{miss}}$.

b-tagging

Furthermore, b-tagging-related uncertainties are considered. Uncertainties on the b-tagging efficiency are derived for jets from b-, c- and light-quarks [123]. The b-tagging efficiency is calculated and uncertainties on jets from b-quarks are derived using mainly $t\bar{t}$ events [142]. The b-tagging efficiency can be chosen using a specific working point, which is a cut value on the MV2 classifier, discussed in Section 5.7. In the following analyses, a signal efficiency of 77% is chosen, which is averaged over a p_{T} range from $20\,\mathrm{GeV}$ to $300\,\mathrm{GeV}$ in the efficiency measurement. In order to scale the b-tagging efficiency in simulation to data, scale factors are derived, which come with an uncertainty and need to be applied on an event-by-event basis. The scale factors take different sources of systematic uncertainties into account, which affect the $t\bar{t}$ sample, used for the measurement: First of all, the MC generator modelling is varied in order to get an estimate of the dependence on the modelling of the hard-scattering simulation. Furthermore, cross-section assumptions are varied by changing the normalization. Additionally, all other detector and object reconstruction uncertainties are taken into account to derive an uncertainty on the scale factors. The scale factors are applied to all jet flavors (light, c, b). Furthermore, an extrapolation beyond the calibration of the efficiency measurement is done to a higher kinematic regime and from c-quark initiated jets.

Modelling uncertainties

The uncertainties assigned to the use of physics objects and b-tagging belong to the group of experimental uncertainties, since they aim at giving an accurate detector response with respect to data. On the other hand, there are uncertainties connected to theoretical assumptions entering the MC simulations. The modelling of MC simulations is one large part of theoretical uncertainties. For each of the larger backgrounds (Z + jets, $t\bar{t}$, $t\bar{t} + X$) the generator for the simulation of the hard-scattering and the parton shower simulation is varied. For the Z + jets process, the nominal sample, simulated with SHERPA 2.2.1, is exchanged with MADGRAPH using the NNPDF3.0 NLO PDF set. Subsequently, as a shower generator PYTHIA 8 with the A14 NNPDF2.3LO tune is used. This gives a handle on both the generator and the shower modelling, and is denoted as the Z + jets modelling uncertainty in the following. For $t\bar{t}$ events, the nominal sample POWHEG-BOX interfaced to PYTHIA 8 is replaced with MADGRAPH5_aMC@NLO for the generator uncertainty and HERWIG 7 [93] for the parton shower uncertainty. Here, generator and shower are varied independently, so that for the generator uncertainty MADGRAPH5_aMC@NLO is used with PYTHIA 8. Accordingly, the parton shower uncertainty is

evaluated by using POWHEG-BOX interfaced to HERWIG 7. The same procedure is also applied to $t\bar{t} + X$, where the nominal sample is replaced with SHERPA 2.1 using the NNPDF set in order to estimate the generator uncertainty. For the shower uncertainty of the $t\bar{t} + X$ process, the parameters of the A14 tune are varied.

An additional modelling uncertainty arises from assumptions about the renormalization scale μ_R and factorization scale μ_F. These scales are introduced in perturbative QCD processes where divergences occur and not all orders can be taken into account. When handling (regularizing and renormalizing) divergent integrals of cross-section calculations with ultra-violet singularities, the renormalization scale μ_R is introduced. According to the factorization theorem, the cross-section of a specific process is derived with Eq. 4.2. The soft term, which is described by the PDF, absorbs infrared divergences and the factorization scale μ_F is introduced. Thus, it can be deducted that the variation of both scales introduces deviations in the physics modelling and needs to be studied. As a standard procedure, so-called *7-point variations* are derived, which simultaneously vary μ_F and μ_R from 0.5 to 2. All combinations between 0.5 and 2 are built except the largest deviations of $\mu_{F/R} = 0.5$ and $\mu_{R/F} = 2.0$, which are avoided due to scale differences that are believed to be too large. From all scale variations, the envelope, which is the maximum up and down variation from the nominal sample with $\mu_F = \mu_R = 1$ in each bin, is computed and taken as a one σ up and down variation. These scale variations are derived for the Z + jets and $t\bar{t}$ processes.

Furthermore, the use of a specific PDF set influences the simulation of the hard-scattering. The PDF used as a standard configuration is the NNPDF set. It is exchanged with the MMHT2014 NLO [174] and CT14 NLO [175] PDF sets, the envelope is built with respect to the nominal PDF, and the up and down variations are derived. PDF uncertainties are only estimated for the Z + jets and $t\bar{t}$ processes. Their impact is small and even more insignificant for subleading background contributions.

Other uncertainties

Some additional uncertainties are taken into account. Since the cross-sections of the background processes are only known up to a certain accuracy, uncertainties are considered. They are listed in Table 8.6. Furthermore, the luminosity has to be measured in order to know the number of pp collisions, which occurred. The measurement comes with its own uncertainty. For 2015+2016 data, an integrated luminosity of $36.1\,\text{fb}^{-1}$ was found to be "good for physics" and comes with an uncertainty of 2.1%, derived with a technique similar to the one described in Ref. [176].

Additionally, the pile-up conditions varied depending on the number of bunch-crossings and the instantaneous luminosity. In order to correct for different pile-up environments between simulation and data, a reweighting is applied to each event. The reweighting procedure introduces an additional systematic uncertainty.

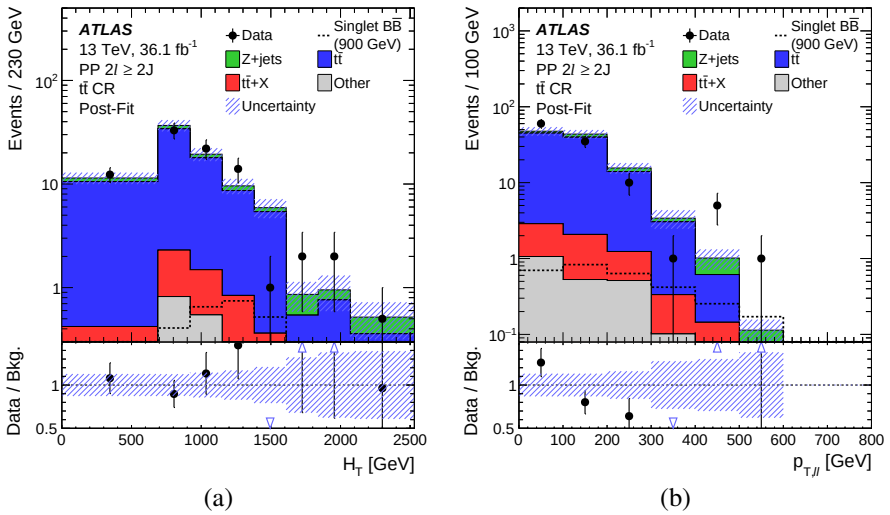
Moreover, an uncertainty of 25% for so-called *fakes* is introduced for $t\bar{t}$ events with $p_{\text{T}}\,(Z) < 200\,\text{GeV}$ in the CRs. It covers the case where a jet mimics an electron and therefore $t\bar{t}$ contributions enter the analysis due to two reconstructed leptons and high hadronic activity, which would not be present for a dileptonic $t\bar{t}$ decay.

Table 8.6: Cross-section uncertainties assigned to the processes Z+jets [177], $t\bar{t}$ [178, 179], $t\bar{t}+X$ and VV [177].

Process	Uncertainty [%]
Z+jets	± 5
$t\bar{t}$	$+5.6/-6.1$
Diboson	± 6
$t\bar{t}$+W	$+13.3/-11.9$
$t\bar{t}$+Z	$+10.4/-11.9$
$t\bar{t}+\ell\ell$	$+10.4/-11.9$

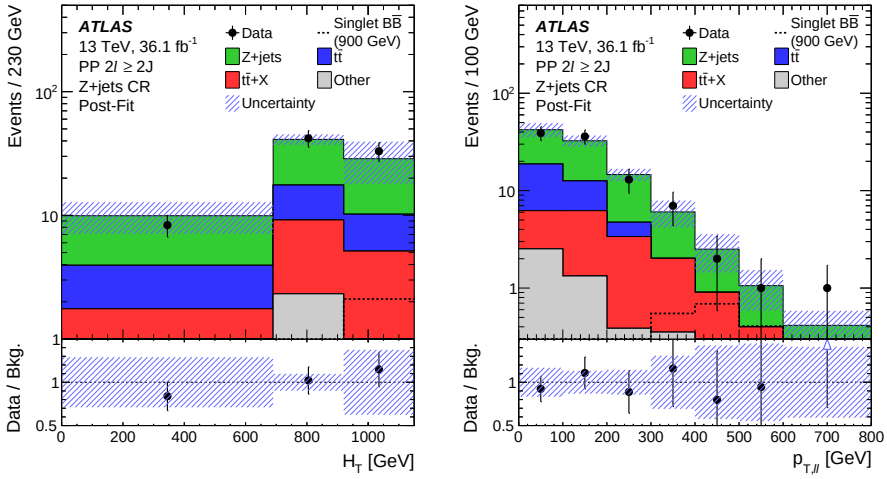
8.6 Data and MC simulation agreement

After the SR and CRs are defined and the systematic uncertainties are estimated, the modelling of MC simulations is compared to data. For the following comparisons, a binned profile likelihood fit with the B-only hypothesis was performed. The obtained post-fit distributions give insight into the background modelling without the presence of signal and take the systematic uncertainties, discussed in Section 8.5, into account. By allowing the fit to adjust the background modelling within the uncertainties, a better agreement between data and simulation is obtained and the systematic uncertainties can be constrained to the statistical uncertainty of the observed data. It is important that the post-fit modelling is reliable, especially for the final discriminant $m(Zb)$. The following distributions are also part of the results when testing for a discovery, described in Section 8.8.1. If mismodelling is debatable, a χ^2 test is performed, which takes the correlations between systematic uncertainties and bins into account. If the χ^2 probability is larger than 5%, the MC distribution is assumed to be modeled sufficiently.



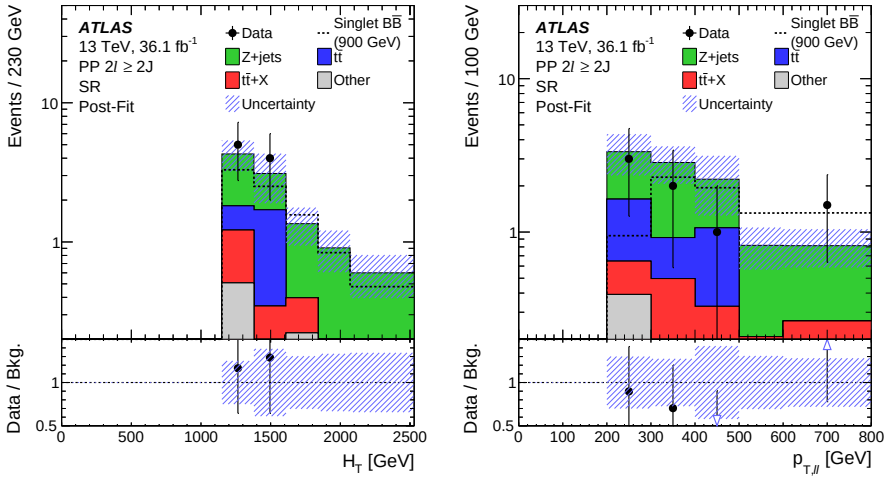
(a) (b)

Figure 8.6: Post-fit distributions for a) $H_{\mathrm{T}}(\mathrm{jets})$ [3] and b) $p_{\mathrm{T}}\,(Z)$ [154] in the $t\bar{t}$ CR. Signal according to the pre-fit hypothesis is overlaid. All systematic uncertainties are included. The background is modeled according to the B-only hypothesis. The lower plot shows the ratio of data over the background expectation.



(a) (b)

Figure 8.7: Post-fit distributions for a) $H_{\mathrm{T}}(\mathrm{jets})$ [3] and b) $p_{\mathrm{T}}\,(Z)$ [154] in the $Z+\mathrm{jets}$ CR. Signal according to the pre-fit hypothesis is overlaid. All systematic uncertainties are included. The background is modeled according to the B-only hypothesis.The lower plot shows the ratio of data over the background expectation.



(a) (b)

Figure 8.8: Post-fit distributions for a) $H_{\mathrm{T}}(\mathrm{jets})$ [3] and b) $p_{\mathrm{T}}\,(Z)$ [154] in the SR. Signal according to the pre-fit hypothesis is overlaid. All systematic uncertainties are included. The background is modeled according to the B-only hypothesis.The lower plot shows the ratio of data over the background expectation.

In Figure 8.6, 8.7 and 8.8, $H_{\mathrm{T}}(\mathrm{jets})$ and $p_{\mathrm{T}}\,(Z)$ are illustrated for the SR and CRs. In Figures 8.6(a) and 8.6(b), the $t\bar{t}$ CR is shown. It is visible that the region is quite pure in $t\bar{t}$ events and that the modelling of background is mostly within the systematic uncertainties. The fifth bin of the $H_{\mathrm{T}}(\mathrm{jets})$ distribution shows a downward fluctuation. The χ^2 probability to find this level of agreement or less between data and simulation in all bins is 41%. It can be deduced that the distribution shows decent MC modelling. On the contrary, the χ^2 probability to find the $p_{\mathrm{T}}\,(Z)$ distribution in Figure 8.6(b) is 0.02%. The data and MC simulation agreement was investigated for the cut sequence in the $t\bar{t}$ CR. No mismodelling is observed for all cuts except the last one, which is the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. Nevertheless, the

observed mismodelling is found in a variable, which is not the final discriminant. Additionally, $p_{\mathrm{T}}\,(Z)$ shows a low correlation with $m(Zb)$, and the total number of MC events in the $t\bar{t}$ CR is in agreement with the data observation within the uncertainties. Therefore, the mismodelling is assumed to have no strong impact on the final fit results.

For the Z + jets CR in Figures 8.7(a) and 8.7(b), it can be seen that data and simulation are well within the uncertainties. Furthermore, the calculated χ^2 probabilities are above 93%.

Modelling of the $H_{\mathrm{T}}(\mathrm{jets})$ and $p_{\mathrm{T}}\,(Z)$ distributions in the SR is depicted in Figures 8.8(a) and 8.8(b). What is suggested by the coverage of data within the uncertainties is supported by χ^2 probabilities above 92%: Data is well modeled by the MC background expectation.

After investigating the modelling of the kinematic variables used in the SR optimization, the final discriminant $m(Zb)$ is examined. In Figure 8.9, all three regions are illustrated. The modelling of

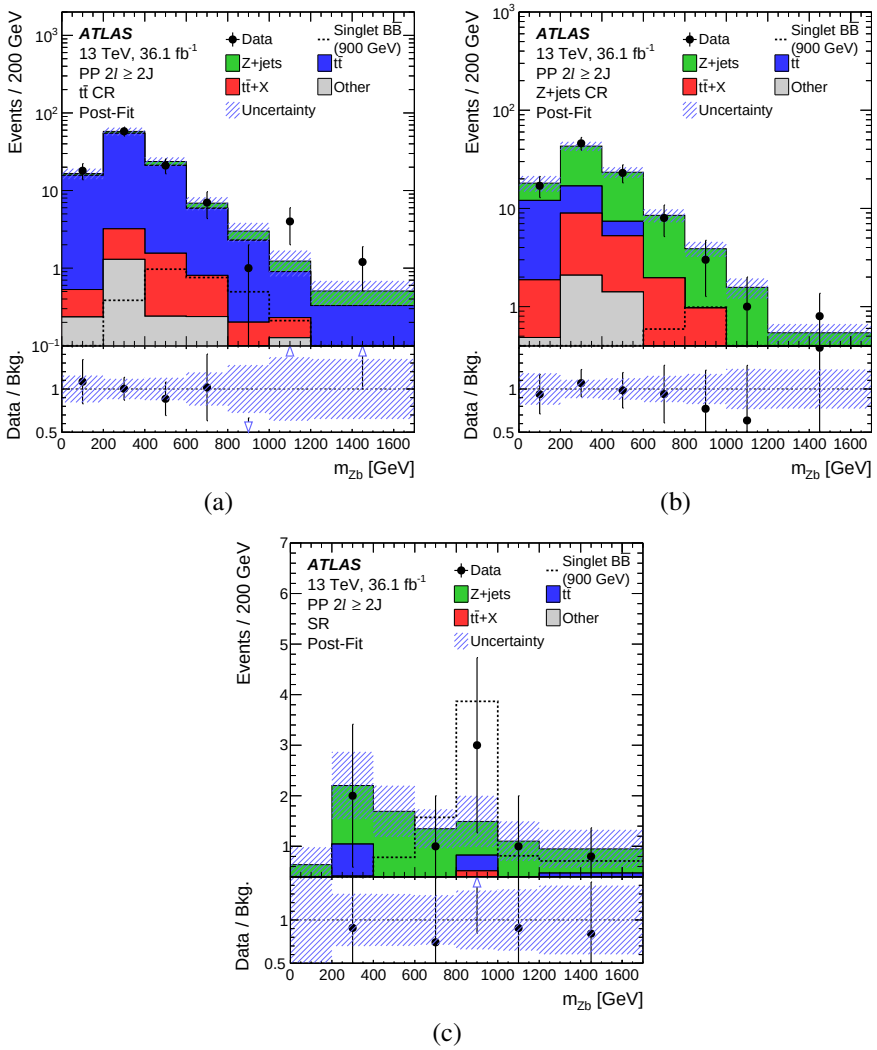


Figure 8.9: Post-fit distributions [3] for the final discriminant $m(Zb)$ in the a) $t\bar{t}$ CR, b) Z + jets CR and c) SR. Signal according to the pre-fit hypothesis is overlaid. All systematic uncertainties are included. The background is modeled according to the B-only hypothesis. The lower plot shows the ratio of data over the background expectation.

CRs is within the uncertainties for most of the bins. The χ^2 probability of 26% in the $t\bar{t}$ CR supports the compatibility of data and background as well. In the SR in Figure 8.9(c), it can be seen that all observed data points are within uncertainties and are compatible with the B-only hypothesis.

A summary of the yields at post-fit level is presented in Table 8.7. The number of background events is compatible with the data observation in the CRs and SR.

Table 8.7: Observed number of events [3] for data and the *post-fit* result for background in the CRs and the SR. A fit of the background expectation to the data $m(Zb)$ distributions under the B-only hypothesis is performed. The uncertainty on the post-fit event yields is the full uncertainty from the fit.

	$t\bar{t}$ CR	Z + jets CR	SR
Z + jets	9.0 ± 2.3	60 ± 10	6.5 ± 2.2
$t\bar{t}$	95 ± 12	20 ± 6	2.2 ± 1.5
Single top	1.5 ± 0.6	0.63 ± 0.28	0.016 ± 0.011
$t\bar{t} + X$	4.5 ± 0.8	14.7 ± 2.7	1.3 ± 0.4
Diboson	0.74 ± 0.20	4.2 ± 0.8	0.9 ± 0.4
Total Bkg.	111 ± 12	100 ± 10	10.9 ± 2.7
Data	112	100	9
Data/Bkg.	1.01 ± 0.11	1.00 ± 0.10	0.83 ± 0.21

8.7 Fit model

The model of the B-only fit which is performed in Section 8.6 is studied closer in the following. It is important to study the adjustments made by the fit on the pre-fit background modelling in order to understand the origin of the changes in shape and normalization.

In Figure 8.10, the pull plot for a B-only fit is shown. The black dots grouped around a value of zero show, how much the post-fit means are shifted in comparison to their pre-fit value divided by their pre-fit uncertainty. These deviations from a mean of zero are called *pulls*. The green and yellow bands show the one- and two-σ intervals of each Gaussian distribution at pre-fit level. The horizontal black line, attached to the black dots, gives insight into the size of the post-fit uncertainties. If the length of the line is smaller than the one-σ band, it means that a systematic uncertainty is constrained by the statistical power of the data. Pulls with a post-fit impact equal or smaller than one σ are likely, since the fit model is adjusted to improve the agreement of data and simulation. Constraints occur, if the pre-fit uncertainty decreases due to the presence of a large number of data events. Furthermore, a pre-fit uncertainty, which is chosen larger than required, may be constrained by the data. The pull plot in Figure 8.10 demonstrates that no large pulls are observed. Furthermore, the NP constraints are small, with the $t\bar{t}$ generator uncertainty showing the largest constraint. This is induced by large uncertainties for the $t\bar{t}$ generator. The uncertainty can reach up to a 100% deviation from the nominal prediction, which is incompatible with the data in the $t\bar{t}$ CR.

The correlation of all systematic uncertainties above a threshold of 20% is depicted in Figure 8.11. Most of the correlations are rather small and only the largest correlation, also corresponding to the

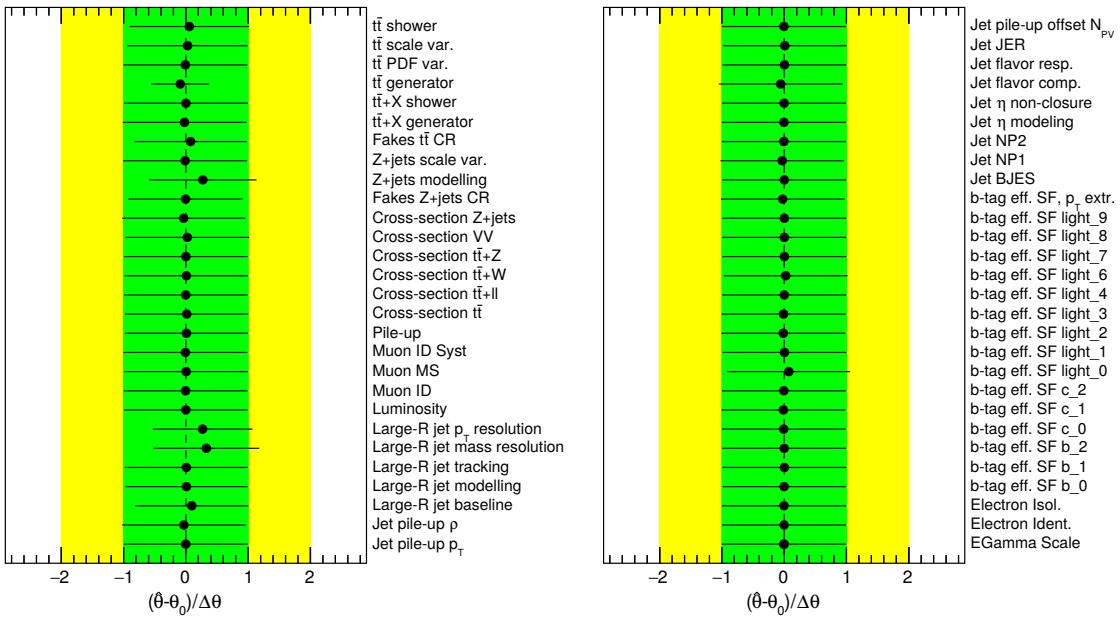


Figure 8.10: Pull plot for the B-only hypothesis of the VLB quark mass hypothesis of 900 GeV. Pulls of the post-fit mean from the pre-fit value and post-fit one-σ bands of Gaussian distributed NPs are illustrated as black dots and horizontal lines on the dots, respectively.

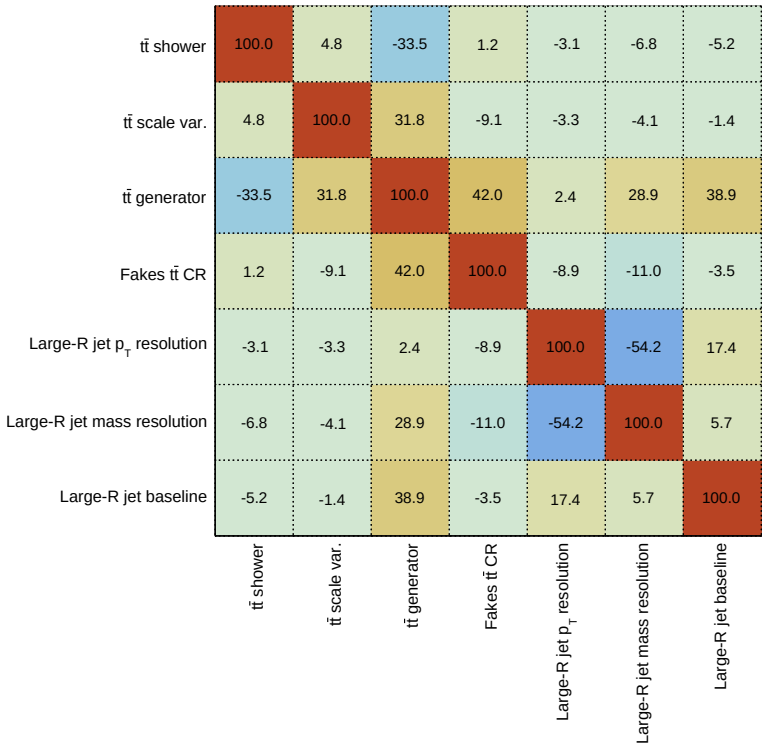
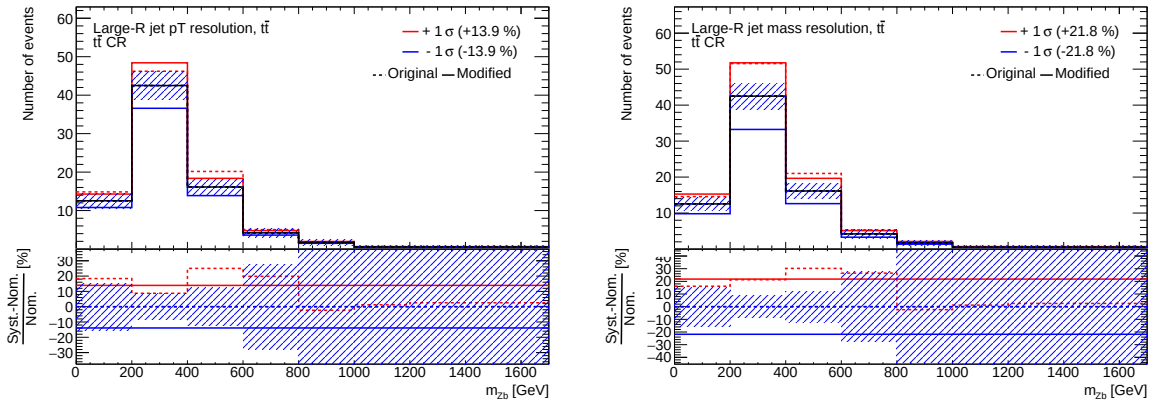


Figure 8.11: Correlations between the systematic uncertainties for a B-only fit of the VLB quark mass hypothesis of 900 GeV. Systematic uncertainties with a magnitude of the correlation of at least 20% to any other systematic uncertainty are depicted.



(a) (b)

Figure 8.12: Systematic shape plots for the large-R jet a) p_{T} and b) mass resolution for the $t\bar{t}$ background in the $t\bar{t}$ CR. The shaded bands depict the total statistical uncertainty on the nominal $t\bar{t}$ background in the $t\bar{t}$ CR.

largest pulls, is discussed. The largest correlation can be found between the large-R jet p_{T} and mass resolution uncertainties, which are anticorrelated with -54.2%. In Figure 8.12, their systematic shape variations are shown for the $t\bar{t}$ background in the $t\bar{t}$ CR. The red dashed line shows the original shape of the systematic, given to the fit. The systematic uncertainties are smoothed and the solid red and blue lines correspond to the symmetrized uncertainty entering the fit. For all backgrounds and regions, the shape of both systematic uncertainties is very similar, both before and after smoothing, and only the normalization is different. Therefore, a correlation is expected. The reason for the similarity can also be understood: Both resolution uncertainties introduce a smearing of the large-R jet mass and p_{T} distributions according to a Gaussian distribution with a larger width of 20% compared to the nominal prediction. The requirements for the selection with at least two large-R jets cause some of the events, drawn according to the smeared and wider distribution, to fail these cuts. Since the cuts are present for large-R jet p_{T} and mass, the smearing of both distributions causes the same kind of up and down fluctuations for both systematic uncertainties.

8.8 Results

A binned profile likelihood fit is used in order to test for the existence of VLT and VLB quarks or set exclusion limits in case of their absence. The final discriminant $m(Zb)$ is used in the fit in three regions: one SR with high sensitivity and two CRs. The fit adjusts the nominal background modelling within the freedom introduced by Gaussian-shaped NPs, with their width set according to the $\pm 1\sigma$ variations described in Section 8.5. Further details about the statistical procedure are given in Chapter 7.

8.8.1 Test for discovery

In order to test for a discovery, the B-only hypothesis is used. The incompatibility of data with the background expectation would suggest that the B-only hypothesis is rejected and therefore, new physics is found. The distributions, corresponding to a B-only fit, were closer investigated in Section 8.6. Judging by eye, it can already be seen that no obvious deviation from the Standard Model (SM) prediction is found. The calculation of the observed significances Z for the VLB and VLT quark with different mass hypotheses is shown in Table 8.8. These significances support the observation that no new resonance is detected, since their magnitude is within one or two standard deviations.

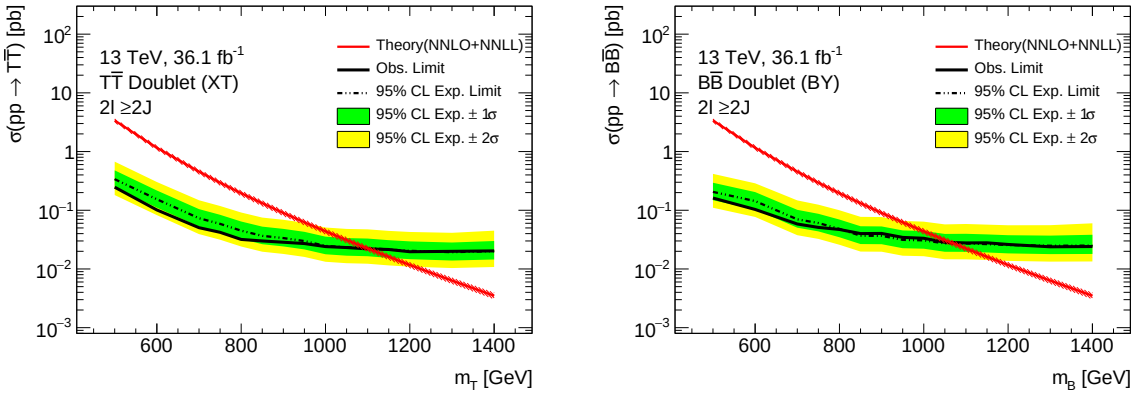
Table 8.8: Observed significance Z in units of the standard deviation σ for the singlet BRs for different mass hypotheses of VLT and VLB quarks for the $2\ell \geq 2J$ channel using 36.1 fb^{-1}.

Mass [GeV]	600	700	800	900	1000	1100
$T\bar{T}$	-1.11	-1.32	-0.89	-0.28	0.19	0.15
$B\bar{B}$	-1.60	-0.50	-0.04	0.61	0.45	0.13

8.8.2 Exclusion limits on the mass and the production cross-section

Although no VLT or VLB quarks are found, the analysis provides valuable results by setting exclusion limits on the production cross-section and mass of the VLQs. These results have direct consequences on the theories predicting VLQs, which are hence constrained. The results can also be used by theorists to draw conclusions for other new physics models and thus the input is important to the theory community.

In order to obtain exclusion limits, an S+B fit is performed. In Figures 8.13(a) and 8.13(b), the observed cross-section limits at 95% CL for $T\bar{T}$ and $B\bar{B}$ production are depicted assuming the doublet BRs, respectively. The red line indicates the theory curve, corresponding to pair production according to the strong interaction, and the observed limit is drawn as a solid black line. The expected limit is shown as a dashed line, and the one- and two-σ intervals are depicted for the 95% CL limit. At the intersection of the theory line with the observed limit curve, the excluded mass limit is set. For VLT quarks, masses below $m_T < 1102$ GeV are excluded, while for VLB quarks the limit is set at $m_B < 1063$ GeV. The limits for VLB and VLT quarks in the singlet model and with a 100% BR assumption for a decay into a Z-boson, are listed in Table 8.9. As expected, the limits increase for



(a) (b)

Figure 8.13: Expected and observed limits on the cross-section for a) $T\bar{T}$ and b) $B\bar{B}$ production according to the doublet BRs for the $2\ell \geq 2J$ channel of the $Z(\ell\ell)t/b + X$ search at 95% CL. The $\pm 1(2)\sigma$ bands of the expected limit at 95% CL are depicted in green (yellow). The theory curve is shown in red.

Table 8.9: Observed (expected) 95% CL mass limits for the singlet, doublet and 100% BR into $T \to Zt$ and $B \to Zb$ benchmark models for the $2\ell \geq 2J$ channel and the combination of the $Z(\ell\ell)t/b + X$ search. In addition, the $\pm 1\sigma$ bands around the expected limit are listed for the $2\ell \geq 2J$ channel at 95% CL. The limits for the combination are taken from Ref. [3].

	Observed (exp.) limits [GeV]	
Model	$2\ell \geq 2J$ channel	Combination
$T\bar{T}$ singlet	946 (927^{+67}_{-76})	1030 (1060)
$T\bar{T}$ doublet	1102 (1095^{+60}_{-62})	1210 (1210)
100% $T \to Zt$	1212 (1206^{+67}_{-52})	1340 (1320)
$B\bar{B}$ singlet	929 (952^{+71}_{-50})	1010 (1030)
$B\bar{B}$ doublet	1063 (1068^{+70}_{-52})	1140 (1120)
100% $B \to Zb$	1120 (1126^{+52}_{-49})	1220 (1180)

increasing BR into the Zt/b final state. The limits exceed the TeV-scale between the 25% and 50% BR into Zt/b final states. The highest limits can be set at 1212 GeV for VLT quarks and 1120 GeV for VLB quarks with a 100% BR into Zt/b final states. The illustration of the exclusion limits for the singlet model and the exclusive decay into Zt/b are depicted in Appendix B.2.

The most important systematic uncertainties, influencing an analysis can be evaluated with a ranking plot, illustrated in Figure 8.14. Not only the pulls and constraints corresponding to the pull plot are shown, but also the pre- and post-fit impact on the signal strength μ. With respect to the pre- and post-fit uncertainty, all systematic uncertainties and γ factors are independently varied up and down by one σ around the best-fit value and the induced deviation on μ is evaluated in an S+B fit. It can be seen that for VLT and VLB quarks the modelling uncertainties for the main backgrounds, which are Z + jets modelling, $t\bar{t}$ generator and showering, are among the five most important uncertainties.

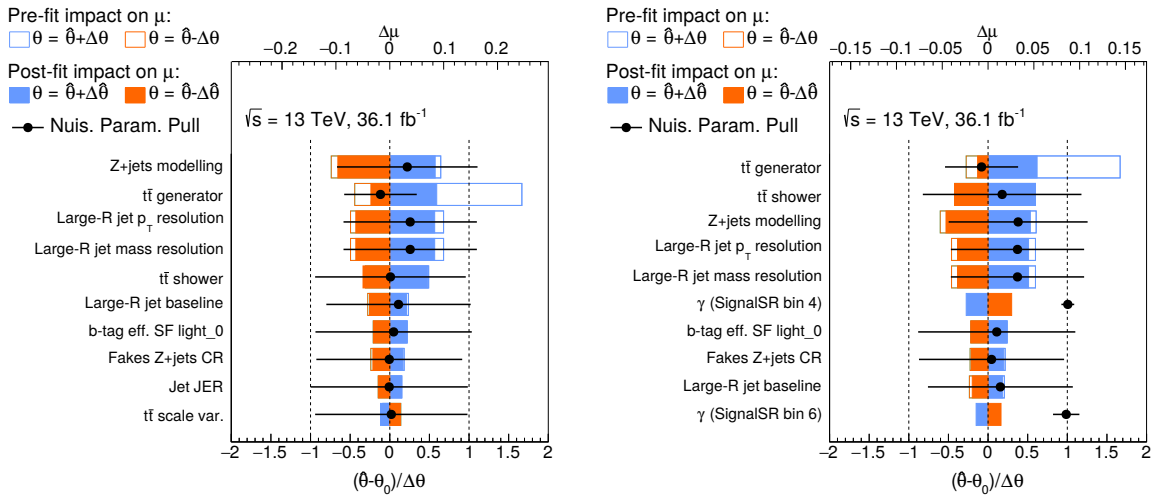


Figure 8.14: Ranking plots for a) VLT and b) VLB quarks for a mass of 900 GeV assuming BRs according to the singlet model. The impact of each NP and γ factor is evaluated on the signal strength μ in an S+B fit of the signal and background prediction to the observed data. Each NP is fixed to $\pm 1\sigma$ of the pre- and post-fit estimates around the best-fit value. A fit of all other systematic uncertainties and the statistical uncertainty is performed, which derives the impact of one specific NP with respect to the usual S+B fit and its signal strength μ. Furthermore, the pulls and constraints of the NPs are depicted.

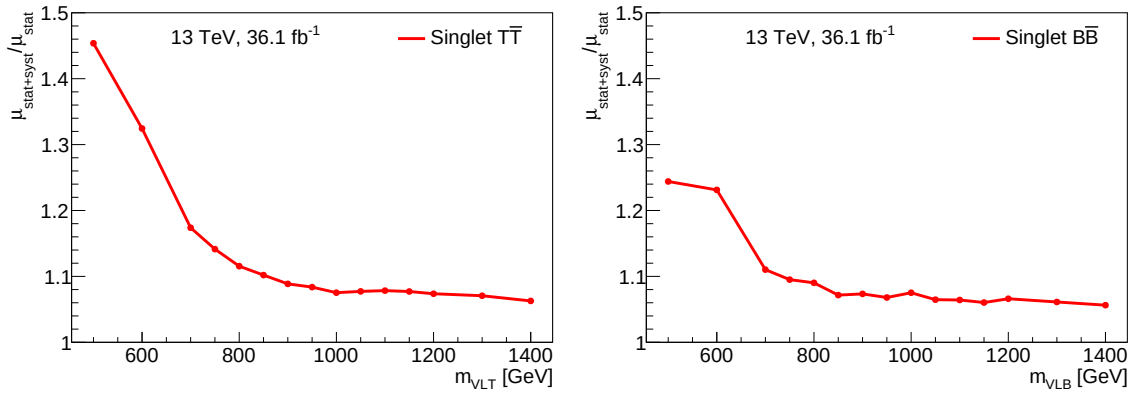


Figure 8.15: Ratio of expected exclusion limits at 95% CL for a fit with statistical together with systematic uncertainties and a fit with statistical uncertainties only. The limits for a) VLT and b) VLB quarks are illustrated for BRs according to the singlet models.

Furthermore, the variations of the large-R jet p_{T} and mass resolution uncertainties influence the exclusion limits almost as significantly.

In general, the impact of systematic uncertainties is negligible in this analysis, since the analysis is dominated by statistical uncertainties. As can be seen in Figures 8.15(a) and 8.15(b), the ratio of exclusion limits for a fit with statistical together with systematic uncertainties and a fit with statistical uncertainties only drops for low VLT and VLB quark masses, respectively. The ratio approaches a constant value for high VLQ masses. This indicates that the search becomes more independent of systematic uncertainties for higher VLQ masses. This conclusion can also be obtained by looking

at the fitted $\hat{\mu}$ from the S+B fit of the ranking plot and its total uncertainty for a mass hypothesis of 900 GeV for VLT and VLB quarks, assuming singlet BRs: For the VLT and VLB quark search, values of $\hat{\mu}_{\text{T,900}} = -0.09 \pm 0.34$ and $\hat{\mu}_{\text{B,900}} = 0.17 \pm 0.30$ are derived, respectively. Comparing $\Delta\mu$ of the highest ranked systematic uncertainties from the ranking plot at post-fit level to the total uncertainty on μ, it can be concluded that the statistical uncertainty is the largest individual uncertainty.

8.9 Combination of channels sensitive to Z-boson final states

The dilepton channel with at least two large-R jets ($2\ell \geq 2J$) is part of a larger search program within the ATLAS collaboration. Together with the dilepton channel for zero and one large-R jet (2ℓ 0–1J) and the trilepton channel ($\geq 3\ell$), three orthogonal channels for the analysis $Z(\ell\ell)t/b + X$ [3] exist and are combined using 36.1 fb^{-1} in order to improve the sensitivity of the individual channels. The main selection criteria for the SRs are presented in the sketches in Figure 8.16.

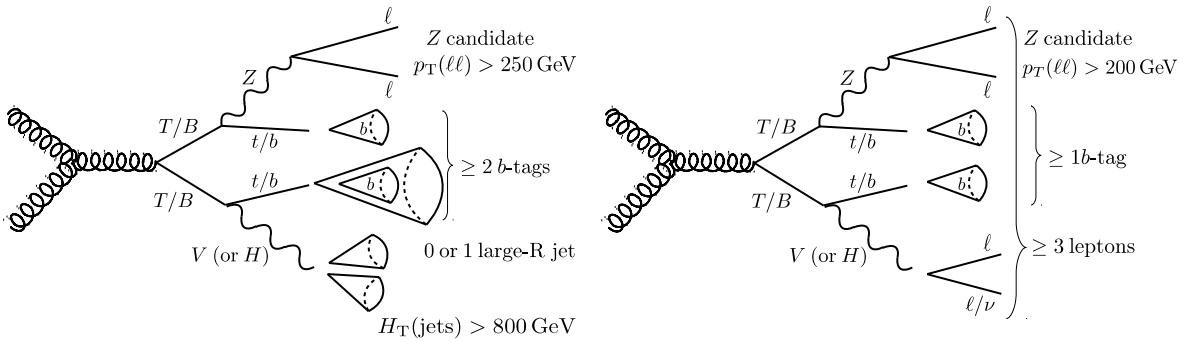


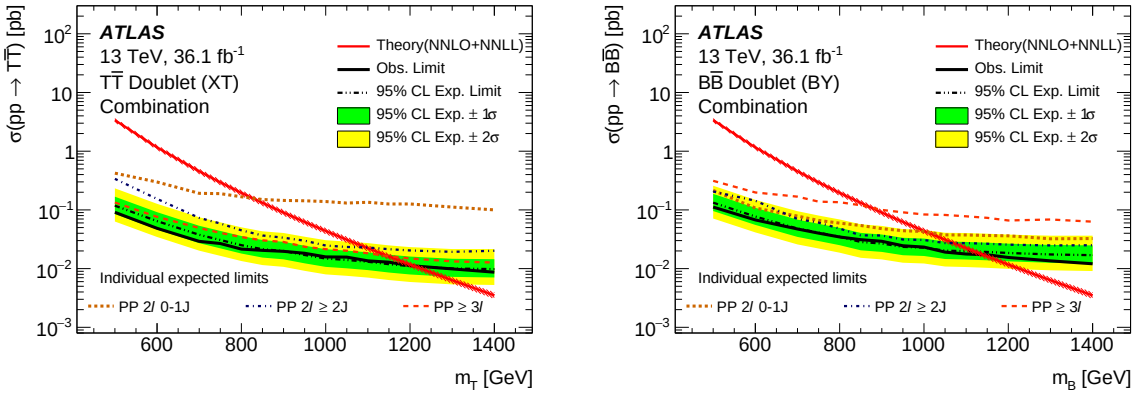
Figure 8.16: Illustration of the final states for the a) 2ℓ 0–1J and b) $\geq 3\ell$ channels [3] of the $Z(\ell\ell)t/b + X$ search with the most important selection criteria.

For the 2ℓ 0–1J channel, requirements on at least two b-tagged jets, high p_{T} (Z) and high H_{T}(jets) are made. The phase-space with high sensitivity is split into two SRs for each large-R jet multiplicity. As discriminant in the profile likelihood fit, H_{T}(jets) is used. Two CRs for the main backgrounds, which are the Z + jets and $t\bar{t}$ processes, are constructed.

In case of the $\geq 3\ell$ channel, at least one b-tagged jet is required in order to keep a high signal efficiency. Additionally, the Z-boson candidate is required to carry a p_{T} of at least 200 GeV. In the end, the variable S_{T}, which is the sum of H_{T}(jets) and the scalar p_{T} of all leptons, is used in the profile likelihood fit. The main backgrounds are VV and $t\bar{t} + X$, and one CR for each is built.

All three channels are combined in one binned profile likelihood fit. As for the $2\ell \geq 2J$ channel, no deviation from the SM prediction in the B-only fit is observed. Therefore, exclusion limits are set on the production cross-section and the VLQ masses. In Figures 8.17(a) and 8.17(b), limits for the doublet BRs for VLT and VLB quarks are depicted, respectively. The distributions show the expected and observed limits at 95% CL from the combination, as well as the expected limits from all individual channels of the $Z(\ell\ell)t/b+X$ search. Masses below $m_T < 1210$ GeV and $m_B < 1140$ GeV are excluded. For VLT quarks, the limits of the $\geq 3\ell$ channel are the strongest across the whole mass range with significant contributions from the $2\ell \geq 2J$ channel. For low VLT quark masses, the

$2\ell \geq 2J$ channel becomes less sensitive due to the absence of boosted final-state objects. For high masses, their number increases, producing more large-R jets from top-quarks, V- or H-bosons. Since events with boosted final states are primarily selected by the $2\ell \geq 2J$ channel, the sensitivity of the 2ℓ 0–1J channel is not competitive with the other two channels. In case of VLB quarks, the $2\ell \geq 2J$ channel contributes the largest sensitivity starting at $\approx 700\,\text{GeV}$. For $B\bar{B}$ production in the doublet or the 100% BR case, only neutral currents are possible as decay modes. This leads to an absence of top-quarks and therefore to less leptonic decays, reducing the sensitivity for the $\geq 3\ell$ channel. The 2ℓ 0–1J channel contributes to the sensitivity as well. Since for $B\bar{B}$ production the production of high-mass large-R jets is less likely than for $T\bar{T}$, the sensitivity is split more evenly between the 2ℓ 0–1J and $2\ell \geq 2J$ channels. The mass limits from the combination are listed in Table 8.9. The limits on the cross-section for $B\bar{B}$ and $T\bar{T}$ production for the singlet BRs and the 100% BR into Zt/b are depicted in Appendix B.3.



(a) (b)

Figure 8.17: Expected and observed cross-section limits at 95% CL for a) $T\bar{T}$ and b) $B\bar{B}$ production with BRs according to the doublet model for the combination of all $Z(\ell\ell)t/b + X$ channels [3]. Furthermore, the expected limits of the individual channels 2ℓ 0–1J, $2\ell \geq 2J$ and $\geq 3\ell$ of the analysis are shown. The $\pm 1(2)\sigma$ bands of the expected limit at 95% CL are depicted in green (yellow). The theory curve is shown in red.

Since the assumption of BRs is crucial for the exclusion limits and no information about intermediate BRs is given, a BR scan is performed. It visualizes the sensitivity in the whole BR plane with various BR assumptions into Z-, W- or H-boson final states. All BRs from zero to one for W-, Z- and H-boson decays with the requirement of $\text{BR}(Q \to Wq') + \text{BR}(Q \to Zq) + \text{BR}(Q \to Hq) = 1$ are scanned with a step size of 5%. The expected and observed 95% CL limits on the VLT quark masses are illustrated in Figures 8.18(a) and 8.18(b) and for VLB quarks in Figures 8.18(c) and 8.18(d). While for VLT quarks, a slightly better exclusion is possible in the Ht corner, the VLB quark search is more sensitive to the Wt corner. In both cases the top-quark is the reason for an increased sensitivity, either leading to a more massive large-R jet for the $2\ell \geq 2J$ channel or an additional lepton for the 3ℓ channel. Clearly, the highest sensitivity is achieved for the Z-boson corner with limits of $m_B > 1220\,\text{GeV}$ and $m_T > 1340\,\text{GeV}$ in the 100% BR case.

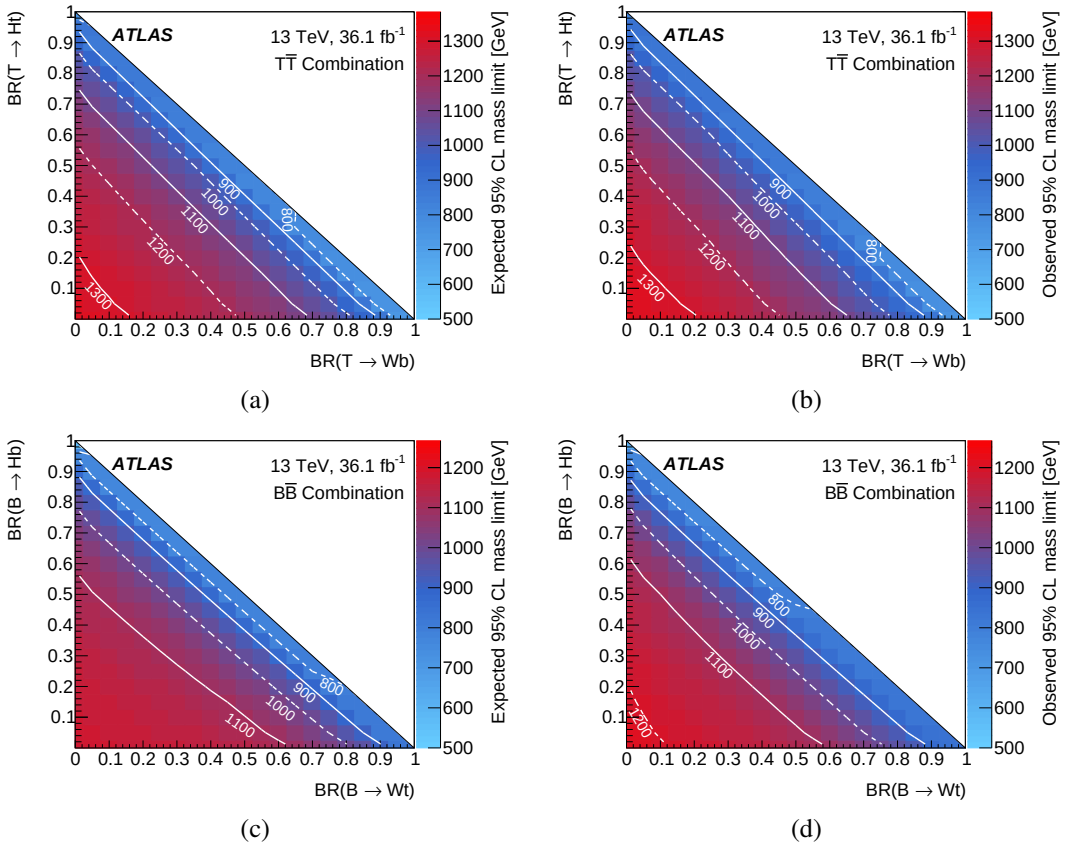


Figure 8.18: BR scan for the $Z(\ell\ell)t/b + X$ search. The mass limits at 95% CL for $T\bar{T}$ and $B\bar{B}$ production are shown on the top and bottom, respectively. Expected limits can be seen on the left side, while observed limits are on the right side [3].

8.10 Combination of pair production searches at the ATLAS experiment

A combination can not only be done for a VLQ search with $Z(\ell\ell)$ final states, as presented in Section 8.9. A combination of all pair production searches performed by the ATLAS collaboration with 36.1 fb^{-1} [4] was completed after the 2015+2016 data taking.

VLQ searches targeting different final states are used: One analysis targeting the H-boson corner searches for $H(bb)t + X$ final states [180] and only for VLT quarks. A complex categorization based on b-, top- and H-tagging is carried out and the scalar sum, S_{T}, of transverse momenta of jets, leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ is used as final discriminant.

Furthermore, in the W-boson corner an analysis looking for $T\bar{T} \to W(\ell\nu)b + X$ [181] is done. The analysis is specifically sensitive to $WbWb$ final states due to the use of W-boson tagging and the invariant mass $m(\ell\nu b)$ as final discriminant. Another analysis is performed in this corner looking for $B\bar{B} \to W(\ell\nu)t + X$ [126]. In this case, a BDT approach for the SR selection is combined with the reconstruction of the $B\bar{B}$ initial state.

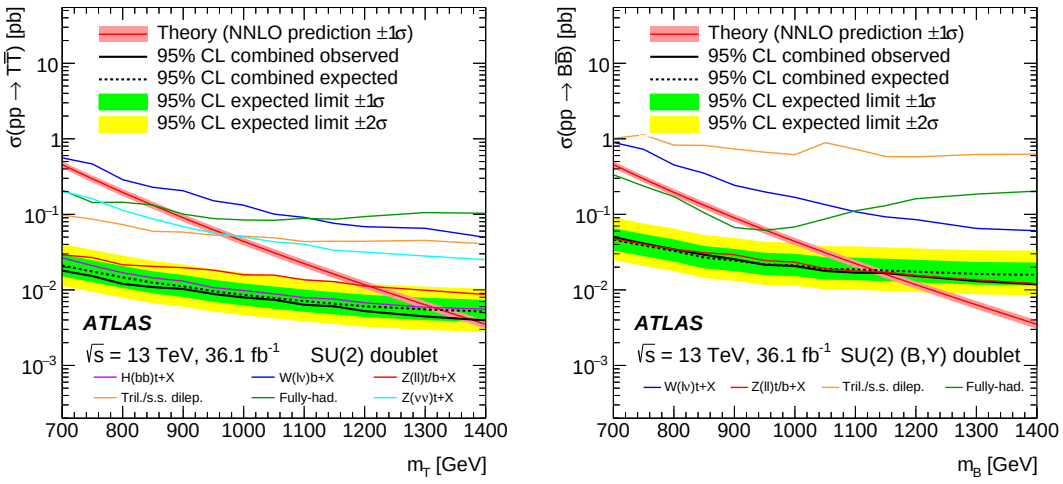
The Z-boson corner is covered by the $Z(\ell\ell)t/b + X$ search [3], described in this chapter. The analysis is sensitive to the Z-boson corner, but also gives sensitivity in the H- and W-boson corners

due to the inclusive topology for the second VLQ decay. Either the invariant mass $m(Zb)$, $H_{\mathrm{T}}(\mathrm{jets})$ or S_{T} are used as final discriminants in the different channels, described in Section 8.9.

Additionally, the analysis $Z(\nu\nu)t + X$ [182] targets the Z-boson corner and searches for VLT quarks only. The analysis requires final states with large $E_{\mathrm{T}}^{\mathrm{miss}}$ in a one-bin SR, which further exploits the boosted final states by using criteria like the requirement of at least two reclustered (RC) jets with high mass.

Another analysis, used for the ATLAS-wide combination, is the same-sign dilepton analysis [183] looking for VLB and VLT quarks. Two of the final-state charged leptons have to have the same charge, and eight SRs with one bin are used.

Last but not least, there is a search for VLT and VLB quarks investigating purely hadronic final states [1], which gives sensitivity to a broad number of VLQ final states. The analysis makes use of variable-R reclustered (vRC) jets and a neural-network based tagger for the unambiguous identification of V-boson, H-boson and top-quark jets. Events are sorted into different categories based on the number of V-, H-, top- and b-tags, and a signal probability, calculated with the matrix element method [184], is used in a binned profile likelihood fit.



(a) (b)

Figure 8.19: Observed and expected exclusion limits [4] from the ATLAS pair production combination on the cross-section of a) $T\bar{T}$ and b) $B\bar{B}$ production. For VLB quarks, the doublet BRs according to (B, Y) are assumed. For VLT quarks, the SU(2) doublet corresponds to the (X, T) and (T, B) doublets. Solid lines show the observed limits of the combination and all contributing analyses. The important contribution of the $Z(\ell\ell)t/b + X$ search is visible. The $\pm 1(2)\sigma$ bands of the expected limit at 95% CL are depicted in green (yellow). The theory curve is shown in red.

Since no deviation from the SM expectation is found in the combination, limits on the production cross-section and the VLQ masses are set at 95% CL. For the singlet BRs, VLT and VLB quark masses with $m_T < 1.31\,\mathrm{TeV}$ and $m_B < 1.22\,\mathrm{TeV}$ are excluded, respectively. In Figure 8.19, the cross-section limits for $B\bar{B}$ and $T\bar{T}$ production are illustrated for the doublet BRs. In case of VLT quarks, the BRs are the same for both possible doublets (T, B) and (X, T), while for VLB quarks they are very different. For the (B, Y) doublet, only decays via a neutral boson can occur, which is inverted for the (T, B) doublet with charged-current decays only. Figure 8.19(a) shows the VLT quark limits,

and the intersection with the theory line excludes VLT quark masses below 1.37 TeV for both doublet models. The $Z(\ell\ell)t/b + X$ search shows the second best observed limit. In Figure 8.19(b), the VLB quark exclusion limits are presented and the $Z(\ell\ell)t/b + X$ search drives the limit. In combination with Figure 8.17(b), showing the limits of the three channels for the (B, Y) doublet model, it can be concluded that the $2\ell \geq 2J$ channel presented as a main part of this chapter is the channel with the highest exclusion limits of all ATLAS searches in this combination. For BRs according to the (B, Y) model, VLB quark masses below 1.14 TeV are not compatible with the SM. Depictions of exclusion limits for the additional doublet model for VLB quarks and the singlet BRs for VLB and VLT quarks are shown in Appendix B.4.

8.11 Other VLQ searches with 2015+2016 data

Besides the pair production searches performed by the ATLAS collaboration and combined in one combination with the highest sensitivity, other VLQ searches were performed by the ATLAS and CMS collaborations with the dataset taken in 2015 and 2016.

The CMS collaboration published a variety of pair production searches with different final state topologies for VLT and VLB quarks. A search similar to the $Z(\ell\ell)t/b + X$ analysis was performed with the 2016 dataset of $35.9\,\text{fb}^{-1}$. VLT quark masses below 1280 GeV [185] were excluded for the 100% BR assumption into $T \to Zt$, while for VLB quark masses below 1130 GeV [185] are not compatible with the SM expectation for a BR of 100% into Zb. The $Z(\ell\ell)t/b + X$ search performed by the ATLAS collaboration, which includes the $2\ell \geq 2J$ channel discussed as in this chapter, was able to exclude VLT and VLB quarks at higher values with these BR assumptions: The limits [3] were set at 1340 GeV and 1220 GeV for VLT and VLB quarks, respectively.

The highest limit on pair production of VLT quarks by the CMS collaboration was set by the all-hadronic analysis using $35.9\,\text{fb}^{-1}$ and excluded masses below 1370 GeV [186] using a neural network approach. For VLB quarks, an analysis with leptons was able to exclude $m_b < 1240$ GeV [187]. Both limits are based on the assumption of a 100% BR into the decay mode with the highest sensitivity. Thus, they cannot directly be compared to the combination of pair production searches by the ATLAS collaboration which does not quote a 100% BR case, but rather assumes decays according to singlet and doublet models. However, the results with approximately $36\,\text{fb}^{-1}$ by the ATLAS and CMS collaborations make the same observations: The pair production of VLT and VLB quarks is excluded at and well above the TeV scale depending on the BR assumptions.

Multiple searches for single production were performed by the ATLAS and CMS collaborations with the data taken in 2015 and 2016 with $\sqrt{s} = 13$ TeV. The final state topologies $B \to Zb$ [188], $B \to Hb$ [189], $B \to Wt$ [190], $T \to Zt$ [3, 127, 188, 191], $T \to Wb$ [192, 193] and $T \to Ht$ [194–196] were targeted with different analysis strategies. Mainly limits on the production cross-section and the coupling to SM quarks were set. Some searches also calculate the exclusion mass limit under specific coupling assumptions. For example in Ref. [3], VLT quarks with masses below 1600 GeV are excluded for large couplings c_W. The results show that no deviation from the SM expectation was found in single production searches and that masses significantly above 1 TeV might be excluded depending on the model parameters. The complementary search programs for single and pair production contribute to larger constraints on the phase-space for VLQ existence.

Chapter 9
Full Run-2 search for vector-like bottom and top quarks with 139 fb^{-1}

In Chapter 8, a search for vector-like bottom (VLB) and vector-like top (VLT) quarks was performed with $\int \mathcal{L}\mathrm{d}t = 36.1\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$. A strategy with one signal region and two control regions was chosen in the dilepton channel with at least two large-R jets. The analysis in the following chapter also targets final states with a leptonic Z-boson candidate as in the previous analysis and is thus again denoted as the $Z(\ell\ell)t/b + X$ analysis. The following search is also performed in the dilepton channel; however, a more advanced strategy is chosen by making use of the Multi-Class Boosted-Object Tagger (MCBOT), presented in Chapter 6. Furthermore, the whole Run-2 dataset is used with an integrated luminosity of $139\,\text{fb}^{-1}$, which is almost four times larger than the data used for the previous analysis, denoted as the 2015+2016 analysis in the following. Therefore, it is expected that higher vector-like quark (VLQ) masses beyond the sensitivity boundaries of the 2015+2016 results can be tested for a discovery or an exclusion.

At first, a novel strategy with an MCBOT categorization is introduced in Section 9.1. It is followed by the presentation of the Monte Carlo (MC) samples used for signal and background modelling in Section 9.2. An extensive analysis optimization is then described in Section 9.3, which addresses the MCBOT working point (WP) optimization, the kinematic cuts, the search for a final discriminant, the categorization and more. A presentation of relevant systematic uncertainties, obtained in a different way to the 2015+2016 analysis, is given in Section 9.4. Subsequently, the data and MC simulation agreement is presented for the optimized analysis strategy in Section 9.5. In Section 9.6, the compatibility of VLQ signal with the Standard Model (SM) expectation is tested, and subsequently the derivation of exclusion limits is presented.

9.1 Novel strategy: Event categorization with MCBOT

This search for VLB and VLT quarks targets dileptonic final states from Z-boson candidates with high p_{T}. Alongside the Z-boson, multiple other quarks and bosons are created, which was discussed in Section 8.1. Because the two leptons are expected to origin from the Z-boson candidate, all other visible decays are assumed to be of hadronic nature. As a consequence, the identification of V-bosons, H-bosons and top-quarks with MCBOT is possible.

In Figure 9.1, an exemplary Feynman diagram of the signal process and the resulting signature in the detector is illustrated, which utilizes reclustered (RC) jets in contrast to the use of large-R jets for the 2015+2016 analysis. The first generation of possible decay products in the decay chain of the pair production of VLT and VLB quarks with a $Z(\ell\ell)$-boson candidate are listed and the colors are chosen according to the colors for the DNN output classes in Chapter 6. Due to the boosted final states

of VLQ decays and the presence of multiple final states with V-bosons, H-bosons and top-quarks, MCBOT is very suitable for unambiguous tagging of these final states. A short study of tagging with MCBOT on final states from VLT and VLB quark decays is presented in Appendix D.

A categorization according to MCBOT tags is expected to exceed the sensitivity of the analysis compared to the previous strategy: A dilution of the signal in regions with higher background contribution is avoided. Thus, a multitude of categories, of which a part is extremely sensitive to signal, increases the significance or limit on the signal strength, derived in hypothesis tests using the profile likelihood ratio (PLR). Furthermore, a correlation of systematic uncertainties across multiple regions might reduce the uncertainties due to the statistical power of the data in these regions. Thus, the background prediction is improved. The use of MCBOT can be considered as a more complex kinematic cut used for the categorization of events compared to a simple cut-based analysis. The terms *region* and *category* are used interchangeably in this chapter to describe orthogonal sets of events with specific properties.

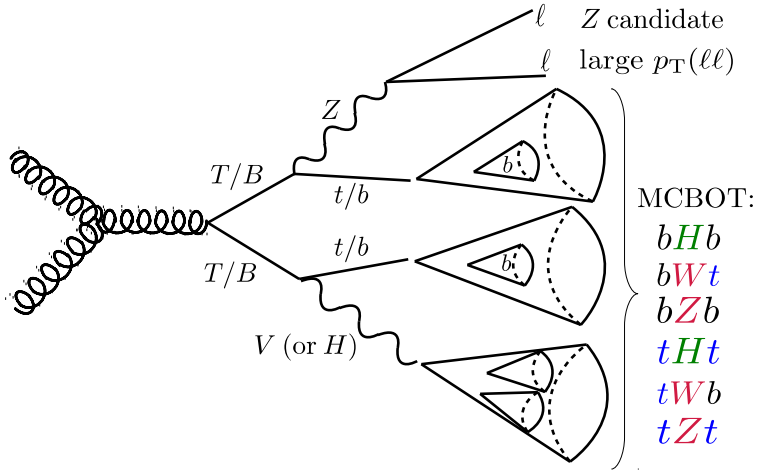


Figure 9.1: Illustration of the production mechanism and possible decays of the pair production of VLT and VLB quarks. An example for reconstructed physics objects is indicated. The expected hadronic decay combinations are listed for VLT and VLB decays, whose identification is attempted with MCBOT. The colors are chosen analogous to the DNN output class colors in Chapter 6.

Throughout the following sections, different categorizations are chosen and compared. Some contain a subset of the events in the dilepton channel and some are more inclusive. The purpose of the different categorizations is always explained, but in the end, categories with all possible combinations of MCBOT tags or no tags are included in the analysis strategy. Thus, the maximum sensitivity is reached and systematic uncertainties are constrained in regions with a high number of events. In each category, a final discriminant is used for further separation between signal and background, which is then used in a shape fit.

An MCBOT categorization is based on the number of V-boson, H-boson and top-quark RC jet tags in different categories. In the following, these tags are abbreviated as V-tag, H-tag and top-tag. A tag only exists, if an RC jet is present. The RC jet multiplicity, by itself, is already a discriminating variable, as the number of large-R jets in the 2015+2016 analysis, which was used to define the very sensitive $2\ell \geq 2J$ channel. In Figure 9.2(a), the number of RC jets per event is illustrated in a histogram normalized to unity after the primary selection, which is defined in Table 9.1. This selection

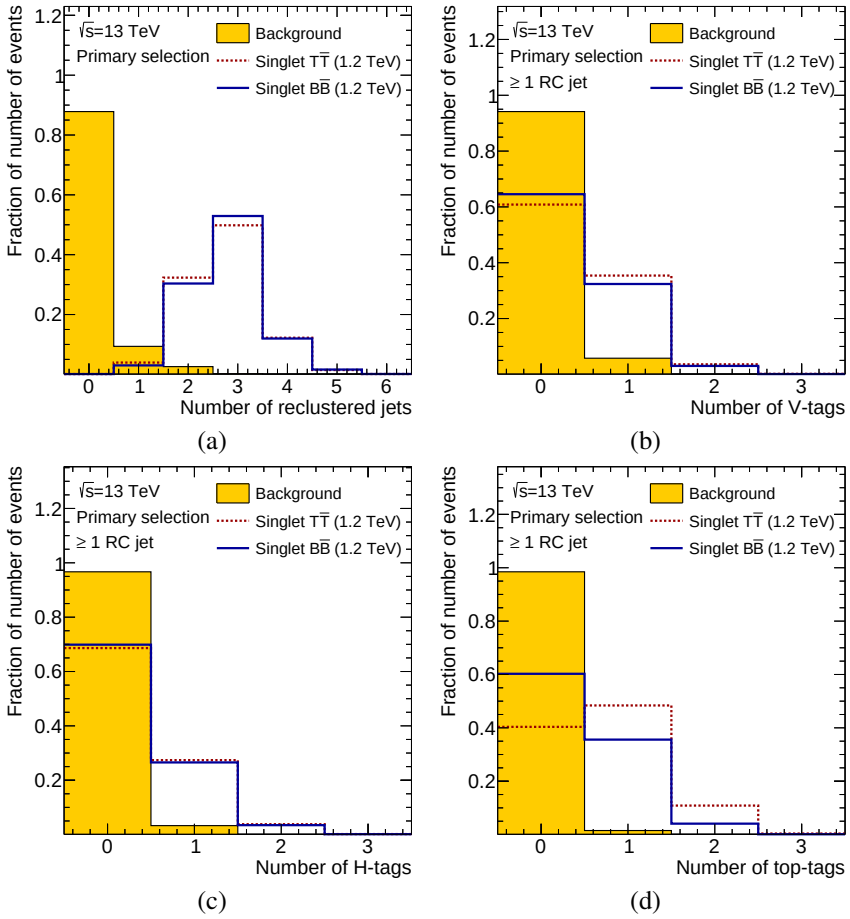


Figure 9.2: Shape comparison for signal and background events for the a) RC jet and b) V-tagged RC jet multiplicity. In c) and d), the H-tagged and top-tagged jet multiplicity is illustrated, respectively. All distributions include the requirements of the primary selection. The V-, H- and top-tagged multiplicities include an additional requirement of at least one RC jet. The MCBOT WPs for tagging correspond to cut values on the DNN output distributions of $c_V = 0.3$, $c_H = 0.35$ and $c_t = 0.3$. Both pair-production benchmark signal masses for VLT and VLB quarks at 1200 GeV are shown, as well as a sum of all background processes. The last bin contains all overflow events.

requires at least two central small-R jets and exactly two opposite-sign same-flavor (OSSF) leptons. In addition, the Z-boson candidate mass, $m_{\ell\ell}$, is required to be within 10 GeV of m_Z. In this figure, it can be seen that most background events have no RC jet, while the number of RC jets for VLB and VLT quark events, pair produced according to the singlet branching ratio (BR) at a mass of 1200 GeV, has its maximum at three RC jets. Thus, the boosted topology of the signal events becomes apparent.

In Figure 9.2(b) to 9.2(d), the number of V-, H- and top-tags is depicted after a requirement of at least one RC jet in addition to the primary selection. For all distributions, a good separation of signal and background is present. For the number of V-tags in Figure 9.2(b), most final states from VLB and VLT quarks have zero tags, but there are around 35–40% of events with at least one V-tag. Events originating from VLB and VLT quarks have a similar number of V-tags in each bin. A similar statement holds for the H-tags, depicted in Figure 9.2(c), where around 30% of RC jets have at least one H-tag. In Figure 9.2(d), the number of top-tags shows a maximum for one top-tag in case of

Table 9.1: Definition of primary selection for the dilepton channel in the full Run-2 analysis.

Primary selection
Preselection (see Table 8.1)
≥ 2 central small-R jets
2 OSSF leptons (e/μ)
$\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$

VLT quark events, while for VLB quarks 60% of the events have no top-tag. Thus, the presence of one top-tag is less frequent for VLB quark events compared to VLT quark events. Due to the higher number of top-quarks emerging from the VLT quark decays in events with at least one Z-boson, as illustrated in Figure 9.1, the larger number of top-tags for VLT quarks is explainable.

The number of V-, H- and top-tags changes significantly depending on the BR assumptions. The singlet model is used as benchmark, because it has the lowest BR into Z-boson final states with $Q \to H/Zq \approx 25\%$ and $Q \to Wq' \approx 50\%$. Thus, an optimization on these BRs leads to the highest possible sensitivity for the singlet model after the analysis optimization. The sensitivity for the doublet and 100% BRs into Zq is significantly higher due to more favorable BRs into Z-boson final states. A good performance of the analysis across the whole BR plane is attempted.

9.2 Signal and background modelling

For signal and background modelling, MC simulations are generated for relevant background processes and a range of VLB and VLT quark mass hypotheses. The strategy and technical aspects mirror the procedure of the 2015+2016 analysis, presented in Chapter 8. A summary of the generator and shower programs including the tunes and PDF sets is given in Table 8.5. However, some differences are present, which are discussed in this section, assuming that every not-mentioned modelling feature remains the same.

Overall, new recommendations for, e.g. object reconstruction, are incorporated in the new simulations for full Run-2 and the pile-up profile is adjusted to the data taking conditions in 2017 and 2018. For $t\bar{t}$, the recommendation for isolation criteria of electrons changed and the *FixedCutTightTrackOnly* WP with high signal acceptance became unavailable. Thus, electron isolation is dropped, and an uncertainty on the fake estimate of $t\bar{t}$ is applied, as described in Section 9.4. The VVV background was negligible in the 2015+2016 analysis and therefore is not considered in the following analysis. For the VV background, SHERPA 2.2.2 is used as generator instead of a mixture of SHERPA 2.2.1 and SHERPA 2.2.2. The simulation of signal samples is performed for VLQ mass hypotheses from 600 GeV to 2000 GeV. Within the range 1000–1800 GeV, samples with a step size of 100 GeV are produced, while 200 GeV are chosen outside of the region of high expected sensitivity. The optimization strategy, presented in the next section, only includes a subset of VLQ mass hypotheses due to the unavailability of a varying number of samples during the early stages of the analysis.

9.3 Analysis optimization

The optimization of the analysis is performed in different steps. First, a categorization in MCBOT tags and RC jet multiplicities is performed in Subsection 9.3.1. As a second step, kinematic cuts are added to the selection and the WPs of the V-tagger, H-tagger and top-tagger are optimized in Subsection 9.3.2. As a next step, the results of the optimization are used to perform a search for the final discriminant with the highest expected sensitivity in Subsection 9.3.3. In Subsection 9.3.4, a different approach for the categorization, based on a splitting in b-tagged jet multiplicities, is presented, which enhances the sensitivity further. Finally, the analysis strategy is simplified by merging some categories with low event numbers, presented in Subsection 9.3.5. Subsection 9.3.6 is focused on the handling of data and MC simulation mismodelling in the phase-space of this VLQ search, which introduces two control regions (CRs). The analysis strategy is finalized at this stage, and the event classification including all categories is summarized in Subsection 9.3.7. If not specified otherwise, the studies during the optimization are done with statistical uncertainties only.

9.3.1 Categorization split according to RC jet multiplicity

The idea for the starting point of the optimization is taken from the 2015+2016 analysis. It has proven beneficial to split the analysis with respect to the large radius jet multiplicity. The dilepton channel was split into a resolved ($< 2J$) and a boosted ($\geq 2J$) channel before. In this fashion, a splitting of RC jet multiplicities in combination with MCBOT tags is taken as a first approach. Furthermore, $m(Zb)$ is taken as the final discriminant and other choices are studied later in Subsection 9.3.3. Previously, the analysis selection ensured the presence of at least one b-tagged jet for the definition of $m(Zb)$. For the following studies, no b-tagged jet might be present in an event and thus, a value of zero is assigned to $m(Zb)$ in these cases. This procedure becomes obsolete after changing the strategy to a minimum number of b-tagged jets, defined in Subsection 9.3.4.

The following notation is kept throughout the chapter: For n RC jets, MCBOT assigns m_V V-tags, m_H H-tags and m_t top-tags with $n, m_V, m_H, m_t \in \mathbb{N}_0$ and $m_V + m_H + m_t \leq n$. Then, the abbreviation nRC-m_VV-m_HH-m_tt is chosen. An event with two RC jets, one H-tag and one top-tag - and thus no V-tag - is abbreviated as 2RC-0V-1H-1t.

The strategy is focused on the optimization of the VLT quark at first, assuming singlet BRs. The optimization of the search for the VLB quark is considered at a later stage, and its strategy is not expected to deviate from the optimization for the VLT quark significantly. For a VLT quark with a mass of $1400\,\text{GeV}$, all possible categories with and without splitting in RC jet multiplicities are built. The significance Z is calculated for each category for different choices of the three MCBOT WPs, and the seven most significant categories are combined to a collection, called the *basic* categories.

In a second step, six additional categories are added to the basic RC jet categories. Again, the categories are selected depending on the highest sensitivity. These categories are summarized as the *standard* RC jet splitting and are also based on a splitting according to the number of MCBOT tags. The categories belonging to the basic and standard RC jet splitting are listed in Table 9.2. The basic categories are a subset of the standard splitting and are highlighted. It should be noted that these categories do not mimic the truth final states listed in Figure 9.1, but can rather correspond to any combination of MCBOT tags. Multiple reasons can lead to no or a wrong identification of an RC jet:

Table 9.2: Set of basic and standard RC jet categories. The basic categories are highlighted. The category "≥ 3 tags" summarizes all categories with any combination of V-, H- and top-tags yielding at least three MCBOT tags in total. The standard RC jet categories include all of the 13 regions.

≥ 3 RC jets	2 RC jets
≥ 3 tags	
0V-0H-2t	0V-0H-2t
1V-0H-1t	1V-0H-1t
0V-1H-1t	0V-1H-1t
0V-0H-1t	0V-0H-1t
	1V-1H-0t
	2V-0H-0t
	0V-1H-0t
	1V-0H-0t

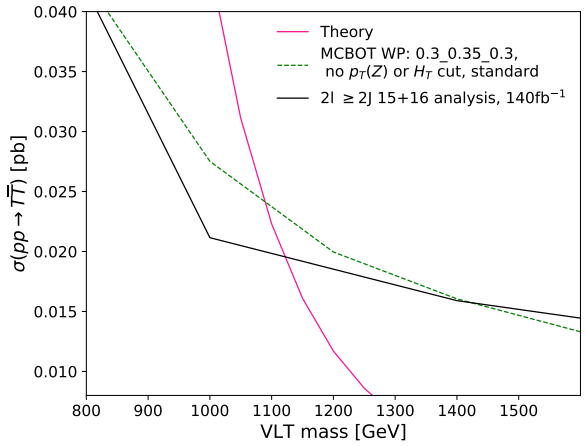


Figure 9.3: Expected cross-section limits at 95% CL for $T\bar{T}$ production with $m(Zb)$ as final discriminant and a categorization according to the standard RC jet splitting. The cuts of the primary selection are applied. Furthermore, the expected limits from the $2l \geq 2J$ channel of the 2015+2016 analysis are depicted. The number of MC events corresponds to an integrated luminosity of 140 fb^{-1}.

A misidentification between the four MCBOT classes occurs, if RC jet properties are similar or if e.g. the decay products of a top-quark decay are only partially contained in the RC jet. Furthermore, RC jets might not reach the minimum p_{T} threshold for an MCBOT tag, which is $p_{\mathrm{T}} > 150$ GeV. Additionally, the decay products from the two VLQ decays can overlap and thus the RC jet content can hardly be identified correctly. Therefore, an inclusive strategy is chosen, where all possible combinations of MCBOT tag multiplicities are considered.

In order to compare the sensitivity of the 2015+2016 analysis with the sensitivity of the new strategy, the $2\ell \geq 2J$ channel of the previous analysis is scaled to 140 fb^{-1}, which was the first integrated luminosity estimate for the full Run-2 dataset. In Figure 9.3, the expected limits at 95% CL on the cross-section for $T\bar{T}$ production are depicted assuming singlet BRs for both strategies and using statistical uncertainties only. As cut values for the MCBOT WPs c_V, c_H and c_t, values of $c_V > 0.3$,

$c_H > 0.35$ and $c_t > 0.3$ are chosen for this comparison. It can be seen that the strategy of the $2\ell \geq 2J$ channel from the 2015+2016 analysis is more sensitive up to a VLT mass of 1400 GeV, when the MCBOT categorization cross-section limits become lower. However, more information about the kinematics in an event can be exploited for the MCBOT categorization in order to improve the sensitivity over the whole VLQ mass range. Therefore, the analysis requirements are expanded in a next optimization step and minimum requirements on $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ are considered, since these kinematic variables have shown great discrimination power, as discussed in Section 8.2.

9.3.2 Kinematic cuts and MCBOT working point optimization

In order to find the optimal $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts, an optimization of the limits on the pair production cross-section is performed for VLT quarks. Furthermore, the WPs of MCBOT influence the limits, since their choice determines how many signal and background events are contained in each category. Therefore, a five-dimensional optimization is performed where the $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$, as well as the cut values of the WPs c_V, c_H and c_t of the V-, H- and top-tagger are varied.

A variety of different limit curves is depicted for the standard RC jet splitting in Figure 9.4. The family of curves with solid, different-color lines except the pink theory line depicts the limit change from lower to higher $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts. Different line styles correspond to different MCBOT WPs, denoted as cuts on the three taggers in the order c_V_c_H_c_t. It can be seen that the influence of the different WPs is negligible compared to the improvements achieved by $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts. Only a subset of the studied MCBOT WPs is shown here, which were varied in any combination between 0.25 and 0.5 in steps of 0.05. For VLT quarks simulated according to the singlet model, the sensitivity is expected to almost reach a mass limit of 1200 GeV for the best cut combinations. One of the most sensitive cut combination with MCBOT WPs 0.3_0.35_0.3, $p_{\mathrm{T}}(Z) > 300$ GeV and $H_{\mathrm{T}}(\mathrm{jets}) > 1150$ GeV is chosen as part of a preliminary analysis selection.

As discussed in Chapter 6, two different versions of MCBOT were trained: The standard and the b-tag enriched FLat-In-p_{T} (FLIP) taggers. In Figure 9.5, a comparison of both taggers with the WP combination of 0.3_0.35_0.3 is made for the basic categorization. A significant difference is observed: The standard MCBOT reaches significantly smaller and thus stronger limits than the b-tag enriched tagger. For the 0.3_0.35_0.3 WPs, the signal efficiencies of the standard MCBOT for VLT quarks with $m = 1400$ GeV and singlet BRs are 55%_41%_54%. For the b-tag enriched MCBOT, the efficiencies correspond to 55%_35%_39%. Although the efficiencies are higher for standard MCBOT, the large difference in limits cannot be explained with the different efficiencies: It was shown in Figure 9.4 that the different MCBOT WPs lead to comparable limits. The signal efficiencies for the standard MCBOT at the WPs 0.35_0.5_0.4 are 51%_32%_47%. Changes of 7% in signal efficiency lead to small limit deviations. The statement can be extended to a change of 15% in the top-quark tagging efficiency between the b-tag enriched and standard tagger at $c_t = 0.3$. However, a worse background rejection at this cut for the b-tag enriched tagger is likely to cause the decreased sensitivity. The background composition in the analysis, especially in the absence of requirements on the b-tagged jet multiplicity as in this comparison, is likely to be more similar to the standard test sample, used for evaluation in Section 6.8. Thus, the b-tag enriched tagger is expected to lead to a categorization with more background in the categories with a top-tag, which diminishes the signal purity in the categories and leads to a worse sensitivity. In the following, only the standard MCBOT tagger is considered.

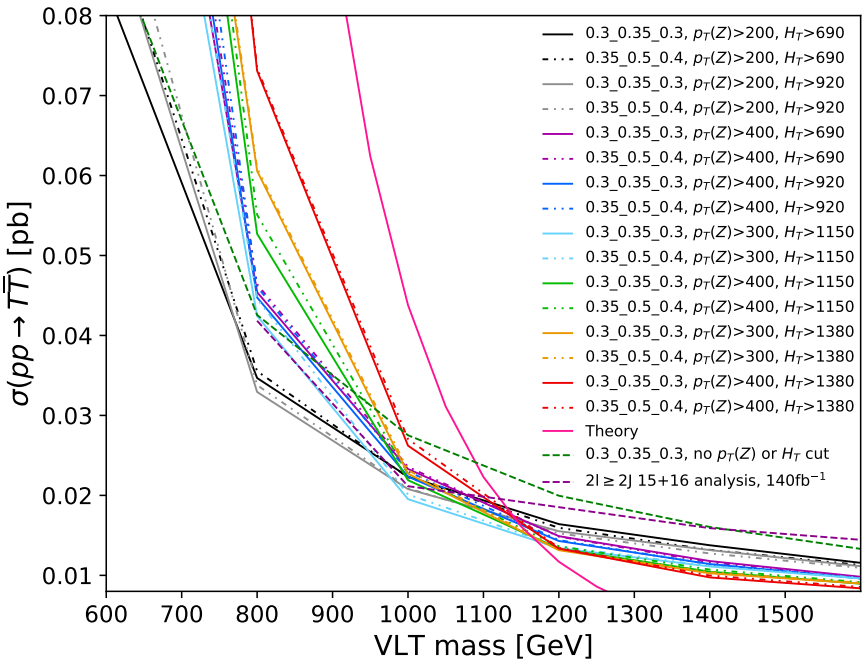


Figure 9.4: Expected cross-section limits for $T\bar{T}$ production assuming singlet BRs with $m(Zb)$ as final discriminant and the standard RC jet splitting at 95% CL. With the family of curves, the optimization of the three MCBOT WPs, as well as the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts is illustrated. Pairs of curves with the same color correspond to the same $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts with changed MCBOT WPs in different line styles. For reference, the limits from the $2\ell \geq 2J$ channel of the 2015+2016 analysis using $140\,\mathrm{fb}^{-1}$, as well as the standard RC jet splitting without $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts are depicted. The number of MC events corresponds to $140\,\mathrm{fb}^{-1}$. The units of the cuts on $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ in the legend are given in GeV.

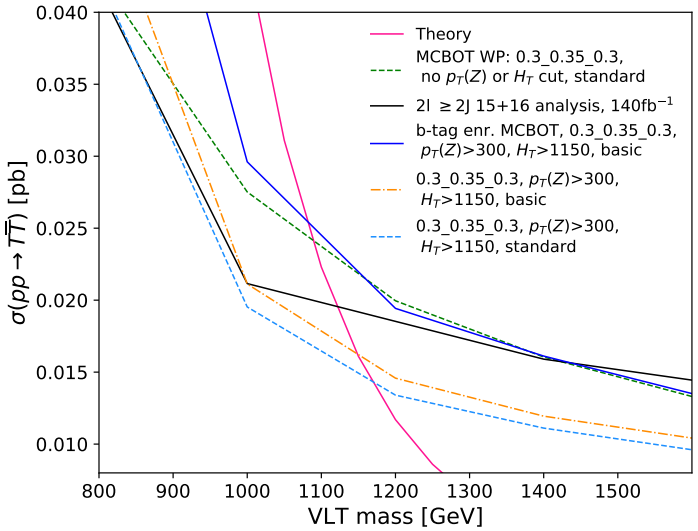


Figure 9.5: Expected cross-section limits for $T\bar{T}$ production with singlet BRs and the standard RC jet splitting at 95% CL. The limits for the b-tag enriched training of MCBOT in comparison to the standard MCBOT are illustrated. A comparison of the basic and the standard RC jet splitting is given. The limits from the $2\ell \geq 2J$ channel of the 2015+2016 analysis using $140\,\mathrm{fb}^{-1}$, as well as the standard RC jet splitting without $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts are illustrated for reference. The number of MC events corresponds to $140\,\mathrm{fb}^{-1}$ and the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts are given in units of GeV.

In Figure 9.5, also a comparison of the limit improvement of the basic and the standard RC jet splitting is made. Adding six categories to the basic selection for the standard selection leads to an improved sensitivity. For $m_T = 1200\,\text{GeV}$, the cross-section limit for singlet BRs decreases by 8% when the standard categorization is used. Adding more categories improves the limit and motivates the idea to include as many of the events of the dilepton channel as possible.

9.3.3 Final discriminant search

As final discriminant, the variable $m(Zb)$ was chosen as a starting point, as described in Subsection 9.3.1. A comparison of different discriminants is made in this subsection in order to improve the analysis strategy or ensure that the best choice was made with $m(Zb)$. First, a shape comparison of the variables is performed and secondly, the expected mass limits are compared among different variable choices.

$m(Zb)$

The invariant mass of the leptonic Z-boson candidate and the leading b-tagged jet, $m(Zb)$, has large discrimination power between signal and background and was used as final discriminant in Chapter 8. Without the presence of a b-tagged jet, the invariant mass is set to zero. Therefore, the shape of $m(Zb)$ varies significantly depending on the requirement of no or at least one b-tagged jet. In order to compare the shape of $m(Zb)$ without a bias to lower values due to a lack of a b-tagged jet, a cut on at least one b-tagged jet at 77% signal efficiency is included in the illustration of $m(Zb)$ in Figure 9.6(a). The mass of $1200\,\text{GeV}$ is taken as benchmark for VLT and VLB quarks, since the expected limits for VLT quarks with singlet BRs showed the expected significance to be in this mass range. Besides the requirement of at least one b-tagged jet, the primary selection and the requirements on $p_\text{T}\,(Z) > 300\,\text{GeV}$ and $H_\text{T}(\text{jets}) > 1150\,\text{GeV}$ are applied. It can be seen that $m(Zb)$ is able to reconstruct the VLB quark resonance at the mass of $1200\,\text{GeV}$, while slightly lower masses are present for VLT quarks, as discussed for Figure 8.4. After the application of the high kinematic thresholds on $p_\text{T}\,(Z)$ and $H_\text{T}(\text{jets})$, the remaining background events become similar to the VLB and VLT quark signals, but a significant separation power between signal and background remains for $m(Zb)$, as depicted in Figure 9.6(a).

$m(Z$+subleading b-tagged jet)

During the production of a VLQ pair, each VLQ is characterized by at least one b-quark in the decay chain, which is either produced directly or can occur as product of a top-quark, H- or V-boson decay. When constructing $m(Zb)$, the b-tagged jet with the highest p_T is chosen. As illustrated in Figure 9.6(a), the resonance can be reconstructed with this variable. However, it is not clear, if this is the only possible choice to obtain the resonance: The b-tagged jet with the second largest p_T, called the subleading b-tagged jet or 2^nd b-tagged jet, could also yield the right invariant mass combination for the VLQ. In Figure 9.6(b), the invariant mass of the Z-boson candidate and the 2^nd b-tagged jet, denoted as $m(Z\text{+ }2^\text{nd}\,b)$, is illustrated for background and signal. In addition to the primary selection and the previously defined $p_\text{T}\,(Z)$ and $H_\text{T}(\text{jets})$ cuts, a requirement of at least two b-tagged jets is made, which ensures that the discrimination power is only compared among events with high b-tagged

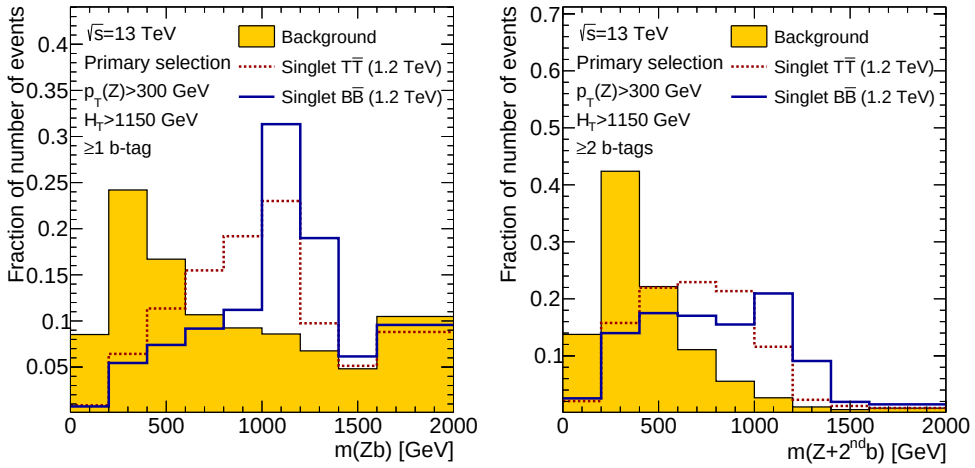


Figure 9.6: Depiction of a) $m(Zb)$ for at least one b-tagged jet and (b) $m(Z+2^{\text{nd}}\,b)$ for at least two b-tagged jets in the event. Furthermore, the primary selection and the requirements on $p_{\text{T}}\,(Z) > 300\,\text{GeV}$ and $H_{\text{T}}(\text{jets}) > 1150\,\text{GeV}$ are applied. Both pair-production benchmark signal masses for VLT and VLB quarks at 1200 GeV are shown, as well as a sum of all background processes. The under- and overflow bins are included.

jet multiplicities. From the VLB quark peak at 1200 GeV it can be seen that the mass resonance of the VLB quark is reconstructed in some events. Overall, both signal processes yield lower invariant masses for the variable $m(Z+2^{\text{nd}}\,b)$ than for $m(Zb)$. However, the background distribution shows a significant fraction of events with low $m(Z+2^{\text{nd}}\,b)$ values. Therefore, the separation power for signal and background is larger and the sensitivity is likely to exceed the sensitivity in an analysis which uses $m(Zb)$.

H_{T}(jets)

The hadronic activity in an event is contained in the variable $H_{\text{T}}(\text{jets})$ and has shown to have good discrimination power. For completeness, a comparison of the expected limits also includes $H_{\text{T}}(\text{jets})$. In Figure 9.7(a), $H_{\text{T}}(\text{jets})$ is illustrated after the primary selection and the requirement of $p_{\text{T}}\,(Z) > 300\,\text{GeV}$. The good discrimination power is shown once more and the previously optimized cut on the standard RC jet splitting of $H_{\text{T}}(\text{jets}) > 1150\,\text{GeV}$ is well motivated in this illustration: Above this threshold more than 95% of background events are suppressed.

$m(Zt)$

Another possibility for the definition of a final discriminant becomes available with the use of MCBOT. As for most signal events more than one b-tagged jet is expected, the number of expected top-quarks is also higher for signal than for background. For example, there is no hadronic top-quark decay expected in $t\bar{t}$ events, when a requirement of two high-p_{T} leptons is made. Therefore, an approach for the reconstruction of the $T \to Z(\ell\ell)t$ decay could lead to better discrimination power, especially for the VLT quark search. The RC jet with the highest top-tagger output is combined with the Z-boson candidate to build the invariant mass $m(Zt)$. The $m(Zt)$ distribution is visualized in Figure 9.7(b). A

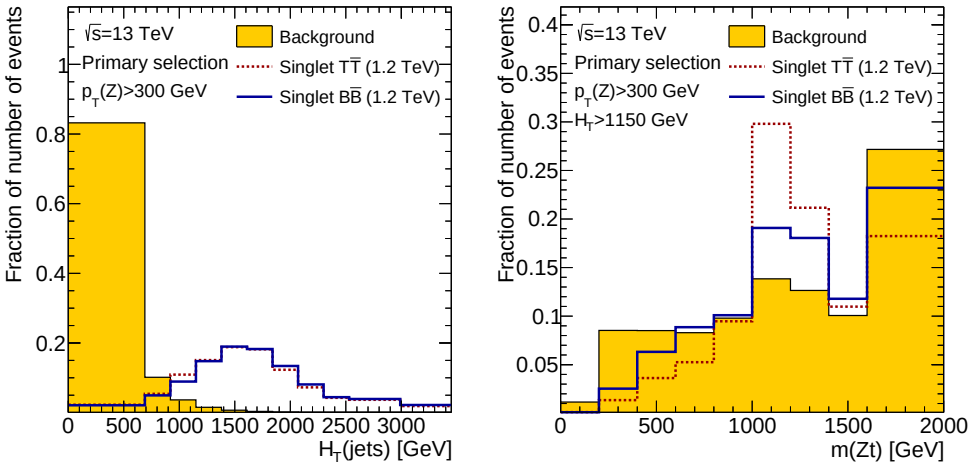


Figure 9.7: Depiction of a) $H_{\mathrm{T}}(\mathrm{jets})$ after applying the $p_{\mathrm{T}}(Z) > 300\,\mathrm{GeV}$ cut and b) $m(Zt)$ after applying the $p_{\mathrm{T}}(Z) > 300\,\mathrm{GeV}$ and $H_{\mathrm{T}}(\mathrm{jets}) > 1150\,\mathrm{GeV}$ cuts. Both pair-production benchmark signal masses for VLT and VLB quarks at 1200 GeV are shown, as well as a sum of all background processes. The under- and overflow bins are included.

high peak is observed for VLT quarks at the benchmark mass, which shows that the reconstruction of $m(Zt)$ works well. Additionally, VLB quarks show a distinct peak close to 1200 GeV, which shows that the b-tagged jet from the $B \to Z(\ell\ell)b$ decay is occasionally the RC jet with the highest top-tagger output. However, the background distribution also shows a peak at the VLQ masses and is located at high values of $m(Zt)$. Thus, the discrimination power is worse than for $m(Zb)$ and $m(Z+2^{\mathrm{nd}}\,b)$. As discussed in the feature investigation for MCBOT in Section 6.8, the mass of background jets is shaped by the tagger and thus a high top-tagger output corresponds to a high RC jet mass. Therefore, background jets with high masses are more likely to be mistagged as top-quark jets and thus, might enter the $m(Zt)$ distribution at high values, since the definition of $m(Zt)$ does not necessarily include the presence of a true top-quark.

Other discriminating variables

In the previous subsection, the most promising choices for final discriminants were discussed in more detail. Nevertheless, other variables are studied as well. Instead of using only the hadronic activity $H_{\mathrm{T}}(\mathrm{jets})$ in an event, the scalar sum of all transverse jet and lepton momenta is tested, denoted as S_{T}. Lower expected limits with S_{T} as final discriminant are found compared to the use of $H_{\mathrm{T}}(\mathrm{jets})$. Furthermore, the use of S_{T} prevents the use of $p_{\mathrm{T}}(Z)$ or $H_{\mathrm{T}}(\mathrm{jets})$ cuts due to large correlations and thus the variable is dropped as possible choice for a final discriminant.

Since the definition of $m(Zb)$ and $m(Z+2^{\mathrm{nd}}\,b)$ is only properly defined in the presence of b-tagged jets, a combination of these variables with the invariant mass of the Z-boson candidate and the subleading jet is investigated. This variable might contain more event information than $m(Zb)$ and $m(Z+2^{\mathrm{nd}}\,b)$ without the requirement of b-tagged jets. Since the discrimination power of $m(Z+2^{\mathrm{nd}}\,b)$ seems to be most promising at this point of the optimization, the variable $m(Z+2^{\mathrm{nd}}\mathrm{jet})$ could be discriminating as well, if the corresponding jet is not b-tagged, but belongs to the VLQ decay with

the leptonic $Z(\ell\ell)$-boson. A limit comparison of these two variables, $m(Zb)/m(Z + 2^{\text{nd}}\text{jet})$ and $m(Z\text{+ }2^{\text{nd}}\, b)/m(Z + 2^{\text{nd}}\text{jet})$, is done in the following. The notation means that the first variable is used, if at least one or two b-tagged jets are present for $m(Zb)$ and $m(Z\text{+ }2^{\text{nd}}\, b)$, respectively. If this requirement is not fulfilled due to a lack of b-tagged jets, $m(Z + 2^{\text{nd}}\text{jet})$ is used as discriminant.

Instead of the previously mentioned definition of $m(Zt)$, a different $m(Zt)$ definition with a minimum requirement on the top-tagger output of 0.5 is created. The idea is to avoid the selection of RC jets in an $m(Zt)$ definition, if the event does not contain a top-jet-like RC jet. A threshold of 0.5 ensures that no other MCBOT class has a larger DNN output. For events with RC jets not passing this threshold, $m(Z\text{+ }2^{\text{nd}}\, b)$ is used as final discriminant, if at least two b-tagged jets are present. If not, the subleading jet is taken for the reconstruction of the invariant mass. This variable is denoted as $m(Zt)/m(Z\text{+ }2^{\text{nd}}\, b)/m(Z + 2^{\text{nd}}\text{jet})$.

Choice of final discriminant

The expected mass limits at 95% CL between the most promising discriminants are compared for VLT quarks with the singlet BRs in order to choose a final discriminant. Table 9.3 shows different comparisons at different stages of the selection for the standard RC jet splitting with 13 categories. If no value is listed, the variable was either dropped due to lower limits or was studied afterwards. At primary selection level or with the additional requirement of $p_{\text{T}}\,(Z) > 300\,\text{GeV}$, the comparison of $m(Zb)$ and $H_{\text{T}}(\text{jets})$ shows that $m(Zb)$ performs at least as good as $H_{\text{T}}(\text{jets})$. Since an additional requirement on $H_{\text{T}}(\text{jets})$ is beneficial for the optimization, $H_{\text{T}}(\text{jets})$ is dropped as a candidate for the final discriminant. Furthermore, $m(Zt)$ is discarded at this stage, since the limits do not show a better performance than $m(Zb)$. It should be noted that these comparisons are done with a different MCBOT WP choice ($0.35\_0.5\_0.37$) compared to the choice presented in Subsection 9.3.2 ($0.3\_0.35\_0.3$). However, the comparison is consistent, since the same WPs are chosen for the categorizations with different final discriminants, which are then compared. Additionally, it was shown in Subsection 9.3.2 that a WP difference of this size leads only to small limit deviations.

Table 9.3: Expected mass limits at 95% CL for the VLT quark with singlet BRs for different final discriminants and different selections. As MCBOT WP $0.3\_0.35\_0.3$ is used. For limits with the superscript *) a different MCBOT WP is applied: $0.35\_0.5\_0.37$. The cuts above include all cuts from the previous columns in addition to the one mentioned. Note: If no b-tagged jet is present for the definition of $m(Zb)$ and $m(Z\text{+ }2^{\text{nd}}\, b)$, the variable's value is set to zero.

	Primary sel.	+$p_{\text{T}}\,(Z)$ cut	+$H_{\text{T}}(\text{jets})$ cut	+$\geq 2b$-tags
$m(Zb)$	$1036^{*)}$	$1052^{*)}$	-	1167
$m(Zb)/m(Z + 2^{\text{nd}}\text{jet})$	-	-	1126	-
$m(Z\text{+ }2^{\text{nd}}\, b)$	-	-	-	1186
$m(Z\text{+ }2^{\text{nd}}\, b)/m(Z + 2^{\text{nd}}\text{jet})$	-	-	1115	-
$m(Zt)$	-	$1050^{*)}$	-	-
$H_{\text{T}}(\text{jets})$	$1022^{*)}$	$1054^{*)}$	-	-
$m(Zt)/m(Z\text{+ }2^{\text{nd}}\, b)/$ $m(Z + 2^{\text{nd}}\text{jet})$	-	-	1132	1172

With the additional requirement of $H_{\mathrm{T}}(\mathrm{jets}) > 1150\,\mathrm{GeV}$, the three most promising discriminants are compared in Table 9.3: The discriminants $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ are combined with $m(Z + 2^{\mathrm{nd}}\mathrm{jet})$, since no requirement on b-tagged jets is made. The same argumentation holds for the third variable in the comparison, $m(Zt)/m(Z\text{+ }2^{\mathrm{nd}}\,b)/m(Z + 2^{\mathrm{nd}}\mathrm{jet})$. It is observed that the variable $m(Zt)/m(Z\text{+ }2^{\mathrm{nd}}\,b)/m(Z + 2^{\mathrm{nd}}\mathrm{jet})$ performs best among these three choices. Thus, it is compared to the variables $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ in a region with at least two b-tagged jets, which ensures the proper definition of $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$. It should be noted that the previous definitions of $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ were also meaningful in the absence of b-tagged jets, since a good separation for events with at least one or two b-tagged jets could still be achieved for $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$, respectively. However, a less nested variable definition is desired and an additional requirement on the number of b-tagged jets has shown to have good discrimination power in Chapter 8. From Table 9.3 it can be seen that the best expected limit is obtained for $m(Z\text{+ }2^{\mathrm{nd}}\,b)$. Thus, it can be concluded that $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ shows the best performance and is chosen as final discriminant.

An illustration of the limit comparison between $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ with the standard RC jet splitting is shown in Figure 9.8. In Figure 9.8(a), the expected limits for VLT quarks with singlet BRs are shown. It can be seen that the use of $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ gives lower and thus stronger limits than $m(Zb)$ over the whole mass range from 800 GeV to 1400 GeV.

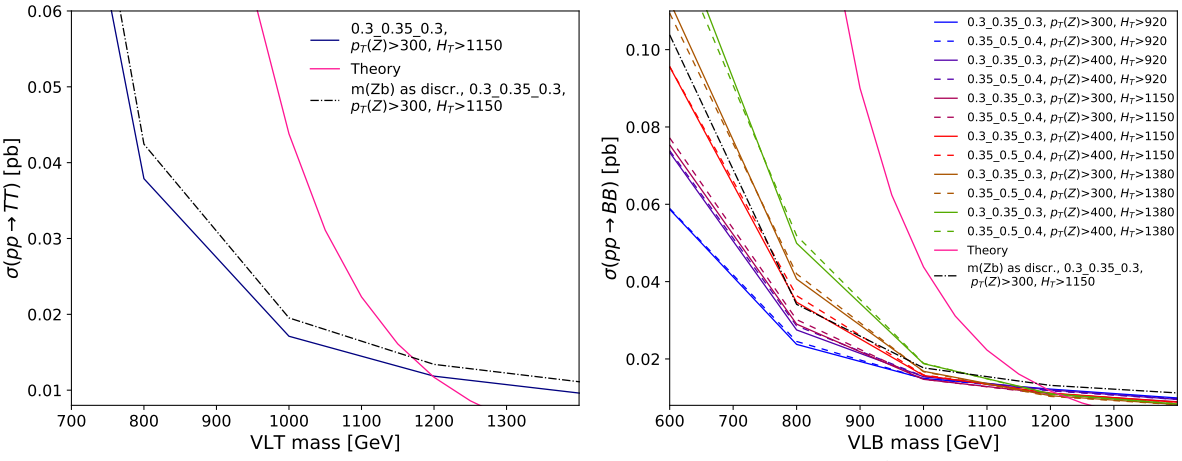


Figure 9.8: Expected cross-section limits at 95% CL for a) $T\bar{T}$ and b) $B\bar{B}$ production assuming singlet BRs. The variable $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ is used as final discriminant for the standard RC jet splitting categories, starting in the legend with 0.3_0.35_0.3. Furthermore, the same categorization and selection is depicted for the WP choice of 0.3_0.35_0.3 with the alternative variable $m(Zb)$. For the VLB quark optimization, a few five-dimensional cut combinations including the most sensitive are illustrated in addition. The selection applied in both figures contains the primary selection and the optimized cuts for $p_{\mathrm{T}}\,(Z)$, $H_{\mathrm{T}}(\mathrm{jets})$ (given in units of GeV) and the MCBOT WPs. The number of MC events corresponds to $140\,\mathrm{fb}^{-1}$.

In Figure 9.8(b), the same setup as in Figure 9.8(a) is shown for VLB quarks with singlet BRs. Compared to $m(Zb)$, the limits for $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ are also stronger for VLB quarks across the whole mass range. In addition, the different cut combinations for the MCBOT WPs, the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ are illustrated. The set of cuts, optimized for VLT quarks, is very close to optimal for VLB quarks and thus the strategy for both VLQ searches is harmonized with the same MCBOT WPs, $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ choices of $c_V = 0.3$, $c_H = 0.35$, $c_t = 0.3$, $p_{\mathrm{T}}\,(Z) > 300\,\mathrm{GeV}$ and $H_{\mathrm{T}}(\mathrm{jets}) > 1150\,\mathrm{GeV}$.

9.3.4 Categorization split according to b-tagging information

The MCBOT categorization according to the standard RC jet splitting with 13 categories has shown to improve the expected VLT quark limits for singlet BRs in comparison to the $2\ell \geq 2J$ limits from the 2015+2016 analysis. The $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts are necessary to reach this improvement. However, the current categorization only covers a small proportion of the events in the dilepton channel of the $Z(\ell\ell)t/b + X$ analysis. More, but less sensitive categories than the ones listed in Table 9.2 can be built with a splitting in RC jet multiplicities: Categories with zero and one MCBOT tag, as well as a part of the categories with two MCBOT tags are missing. However, additional categories would complicate the profile likelihood fit significantly and a better approach is sought.

The studies in Subsection 9.3.3 showed that $m(Zb)$ and $m(Z\text{+ }2^{\text{nd}}\,b)$ are (among) the best discriminating variables. The presence of at least one or two b-tagged jets is beneficial for the definition of $m(Zb)$ and $m(Z\text{+ }2^{\text{nd}}\,b)$, respectively. Thus, a reasonable deduction in the multiplicity of categories includes an MCBOT categorization on top of a splitting by b-tagged jet multiplicities. The b-tagged jet multiplicity has shown to discriminate well between signal and background in Figure 8.2(b). Over 80% of background events were suppressed by requiring at least one b-tagged jet. The events with at least one b-tagged jet are arranged in two categories: Events with one b-tagged jet and events with at least two b-tagged jets. For both groups, a categorization according to MCBOT-tags is defined. Regions with one b-tagged jet use $m(Zb)$ as final discriminant, while categories with at least two b-tagged jets are able to exploit the advantages of the most sensitive final discriminant $m(Z\text{+ }2^{\text{nd}}\,b)$.

The new categorization is summarized in Table 9.4. Eleven categories are defined for each b-tagged jet multiplicity using the 77% signal efficiency WP, which are denoted as $1b$ and $\geq 2b$. All categories with at least three MCBOT tags are merged in the "≥ 3 tags" category due to extremely low background yields in any further splitting of categories, making a shape fit in more than one bin meaningless. The categories with one and two MCBOT tags are defined such that all possible tag combinations for V-, H and top-tags are included. Due to a possible association of an RC jet to the wrong MCBOT class with respect to the true jet origin, it is reasonable to make use of all events in the dilepton channel above certain kinematic requirements, like the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts. Therefore, also two categories with zero MCBOT tags are built for $1b$ and $\geq 2b$. The sensitivity of the categories used in the final categorization is compared in Subsection 9.3.7.

An additional change is applied to the analysis strategy in order to improve the sensitivity further: Instead of a cut on $H_{\mathrm{T}}(\mathrm{jets})$, a cut on $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ is tested, which describes the sum of $H_{\mathrm{T}}(\mathrm{jets})$ and the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. Table 9.5 shows the expected cross-section limits and mass limit for the hypothesis of a VLT quark with singlet BRs and with $m_T = 1200\,\mathrm{GeV}$. Cut values of $H_{\mathrm{T}}(\mathrm{jets})$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ according to both variables' bin widths of 230 GeV are chosen for the sensitivity test. The mass limits for minimum requirements on $H_{\mathrm{T}}(\mathrm{jets})$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ of $1150\,\mathrm{GeV}$, $1380\,\mathrm{GeV}$, $1610\,\mathrm{GeV}$ and $1840\,\mathrm{GeV}$ are compared. The study shows that a requirement of at least $1380\,\mathrm{GeV}$ on $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ increases the expected limits further compared to a cut on $H_{\mathrm{T}}(\mathrm{jets})$, and therefore it is used as replacement for the cut on $H_{\mathrm{T}}(\mathrm{jets}) > 1150\,\mathrm{GeV}$ in the following analysis. Any of the cut values higher than $1380\,\mathrm{GeV}$ for $H_{\mathrm{T}}(\mathrm{jets})$ or $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ resulted in lower expected mass limits and is thus not listed in Table 9.5. A selection including the primary selection, the requirements of $p_{\mathrm{T}}\,(Z) > 300\,\mathrm{GeV}$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}} > 1380\,\mathrm{GeV}$, as well as a cut on at least one b-tagged jet is denoted as the *full selection* in the following.

Table 9.4: MCBOT categorization in b-tagged jet multiplicities with 22 categories. The number of MCBOT tags and RC jets for each pair of categories is also listed.

# MCBOT tags	# RC jets	MCBOT categories in b-tag categories
≥ 3	≥ 3	≥ 3 tags $\cdot$ $\{1\,b, \geq 2\,b\}$
2	≥ 2	1V-0H-1t $\cdot$ $\{1\,b, \geq 2\,b\}$
		0V-1H-1t $\cdot$ $\{1\,b, \geq 2\,b\}$
		1V-1H-0t $\cdot$ $\{1\,b, \geq 2\,b\}$
		2V-0H-0t $\cdot$ $\{1\,b, \geq 2\,b\}$
		0V-2H-0t $\cdot$ $\{1\,b, \geq 2\,b\}$
		0V-0H-2t $\cdot$ $\{1\,b, \geq 2\,b\}$
1	≥ 1	1V-0H-0t $\cdot$ $\{1\,b, \geq 2\,b\}$
		0V-1H-0t $\cdot$ $\{1\,b, \geq 2\,b\}$
		0V-0H-1t $\cdot$ $\{1\,b, \geq 2\,b\}$
0	≥ 0	0 tags $\cdot$ $\{1\,b, \geq 2\,b\}$

Table 9.5: Expected cross-section limits at 95% CL for $m_T = 1200\,\text{GeV}$ and mass limits for the VLT quark assuming the singlet model and applying different cuts on $H_\text{T}(\text{jets})$ and $H_\text{T}(\text{jets}) + E_\text{T}^\text{miss}$. As MCBOT WPs, the combination 0.3_0.35_0.3 is used. In addition to the primary selection, $p_\text{T}\,(Z) > 300\,\text{GeV}$ and the categorization described in Table 9.4 is applied. The number of MC events corresponds to $140\,\text{fb}^{-1}$.

Limit	$H_\text{T}(\text{jets})$ $> 1150\,\text{GeV}$	$H_\text{T}(\text{jets})$ $> 1380\,\text{GeV}$	$H_\text{T}(\text{jets}) + E_\text{T}^\text{miss}$ $> 1150\,\text{GeV}$	$H_\text{T}(\text{jets}) + E_\text{T}^\text{miss}$ $> 1380\,\text{GeV}$
Mass [GeV]	1206	1210	1207	1220
Cross-section [pb]	0.011	0.011	0.011	0.010

An illustration of the expected singlet limits for the MCBOT categorization using b-tagged jet splitting is shown in Figure 9.9. For VLT and VLB quarks, a comparison is made to the previously optimized selection with the standard RC jet splitting in Figures 9.9(a) and 9.9(b), respectively. Furthermore, a selection purely based on the RC jet multiplicity and the b-tagged jet splitting is defined: Instead of an MCBOT categorization, the number of RC jets in an event with exactly one or at least two b-tagged jets is used to define categories. A splitting into zero, one, two and at least three RC jets is performed, which leads to eight categories for the "no MCBOT" categorization in Figure 9.9.

It can be seen that MCBOT in combination with the b-tagged jet splitting performs best, which shows that the analysis strategy performs better than the strategy chosen for the 2015+2016 analysis in the $2\ell \geq 2J$ channel using $140\,\text{fb}^{-1}$. Furthermore, a significant improvement of the sensitivity is reached due to the use of MCBOT, which is shown in the ratio plot by comparing the new strategy with the "no MCBOT" limits. For the singlet BRs, the improvement starts at 26% and reaches around 34% for VLT quarks at a mass hypothesis of 2 TeV. The sensitivity for VLB quarks with singlet BRs reaches from around 13% at low to almost 17% at high masses. For VLT quarks, the improvement is even larger for the doublet and 100% BR into Z-boson final states, where the performance

gain reaches up to 45% for high masses. The corresponding plots are shown in Appendix C.1. The reason for a larger limit improvement with the MCBOT categorization for VLT quarks compared to VLB quarks is the multiplicity of boosted objects in the final states. As mentioned earlier, the pair production of VLT quarks produces more top-quarks, and they are always present in the final states, independent of the BR assumptions for singlet, doublet or 100%. For VLB quarks, the situation is different: A significant number of top-quarks is produced in the singlet BRs with $B \to Wt$, but this channel vanishes for the (B, Y) doublet model and the 100% BR into Zb final states. Thus, the limits for VLB quarks are closer to the sensitivity of a categorization without MCBOT compared to the VLT quark analysis.

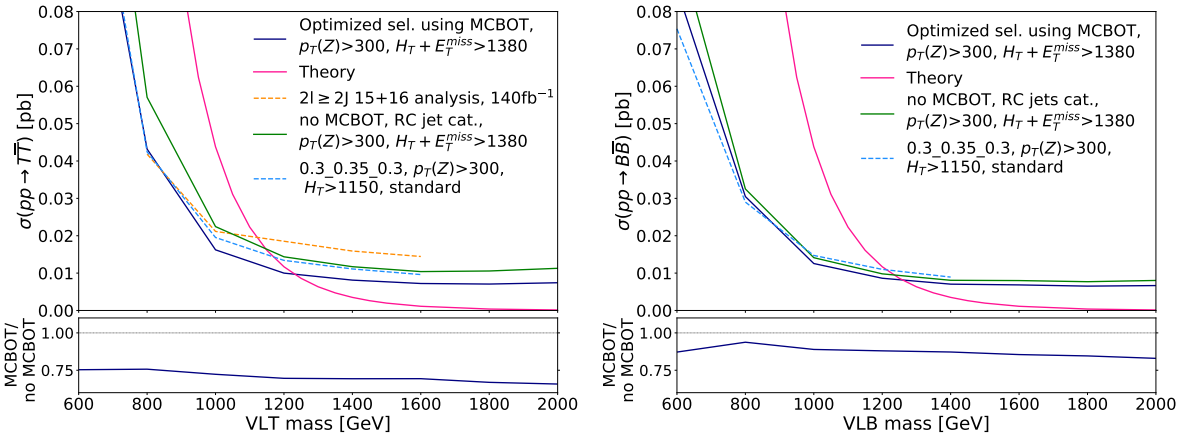


Figure 9.9: Expected cross-section limits at 95% CL for a) $T\bar{T}$ and b) $B\bar{B}$ production assuming singlet BRs. The MCBOT categorization with splitting in b-tagged jet multiplicities and the optimized $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ cuts is shown as "optimized selection". In addition, a categorization solely based on RC jet multiplicity splitting without the use of MCBOT is depicted, denoted as "no MCBOT". Furthermore, the standard categorization is shown with the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets})$ cuts optimized in Section 9.3.2. For VLT quarks, the expected limit for the selection of the $2\ell \geq 2J$ channel of the 2015+2016 analysis is shown. The ratio plot shows the optimized selection divided by the categorization without MCBOT. The number of MC events corresponds to $140\,\mathrm{fb}^{-1}$. Kinematic cuts are given in units of GeV.

9.3.5 Merging of categories

An MCBOT categorization with a splitting in b-tagged jet multiplicities reduces the number of categories in comparison to the standard RC jet splitting. However, still a multitude of 22 categories remains. Some of these regions have a background expectation of one event or less. A shape fit in such a region is not meaningful. Furthermore, the application of a profile likelihood fit is simplified significantly if the asymptotic approximation is valid, which requires a minimum of a few expected events (ranging from two to ten events depending on the expected significance [143]) to approach the result of pseudo-experiments. Therefore, the categories with the lowest yields are combined. The sensitivity of many category combinations is checked and it is observed that the combination of high sensitivity regions with each other gives the highest sensitivity of the combination. When comparable mass limits occur, the category combination with the highest background yield is chosen. The new categories are presented in Table 9.6. It can be seen that the zero and one MCBOT tag categories

are not changed, while most of the categories with at least two MCBOT tags are combined. The loss in expected mass limits by combining these categories in comparison to the categories introduced in Table 9.4 is marginal for both VLB and VLT quarks. The expected limit for VLB quarks decreases from 1241 GeV to 1239 GeV, while for VLT quarks the limit drops from 1219 GeV to 1214 GeV.

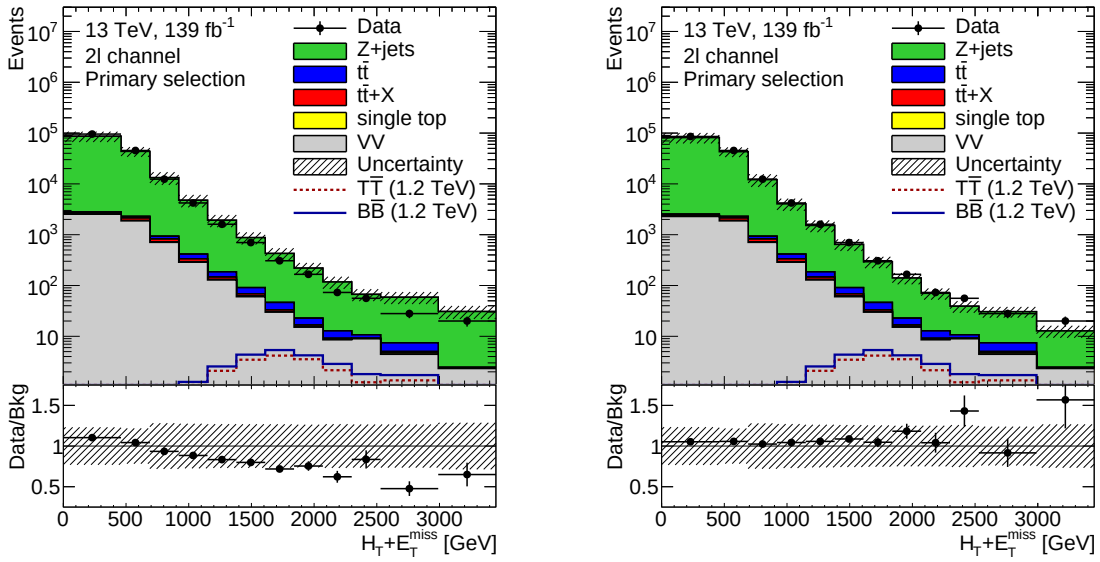
Table 9.6: Optimized set of categories for the VLT and VLB quark searches. All the MCBOT categories are split into one and at least two b-tagged jets. For one b-tagged jet $m(Zb)$ is chosen as discriminant, while for at least two b-tagged jets $m(Z+\ 2^{\text{nd}}\ b)$ is used.

# MCBOT tags	$1b$-tag	$\geq 2b$-tags
0	0V-0H-0t	0V-0H-0t
1	1V-0H-0t	1V-0H-0t
1	0V-1H-0t	0V-1H-0t
1	0V-0H-1t	0V-0H-1t
2	2V-0H-0t + 0V-2H-0t + 1V-0H-1t = 2V+2H+1V1t	1V-1H-0t + 2V-0H-0t + 0V-2H-0t + 0V-0H-2t = 1V1H+2V+2H+2t
2	0V-0H-2t + 1V-0H-1t = 2t+1V1t	0V-1H-1t
≥ 2	1V-1H-0t + $\geq$ 3 MCBOT tags = 1V1H+$\geq$3tags	1V-0H-1t + $\geq$ 3 MCBOT tags = 1V1t+$\geq$3tags

9.3.6 Definition of control regions

The study of agreement between data and MC simulation, which is discussed in more detail in Section 9.5, shows a mismodelling of variables sensitive to jet activity. Examples are the small-R jet multiplicity or the momenta of the leading and subleading small-R jets. Hence, the variable $H_{\mathrm{T}}(\mathrm{jets})$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$, which are highly correlated to the jet momenta, are affected. In Figure 9.10(a), the $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ variable is illustrated at primary selection level including systematic uncertainties, which are discussed further in Section 9.4. A slope for increasing values of $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ is observed, which exceeds the one-sigma band for high values; however, not all systematic uncertainties are included as stated in the caption. The final result for the luminosity measurement of $139\,\mathrm{fb}^{-1}$ is used from now on in the normalization of the MC simulation.

In order to study the impact of the mismodelling on the phase-space of the analysis, an $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ reweighting is applied at primary selection level and inclusive in b-tagged jet multiplicities, which also includes zero b-tagged jets. The background Z + jets is used as an input, since it is the main background contribution and likely to cause a mismodelling due to an overprediction of activity in the parton shower [197]. The result of the reweighted $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shown in Figure 9.10(b). The slope of $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ which is visible in the ratio plot is corrected, while the



(a) (b)

Figure 9.10: $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ distribution before and after the reweighting of the $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the $Z+$jets background. The b-tag inclusive $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ distribution at primary selection level is used as input for the reweighting. All systematic uncertainties introduced in Section 9.4 are included except the modelling uncertainties of $Z+$jets, $t\bar{t}$ and $t\bar{t}+X$ and the $t\bar{t}$ ISR and FSR uncertainties. The last bin contains all overflow events.

normalization of the $Z+$jets process is kept constant. Thus, an additional contribution to the $Z+$jets normalization by the reweighting is avoided. The designated uncertainties in the profile likelihood fit should be used to adjust the normalization, if a correction of the $Z+$jets contribution improves the data and MC simulation agreement.

As a next step, the influence of the reweighting on three MCBOT categories with the highest number of MC events is investigated. The idea is to evaluate, if the reweighting has a sizable effect on the shape of the final discriminant in these regions. If the shape of $m(Zb)$ or $m(Z+2^{\mathrm{nd}}\,b)$ is not influenced by the $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ reweighting significantly, a reweighting is redundant: Only the normalization of MC events would be influenced by a reweighting and the normalization could be corrected by the normalization uncertainties in the fit.

In Figure 9.11, the three categories 1b-0V-0H-0t, 1b-0V-0H-1t and 1b-0V-1H-0t are shown before and after the application of the reweighting. It can be seen that the normalization of the total background expectation is lower for the setup with an $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ reweighting. This can be understood when looking at the tail of the $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in Figure 9.10(b): Due to the abundance of MC background in this region, the number of background events in the phase-space of the search with $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}} > 1380\,\mathrm{GeV}$ is too high. After the application of the reweighting, the MC expectation is closer to the observed number of data events. However, the shape of $m(Zb)$ is only changed marginally by the reweighting, which is supported by the fact that the correlation of $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ with $m(Zb)$ is small: For the $Z+$jets process, the correlation coefficient at primary selection level between $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ and $m(Zb)$ is 0.1. For $m(Z+2^{\mathrm{nd}}\,b)$, the correlation coefficient is even smaller with 0.04. Therefore, no other effect than a normalization difference is expected for the MCBOT categorization with at least two b-tagged jets. With only the background normaliza-

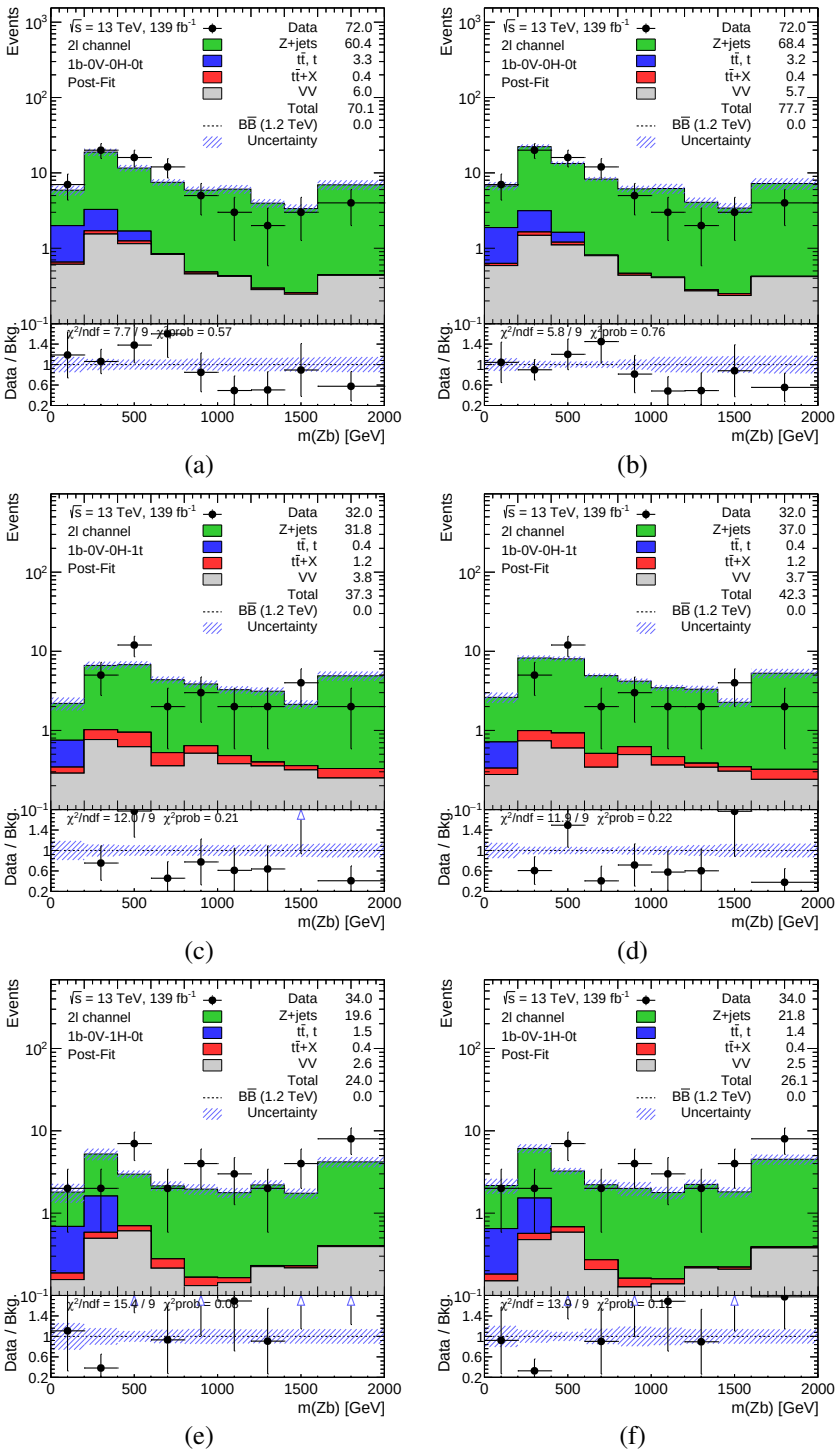


Figure 9.11: Post-fit distributions for the three MCBOT regions 1b-0V-0H-0t, 1b-0V-0H-1t and 1b-0V-1H-0t. Statistical and systematic uncertainties, described in more detail in Section 9.4, are included with the following deviations: The modelling uncertainty of the Z + jets process only includes the MC campaign for 2015+2016. The modelling uncertainties of the $t\bar{t}$ and $t\bar{t}+X$ processes are not included yet, as well as the $t\bar{t}$ ISR and FSR uncertainties. On the left, a reweighting in the $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ variable is applied to the Z + jets background according to the procedure described in Subsection 9.3.6, while on the right no reweighting is applied. The subplot shows the ratio of data over the background prediction. Furthermore, the results of a χ^2 test over the NDF are depicted, as well as the χ^2 probabilities.

tion being affected in all MCBOT categories, the effect of a reweighting is expected to be compensated by a change of background normalization which can be adjusted during the fitting procedure.

Additionally, the acceptance of signal events in all MCBOT categories is compared with and without reweighting and found to be almost constant across categories. Therefore, it can be concluded that the main effect of the reweighting is a normalization difference. Furthermore, the decision against a reweighting does not mask the underlying mismodelling of jet properties or shift them to other variables. Additionally, a reweighting should only be applied with a corresponding systematic uncertainty, which can decrease the performance of the search.

Instead of a reweighting, another approach is chosen: Two CRs, which are orthogonal to the MCBOT categorization, are defined in order to obtain a handle on the normalization of the Z + jets background. Therefore, the variable $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ is used as variable in the binned profile likelihood fit. One region is defined for each b-tagged jet multiplicity of the MCBOT categorization with either exactly one or at least two b-tagged jets. The $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ region adjacent to the search phase-space, which ranges from 920 GeV to 1380 GeV, is chosen for the CRs. Due to large statistics of expected background events in the two CRs, the normalization can be adjusted during the fitting procedure and might lead to a better data and MC modelling. With regard to the kinematics, the CRs are very close to the MCBOT categorization, which ensures a valid extrapolation of the background normalization from the CRs to the MCBOT categories.

9.3.7 Summary of event classification categories

An improved strategy for the search for pair production of VLT and VLB quarks with a leptonic Z-boson candidate in the final state was presented. It utilizes MCBOT in order to build a categorization dependent on the number of V-, H- and top-tags. Furthermore, an additional splitting in exactly one and at least two b-tagged jets improves the performance. The number of categories is decreased from 22 to 14 in order expect a sufficient amount of background events for the application of the asymptotic approximation in the following hypothesis tests. At the same time the sensitivity is kept high, since the combination of categories is optimized on the mass and cross-section limits. These categories consist of signal enriched and signal depleted regions, and cover all events in the dilepton channel with kinematic requirements above $p_{\mathrm{T}}\,(Z) > 300\,\mathrm{GeV}$, $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}} > 1380\,\mathrm{GeV}$ and $n(b\text{-tagged jets}) \geq 1$ in addition to the primary selection, together defined as the full selection. The sensitivity of the individual categories is summarized in Figure 9.12. It shows that categories with a larger number of MCBOT tags and at least two b-tagged jets, have larger sensitivity, measured with $\frac{S}{\sqrt{B}}$, and are denoted as signal enriched regions. Regions without or with only one MCBOT tag are summarized as signal depleted regions, and are used to control the background normalization. Three signal depleted regions were shown with unblinded data in a study of the $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling. These mismodelling studies of the variable $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ suggested the use of two additional regions, which are used as CRs for the background normalization. The selection criteria for all of the categories, including CRs, are outlined in Table 9.7. Furthermore, the optimized final discriminants used in different regions are summarized: While regions with exactly one b-tagged jet in an event use $m(Zb)$, $m(Z\text{+ }2^{\text{nd}}\,b)$ is used for events with at least two b-tagged jets. On the other hand, the CRs utilize $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ as variable in both regions of the b-tagged jet splitting in order to improve the background normalization for exactly one and at least two b-tagged jets separately.

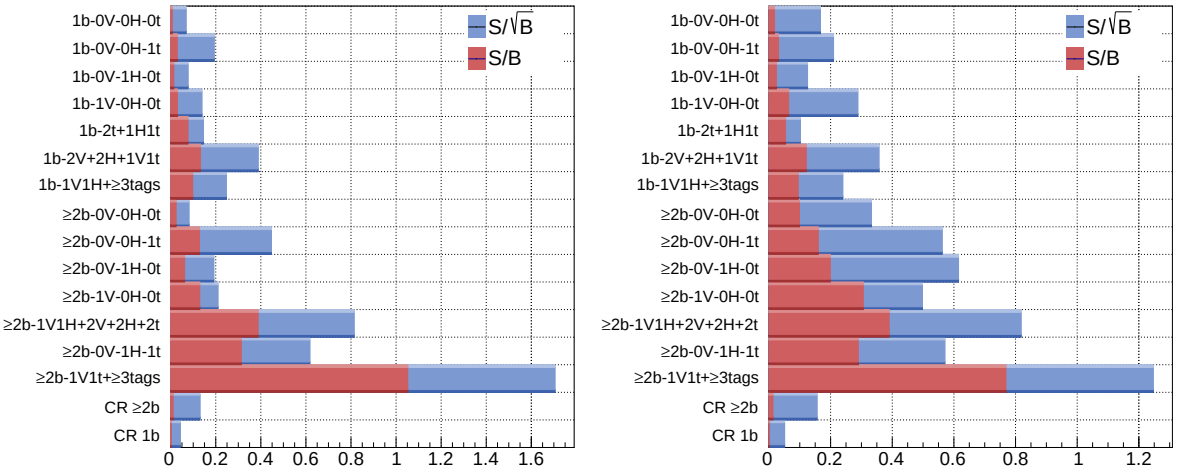


Figure 9.12: Overview of the signal over background ratio, $\frac{S}{B}$, and the sensitivity, $\frac{S}{\sqrt{B}}$ for all MCBOT categories and the two CRs. The a) VLT and b) VLB quark signals with a mass of 1200 GeV are depicted assuming the singlet model, respectively.

Table 9.7: Full selection for the dilepton channel of the $Z(\ell\ell)t/b + X$ analysis. The events are split into one and at least two b-tagged jets and the MCBOT categorization listed in Table 9.6, which contains 14 MCBOT regions. As final discriminant, MCBOT regions with one b-tagged jet use $m(Zb)$, while regions with at least two b-tagged jets use $m(Z+2^{\text{nd}}\,b)$. Two CRs are defined as well, which use $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}}$ as variable in the following profile likelihood fit.

Full selection for all categories				
Primary selection: Preselection ≥ 2 central small-R jets exactly two OSSF leptons (e/μ) $\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$				
≥ 1 b-tagged jet $p_{\text{T}}(Z) > 300\,\text{GeV}$				
MCBOT regions (14) $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}} > 1380\,\text{GeV}$			**CRs (2)** $920\,\text{GeV} < H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}} < 1380\,\text{GeV}$	
Discr.:	1 b-jet: $m(Zb)$	≥ 2 b-jets: $m(Z+2^{\text{nd}}b)$	1 b-jet: $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}}$	≥ 2 b-jets: $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}}$

9.4 Systematic uncertainties

An extensive overview of the relevant systematic uncertainties in the $Z(\ell\ell)t/b + X$ search was given for the 2015+2016 analysis in Section 8.5. A similar strategy is adopted in the full Run-2 search, and only differences are highlighted here. The newest recommendations within the ATLAS experiment were implemented in the MC simulations used for the estimation of the up and down variations by one standard deviation for all uncertainty sources.

For Z + jets, a new uncertainty is implemented, which covers the uncertainty in the heavy flavor (HF) composition of the produced jets. Whether b-tagged jets are present or not can lead to deviations in the normalization of the order of 30-50%, which was observed in several ATLAS analyses using the Z+jets process with a requirement on b-tagged jets [197]. It should be noted that no mismodelling in the Z+heavy-flavor-jet background was seen in the 2015+2016 analysis described in Chapter 8 and thus the corresponding uncertainty was not included previously. However, in the full Run-2 analysis an uncertainty of 30% on events originating from Z-boson production in association with c- and b-quarks is applied in order to cover a potential mismodelling. This procedure also improves the statistical model for the following profile likelihood fits with the knowledge from other 2015+2016 ATLAS analyses. The same procedure is applied to VV events in which light or HF jets occur in hadronic W- or Z-boson decays. Both the Z + jets and VV HF uncertainties are added to the MCBOT categories with exactly one and at least two b-tagged jets separately, since the production of events with different b-tagged jet multiplicities is assumed to be uncorrelated.

Furthermore, uncertainties on the strong coupling constant α_S are estimated for the main background, Z + jets, by a variation of its value. Additionally, ISR and FSR uncertainties are included for $t\bar{t}$ events. In the 2015+2016 analysis, $t\bar{t}$ events with $p_{\mathrm{T}}(Z) < 200$ GeV were assigned with a fake uncertainty. In the new analysis strategy, no events with $p_{\mathrm{T}}(Z)$ below 300 GeV are present. However, the removed electron isolation might lead to a larger fraction of jets which could mimic an electron in the detector. Thus, a 25% uncertainty is applied to the normalization of $t\bar{t}$ events.

In the 2015+2016 analysis, a set of systematic uncertainties was assigned due to the use of large-R jets. As a result of using RC jets instead, these uncertainties become redundant, since only calibrated inputs are used for the reconstruction of RC jets, as described in Chapter 5. Furthermore, MCBOT uses RC jets reclustered from small-R jets to assign V-, H- and top-tags. Because the uncertainties of the small-R jets can be propagated through the tagging algorithms, there is no need for dedicated MCBOT tagging uncertainties.

The luminosity uncertainty of the full Run-2 dataset is estimated to 1.7%.

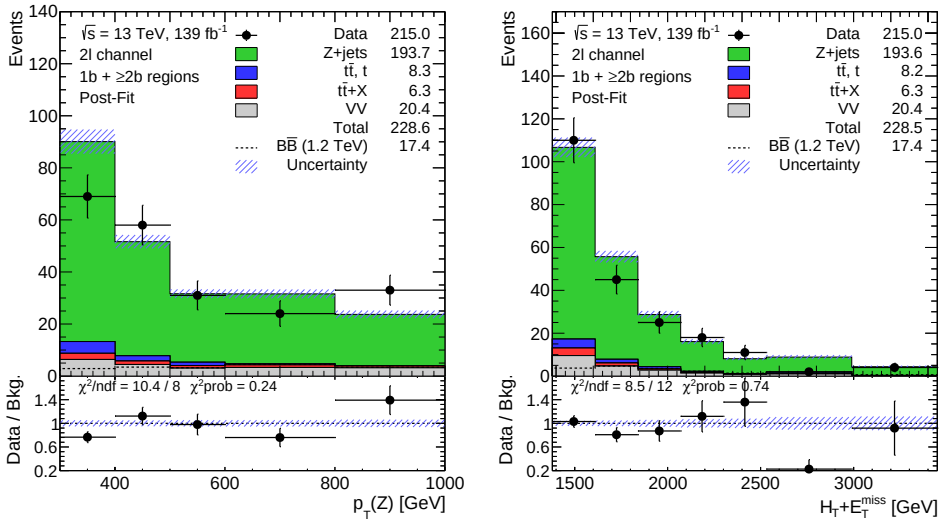
During the unblinding procedure, unexpected results of the $t\bar{t} + X$ shower and generator uncertainties were found. After performing the profile likelihood fit, the post-fit uncertainty of these two uncertainties showed a contribution larger than $\pm 1\sigma$ of the pre-fit uncertainty, which is an indication for an unstable fit. A more accurate estimation of the minimization procedure was used in order to reduce the post-fit uncertainties, but was not successful. The problem was pinpointed to two regions, $\geq$2b-0V-0H-0t and $\geq$2b-0V-1H-0t, with extremely high contributions of the $t\bar{t} + X$ generator and shower uncertainties, where the $\pm 1\sigma$ pre-fit uncertainty reaches up to $\approx 600\%$ compared to the nominal prediction. The large uncertainties are assumed to be unphysical and are caused by a small number of MC events. Furthermore, the $t\bar{t} + X$ contribution to the total background in both regions is negligible. Therefore, both uncertainties are removed from these two regions. Additionally, the

contribution of the $t\bar{t} + X$ process to the total background was checked in all of the search regions and removed from those with very small $t\bar{t} + X$ proportions and a negligible impact on the results. These regions are $\geq$2b-1V-0H-0t, 1b-1V1H+$\geq$3tags, 1b-2V+2H+1V1t, 1b-1V-0H-0t, 1b-0V-1H-0t, 1b-0V-0H-1t and 1b-0V-0H-0t in addition to the already mentioned regions.

9.5 Data and MC simulation agreement

The definition of a novel analysis strategy in Section 9.3 introduced a multiplicity of MCBOT categories in addition to two CRs, where the CRs are motivated by a $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling. A closer look at the data and MC agreement is taken in this section in order to get more insight into the modelling of relevant variables. As already motivated in the previous chapter in Section 8.4, a binned profile likelihood fit of the MC expectation to the measured data is performed with the B-only hypothesis. This procedure improves the agreement of data and MC simulation. The fit is performed simultaneously in the categorization with 16 regions, which is discussed in Section 9.6. The change of background normalization and shape, and constraints of systematic uncertainties are then extrapolated to other variables in more inclusive regions. A study of the agreement of data and background allows to draw conclusions about the validity of the background model.

At first, the modelling of the $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ variables after the full selection is studied in Figures 9.13(a) and 9.13(b). Both distributions show that the number of total background events exceeds the number of observed data events. Thus, the overprediction of MC simulation caused by the $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling can not fully be compensated by the adjustments made by the profile likelihood fit. The main background process in this phase-space is Z + jets, which is followed by $t\bar{t}$ and



(a) (b)

Figure 9.13: Post-fit distributions of a) $p_{\mathrm{T}}\,(Z)$ and b) $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ after the full selection. The VLB quark signal at 1.2 TeV with singlet BRs according to the pre-fit hypothesis is overlaid. Systematic and statistical uncertainties are included in the uncertainty band. The background is modeled according to the B-only hypothesis. The subplot shows the ratio of data over the background prediction. Furthermore, the results of a χ^2 test over the NDF are depicted, as well as the χ^2 probabilities.

$t\bar{t}+X$ production in significantly smaller fractions. All systematic uncertainties are included, and the subplot below shows the ratio of data divided by the background expectation per bin. Furthermore, the value of the χ^2 test between data and the post-fit background contribution over the number of degrees of freedom is depicted, from which the χ^2 probability is calculated. This measure for the level of agreement takes correlations of bins and correlations between systematic uncertainties into account. For $p_{\mathrm{T}}\,(Z)$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$, a reasonable agreement of data and simulation is observed. As in Chapter 8, a probability of smaller than 5% indicates that the disagreement of data and background should be investigated further.

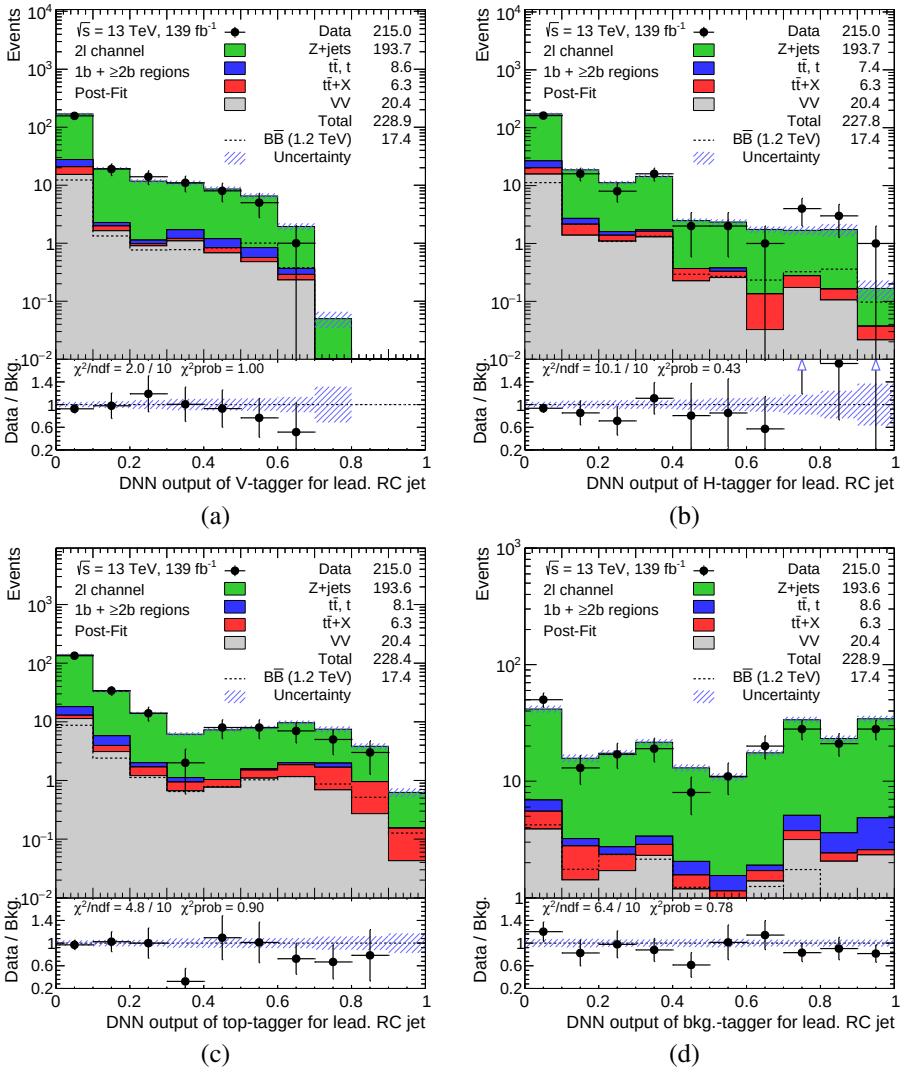


Figure 9.14: Post-fit distributions of the MCBOT DNN output for the a) V-tagger, b) H-tagger, c) top-tagger and d) background-tagger for the leading RC jet. The full selection is applied. The VLB quark signal at 1.2 TeV with singlet BRs according to the pre-fit hypothesis is overlaid. Systematic and statistical uncertainties are included in the uncertainty band. The background is modeled according to the B-only hypothesis. The subplot shows the ratio of data over the background prediction. Furthermore, the results of a χ^2 test over the NDF are depicted, as well as the χ^2 probabilities.

Furthermore, the MCBOT DNN outputs for the V-, H-, top- and background-tagger are important variables in the construction of the analysis strategy. Well modeled distributions ensure that no difference in the acceptance for the MCBOT distributions is introduced for data and MC simulation. The DNN outputs of the leading RC jet for all four MCBOT taggers are illustrated in Figure 9.14 after the full selection. It should be noted that the main backgrounds mainly produce RC jets, which should belong to the background MCBOT class: For the Z+jets process, the Z-boson is most likely the parent particle of the two OSSF leptons, required in the selection. Therefore, the associate production of jets, which are reconstructed as RC jets, are tagged by MCBOT and no V-bosons, H-bosons or top-quarks are expected. Nevertheless, the RC jets are more likely to be tagged as one of the three signal classes, if their properties mimic the features of V-boson, H-boson and top-quark jets. More details about the features of RC jets for each of the four taggers were described in Section 6.8.

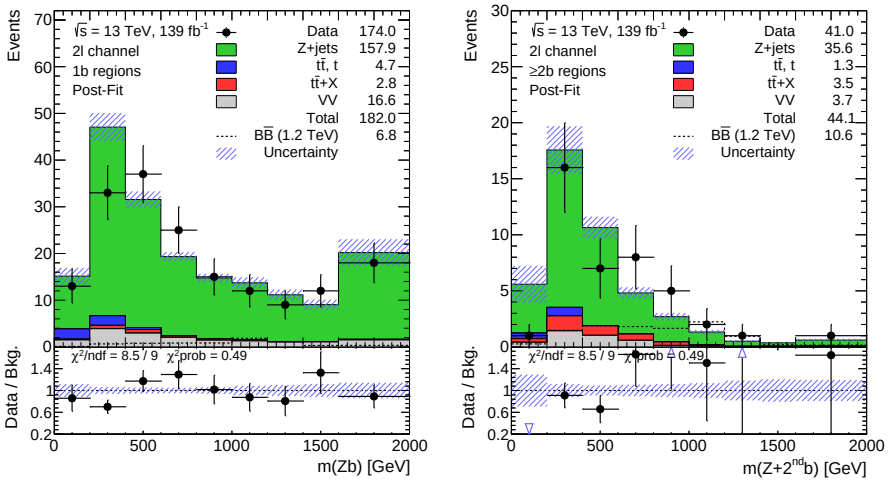
The DNN output of the V-tagger is depicted in Figure 9.14(a) and shows good agreement of data and MC simulation. RC jets with a DNN output larger than 0.3 obtain a V-tag before ambiguous tags are resolved. The DNN output drops below one expected MC event for values larger than 0.7, which illustrates the same trend as the DNN output of the V-tagger in Figure 6.10(b): It is more difficult for MCBOT to classify V-boson jets correctly than it is for H-boson and top-quark jets, since V-boson jet properties are closer to background jet features.

In Figure 9.14(b), the output of the H-tagger is shown for the leading RC jet. The agreement of data and background is reasonable. A requirement on the DNN output of at least 0.35 is made to obtain a H-tag. The MCBOT studies have shown that for the standard MCBOT tagger the decision, whether an RC jet is tagged as a H-boson jet or not, depends largely on the b-tagging information alongside the RC jet mass. RC jets with one or two b-tagged subjets are likely to be identified as H-boson jets in a p_{T} range below 1 TeV. For the full selection of the analysis, at least one b-tagged jet is required in each event and thus there is always at least one b-tagged subjet in one of the RC jets. With the RC jet mass around the H-boson mass, an RC jet is likely to be tagged as a H-boson.

Furthermore, the DNN output of the top-tagger is illustrated in Figure 9.14(c) for the leading RC jet. As for the V- and H-tagger, the simulation is in agreement with the data distribution. A fluctuation is seen in the fourth bin, but the overall χ^2 probability is fairly high and thus is compatible with a statistical nature of the fluctuation. It can be seen that for high top-tagger values a noticeable fraction of $t\bar{t}+X$ events is present. The $t\bar{t}+X$ process is the only background, which distinguishes itself from the other processes by the possibility of hadronic top-quark decays in an event. Thus, a higher fraction of the RC jets from $t\bar{t}+X$ events receives a very top-quark-like DNN output.

Additionally, the DNN output of the background-tagger is depicted in Figure 9.14(d). For the leading RC jet, a larger fraction of RC jets from $t\bar{t}$ events is located at high DNN outputs, which are likely to be caused by jets produced during leptonic $t\bar{t}$ decays. Overall, the data and MC simulation agreement is good. As for all other taggers, the number of MC events exceeds the number of data events for the background-tagger, which is caused by the data and MC simulation discrepancy in the tail of the $H_{\mathrm{T}}(\mathrm{jets})+E_{\mathrm{T}}^{\mathrm{miss}}$ distribution discussed in Subsection 9.3.6.

At last, the final discriminants $m(Zb)$ and $m(Z\text{+ }2^{\mathrm{nd}}\,b)$ are illustrated in Figures 9.15(a) and 9.15(b), respectively. For $m(Zb)$, the requirement of exactly one b-tagged jet is applied in addition to the full selection. The agreement between data and MC simulation is reasonable. For $m(Z\text{+ }2^{\mathrm{nd}}\,b)$, at least two b-tagged jets are required after the full selection, and the χ^2 probability supports the observation that decent data and MC agreement is present.



(a) (b)

Figure 9.15: Post-fit distributions of a) $m(Zb)$ for exactly one b-tagged jet and b) $m(Z+\,2^{\text{nd}}\,b)$ after requiring at least two b-tagged jets. The full selection is applied for both distributions. The VLB quark signal at 1.2 TeV with singlet BRs according to the pre-fit hypothesis is overlaid. Systematic and statistical uncertainties are included in the uncertainty band. The background is modeled according to the B-only hypothesis. The subplot shows the ratio of data over the background prediction. Furthermore, the results of a χ^2 test over the NDF are depicted, as well as the χ^2 probabilities.

9.6 Results

In order to perform a hypothesis test and make a statement about the existence of VLB and VLT quarks with the whole Run-2 dataset, a profile-likelihood-ratio test is performed for the optimized search strategy. The presentation of results follows the procedure in Section 8.8. At first, the significance of different VLQ mass hypotheses is tested, and in case of the absence of a discovery, exclusion limits are set on the production cross-section and the VLQ masses for different BR hypotheses.

9.6.1 Test for discovery

The increased sensitivity due to the larger dataset compared to the 2015+2016 search and the improvements introduced by the use of MCBOT provide new potential to find a deviation of data from the SM expectation in the full Run-2 search. The rejection of the B-only hypothesis would favor the S+B hypothesis. All 16 analysis categories are used in a simultaneous fit and the overall agreement of the B-only hypothesis with data is quoted with the significance Z. The calculation of this significance at 95% CL uses the test statistic q_0, described in Eq. 7.7, and the PLR from Eq. 7.4. Thus, the signal strength μ is still used in the significance calculation for the unconditional maximum likelihood estimator in the denominator of the PLR. The significances for different VLB and VLT quark masses in the singlet model are listed in Table 9.8. The observations show that the data is compatible with SM background, and no VLT or VLB quark signals are found.

The post-fit agreement of data and simulation is investigated with a B-only fit of background to the observed data, which also shows that no discovery can be claimed. The distributions are depicted in Figures 9.16 and 9.17 for all 16 categories of the search. The corresponding pre-fit distributions

Table 9.8: Observed significance Z at 95% CL in units of the standard deviation σ assuming the singlet model for different mass hypothesis of VLT and VLB quarks for $139\,\text{fb}^{-1}$.

Mass [GeV]	600	800	1000	1200	1400	1600	1800	2000
$T\bar{T}$	-1.58	-0.63	0.04	0.39	0.29	0.47	0.56	0.52
$B\bar{B}$	0.61	1.86	1.54	1.20	0.85	0.82	1.00	0.92

are shown in Figures C.2 and C.3 in Section C.2. The categories containing events with exactly one b-tagged jet are shown in Figure 9.16(a) to 9.16(g). The corresponding CR is shown in Figure 9.16(h), and good post-fit agreement is observed in the CR for data and simulation. Again, the agreement for all categories is tested with the χ^2 probability. A value smaller than 5% is rated as an incompatibility with the corresponding hypothesis. All regions show reasonable agreement except for the category 1b-1V1H+$\geq$3tags with a χ^2 probability lower than 5%. The disagreement is studied in more detail in the following dedicated subsection. The post-fit yields are shown in the legend of the post-fit plots and give a quantitative estimate of the background composition. The subleading background contribution varies between contributions from the backgrounds $t\bar{t}$, $t\bar{t}+X$, but also VV, while the majority of events are associated to Z + jets production.

It should also be mentioned that the binning is adjusted in comparison to the studies presented in Section 9.3. Very low event numbers in the highest bins of the final discriminant are avoided in order to prevent large uncertainties on the fit parameter μ and to be closer to the regime, where the asymptotic approximation is valid. Thus, a maximum of four bins is chosen, which is further reduced for the categories with at least two b-tagged jets, which carry less events than the regions with one b-tagged jet.

In Figure 9.17(a) to 9.17(g), the MCBOT categorization for at least two b-tagged jets is depicted. The agreement of data and MC simulation is reasonable in all distributions, except the one in Figure 9.17(c), which is the region $\geq$2b-0V-1H-0t, with a χ^2 probability of less than 1%. Studies concerning this mismodelling are discussed below. It should be noted that the observation of zero events in the region $\geq$2b-0V-1H-t is compatible with a statistical fluctuation, as indicated by the χ^2 probability and thus is not studied further. The agreement of data and background in the CR for at least two b-tagged jets shows good agreement and is illustrated in Figure 9.17(h). Furthermore, a larger amount of $t\bar{t}+X$ events is present in the categories with at least two b-tagged jets compared to the regions with one b-tagged jet. The pre-fit hypothesis of the VLB quark signal at $1200\,\text{GeV}$ assuming singlet BRs is also shown in the distributions and it can be seen that especially the categories with at least two b-tagged jets are highly sensitive, which supports the statements obtained from the sensitivity plots in Figure 9.12.

In order to get insight into the adjustments to the statistical model made by the fit, the pull plot is studied. It is illustrated in Figure 9.18 and only nuisance parameters (NPs) with an influence on the fit of larger than 1% are listed, which is the pruning threshold. Most of the NPs are not changed and remain at their initial mean value and have an approximate post-fit uncertainty of $\pm1\sigma$, where σ is the pre-fit uncertainty. There are five uncertainties, which have a larger impact and are discussed in the following.

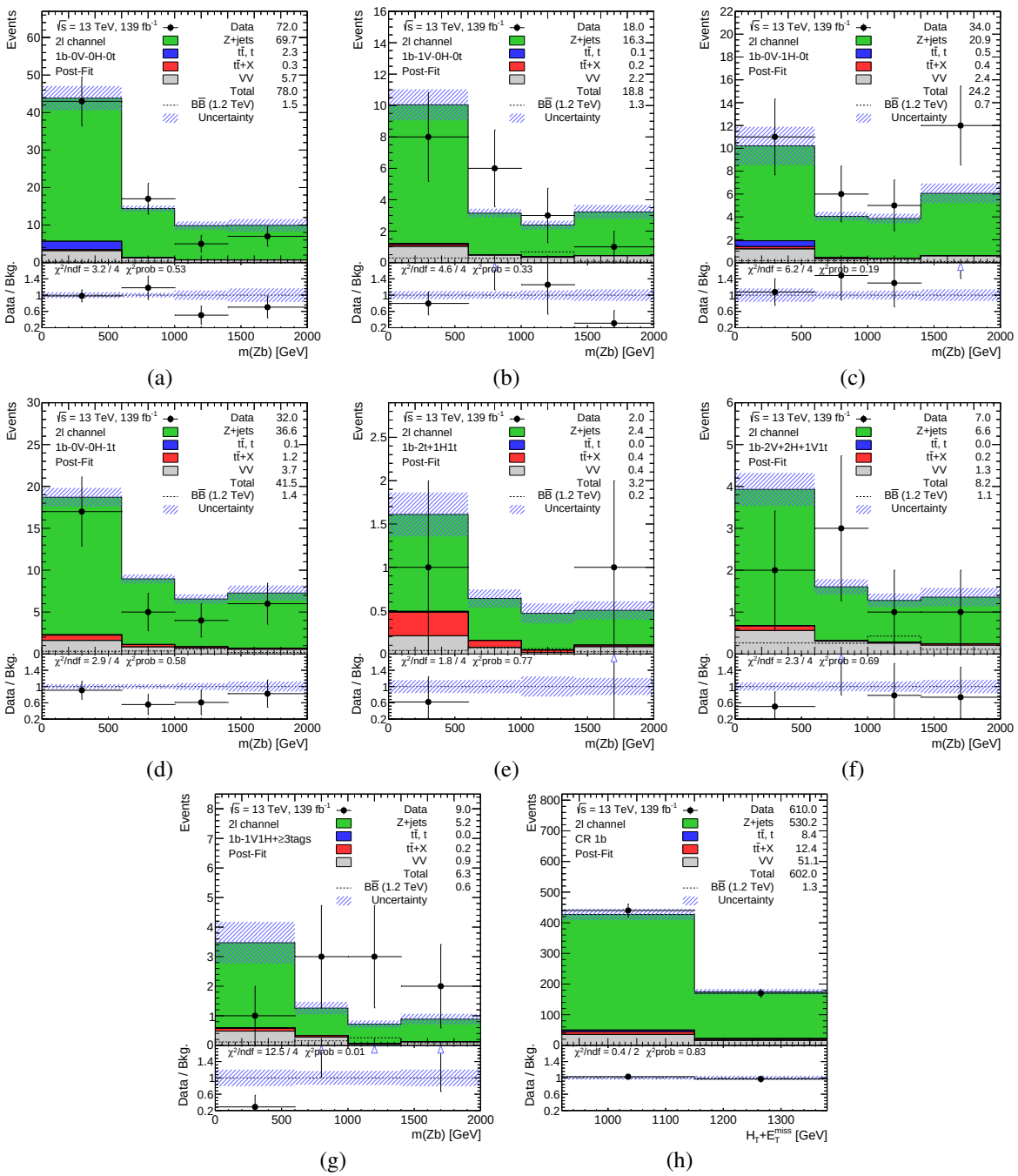


Figure 9.16: Post-fit plots for all categories with exactly 1 b-tagged jet with $m(Zb)$ as final discriminant. In addition, the CR for exactly 1 b-tagged jet with $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ as variable is shown in h). Systematic and statistical uncertainties are included in the uncertainty band. The pre-fit signal for VLB quarks with $m = 1200\,\mathrm{GeV}$ assuming singlet BRs is overlaid. All backgrounds are depicted and stacked on top of each other. The legend contains the yields of post-fit background, pre-fit VLB signal and data. The subplot shows the ratio of data over the background prediction. Furthermore, the result of a χ^2 test over the NDF is depicted, as well as the χ^2 probability.

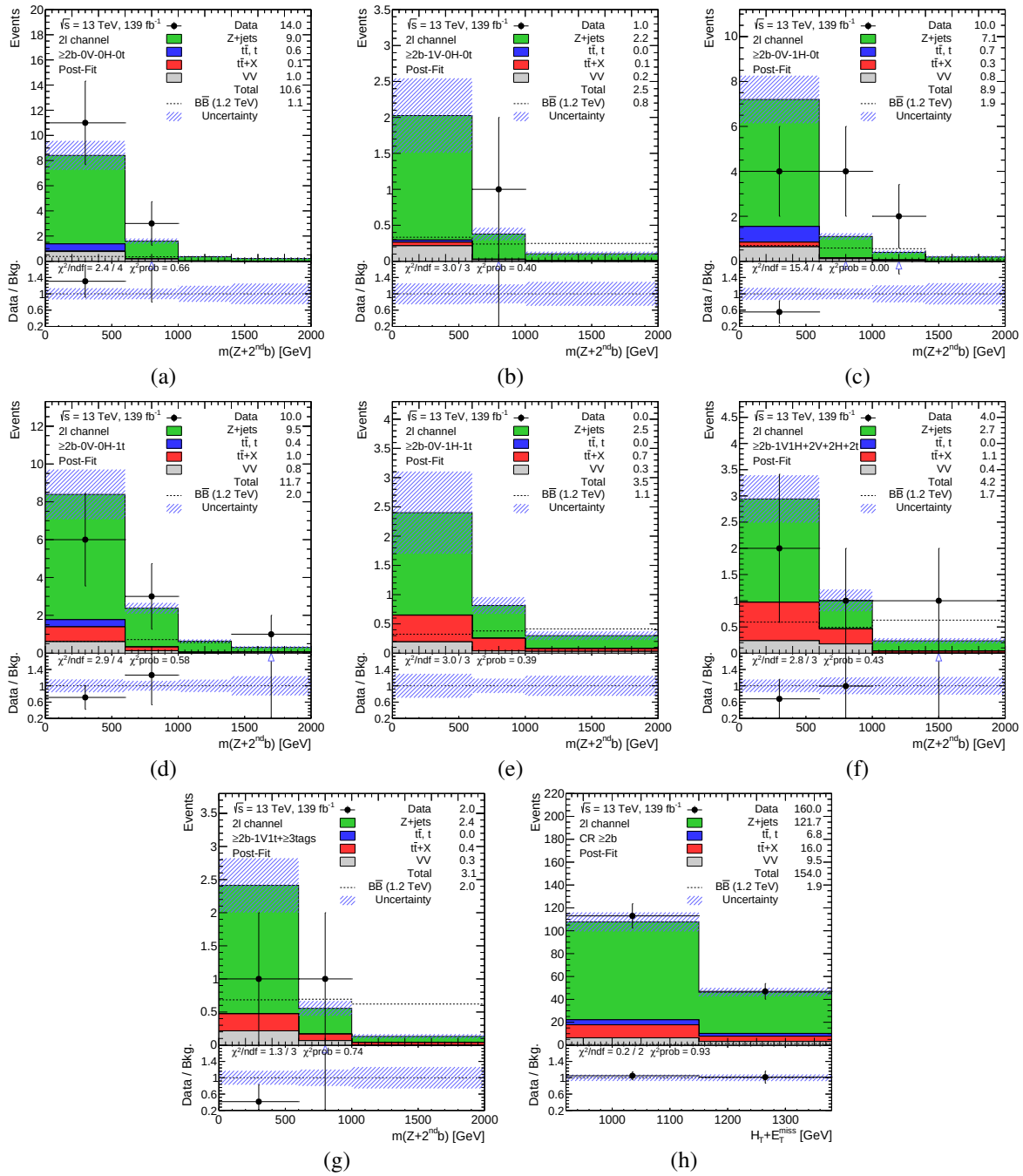


Figure 9.17: Post-fit plots for all categories with at least 2 b-tagged jets with $m(Z+\,2^{\text{nd}}\,b)$ as final discriminant. In addition, the CR for at least 2 b-tagged jet with $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}}$ as variable is shown in h). Systematic and statistical uncertainties are included in the uncertainty band. The pre-fit signal for VLB quarks with $m = 1200\,\text{GeV}$ assuming singlet BRs is overlaid. All backgrounds are depicted and stacked on top of each other. The legend contains the yields of post-fit background, pre-fit VLB signal and data. The subplot shows the ratio of data over the background prediction. Furthermore, the result of a χ^2 test over the NDF is depicted, as well as the χ^2 probability.

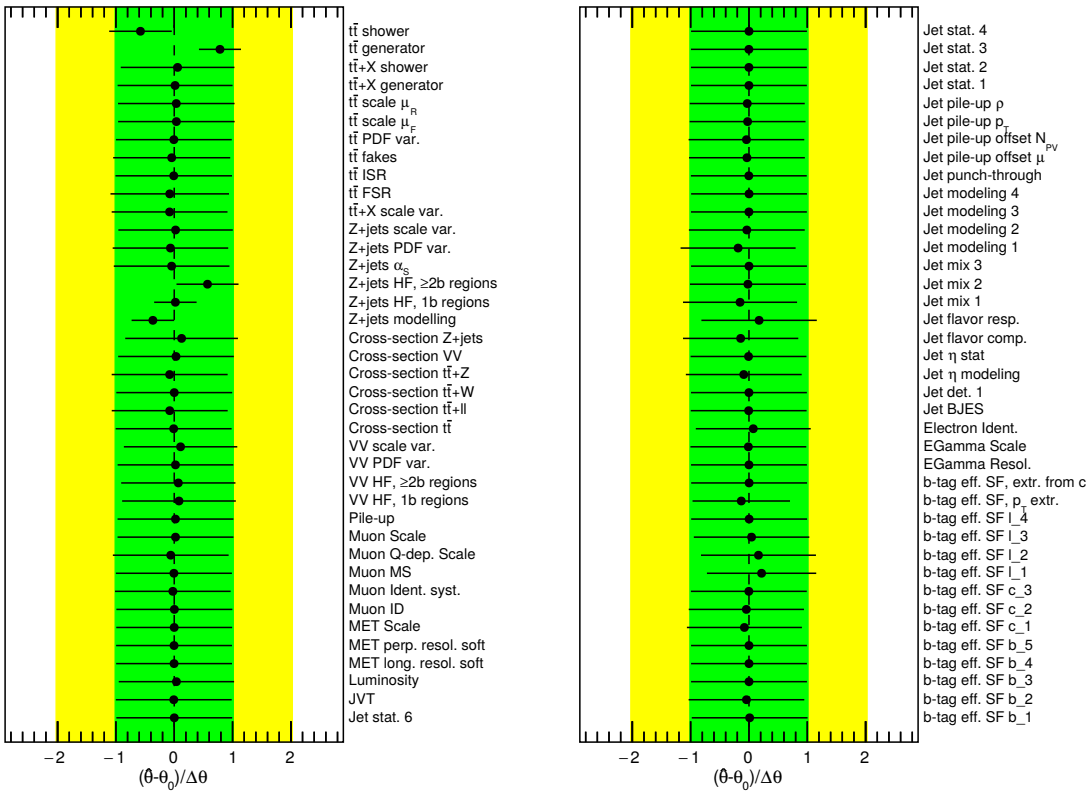


Figure 9.18: Pull plot for the B-only hypothesis used in the observed absence of VLT and VLB quark signal. Pulls of the post-fit mean and post-fit one sigma bands of Gaussian distributed nuisance parameters are illustrated as black dots and horizontal lines on the dots, respectively. A B-only fit is performed for the singlet hypothesis of the VLB quark with $m = 1200\,\text{GeV}$.

The first two of these five uncertainties are associated to the $t\bar{t}$ background. The $t\bar{t}$ background does almost exclusively contribute to the events in the first bin of the $m(Zb)$ and $m(Z+\,2^{\text{nd}}\,b)$ distributions, shown in Figures 9.16 and 9.17, but also in Figures C.2 and C.3. It is adjusted by the $t\bar{t}$ shower and generator uncertainties, which are pulled in opposite directions to change the $t\bar{t}$ normalization. The background expectation of the first bin is larger at pre-fit level and is decreased by these two uncertainties. Both uncertainties are slightly anticorrelated, which is depicted in Figure 9.19. Thus, the effect of the shower and generator uncertainties adjusts this one-bin contribution and reduces the prediction of the $t\bar{t}$ process. Due to high event numbers in the first bin and in the CRs, both systematic uncertainties, which are of the order of 100-200% in some regions, are constrained by the fit.

The three other uncertainties, used to improve the fit model, are uncertainties connected to the main background, $Z+$jets. The $Z+$jets modelling uncertainty in combination with the $Z+$jets HF uncertainties is used to change the shape of $Z+$jets process and to reduce the normalization to a lower background contribution. The data and MC simulation mismatch in the search phase-space, induced by the $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}}$ mismodelling, is improved with these adjustments. Furthermore, these three $Z+$jets uncertainties are constrained by data mainly in the CRs and in bins of MCBOT categories with large event numbers. The opposite direction of the pulls on the modelling and the HF $\geq 2b$ uncertainty is changing the normalization in the same direction due to a slight anticorrelation, shown in Figure 9.19.

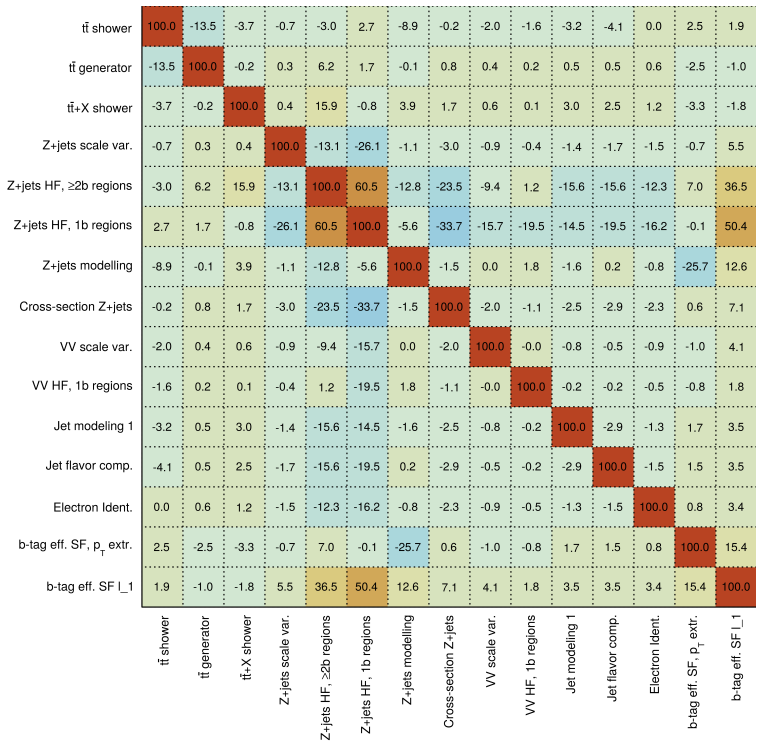


Figure 9.19: Correlation of systematic uncertainties with any correlation exceeding a magnitude of 13.5%. A B-only fit is performed for the singlet hypothesis of the VLB quark with $m = 1200\,\text{GeV}$.

Data and MC simulation disagreement in 1b-1V1H+≥3tags and ≥2b-0V-1H-0t

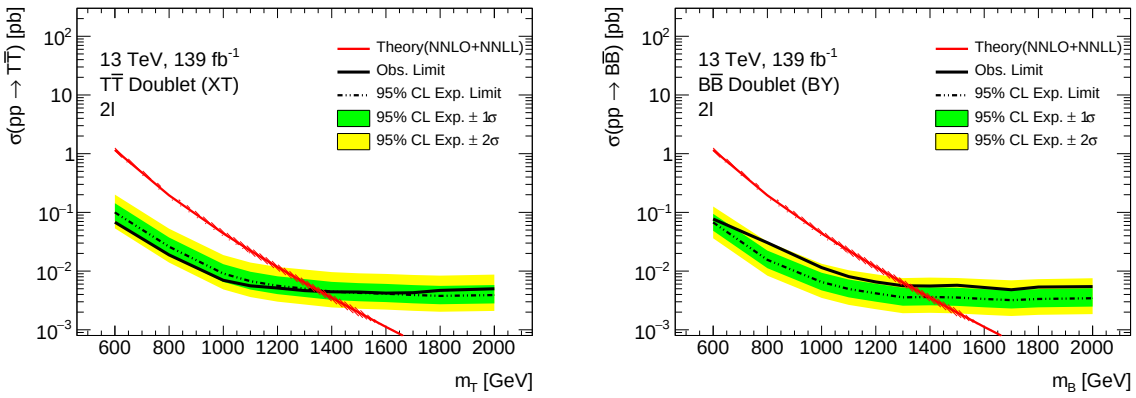
The low χ^2 probabilities of the two MCBOT categories 1b-1V1H+≥3tags and ≥2b-0V-1H-0t were investigated in detail. The comparison of the SM background and the observed data in the 14 other categories showed good agreement. In order to get an overall overview of the level of data and MC agreement achieved by the fit, a global χ^2 probability was calculated. A value of 18% is observed, which stands for reasonable agreement.

Other checks were performed in order to investigate the disagreement further. The fit model was adjusted such that the Z + jets modelling uncertainty gained more freedom, since this main uncertainty might provide sufficient freedom to the fit in order to adjust the shape and the normalization of the main background in order to improve the data and background agreement. The uncertainty was split for categories with one and at least two b-tagged jets, and also the smoothing was removed in a separate step. No significant improvement of the post-fit agreement was found. Furthermore, the DNN output of all four MCBOT classes was investigated, as well as a feature investigation of all 19 data events in the 1b-1V1H+≥3tags and ≥2b-0V-1H-0t categories was done. For all data events, the values of the final discriminants were checked, as well as the variables $p_\text{T}\,(Z)$ and $H_\text{T}(\text{jets}) + E_\text{T}^\text{miss}$. Furthermore, all four DNN outputs were investigated for the three leading RC jets in order to study how unambiguous a tag is. Since the RC jet mass and the b-tagging decision have shown to contribute significantly to the MCBOT tagging decision in Chapter 6, these properties were also studied for the three leading RC jets. The following observations were made, but no observation pointing towards another reason than a statistical fluctuation was found: The H-tags in the ≥2b-0V-1H-0t region are

very H-boson-like, since they have a high DNN output for the H-tagger. One of the data events in the bin from 600 to 1000 GeV, which shows an excess in data, is located directly at the bin border to the adjacent first bin, which cannot be considered by the fit. In the 1b-1V1H+$\geq$3tags region, RC jets with rather low V-boson and H-boson DNN outputs are tagged. Especially for V-tags, the background DNN output is always larger than 0.5. However, no indication for an insufficient fit model or a systematic effect introduced by MCBOT was seen. Therefore, the excess in data in these two regions is assumed to be a statistical fluctuation.

9.6.2 Exclusion limits on the mass and the production cross-section

Since the data observation did not show a significant deviation from the SM background expectation in the test of the B-only hypothesis at 95% CL, exclusion limits are set. For the derivation of exclusion limits, an S+B fit is performed with different BR assumptions for VLB and VLT quarks. As described in the 2015+2016 analysis in Chapter 8, the singlet and doublet BRs, as well as the most sensitive case, the 100% BR into Zt/b, is chosen to quote the sensitivity and the observed exclusion limits at 95% CL.



(a) (b)

Figure 9.20: Expected and observed limits at 95% CL on the cross-section for a) $T\bar{T}$ and b) $B\bar{B}$ production with BRs according to the doublet models (X,T) (or (T,B)) and (B,Y), respectively. The 2ℓ channel of the $Z(\ell\ell)t/b+X$ analysis using 139 fb^{-1} is shown. The $\pm1(2)\sigma$ bands of the expected limit at 95% CL are depicted in green (yellow). The theory curve is shown in red.

In Figures 9.20(a) and 9.20(b), the expected and observed limits at 95% CL for VLT and VLB quarks with the assumption of the doublet models are depicted, respectively. The intersection of the solid line and the theory curve determines the observed limit, which yields 1350 GeV for VLT quarks and 1316 GeV for VLB quarks. The expected and observed limits at 95% CL are summarized in Table 9.9. The observed limits reach from 1159 GeV for the VLT quarks with singlet BRs to 1443 GeV for the 100% BR into Zt. For VLB quarks, the observed limits are slightly lower: They are set at 1146 GeV for the singlet model and 1411 GeV for the 100% BR into the final state Zb. In comparison to the expected limits, it can be seen that the observed limits for VLT quarks are close, while for VLB quarks, a limit deviation of around 80 GeV is present. This deviation between the expected and observed limits is slightly above the $\pm1\sigma$ band for all models and thus is compatible with a statistical

Table 9.9: Observed (expected) mass limits at 95% CL for the singlet and doublet models, as well as the 100% BR into $T \to Zt$ and $B \to Zb$ for the 2ℓ channel for 139 fb^{-1}.

	Observed (exp.) limits [GeV]	
Model	$T\bar{T}$	$B\bar{B}$
singlet	1159 (1188^{+82}_{-63})	1146 (1225^{+78}_{-71})
doublet	1350 (1343^{+76}_{-70})	1316 (1392^{+67}_{-54})
100% $Q \to Zq$	1443 (1457^{+72}_{-65})	1411 (1493^{+75}_{-64})

fluctuation. The depictions of the limits for the singlet model and the 100% BR into Zt/b final states are shown in Appendix C.3.

Due to different sensitivities for VLB and VLT quarks in the different MCBOT categories and their individual bins, deviations of expected and observed limits can occur for one of the signals only. The VLB signal tends to have higher event yields in the higher bins of both final discriminants' distributions, $m(Zb)$ and $m(Z+2^{\text{nd}}\,b)$, since the VLB quark mass resonance is often reconstructed, which is not the case for the VLT quark. Thus, the observation of more data than predicted by MC simulation in some of the second and third bins with high sensitivity for the VLB quark leads to weaker limits on VLB quark masses. This is the case for, e.g. the region $\geq$2b-0V-1H-0t in Figure 9.17(c) or for the region 1b-1V1H+$\geq$3tags in Figure 9.16(g). The former category is more sensitive to VLB quark signal than to VLT quarks, as shown in Figure 9.12. Therefore, the limit contribution on $B\bar{B}$ production from this region is weakened. For the latter category, the sensitivity for both signals is more similar than for the $\geq$2b-0V-1H-0t region, but VLT quarks show a larger proportion of their events in the first bin. Lower values of the final discriminants for VLT quarks are one of the reasons, why the limits for VLT quarks are not weakened significantly. The second reason is the lower observation of data events in the first bin of several MCBOT categories in comparison to the expected background yields. A downward fluctuation causes stronger limits and thus the two effects of lack and excess of data lead to observed limits for VLT quarks, which are close to the expected limits at the sensitivity boundary.

As already mentioned in Subsection 9.6.1, the binning of the MCBOT categories was optimized in order to reduce bins with extremely low event yields. Furthermore, a merging of categories was conducted in Subsection 9.3.5 in order to enlarge the background yields per category for the validity of the asymptotic approximation. The 14 MCBOT regions of the analysis have still low event yields in some bins, and therefore, the influence of the use of the asymptotic approximation on the expected limits is evaluated. The background model for this study differs slightly from the setup, which is used in all results and post-fit plots presented in this chapter since Section 9.5: The main difference originates from the Z + jets modelling uncertainty, which was only estimated using the MC campaign with the pile-up profile of the years 2015+2016 of data taking. Furthermore, the generator and shower uncertainties of the $t\bar{t}+X$ process were kept in all categories. For this study, 10^5 pseudo-experiments are performed in order to sample the probability density functions (pdfs) of the test statistic for the S+B and B-only hypotheses. Three mass hypotheses are chosen for VLB and VLT quarks assuming

Table 9.10: Expected limits on μ at 95% CL using the pdf of the test statistic $\tilde{q}_\mu$ derived with the asymptotic formula and 10^5 pseudo-experiments. Statistical and systematic uncertainties are included. Only the MC campaign corresponding to 2015+2016 is included for the Z + jets modelling uncertainty. The $t\bar{t}+X$ modelling uncertainties are included in all MCBOT categories. Both VLB and VLT are shown for three masses and the singlet BRs. The deviation Δ between μ_{asymp} and μ_{pseudo} is calculated with $1 - \frac{\mu_{\text{asymp}}}{\mu_{\text{pseudo}}}$.

$Q\bar{Q}$	Mass [GeV]	μ_{asymp}	μ_{pseudo}	Δ
$B\bar{B}$	1200	$0.84^{+0.38}_{-0.24}$	$0.84^{+0.38}_{-0.19}$	0.1%
	1600	$7.3^{+3.4}_{-2.0}$	$7.3^{+2.4}_{-1.9}$	0.3%
	2000	61^{+29}_{-17}	62^{+23}_{-15}	2%
$T\bar{T}$	1200	$1.02^{+0.46}_{-0.28}$	$1.02^{+0.33}_{-0.22}$	0.3%
	1600	$7.9^{+3.7}_{-2.2}$	$8.2^{+1.9}_{-2.2}$	3%
	2000	68^{+33}_{-19}	70^{+18}_{-18}	3%

the singlet BRs. The lowest mass corresponds to the expected exclusion of VLQ masses around 1200 GeV. Then, two higher mass points at 1600 GeV and 2000 GeV are chosen, where the signal populates the higher bins of the final discriminants. Thus, a comparison between a limit calculation with the asymptotic approximation and with pseudo-experiments is compared for mass hypotheses, where the largest difference between both approaches is expected. The expected limits on the signal strength μ for the asymptotic approximation and for pseudo-experiments, the $\pm1\sigma$ intervals, as well as the deviation of both approaches, are listed in Table 9.10. At maximum, the deviation is 3% and therefore the asymptotic approximation is valid for this analysis. The $\pm1\sigma$ bands for μ_{asymp} and μ_{pseudo} show that for the asymptotic approximation the $\pm1\sigma$ bands, which estimate the influence of a statistical fluctuation on the limit, are overestimated. The pdfs of the test statistic derived with the asymptotic approximation correspond to the large sample limit and are wider than the pdfs derived with pseudo-experiments [143]. As a cross-check, the observed limits on μ for VLB and VLT quarks with mass hypotheses of 1200 GeV, 1600 GeV and 2000 GeV were compared with the $\pm1\sigma$ and $\pm2\sigma$ bands derived with pseudo-experiments. The limits on μ for $B\bar{B}$ production are within the 2σ bands, while for $T\bar{T}$ production the observed limit is within the 1σ band or close to it. Thus, the derived exclusion limits in this chapter are also compatible with a statistical fluctuation when the assumption of the asymptotic approximation is lifted.

Additionally, the impact of systematic uncertainties on the signal strength μ, which is obtained in an S+B fit of the signal and background expectation to the observed data distribution, is compared in the ranking plots for VLT and VLB quarks in Figures 9.21(a) and 9.21(b), respectively. The benchmark mass of 1200 GeV of both VLQs with the singlet BRs is shown. For VLT and VLB quarks, the Z + jets modelling and HF uncertainties are listed on the top or among the five most significant uncertainties. Since they influence the normalization of the main background in all regions of the fit, their impact is expected to be high.

Other important uncertainties in the ranking plot are related to the uncertainties on the b-tagging efficiency measurement like the NPs "b-tag eff., p_{T} extr." and "b-tag eff. SF l_1", which can influence the analysis selection at various stages. The NP "b-tag eff., p_{T} extr." describes an uncertainty on the b-tagging efficiency measurement, which comes from b-tagged jets, whose p_{T} is outside of the

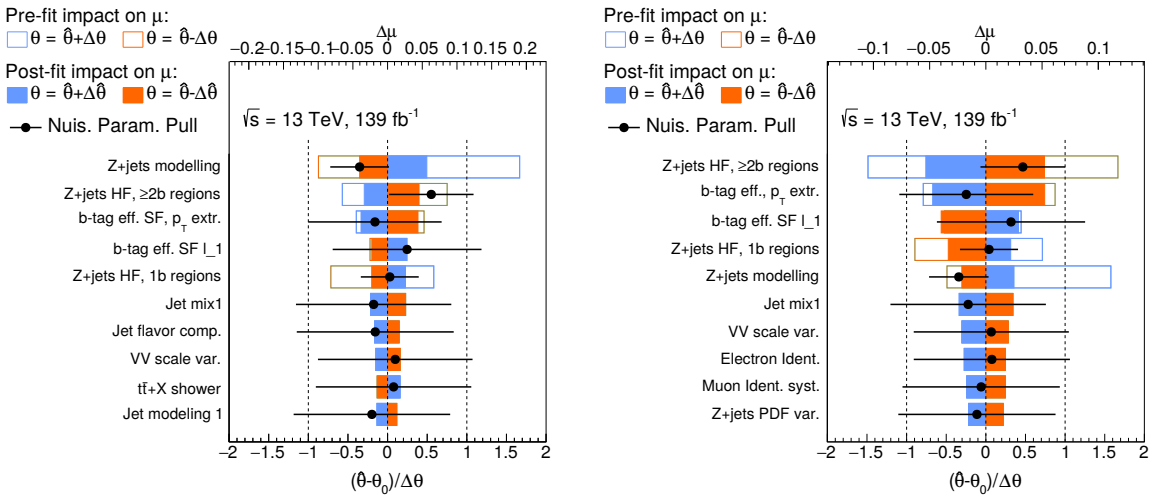


Figure 9.21: Ranking plots for a) VLT and b) VLB quarks for a mass of 1200 GeV assuming singlet BRs. The impact of each NP and γ factor is evaluated on the signal strength μ in an S+B fit of the signal and background prediction to the observed data. Each NP is fixed to $\pm 1\sigma$ of the pre- and post-fit estimates around the best fit value. A fit of all other systematic uncertainties and the statistical uncertainty is performed, which derives the impact of one specific NP with respect to the usual S+B fit and its signal strength μ. Furthermore, the pulls and constraints of the NPs are illustrated.

b-tagging efficiency measurement range and is therefore extrapolated to a higher kinematic regime. A variation of both NPs leads to different acceptance effects in the analysis selection and the MCBOT categorization due to a different tagging performance for b-tagging and MCBOT. These variations can thus influence the fitted value of μ.

The result for the fit including statistical and systematic uncertainties yields a best fit result for the signal strength of $\hat{\mu}_{\mathrm{B},1200} = 0.44 \pm 0.42$ for VLB quarks and $\hat{\mu}_{\mathrm{T},1200} = 0.15 \pm 0.44$. The largest impact of an individual uncertainty reaches a contribution of 0.19 for VLT quarks and 0.12 for VLB quarks at pre-fit level. The post-fit impact is significantly smaller with values of 0.06 for VLT and 0.05 for VLB quarks. In comparison to the total uncertainty on the fitted μ of the S+B fit of 0.42 and 0.44 for VLB and VLT quarks, respectively, the contribution of all systematic uncertainties is $\Delta\mu = 0.15$ for both VLB and VLT quarks. The impact on the total uncertainty from the γ factors, which describe the impact of the limited number of MC events, is 0.03. The impact on $\Delta\mu$ from the statistical uncertainties and γ factors, is derived in the same S+B fit with the NPs fixed to their best fit values and yields $\Delta\mu = 0.39$. Thus, it is concluded that the largest impact of the systematic uncertainties from Figures 9.21(a) and 9.21(b) is comparably small and the analysis is dominated by statistical uncertainties.

Chapter 10
Summary and conclusions

The open questions of the Standard Model (SM) indicate the presence of physics beyond the SM. The naturalness problem describes one of these questions and can be solved by various new theories, involving Little Higgs and Composite Higgs models. From these models, vector-like quarks (VLQs) emerge and are expected around the TeV scale which could make them detectable at the Large Hadron Collider (LHC) with the ATLAS and CMS experiments. Their decay topology into SM particles via $Q \to Z/Hq$ or $Q \to Wq'$ with either a V- or H-boson and a third generation quark q or q' is exploited to build analyses with dedicated final state sensitivity. No evidence for VLQs was found in Run-1. With a larger center-of-mass energy and an increased amount of data in Run-2, a higher sensitivity of VLQ searches is expected in Run-2.

In this context, a search for pair production of vector-like top (VLT) and vector-like bottom (VLB) quarks with a Z-boson in the final state, reconstructed from a pair of e^+e^- or $\mu^+\mu^-$, was designed for $36.1\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$ [3]. The searched decay mode with a Z-boson candidate preexisted in a Run-1 analysis [62] and is denoted as the $Z(\ell\ell)t/b + X$ search. The Run-2 analysis is divided into three channels with two 2ℓ channels, split in events with less and at least two large-R jets J, and a $\geq 3\ell$ channel. This thesis focused on the development of the $2\ell \geq 2J$ channel, whose final state topology is especially sensitive to boosted objects. As a final discriminant, the invariant mass of the leptonic Z-boson and the b-tagged jet with highest p_T, $m(Zb)$, was chosen due to its large separation power for VLT and VLB quark events from background. A profile likelihood fit was performed in the signal region and the two control regions for the main backgrounds, Z + jets and $t\bar{t}$, in order to evaluate the compatibility of the background model with the measured data. No significant deviation from the SM expectation was observed and therefore exclusion limits on the VLT and VLB quark masses and production cross-sections were set. All exclusion limits quoted in this thesis are derived at 95% confidence level. For the (X, T) or (T, B) doublet, VLT quark masses below $1102\,\text{GeV}$ are excluded. VLB quark masses below $1063\,\text{GeV}$ are incompatible with the SM expectation for the (B, Y) doublet. Modelling uncertainties for the main backgrounds Z + jets and $t\bar{t}$ showed to be the most relevant sources of systematic uncertainties; however, the statistical uncertainties dominate this search.

A higher sensitivity towards VLQs is reached by combining multiple analyses. A combination of all three channels of the $Z(\ell\ell)t/b + X$ analysis with doublet branching ratio (BR) assumptions sets higher limits for VLT and VLB quarks by excluding $m_T < 1210\,\text{GeV}$ and $m_B < 1140\,\text{GeV}$, respectively. With the results, the limits of a similar CMS search [185] with the 2016 dataset of $35.9\,\text{fb}^{-1}$ were exceeded by $60\,\text{GeV}$ and $90\,\text{GeV}$ for VLT and VLB quarks for 100% BRs into Zt/b final states, respectively. A subsequent combination of all ATLAS pair production searches for VLQs was able to push the limits higher. VLT quark masses below $1.37\,\text{TeV}$ [4] are not compatible with the

SM for both doublet assumptions, while VLB quark mass limits are set at $1.14\,\text{TeV}$ [4] for the BRs according to the (B,Y) doublet. It was shown that the $2\ell \geq 2J$ channel of the $Z(\ell\ell)t/b+X$ analysis drives the sensitivity of the VLB quark combination assuming the (B,Y) doublet, which highlights the important contribution of this thesis.

The high sensitivity of the $2\ell \geq 2J$ channel of the $Z(\ell\ell)t/b+X$ analysis which utilized boosted final states shows that boosted techniques become more important with increasing VLQ mass hypothesis. Additional knowledge about the origin of a jet allows to tailor an analysis selection to the final state topology of the signal process. A multi-class tagger for simultaneous tagging of V-boson, H-boson, top-quark and background jets was developed in this thesis and represents a novel approach. A similar tagger has first been used in an all-hadronic VLQ search [1] with variable-R reclustered jets. In this thesis, the idea was adapted for a new tagger for fixed-R reclustered jets using a deep neural network, denoted as Multi-Class Boosted-Object Tagger (MCBOT). The training was performed on reclustered (RC) jets with a rather flat p_T spectrum and the preprocessing was advanced. Significant improvements in the rejection of background jets were achieved in comparison to a tagger trained on not-flat-in-p_T samples, which is very similar to the multi-class tagger in Ref. [1]. MCBOT was developed in a standard and b-tag enriched version, where either one outperforms the other, if the analysis dataset shows more similarity to the standard or b-tag enriched training set, respectively. The most discriminating input features of MCBOT are the RC jet mass, the b-tagging information of the subjet with largest p_T and the number of constituents of the RC jet. A large advantage of MCBOT is that all its inputs are calibrated and assigned with systematic uncertainties. Therefore, an additional calibration or derivation of uncertainties is unnecessary in an analysis.

With the full Run-2 dataset of $139\,\text{fb}^{-1}$, which is almost four times larger than the 2015+2016 dataset, the search for VLQs was continued. With MCBOT at hand, an advanced strategy was developed for the 2ℓ channel of the $Z(\ell\ell)t/b+X$ analysis. Event categories were created according to the b-, V-, H- and top-tag multiplicities in order to build high sensitivity regions. The final discriminants $m(Zb)$ and $m(Z+2^{\text{nd}}\,b)$ provided the best separation between VLQ signal and background events, and they were used in categories with exactly and at least two b-tagged jets, respectively. Two control regions, which are close to the kinematics of the MCBOT categorization, were added with the discriminant $H_\text{T}(\text{jets})+E_\text{T}^\text{miss}$ in order to adjust the background normalization in a profile likelihood fit. The full Run-2 search is dominated by statistical uncertainties. The modelling and heavy flavor uncertainties of the main background $Z+\text{jets}$ are the most important systematic uncertainties.

One of the main achievements of this thesis is the improvement of the sensitivity in the 2ℓ channel of the $Z(\ell\ell)t/b+X$ analysis due to the development and the application of MCBOT. The cross-section limits with statistical uncertainties were increased by 13% to 45% depending on the VLQ BR assumption and the VLQ mass hypothesis in comparison to a categorization without MCBOT. The search showed no excess of data events above the SM expectation and exclusion limits were set for the 2ℓ channel including systematic uncertainties. For BRs according to the (X,T) or (T,B) doublet, VLT quark masses below $1350\,\text{GeV}$ are excluded. For VLB quark masses, lower limits of $m_B > 1316\,\text{GeV}$ are set, assuming the BRs of the (B,Y) doublet. In comparison to the pair production combination by the ATLAS collaboration with $36.1\,\text{fb}^{-1}$, the limits for VLB quarks in the (B,Y) model were significantly increased by $176\,\text{GeV}$. Up to the present day, no other VLQ searches of the ATLAS or CMS collaborations have been published with the full Run-2 dataset. The

search for VLT and VLB quarks with $139\,\text{fb}^{-1}$ in the 2ℓ channel of the $Z(\ell\ell)t/b+X$ analysis sets the highest mass limits on $T\bar{T}$ and $B\bar{B}$ production to date assuming the 100% BR into Zt/b final states. In addition, the VLB quark search observes the highest exclusion limit on $B\bar{B}$ production up to now assuming the BRs of the (B, Y) doublet.

The results of this thesis in the 2ℓ channel are used in a combination – currently being prepared for publication – with the $\geq 3\ell$ channel of the $Z(\ell\ell)t/b + X$ search, which also utilizes MCBOT. A combination is expected to improve the exclusion limits further. Additionally, a pair production combination of all ATLAS analyses using $139\,\text{fb}^{-1}$ could exploit the maximum sensitivity of the full Run-2 dataset. The continuing exclusion of phase-spaces for VLQ existence motivates to intensify searches for single production of VLQs, which could have a larger production cross-section. This thesis is part of the ATLAS search program for VLQs, which is currently performed for the full Run-2 dataset. With increasing amounts of data during Run-3 and the High-Luminosity LHC, it will be possible to probe more remote phase-spaces. In addition, the advancement of analysis strategies can significantly contribute to increased search sensitivities like it was shown for MCBOT.

To conclude, the phase-space for the existence of VLQs has been constrained significantly by the two analyses presented in this thesis, supported by the development of MCBOT. The results contribute to the effort to answer the question, if physics beyond the SM is accessible at the LHC.

References

[1] ATLAS Collaboration, *Search for pair production of heavy vector-like quarks decaying into hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **98** (2018) 092005, arXiv: 1808.01771 [hep-ex].

[2] E. Freundlich, *Studies for the search for pair production of vector-like quarks in the dilepton channel $Q\bar{Q} \to Z(\ell\ell)q + X$ at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, MA thesis: TU Dortmund, 2016.

[3] ATLAS Collaboration, *Search for pair- and single-production of vector-like quarks in final states with at least one Z boson decaying into a pair of electrons or muons in pp collision data collected with the ATLAS detector at $\sqrt{s} = 13$ TeV*, Phys. Rev. D **98** (2018) 112010, arXiv: 1806.10555 [hep-ex].

[4] ATLAS Collaboration, *Combination of the searches for pair-produced vector-like partners of the third-generation quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. Lett. **121** (2018) 211801, arXiv: 1808.02343 [hep-ex].

[5] E. M. Freundlich, *Search for pair- and single-production of vector-like quarks in final states with at least one Z boson decaying into a pair of electrons or muons in pp collision data collected with the ATLAS detector at $\sqrt{s}$= 13 TeV*, PoS **LHCP2018** (2018) 017.

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

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

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

[9] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B **44** (1972) 189.

[10] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001 and 2019 update.

[11] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rep. **405** (2005) 279, arXiv: hep-ph/0404175 [hep-ph].

[12] Supernova Search Team, *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, Astron. J. **116** (1998) 1009, arXiv: astro-ph/9805201 [astro-ph].

[13] Supernova Cosmology Project Collaboration, *Measurements of Ω and Λ from 42 high-redshift supernovae*, Astrophys. J. **517** (1999) 565, arXiv: astro-ph/9812133 [astro-ph].

[14] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: hep-ex/9807003.

[15] SNO Collaboration, *Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301, arXiv: nucl-ex/0106015.

[16] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe*, New J. Phys. **14** (2012) 095012, arXiv: 1204.4186 [hep-ph].

[17] 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].

[18] 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].

[19] CMS Collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, Phys. Lett. B **805** (2020) 135425, arXiv: 2002.06398 [hep-ex].

[20] L. Susskind, *Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory*, Phys. Rev. D **20** (1979) 2619.

[21] ATLAS Collaboration, *Search for production of vector-like quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **08** (2015) 105, arXiv: 1505.04306 [hep-ex].

[22] CMS Collaboration, *Search for vector-like charge 2/3 T quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. D **93** (2016) 012003, arXiv: 1509.04177 [hep-ex].

[23] CMS Collaboration, *Search for pair-produced vectorlike B quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. D **93** (2016) 112009, arXiv: 1507.07129 [hep-ex].

[24] R. Contino, L. Da Rold, and A. Pomarol, *Light custodians in natural composite Higgs models*, Phys. Rev. D **75** (2007) 055014, arXiv: hep-ph/0612048 [hep-ph].

[25] ATLAS, CDF, CMS, D0 Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, (2014), arXiv: 1403.4427 [hep-ex].

[26] Wikimedia Commons, *Standard Model of Elementary Particles*, https://commons.wikimedia.org/w/index.php?title=File:Standard_Model_of_Elementary_Particles.svg&oldid=416441632 (accessed on 22.05.2020).

[27] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

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

[29] S. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*,
Phys. Rev. D **2** (1970) 1285.

[30] S. L. Glashow, *The renormalizability of vector meson interactions*,
Nucl. Phys. **10** (1959) 107.

[31] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*,
Phys. Rev. Lett. **13** (1964) 321.

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

[33] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*,
Phys. Rev. Lett. **13** (1964) 508.

[34] G. Guralnik, C. Hagen, and T. Kibble, *Global conservation laws and massless particles*,
Phys. Rev. Lett. **13** (1964) 585.

[35] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*,
Phys. Rev. **145** (1966) 1156.

[36] T. Kibble, *Symmetry breaking in non-Abelian gauge theories*, Phys. Rev. **155** (1967) 1554.

[37] CMS Collaboration, *Observation of $t\bar{t}H$ Production*, Phys. Rev. Lett. **120** (2018) 231801,
arXiv: 1804.02610 [hep-ex].

[38] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173,
arXiv: 1806.00425 [hep-ex].

[39] H. Georgi, H. R. Quinn, and S. Weinberg,
Hierarchy of Interactions in Unified Gauge Theories, Phys. Rev. Lett. **33** (1974) 451.

[40] G. F. Giudice and A. Romanino, *Split supersymmetry*,
Nucl. Phys. B **699** (2004) 65, [Erratum: Nucl. Phys. B 706 (2005) 487],
arXiv: hep-ph/0406088 [hep-ph].

[41] ATLAS Collaboration, *Constraints on mediator-based dark matter and scalar dark energy models using $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector*,
JHEP **05** (2019) 142, arXiv: 1903.01400 [hep-ex].

[42] CMS Collaboration, *Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV*, (2019),
arXiv: 1911.03947 [hep-ex].

[43] CMS Collaboration, *Search for dark photons in decays of Higgs bosons produced in association with Z bosons in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **10** (2019) 139,
arXiv: 1908.02699 [hep-ex].

[44] D. H. Weinberg et al., *Observational probes of cosmic acceleration*, Phys. Rep. **530** (2013) 87.

[45] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **38** (1977) 1440.

[46] C. A. Baker et al., *An Improved experimental limit on the electric dipole moment of the neutron*, Phys. Rev. Lett. **97** (2006) 131801, arXiv: hep-ex/0602020 [hep-ex].

[47] J. D. Wells, *Higgs naturalness and the scalar boson proliferation instability problem*, Synthese **194** (2017) 477, arXiv: 1603.06131 [hep-ph].

[48] A. Djouadi and A. Lenz, *Sealing the fate of a fourth generation of fermions*, Phys. Lett. B **715** (2012) 310, arXiv: 1204.1252 [hep-ph].

[49] O. Eberhardt et al., *Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations*, Phys. Rev. Lett. **109** (2012) 241802, arXiv: 1209.1101 [hep-ph].

[50] ATLAS Collaboration, *Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2013-034, 2013.

[51] CMS Collaboration, *Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV*, CMS-PAS-HIG-12-045, 2012.

[52] ATLAS Collaboration, *Search for pair production of heavy top-like quarks decaying to a high-p_T W boson and a b quark in the lepton plus jets final state at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Phys. Lett. B **718** (2013) 1284, arXiv: 1210.5468 [hep-ex].

[53] CMS Collaboration, *Search for pair produced fourth-generation up-type quarks in pp collisions at $\sqrt{s} = 7$ TeV with a lepton in the final state*, Phys. Lett. B **718** (2012) 307, arXiv: 1209.0471 [hep-ex].

[54] P. Fayet, *Supersymmetry and Weak, Electromagnetic and Strong Interactions*, Phys. Lett. B **64** (1976) 159.

[55] N. Arkani-Hamed, A. Cohen, E. Katz, and A. Nelson, *The littlest Higgs*, JHEP **07** (2002) 034, arXiv: hep-ph/0206021 [hep-ph].

[56] M. Schmaltz and D. Tucker-Smith, *Little Higgs Review*, Ann. Rev. Nucl. Part. Sci. **55** (2005) 229, arXiv: hep-ph/0502182 [hep-ph].

[57] D. B. Kaplan, H. Georgi, and S. Dimopoulos, *Composite Higgs scalars*, Phys. Lett. B **136** (1984) 187.

[58] K. Agashe, R. Contino, and A. Pomarol, *The minimal composite Higgs model*, Nucl. Phys. B **719** (2005) 165, arXiv: hep-ph/0412089 [hep-ph].

[59] H.-C. Cheng, *Little Higgs, Non-standard Higgs, No Higgs and All That*, (2007), arXiv: 0710.3407 [hep-ph].

[60] J. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Perez-Victoria, *Handbook of vectorlike quarks: mixing and single production*, Phys. Rev. D **88** (2013) 94010, arXiv: 1306.0572 [hep-ph].

[61] F. del Aguila, M. Perez-Victoria, and J. Santiago, *Observable contributions of new exotic quarks to quark mixing*, JHEP **09** (2000) 011, arXiv: hep-ph/0007316 [hep-ph].

[62] ATLAS Collaboration, *Search for pair and single production of new heavy quarks that decay to a Z boson and a third-generation quark in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **11** (2014) 104, arXiv: 1409.5500 [hep-ex].

[63] F. del Aguila and M. J. Bowick, *The possibility of new fermions with $\Delta I = 0$ mass*, Nucl. Phys. B **224** (1983) 107.

[64] G. Branco and L. Lavoura, *On the addition of vector-like quarks to the Standard Model*, Nucl. Phys. B **278** (1986) 738.

[65] J. Aguilar-Saavedra, *Identifying top partners at LHC*, JHEP **11** (2009) 030, arXiv: 0907.3155 [hep-ph].

[66] J. Aguilar-Saavedra, *Protos - PROgram for TOp Simulations*, http://jaguilar.web.cern.ch/jaguilar/protos (accessed on 26.05.2020).

[67] M. Chala, *Direct bounds on heavy toplike quarks with standard and exotic decays*, Phys. Rev. D **96** (2017) 015028, arXiv: 1705.03013 [hep-ph].

[68] O. S. Brüning et al., *LHC Design Report Vol.1: The LHC Main Ring*, CERN-2004-003-V-1, 2004.

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

[70] ATLAS Collaboration, *Search for vector-like B quarks in events with one isolated lepton, missing transverse momentum and jets at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) 112011, arXiv: 1503.05425 [hep-ex].

[71] ATLAS Collaboration, *Analysis of events with b-jets and a pair of leptons of the same charge in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **10** (2015) 150, arXiv: 1504.04605 [hep-ex].

[72] ATLAS Collaboration, *Search for the production of single vector-like and excited quarks in the Wt final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **02** (2016) 110, arXiv: 1510.02664 [hep-ex].

[73] ATLAS Collaboration, *Search for single production of vector-like quarks decaying into Wb in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 442, arXiv: 1602.05606 [hep-ex].

[74] W. Stirling, *Parton luminosity ratios for 8 and 13 TeV*, http://www.hep.ph.ic.ac.uk/~wstirlin/plots/lhclumi7813_2013_v1.pdf (accessed on 04.04.2020).

[75] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019.

[76] ATLAS Collaboration, *Luminosity public results Run-2*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2` (accessed on 20.02.2020).

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

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

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

[80] ATLAS IBL Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19, 2010, Addendum: ATLAS-TDR-19-ADD-1 (2012).

[81] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016.

[82] ATLAS Collaboration, *ATLAS Forward Detectors for Luminosity Measurement and Monitoring*, CERN-LHCC-2004-010, 2004.

[83] ATLAS Collaboration, *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*, ATLAS-TDR-18, CERN-LHCC-2008-004, 2008.

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

[85] E. D. Bloom et al., *High-energy inelastic e-p scattering at 6° and 10°*, Phys. Rev. Lett. **23** (1969) 930.

[86] M. Breidenbach et al., *Observed behavior of highly inelastic electron-proton scattering*, Phys. Rev. Lett. **23** (1969) 935.

[87] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rep. **504** (2011) 145, arXiv: `1101.2599 [hep-ph]`.

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

[89] B. Andersson, *The Lund Model*, Cambridge University Press, 1998, ISBN: 9780511524363.

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

[91] B. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nucl. Phys. B **238** (1984) 492.

[92] J.-C. Winter, F. Krauss, and G. Soff, *A Modified cluster hadronization model*, Eur. Phys. J. C **36** (2004) 381, arXiv: `hep-ph/0311085`.

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

[94] Sherpa Collaboration, *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph].

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

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

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

[98] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[99] 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].

[100] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex].

[101] K. Rehermann and B. Tweedie, *Efficient Identification of Boosted Semileptonic Top Quarks at the LHC*, JHEP **03** (2011) 059, arXiv: 1007.2221 [hep-ph].

[102] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, JINST **14** (2019) 03017, arXiv: 1812.03848 [hep-ex].

[103] 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].

[104] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[105] *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-036, 2015.

[106] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex].

[107] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015.

[108] 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].

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

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

[111] Emily Thompson, *What does "boost" mean?*, https://www.quantumdiaries.org/author/emily-thompson (accessed on 02.03.2020).

[112] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, JHEP **02** (2010) 084, arXiv: 0912.1342 [hep-ph].

[113] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. Webber, *New clustering algorithm for multijet cross-sections in e^+e^- annihilation*, Phys. Lett. B **269** (1991) 432.

[114] S. Catani, Y. Dokshitzer, M. Seymour, and B. Webber, *Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions*, Nucl. Phys. B **406** (1993) 187.

[115] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (1993) 3160, arXiv: hep-ph/9305266.

[116] ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035, 2016.

[117] ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **09** (2013) 076, arXiv: 1306.4945 [hep-ex].

[118] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, and C. Wanotayaroj, *Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, JHEP **02** (2015) 075, arXiv: 1407.2922 [hep-ph].

[119] ATLAS Collaboration, *Jet reclustering and close-by effects in ATLAS Run 2*, ATLAS-CONF-2017-062, 2017.

[120] 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].

[121] ATLAS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59, arXiv: 1808.08238 [hep-ex].

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

[123] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex].

[124] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, 2011.

[125] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex].

[126] ATLAS Collaboration, *Search for pair production of heavy vector-like quarks decaying into high-p_T W bosons and top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector*, JHEP **08** (2018) 048, arXiv: 1806.01762 [hep-ex].

[127] ATLAS Collaboration, *Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **05** (2019) 041, arXiv: 1812.09743 [hep-ex].

[128] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, Eur. Phys. J. C **79** (2019) 375, arXiv: 1808.07858 [hep-ex].

[129] ATLAS Collaboration, *Identification of boosted Higgs bosons decaying into b-quark pairs with the ATLAS detector at 13 TeV*, Eur. Phys. J. C **79** (2019) 836, arXiv: 1906.11005 [hep-ex].

[130] W. S. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bull. math. Biophys. **5** (1943) 115.

[131] B. Widrow, *An adaptive 'Adaline' neuron using chemical 'memistors'*, Stanford Electronics Laboratories Technical Report **1553-2** (1960).

[132] F. Rosenblatt, *Principles of neurodynamics: Perceptrons and the theory of brain machanisms*, Spartan Books, 1962, ISBN: 978-3-642-70911-1.

[133] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature **323** (1986) 533.

[134] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, 2006, ISBN: 978-0387-31073-2.

[135] A. Cauchy, *Méthode générale pour la résolution des systèmes d'équations simultanées*, C.R. Acad. Sci. Paris (1847) 536.

[136] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, (2014), arXiv: 1412.6980 [cs.LG].

[137] R. Tibshirani, *Regression Shrinkage and Selection Via the Lasso*, J. Royal Stat. Soc. B **58** (1994) 267.

[138] F. Chollet et al., *Keras*, https://keras.io (accessed on 03.03.2020).

[139] A. Martín et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, https://www.tensorflow.org/ (accessed on 04.03.2020).

[140] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, CoRR (2015), arXiv: 1502.03167.

[141] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014.

[142] 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].

[143] 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 (2013) 2501], arXiv: 1007.1727 [physics.data-an].

[144] J. Neyman, E. S. Pearson, and K. Pearson, *IX. On the problem of the most efficient tests of statistical hypotheses*, Philos. Trans. Royal Soc. A **231** (1933) 289.

[145] ROOT Collaboration, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, (2012).

[146] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60.

[147] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Trans. Am. Math. Soc. **54** (1943) 426.

[148] G. Cowan, *The CL_s method: information for conference speakers*, http://www.pp.rhul.ac.uk/~cowan/atlas/CLsInfo.pdf (accessed on 20.05.2020).

[149] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81.

[150] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, (2003), arXiv: physics/0306116.

[151] W. Verkerke and D. Kirkby, *RooFit Users Manual*, http://roofit.sourceforge.net/ (accessed on 02.04.2020).

[152] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in $e\mu$ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2019), arXiv: 1910.08819 [hep-ex].

[153] CMS Collaboration, *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 011, arXiv: 1711.02547 [hep-ex].

[154] ATLAS Collaboration,
*Auxiliary material of $Z(\ell\ell)t/b + X$ VLQ analysis in Phys. Rev. D **98** (2018) 112010*,
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EXOT-2016-35 (accessed on 04.04.2020).

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

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

[157] S. Schumann and F. Krauss,
A Parton shower algorithm based on Catani-Seymour dipole factorisation,
JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph].

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

[159] 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].

[160] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re,
Top-pair production and decay at NLO matched with parton showers, JHEP **04** (2015) 114,
arXiv: 1412.1828 [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 [hep-ph].

[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] 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].

[164] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*,
JHEP **06** (2011) 128, arXiv: 1106.0522 [hep-ph].

[165] H.-L. Lai et al., *New parton distributions for collider physics*,
Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[166] W.-K. Tung,
New generation of parton distributions with uncertainties from global QCD analysis,
Acta Phys. Pol. B **33** (2002) 2933, arXiv: hep-ph/0206114 [hep-ph].

[167] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*,
Phys. Rev. D **82** (2010) 074018, arXiv: 1005.3457 [hep-ph].

[168] J. P. Araque Espinosa,
Search for heavy fermions with the ATLAS experiment at the LHC collider,
CERN-THESIS-2015-330, PhD thesis: LIP, Minho, 2015.

[169] 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].

[170] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*,
Eur. Phys. J. C **63** (2009) 189, arXiv: 0901.0002 [hep-ph].

[171] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections*,
Eur. Phys. J. C **64** (2009) 653, arXiv: 0905.3531 [hep-ph].

[172] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-015, 2015.

[173] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 135,
arXiv: 1807.09477 [hep-ex].

[174] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne,
Parton distributions in the LHC era: MMHT 2014 PDFs, Eur. Phys. J. C **75** (2015) 204,
arXiv: 1412.3989 [hep-ph].

[175] S. Dulat et al.,
New parton distribution functions from a global analysis of quantum chromodynamics,
Phys. Rev. D **93** (2016) 033006, arXiv: 1506.07443 [hep-ph].

[176] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653,
arXiv: 1608.03953 [hep-ex].

[177] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS analyses*,
ATL-PHYS-PUB-2016-002, 2016.

[178] ATLAS Collaboration,
Simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV,
ATL-PHYS-PUB-2016-004, 2016.

[179] ATLAS Collaboration,
Modelling of the $t\bar{t}H$ and $t\bar{t}V(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses,
ATL-PHYS-PUB-2016-005, 2016.

[180] ATLAS Collaboration, *Search for pair production of up-type vector-like quarks and for four-top-quark events in final states with multiple b-jets with the ATLAS detector*,
JHEP **07** (2018) 089, arXiv: 1803.09678 [hep-ex].

[181] ATLAS Collaboration, *Search for pair production of heavy vector-like quarks decaying to high-p_T W bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **10** (2017) 141, arXiv: 1707.03347 [hep-ex].

[182] ATLAS Collaboration,
Search for pair production of vector-like top quarks in events with one lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector,
JHEP **08** (2017) 052, arXiv: 1705.10751 [hep-ex].

[183] ATLAS Collaboration, *Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **12** (2018) 039, arXiv: 1807.11883 [hep-ex].

[184] K. Kondo, *Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1: Method and Toy Models*, J. Phys. Soc. Jap. **57** (1988) 4126.

[185] CMS Collaboration, *Search for vector-like quarks in events with two oppositely charged leptons and jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 364, arXiv: 1812.09768 [hep-ex].

[186] CMS Collaboration,
Search for pair production of vectorlike quarks in the fully hadronic final state,
Phys. Rev. D **100** (2019) 072001, arXiv: 1906.11903 [hep-ex].

[187] CMS Collaboration,
Search for vector-like T and B quark pairs in final states with leptons at $\sqrt{s} = 13$ TeV,
JHEP **08** (2018) 177, arXiv: 1805.04758 [hep-ex].

[188] CMS Collaboration, *Search for single production of vector-like quarks decaying to a Z boson and a top or a bottom quark in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
JHEP **05** (2017) 029, arXiv: 1701.07409 [hep-ex].

[189] CMS Collaboration,
Search for single production of vector-like quarks decaying to a b quark and a Higgs boson,
JHEP **06** (2018) 031, arXiv: 1802.01486 [hep-ex].

[190] CMS Collaboration, *Search for single production of vector-like quarks decaying to a top quark and a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J. C **79** (2019) 90, arXiv: 1809.08597 [hep-ex].

[191] CMS Collaboration, *Search for single production of a vector-like T quark decaying to a Z boson and a top quark in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
Phys. Lett. B **781** (2018) 574, arXiv: 1708.01062 [hep-ex].

[192] ATLAS Collaboration, *Search for single production of vector-like quarks decaying into Wb in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **05** (2019) 164, arXiv: 1812.07343 [hep-ex].

[193] CMS Collaboration, *Search for single production of vector-like quarks decaying into a b quark and a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
Phys. Lett. B **772** (2017) 634, arXiv: 1701.08328 [hep-ex].

[194] CMS Collaboration, *Search for electroweak production of a vector-like quark decaying to a top quark and a Higgs boson using boosted topologies in fully hadronic final states*,
JHEP **04** (2017) 136, arXiv: 1612.05336 [hep-ex].

[195] CMS Collaboration, *Search for single production of a heavy vector-like T quark decaying to a Higgs boson and a top quark with a lepton and jets in the final state*, Phys. Lett. B **771** (2017) 80, arXiv: 1612.00999 [hep-ex].

[196] CMS Collaboration, *Search for electroweak production of a vector-like T quark using fully hadronic final states*, JHEP **01** (2020) 036, arXiv: 1909.04721 [hep-ex].

[197] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, ATL-PHYS-PUB-2017-006, 2017.

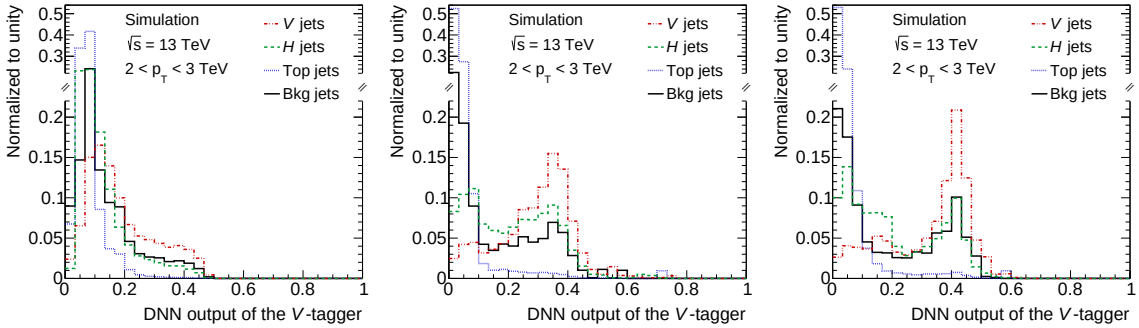
Glossary

BR branching ratio
CL confidence level
CR control region
DNN deep neural network
EM electromagnetic
FLIP FLat-In-p_{T}
HEP high energy physics
HF heavy flavor
ID Inner Detector
JES jet energy scale
JMS jet mass scale
JVT jet vertex tagger
LHC Large Hadron Collider
MC Monte Carlo
MCBOT Multi-Class Boosted-Object Tagger
ML machine learning
MS Muon Spectrometer
NDF number of degrees of freedom
NFLIP Non-FLat-In-p_{T}
NLO next-to-leading order
NN neural network
NP nuisance parameter
OSSF opposite-sign same-flavor
PDF parton distribution function
pdf probability density function
PLR profile likelihood ratio
QCD quantum chromodynamics
QED quantum electrodynamics
RC reclustered
SCT semiconductor tracker
SM Standard Model
SR signal region
SSB spontaneous symmetry breaking
TRT transition radiation tracker
VLB vector-like bottom
VLQ vector-like quark
VLT vector-like top
vRC variable-R reclustered
WP working point

Appendices

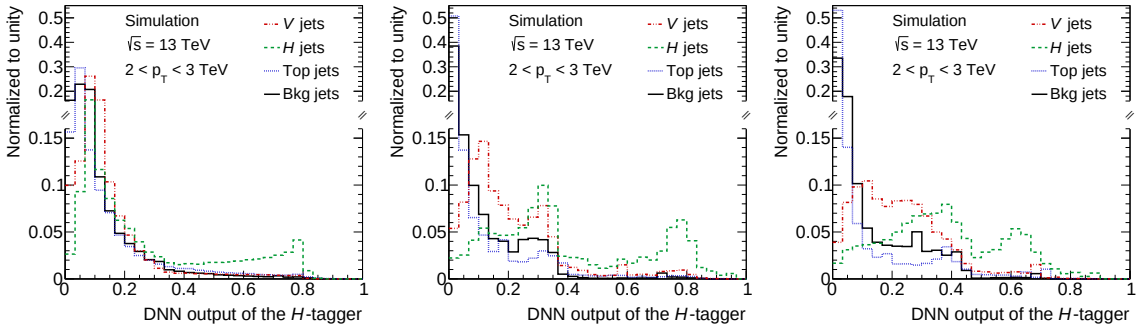
Appendix A
MCBOT figures

A.1 Standard test sample from 2–3 TeV



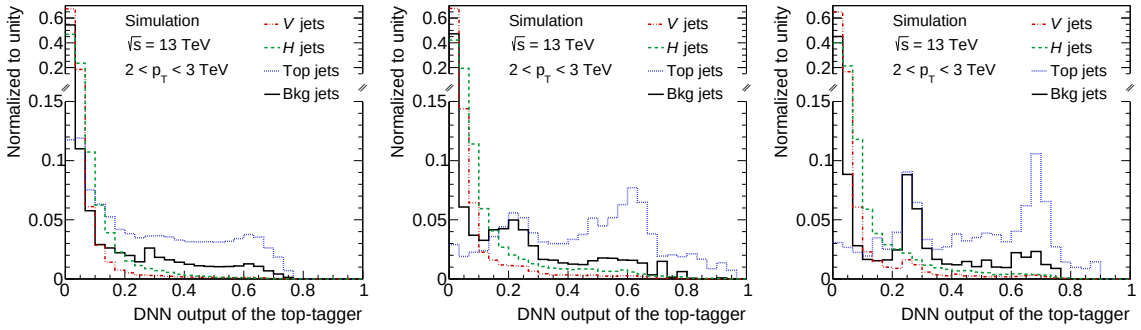
(a) (b) (c)

Figure A.1: DNN output distributions of the V-tagger for a) the NFLIP tagger, b) the standard FLIP tagger and c) the b-tag enriched FLIP tagger from 2–3 TeV. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.



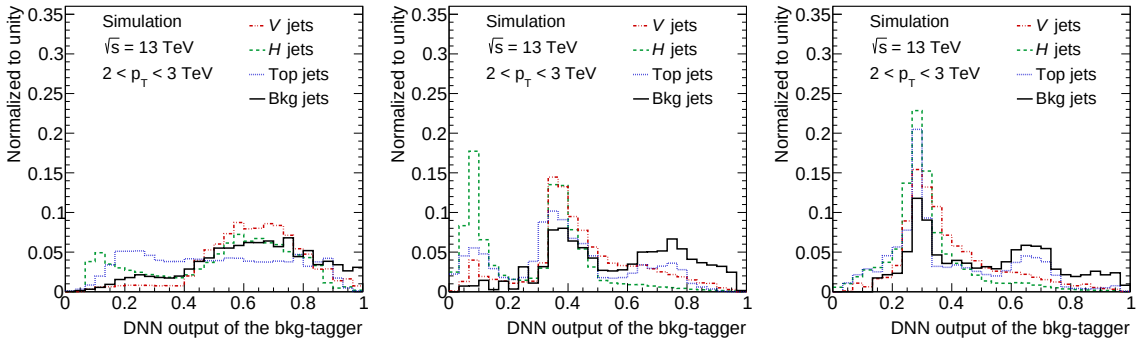
(a) (b) (c)

Figure A.2: DNN output distributions of the H-tagger for a) the NFLIP tagger, b) the standard FLIP tagger and c) the b-tag enriched FLIP tagger from 2–3 TeV. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.



(a) (b) (c)

Figure A.3: DNN output distributions of the top-tagger for a) the NFLIP tagger, b) the standard FLIP tagger and c) the b-tag enriched FLIP tagger from 2–3 TeV. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.



(a) (b) (c)

Figure A.4: DNN output distributions of the background-tagger for a) the NFLIP tagger, b) the standard FLIP tagger and c) the b-tag enriched FLIP tagger from 2–3 TeV. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

A.2 Testing on the b-tag enriched sample

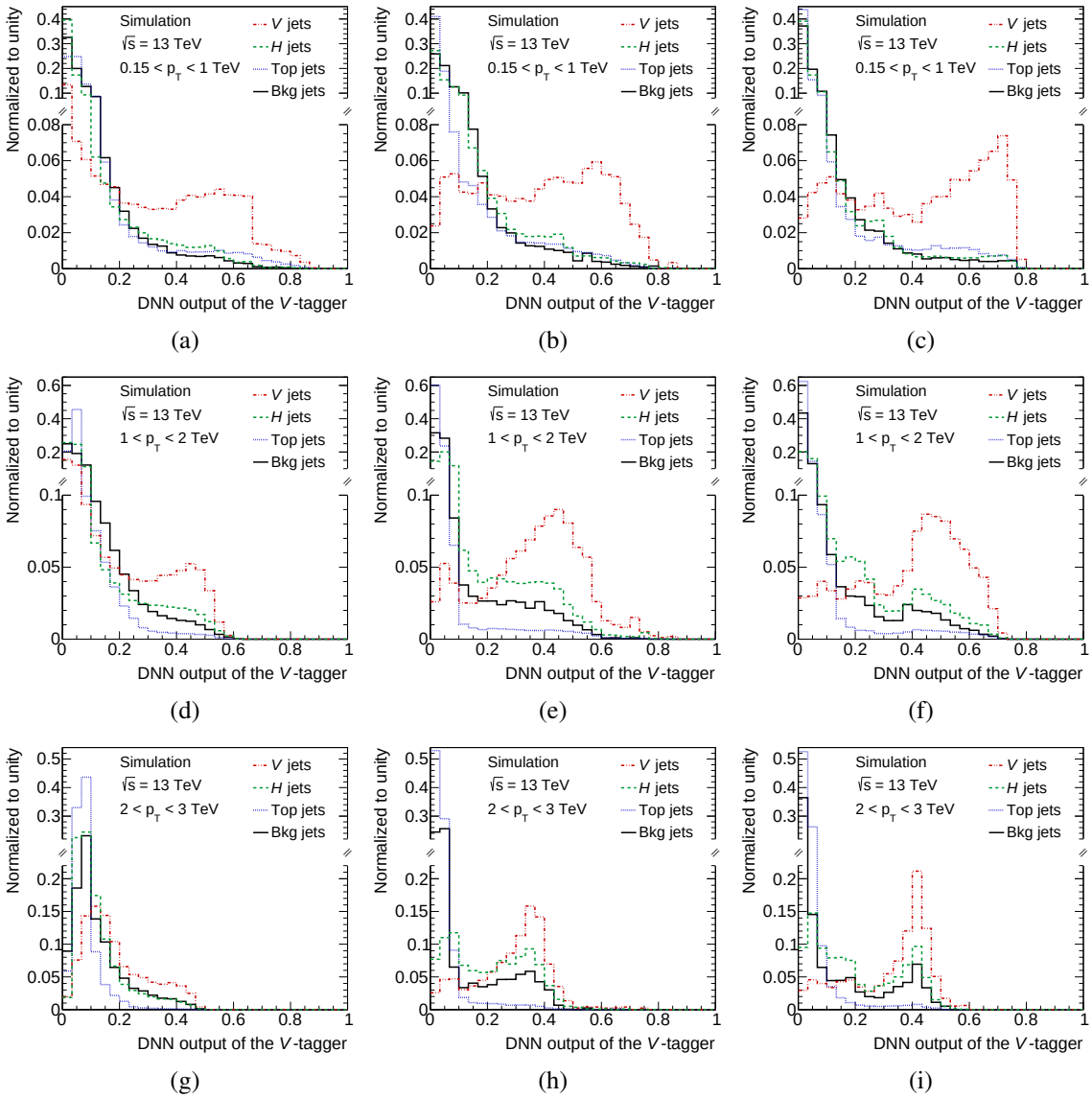


Figure A.5: DNN output distributions of the V-tagger evaluated on the b-tag enriched test sample for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top, 1–2 TeV in the middle and 2–3 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

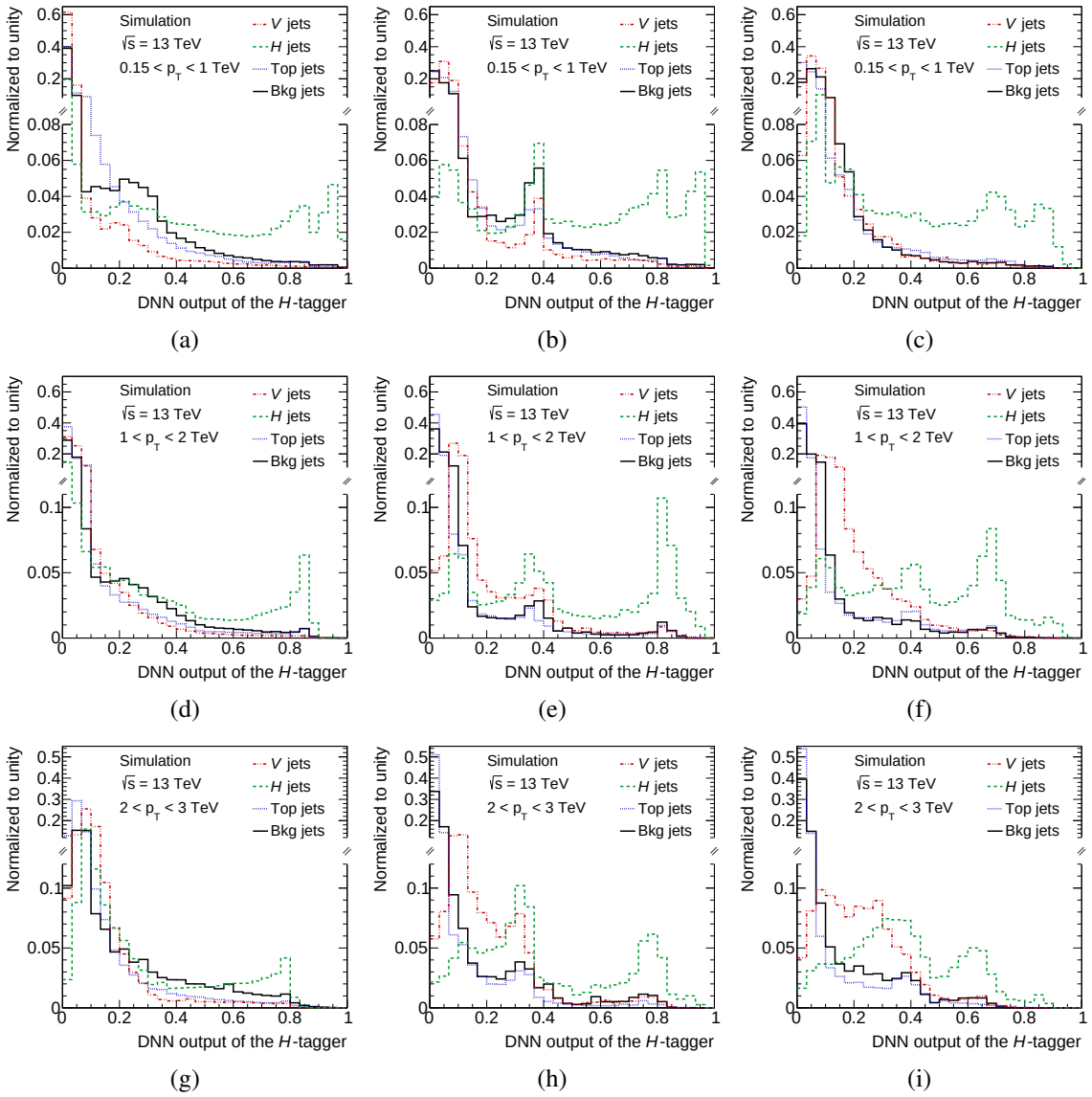


Figure A.6: DNN output distributions of the H-tagger evaluated on the b-tag enriched test sample for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top, 1–2 TeV in the middle and 2–3 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

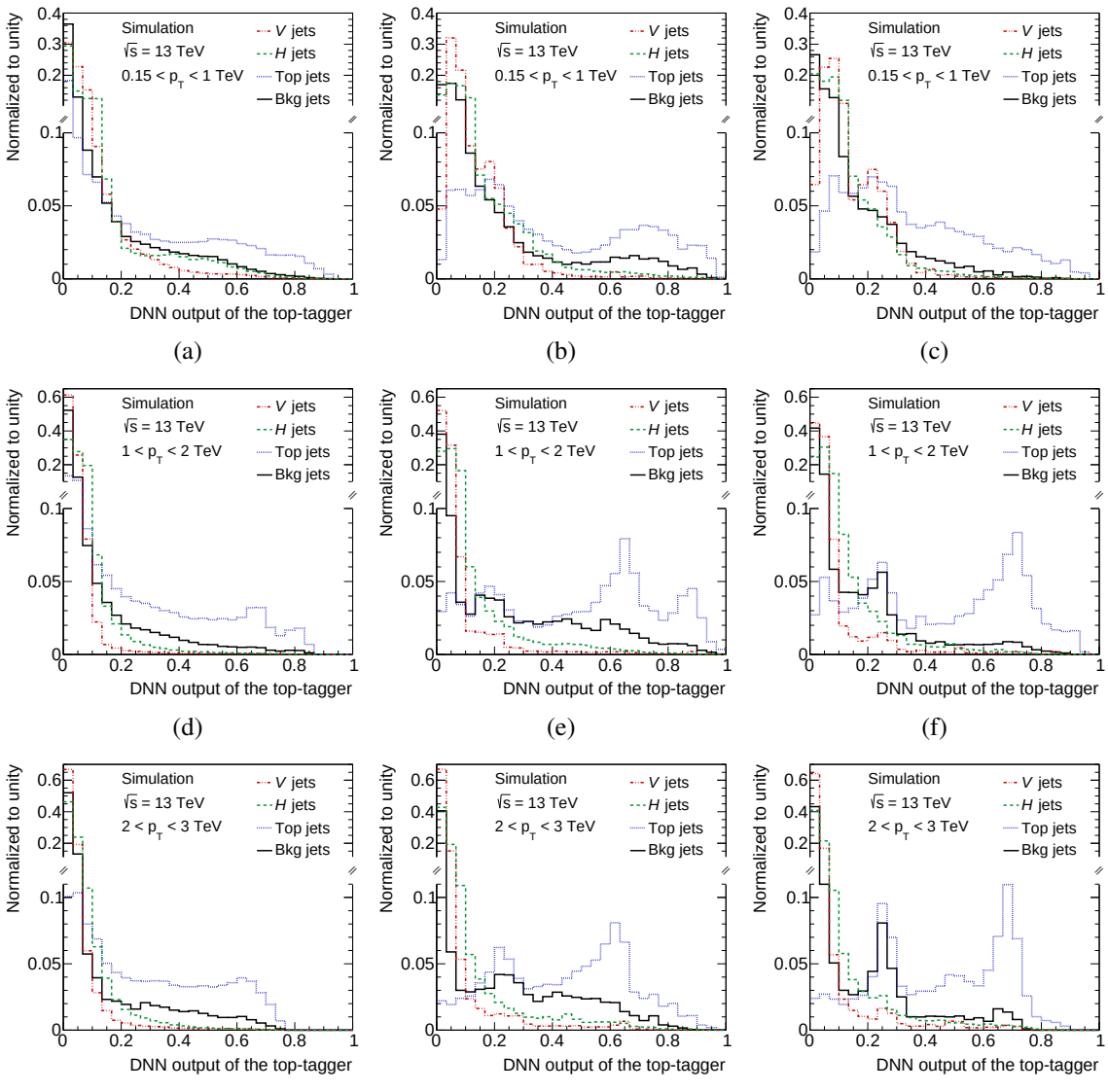


Figure A.7: DNN output distributions of the top-tagger evaluated on the b-tag enriched test sample for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top, 1–2 TeV in the middle and 2–3 TeV on the bottom. All histograms are normalized to unity. For better readability, the vertical axis is split into two intervals.

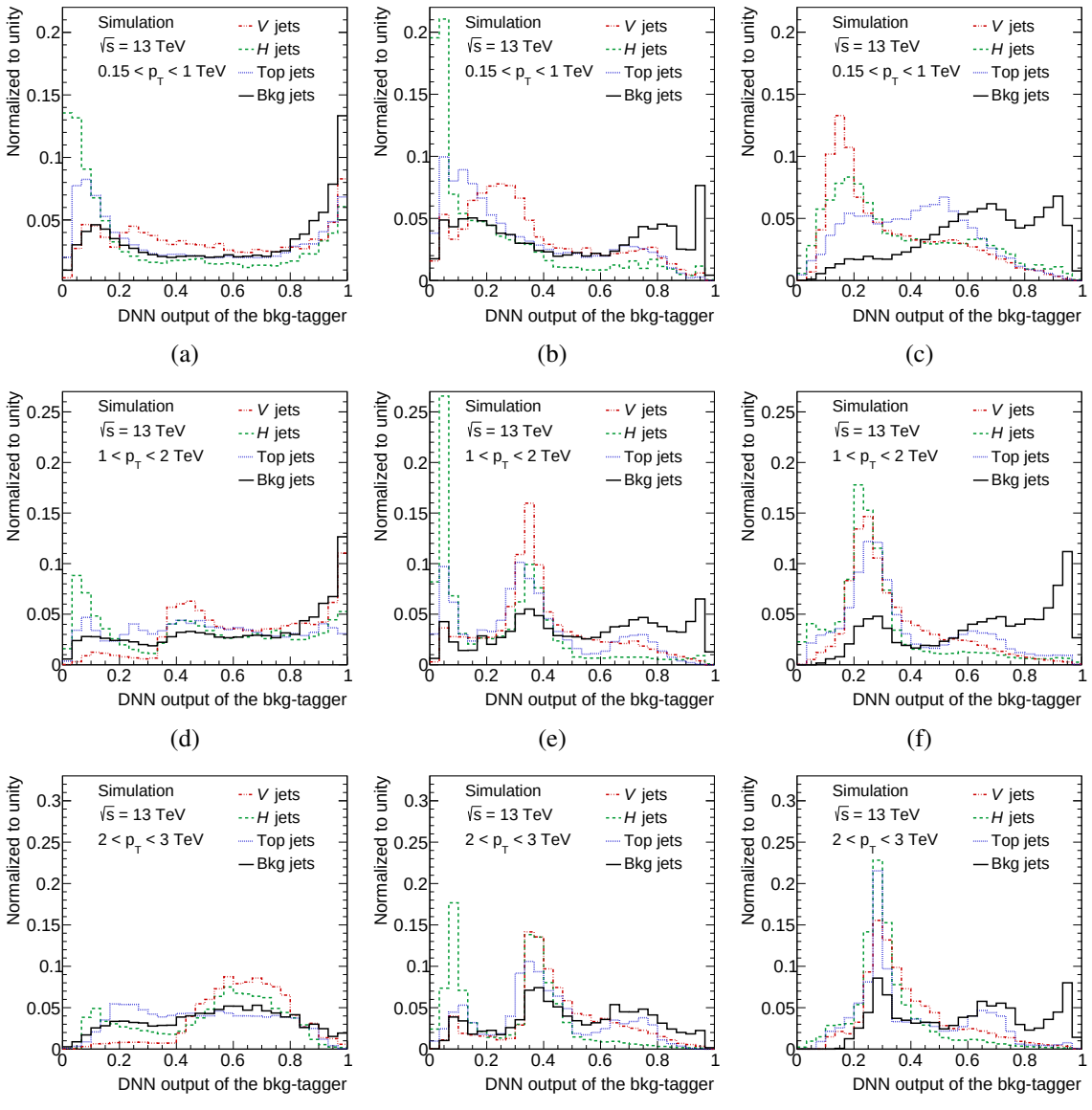


Figure A.8: DNN output distributions of the background-tagger evaluated on the b-tag enriched test sample for the NFLIP tagger on the left, the standard FLIP tagger in the middle and the b-tag enriched FLIP tagger on the right. The first p_{T} bin from 150 GeV to 1 TeV is shown on the top, 1–2 TeV in the middle and 2–3 TeV on the bottom. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

Appendix B

Figures for the $Z(\ell\ell)t/b + X$ search with 36.1 fb^{-1}

B.1 Pre-fit distributions for the $2\ell \geq 2J$ channel

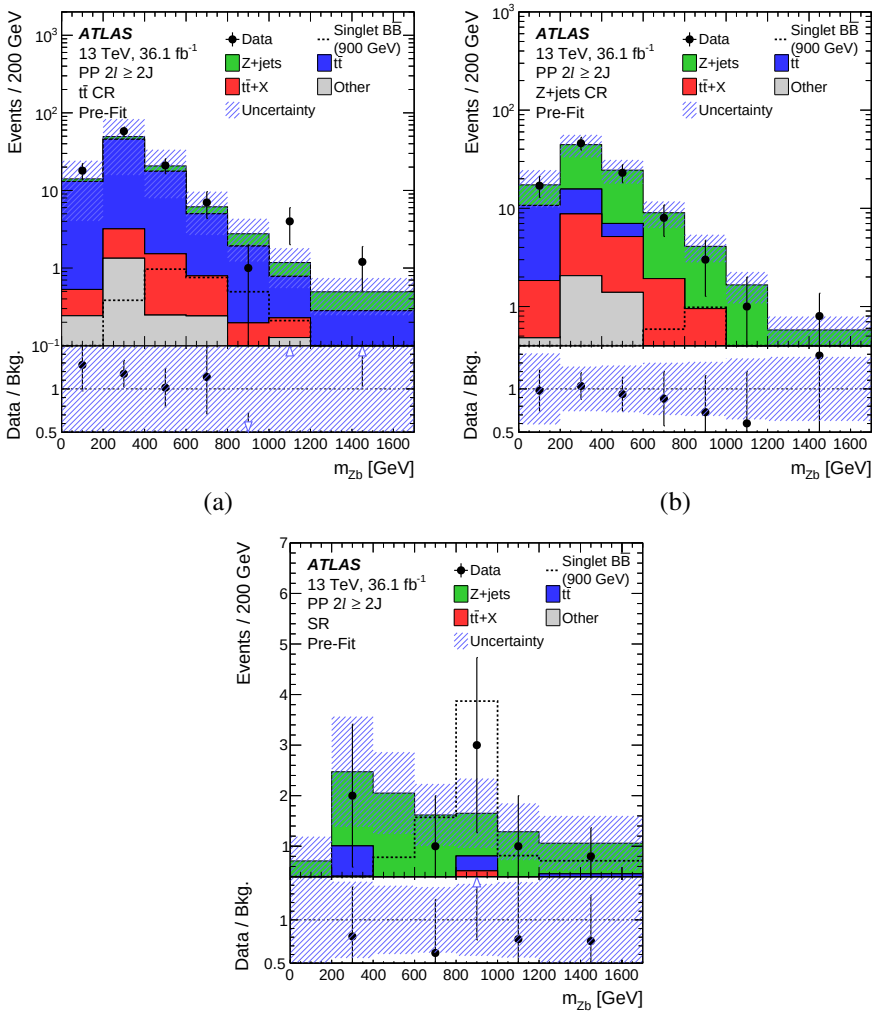


Figure B.1: Pre-fit distributions [154] for $m(Zb)$ in all three regions. The $t\bar{t}$ CR and Z + jets CR are shown in a) and b), respectively, while the SR is shown in c). Signal according to the pre-fit hypothesis is overlaid. All systematic uncertainties are included. The background is modeled according to the B-only hypothesis. The lower plot shows the ratio of data over the background expectation.

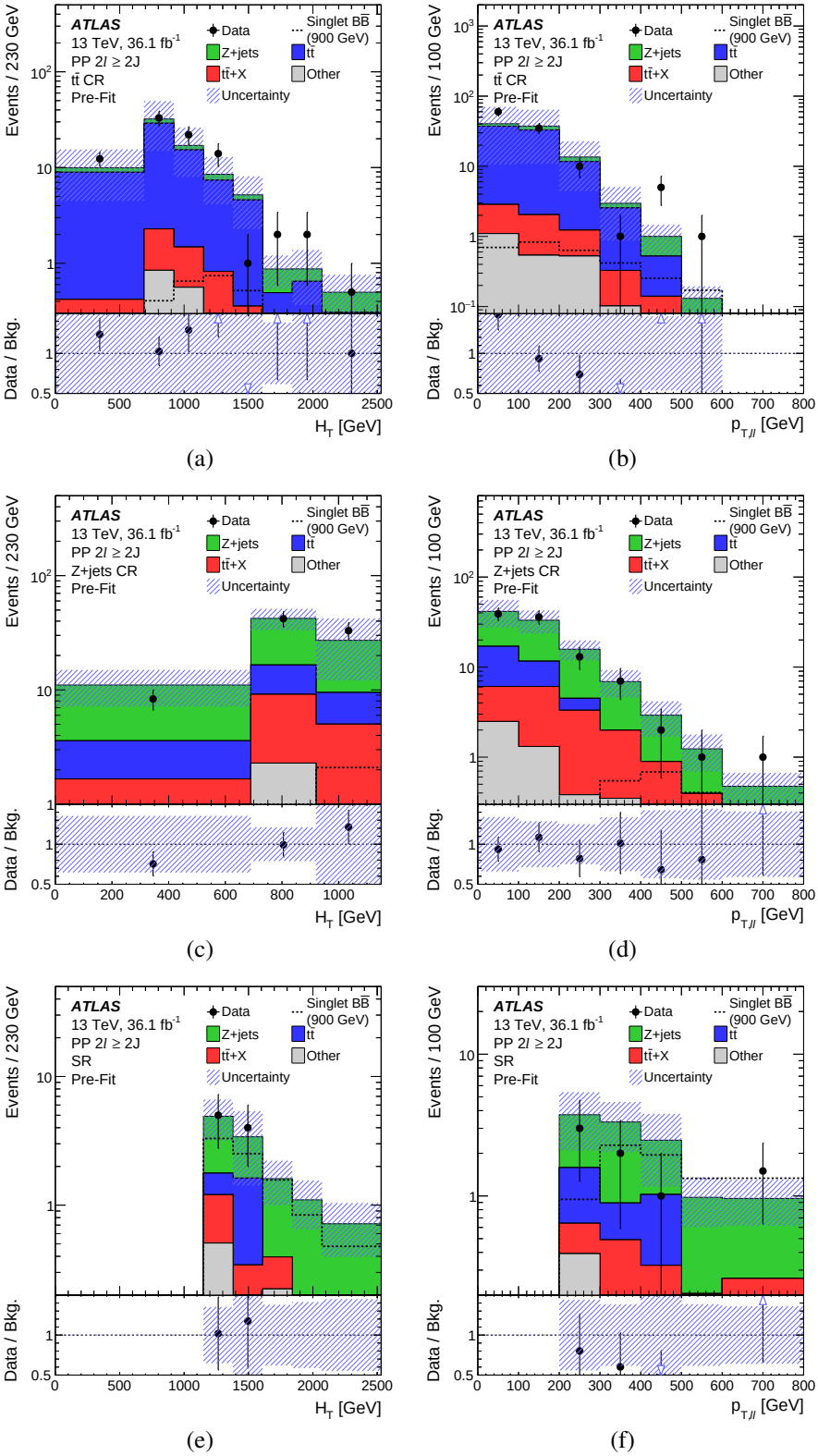


Figure B.2: Pre-fit distributions [154] for $H_{\mathrm{T}}(\mathrm{jets})$ and $p_{\mathrm{T}}\,(Z)$ in all three regions. The $t\bar{t}$ CR is shown in a) and b), while the Z + jets region is illustrated in c) and d). The SR is shown in e) and f). Signal according to the pre-fit hypothesis is overlaid. All systematic uncertainties are included. The background is modeled according to the B-only hypothesis. The lower plot shows the ratio of data over the background expectation.

B.2 Limits from the $2\ell \geq 2J$ channel

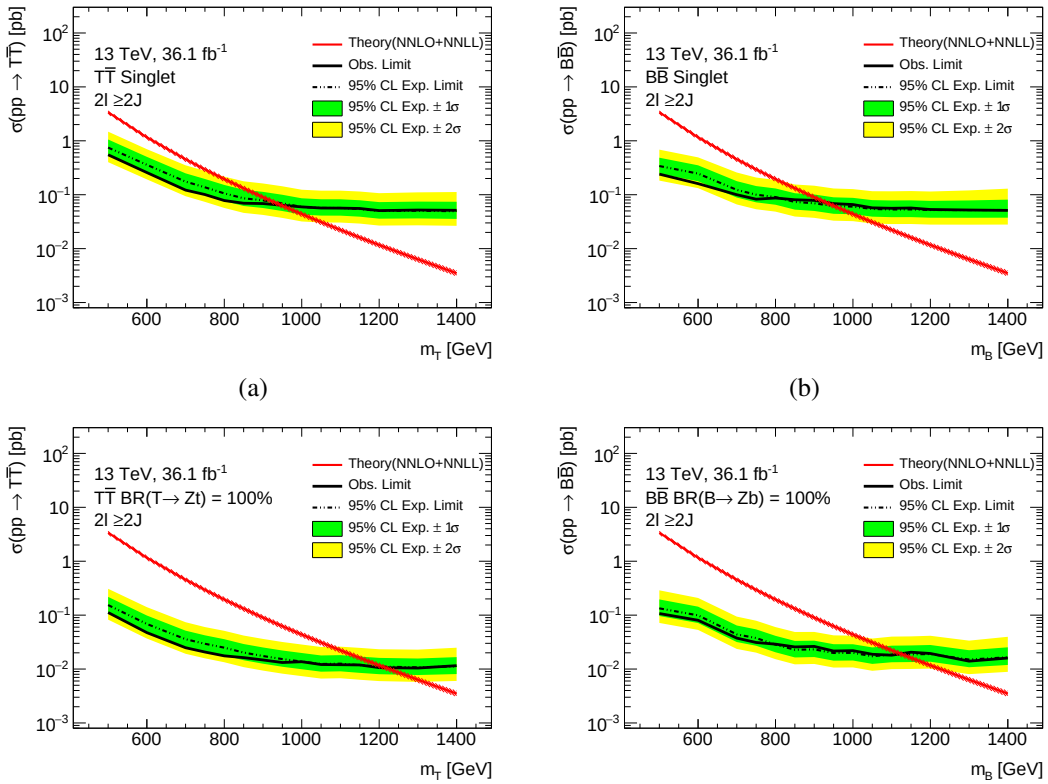


Figure B.3: Expected and observed limits on the cross-section for $T\bar{T}$ and $B\bar{B}$ production from the $2\ell \geq 2J$ channel of the $Z(\ell\ell)t/b + X$ analysis at 95% CL using $36.1\,\text{fb}^{-1}$. The limits for VLT quarks are illustrated for a) the singlet model and c) the 100% BR into Zt. The limits for VLB quarks are depicted for b) the singlet model and d) BR$(B \to Zb) \to 100\%$. The $\pm 1(2)\sigma$ band of the 95% CL limit are depicted in green (yellow). The theory curve is shown in red.

B.3 Limits from the $Z(\ell\ell)t/b + X$ combination

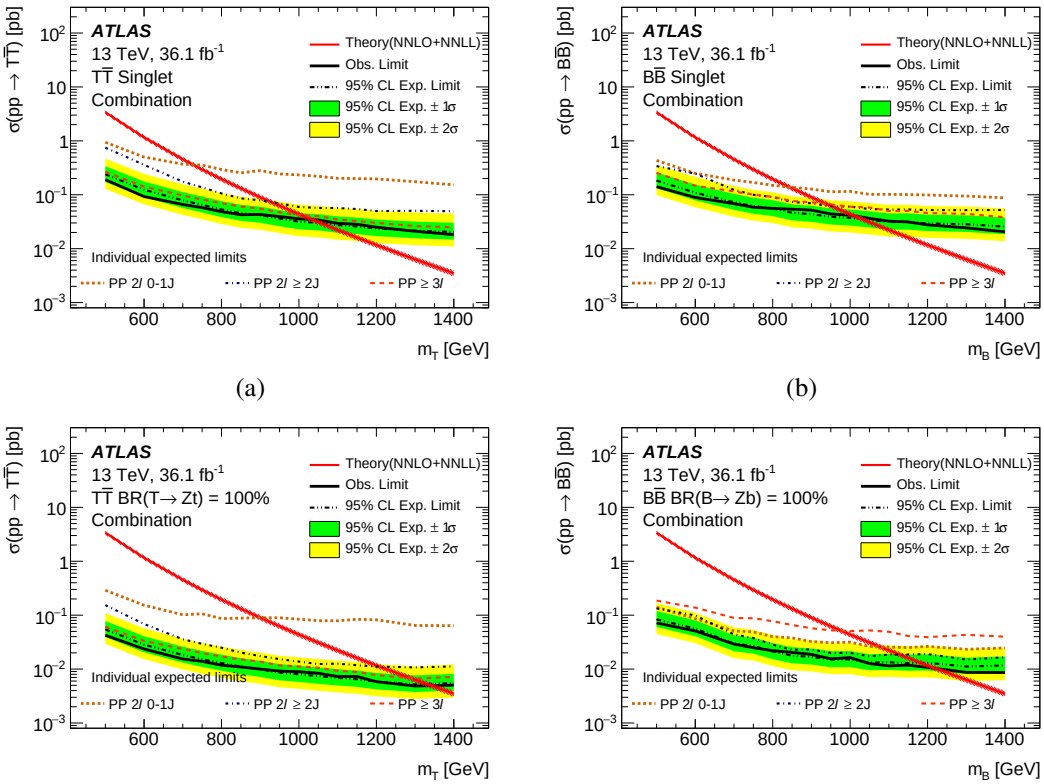


Figure B.4: Expected and observed limits on the cross-section for $T\bar{T}$ and $B\bar{B}$ production assuming the singlet model and 100% BR into Zt/b for the combination of all $Z(\ell\ell)t/b + X$ channels at 95% CL [3]. Furthermore, the expected limits of the individual channels 2ℓ 0–1J, $2\ell \geq 2J$ and $\geq 3\ell$ of the analysis are shown. The limits for VLT quarks are illustrated for a) the singlet model and c) the 100% BR into Zt. The limits for VLB quarks are depicted for b) the singlet model and d) BR$(B \to Zb) \to 100\%$. The $\pm 1(2)\sigma$ band of the 95% CL limit are depicted in green (yellow). The theory curve is shown in red.

B.4 Limits from the combination of pair production searches at the ATLAS experiment

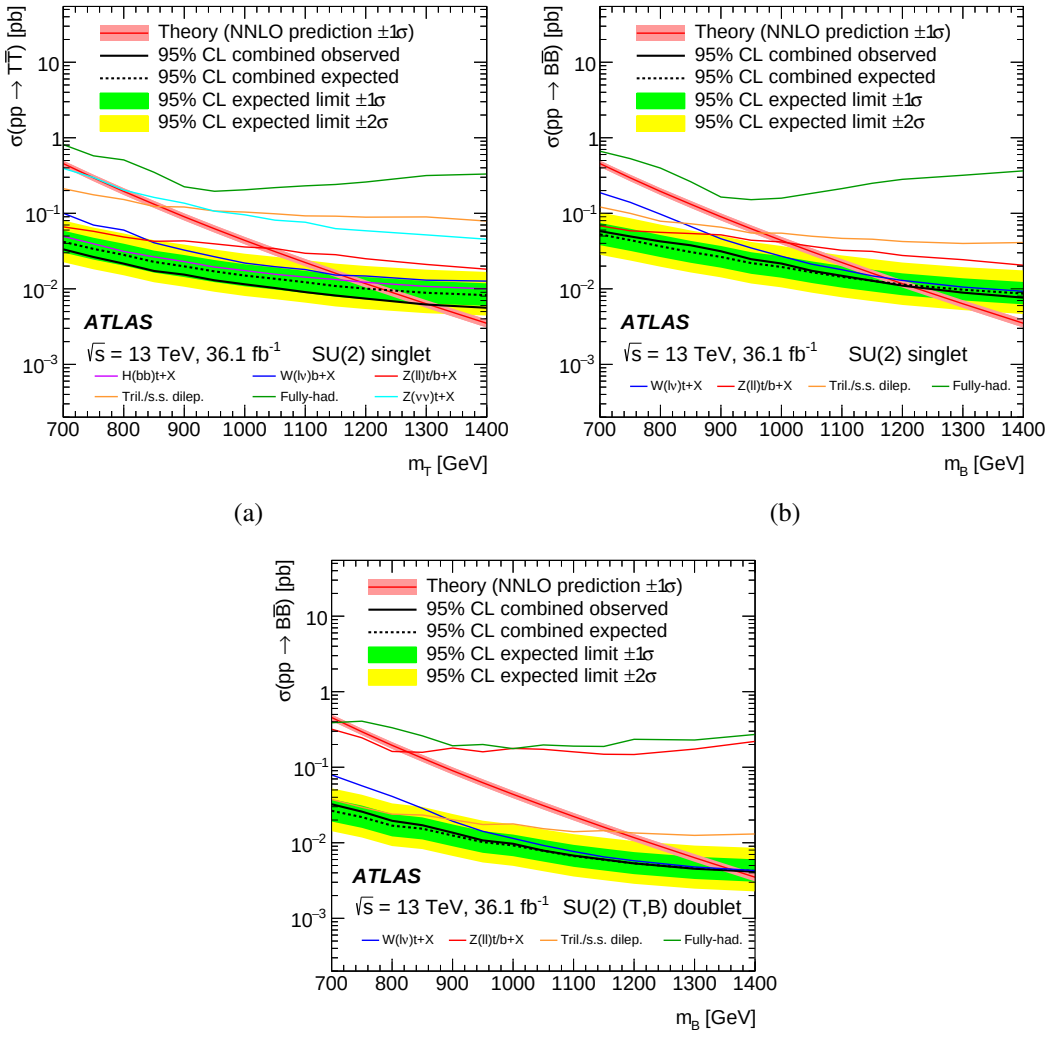


Figure B.5: Observed and expected exclusion limits [4] from the ATLAS pair production combination on the cross-section of a) $T\bar{T}$ and b) $B\bar{B}$ production assuming the singlet model. Furthermore, the limit on the cross-section of $B\bar{B}$ production according to the c) (T, B) doublet is illustrated. Solid lines correspond to observed limits of the combination and all contributing analyses. The expected limits of the combination and their $\pm 1\sigma$ and $\pm 2\sigma$ intervals are depicted.

Appendix C

Figures for the $Z(\ell\ell)t/b + X$ search with 139 fb^{-1}

C.1 Improvements of the analysis strategy using MCBOT

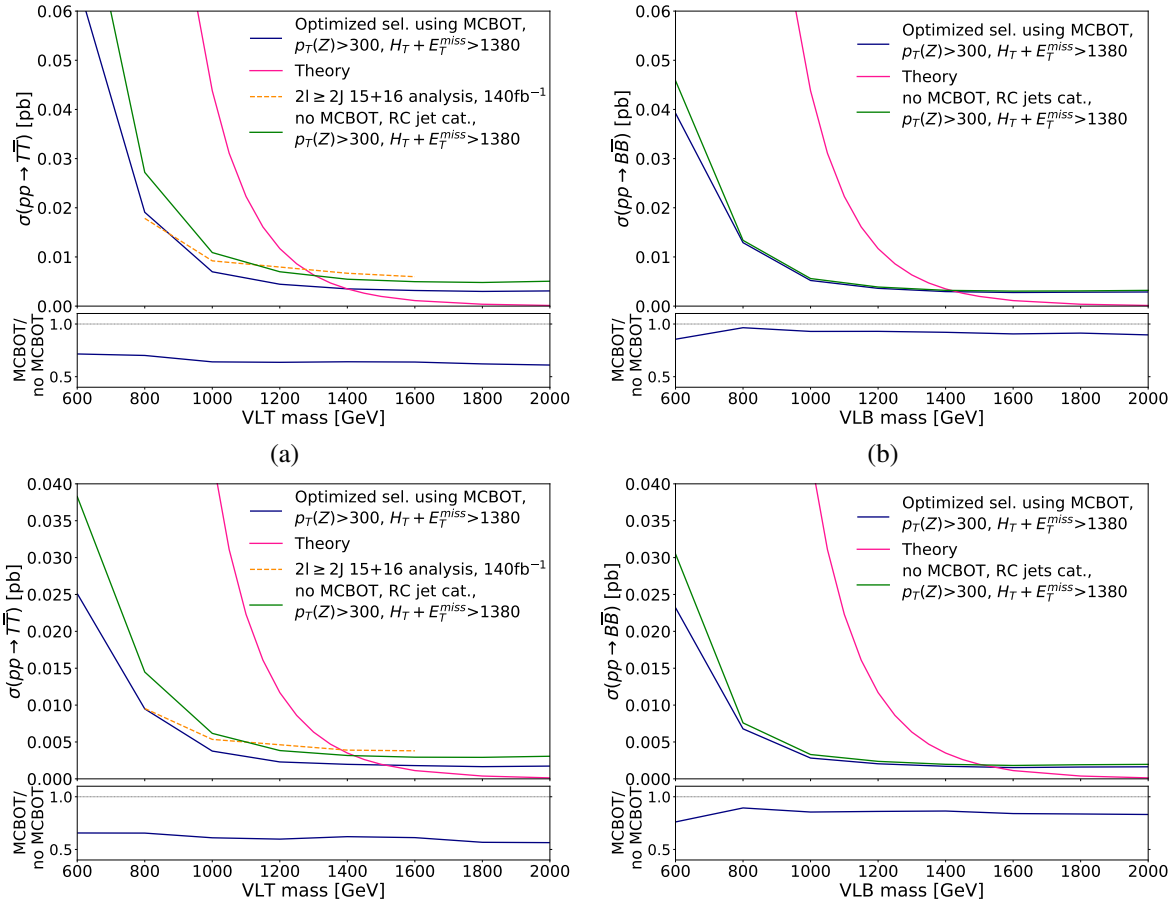


Figure C.1: Expected cross-section limits at 95% CL for $T\bar{T}$ and $B\bar{B}$ production assuming the doublet model in a) and b). In c) and d), the limits for the 100% BR into Zt/b final states are shown. The MCBOT categorization with splitting in b-tagged jet multiplicities and the optimized $p_{\mathrm{T}}(Z)$ and $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ cuts is shown as "optimized selection". In addition, a categorization solely based on RC jet multiplicity splitting without the use of MCBOT is depicted, denoted as "no MCBOT". For VLT quarks, the expected limit for the selection of the $2\ell \geq 2J$ channel of the 2015+2016 analysis is shown using 140 fb^{-1}. The ratio plot shows the optimized selection divided by the categorization without MCBOT. The number of MC events corresponds to 140 fb^{-1}. Kinematic cuts are given in units of GeV.

C.2 Pre-fit distributions in all search regions

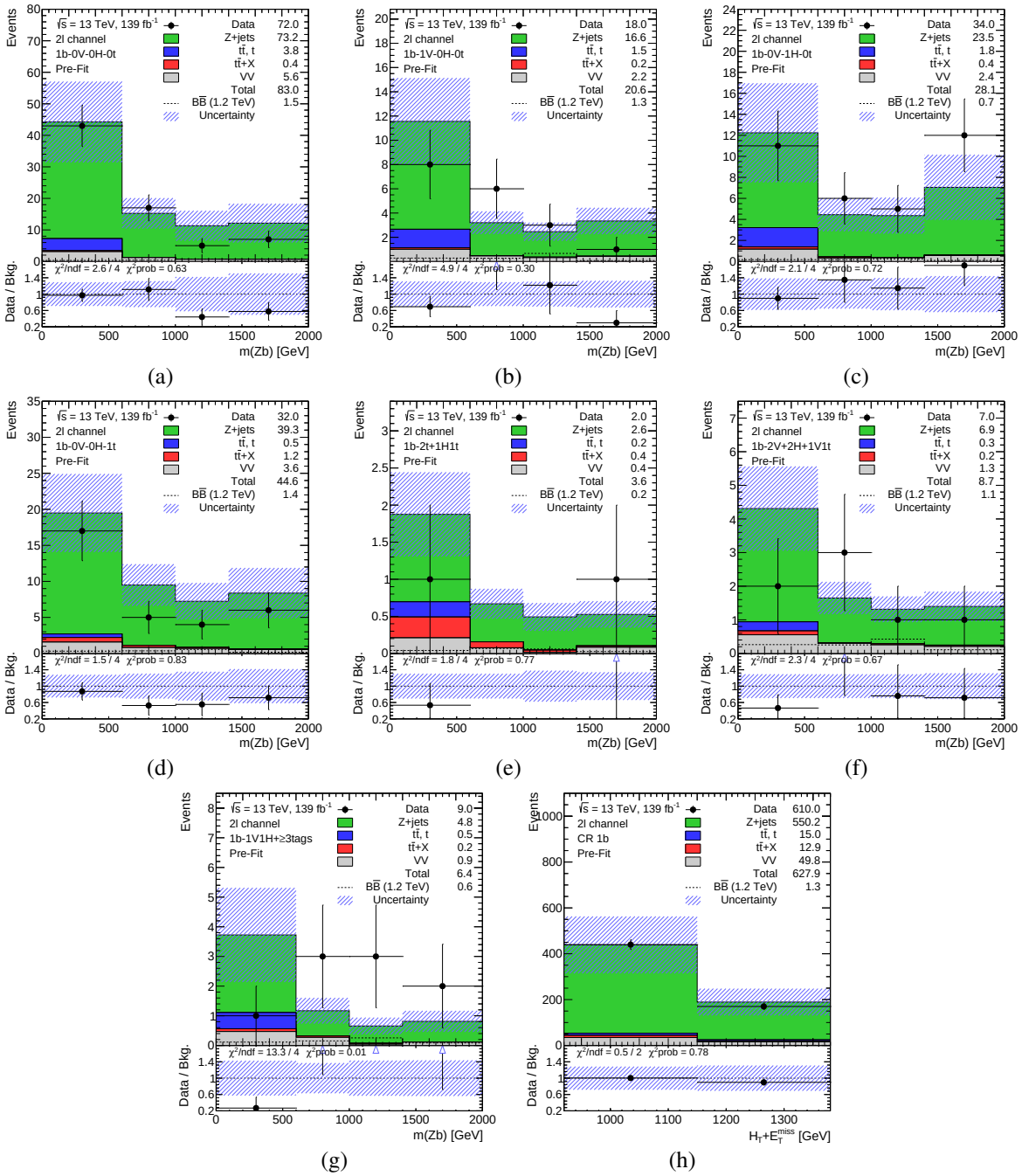


Figure C.2: Pre-fit plots for all categories with exactly 1 b-tagged jet with $m(Zb)$ as final discriminant. In addition, the CR for exactly 1 b-tagged jet with $H_{\mathrm{T}}(\mathrm{jets}) + E_{\mathrm{T}}^{\mathrm{miss}}$ as variable is shown in h). Systematic and statistical uncertainties are included in the uncertainty band. The pre-fit signal for VLB quarks with $m = 1200\,\mathrm{GeV}$ assuming singlet BRs is overlaid. All backgrounds are depicted and stacked on top of each other. The legend contains the yields of pre-fit background, pre-fit VLB signal and data. The subplot shows the ratio of data over the background prediction. Furthermore, the result of a χ^2 test over the NDF is depicted, as well as the χ^2 probability.

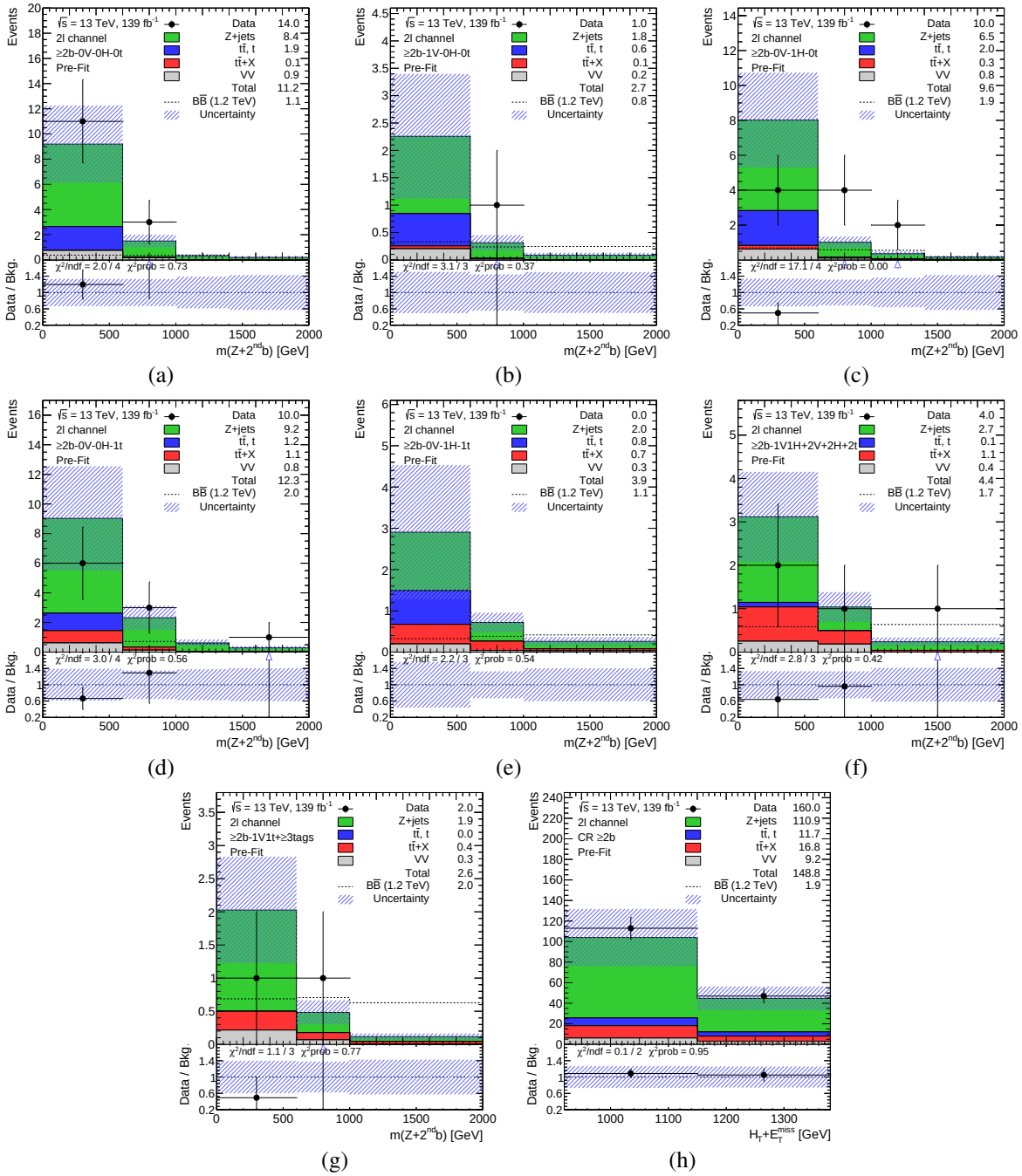


Figure C.3: Pre-fit plots for all categories with at least 2 b-tagged jets with $m(Z+2^{\text{nd}}\,b)$ as final discriminant. In addition, the CR for at least 2 b-tagged jet with $H_{\text{T}}(\text{jets}) + E_{\text{T}}^{\text{miss}}$ as variable is shown in h). Systematic and statistical uncertainties are included in the uncertainty band. The pre-fit signal for VLB quarks with $m = 1200\,\text{GeV}$ assuming singlet BRs is overlaid. All backgrounds are depicted and stacked on top of each other. The legend contains the yields of pre-fit background, pre-fit VLB signal and data. The subplot shows the ratio of data over the background prediction. Furthermore, the result of a χ^2 test over the NDF is depicted, as well as the χ^2 probability.

C.3 Limits for the singlet model and 100% BRs to Zt/b

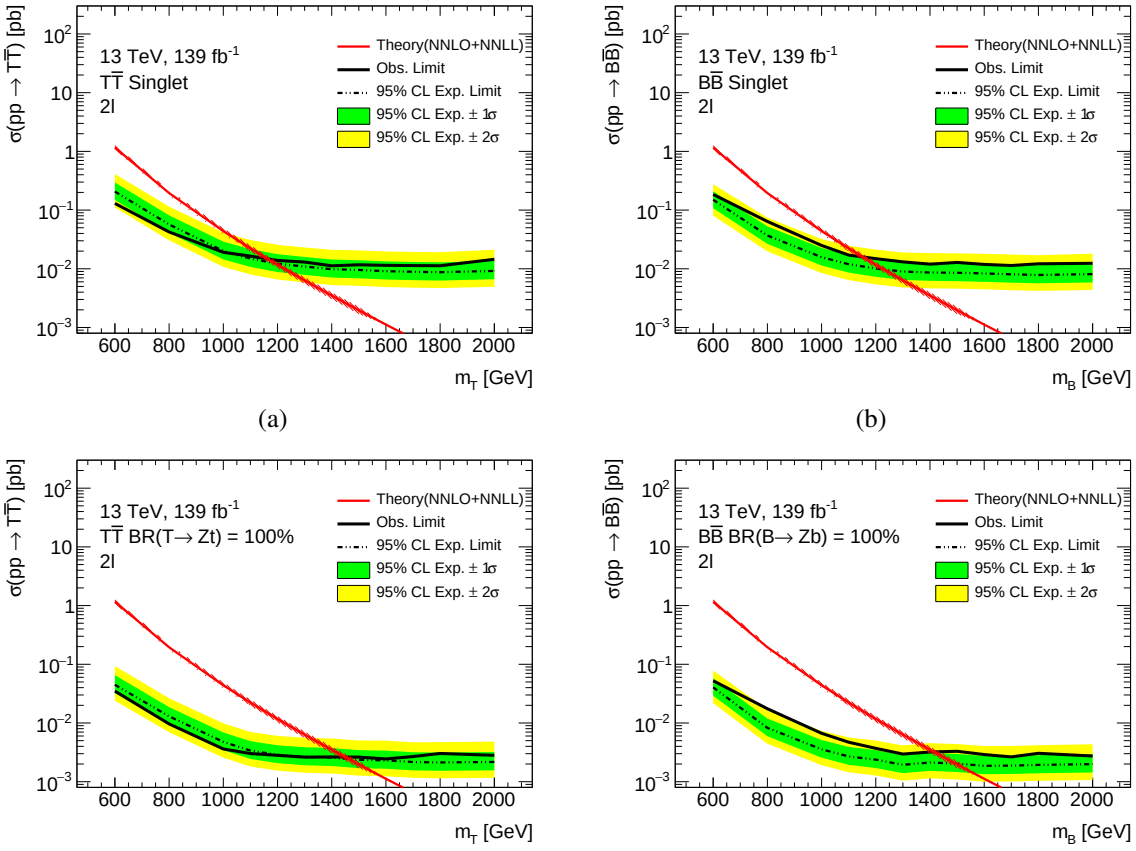


Figure C.4: Expected and observed limits on the cross-section for $T\bar{T}$ and $B\bar{B}$ production from the 2ℓ channel of the $Z(\ell\ell)t/b + X$ analysis at 95% CL using 139 fb^{-1}. The limits for VLT quarks are illustrated for a) the singlet model and c) the 100% BR into Zt. The limits for VLB quarks are depicted for b) singlet model and d) BR$(B \to Zb) \to 100\%$. The $\pm 1(2)\sigma$ band of the 95% CL limit are depicted in green (yellow). The theory curve is shown in red.

Appendix D

Application of MCBOT on VLQ samples

The tagging with the Multi-Class Boosted-Object Tagger (MCBOT) and its performance was studied in detail in Chapter 6. A search for vector-like top (VLT) and vector-like bottom (VLB) quarks with the full Run-2 dataset is presented in Chapter 9 which utilizes MCBOT in order to build event based categories according to the number of V-, H- and top-tags. Since the tagging is performed on an object level, a discrimination between signal and background jets should be achieved on other physics processes than the training composition of MCBOT as well. The following study shows the deep neural network (DNN) output of MCBOT for two vector-like quark (VLQ) samples used in the full Run-2 analysis. In order to investigate the separation power of MCBOT between the four classes of jets originating from VLQs, a truth-matching is applied for each reclustered (RC) jet, which is equivalent to the truth-matching implemented for MCBOT and described in Section 6.4.

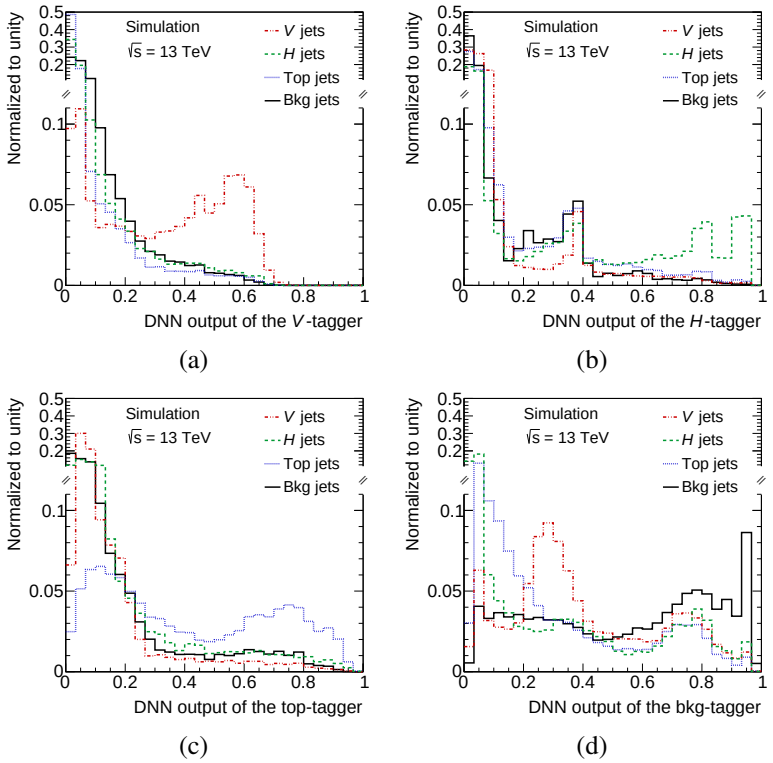


Figure D.1: DNN output distributions of the standard a) V-, b) H-, c) top- and d) background-tagger evaluated on pair-production events of VLT quarks. The BRs according to the singlet model are assumed and the benchmark mass of 1200 GeV from Chapter 9 is used. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

In Figure D.1, the output of the standard V-, H-, top- and background-tagger is shown for a VLT quark sample assuming a mass of 1200 GeV and BRs according to the singlet model. In Figs. D.1(a), D.1(b) and D.1(c), a good separation between the signal class and the three remaining classes is observed. The structures in the DNN output of the four classes are similar to the distributions shown in Section 6.8 for the standard FLat-In-p_{T} (FLIP) tagger in an RC jet p_{T} range of 150 GeV to 1 TeV, which is the p_{T} range of the RC jets from most VLQ decays with a mass of 1.2 TeV. The background-tagger in Figure D.1(d) also shows separation power between signal and background jets; however, it decreases compared to the background-tagger evaluated on the test sample in Chapter 6. The background composition of the VLT quark sample is different compared to the training sample of MCBOT. VLQ decays are rich in b-tagged jets. Therefore, the separation between signal and background jets is more difficult, which can also be seen in Figure D.1(b) at the peak of the background class at H-tagger values around 0.35.

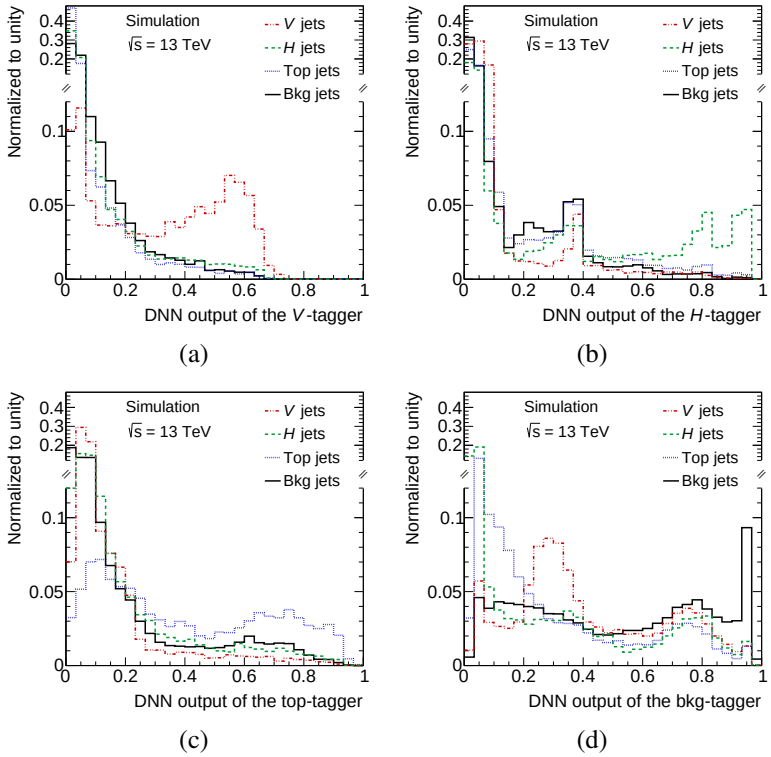


Figure D.2: DNN output distributions of the standard a) V-, b) H-, c) top- and d) background-tagger evaluated on pair-production events of VLB quarks. The BRs according to the singlet model are assumed and the benchmark mass of 1200 GeV from Chapter 9 is used. The four DNN classes V-boson, H-boson, top-quark and background jets are depicted in different colors. All histograms are normalized to unity.

In Figure D.2, the four MCBOT taggers are evaluated on a VLB quark sample assuming a mass of 1200 GeV and the singlet model. The same observations as for the VLT quark sample about all four taggers are made. However, the separation between top-quark and background jets is worse than for the VLT quark sample. The abundance of high-p_{T} b-tagged jets in the final states of VLB quark decays leads to a more difficult environment, where mistags are more likely.

Acknowledgements

The last few years have been an exciting and instructive journey and I would like to use the last words to thank you.

First of all, I would like to thank Prof. Dr. Kevin Kröninger for giving me the opportunity to work at the ATLAS experiment. I very much appreciated the possibilities to attend schools, workshops and conferences during my PhD time. In addition, the year at CERN was an incredible opportunity and I'm very grateful for it.

Many thanks also goes to Prof. Dr. Dr. Wolfgang Rhode for agreeing to review this thesis.

I would like to acknowledge the great support by the Konrad-Adenauer-Foundation. The scholarship has funded this work, and the numerous seminars and continuing education reaching beyond the scope of physics have been extremely valuable.

Furthermore, I would like to thank Priv.-Doz. Dr. Johannes Erdmann for the interesting discussions and the valuable physics feedback throughout the years.

I would also like to thank Assistant Professor Nuno Castro, who provided me with access to the cluster at LIP. This was extremely helpful and allowed me to spend more time on physics. Furthermore, thanks to Dr. Juan Pedro Araque Espinosa for supporting me in the first year of my dissertation by answering many questions, which paved the way for a successful Run-2 search. Also thanks to my colleague Tiago Vale, who has been part of the search for VLQs since the beginning of my master thesis. Thanks a lot also to Dr. Nikiforos Nikiforou for creating the ntuples for the VLQ searches.

Additionally, I would like to thank Dr. Olaf Nackenhorst for his support in the development of MCBOT and the critical feedback on talks and especially my thesis draft.

I'm very glad that I met so many great and supporting people at E4. Thanks to Andrea Teichmann who not only helped with organizational issues, but was also there for nice chats. Thanks also to Markus Alex for promptly fixing my laptop, which broke twice. I would like to thank Sonja Zeißner for the constant support and many great conversations. Also thanks to Björn Wendland for good discussions about statistics, which were very helpful. Thanks to Cornelius Grunwald for celebrating the awesomeness of the Mensa every single day. My appreciation goes to Sonja, Björn, Cornelius and Kevin Sedlaczek for proofreading my thesis. I would also like to thank Lennart Rustige for technical discussions and for your friendship. Thanks to Dr. Ingo Burmeister and Dr. Isabel Nitsche for many fun discussions. The time at E4 wouldn't have been the same without the fun snowboarding holidays. For the great time, I'd like to thank Dr. Robert Theinert, Tobias Kupfer, Dr. Gregor Geßner, Myriam Heiny and Florian Mentzel. Finally, thank you everybody at E4 for the positive working environment.

I was very lucky to spend one year of my PhD at CERN and got to feel as a part of the ATLAS collaboration by first hand experiences. I'm very glad I shared an office with Dr. Diane Cinca and Sonja where we shared language knowledge and lots of chocolate. Also thanks to Sascha Dungs for being part of the great time at CERN and also for allowing me to escape to your office with air-conditioner. Big thanks go to my rugby team. It was an amazing experience to be one of you and

share so many great moments like the Swiss championship. I will never forget which achievements are possible if you work as a team.

I would like to express deep gratitude to my friends which have given me so many great experiences over the past few years. I cannot name all of them, but special thanks go to Julia, Yasmine, Felix and Melanie. I'm glad to have you in my life.

I'm very glad to have met you, Bane. Getting to know CERN with you has been a unique experience. We've spend so much valuable time together and although time was short during the past two years, we've made the best of it - continuously crossing borders and sharing great adventures. I am so grateful for your kindness and support - especially, I would like to thank you for feeding me during times of extreme pressure.

Last but not least, I would like to thank my family for always being there and for giving me a source of strength I can always rely on. Thanks for encouraging and supporting me in every way you can. I'm so lucky to call you my family.