Measurement of the production cross section of a top-antitop quark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector

FLORIAN CHRISTOPH MARIA FISCHER

München 2020

Measurement of the production cross section of a top-antitop quark pair in association with a Z boson at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector

DISSERATION AN DER FAKULTÄT FÜR PHYSIK
DER
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

vorgelegt von

FLORIAN CHRISTOPH MARIA FISCHER

geboren in Annweiler am Trifels

München, den 3. November 2020

Erstgutachter: Prof. Dr. Otmar Biebel
Zweitgutachter: Prof. Dr. Thomas Kuhr
Tag der mündlichen Prüfung: 14.12.2020

"Measure what can be measured,
and make measurable what cannot be measured."

Galileo Galilei (1564–1642)

Zusammenfassung

Die Kopplung des Topquarks an das Z-Boson ist ein elektroschwacher Prozess und im Standardmodell der Teilchenphysik genau vorherbestimmt. Experimentell konnte die Stärke dieser Kopplung bislang nicht genau genug bestimmt werden, um mögliche Beiträge von Prozessen, die von über das Standardmodell hinausgehenden theoretischen Modellen vorhergesagt werden, sicher ausschließen zu können. Ein für diesen Zweck geeigneter Prozess ist die assoziierte Produktion von Paaren von Topquark und Antitopquark mit einem Z-Boson. Aufgrund des kleinen Produktionswirkungsquerschnitts ist es erst mit der hohen Schwerpunktsenergie des LHC und der großen Datenmenge, die innerhalb der letzten Jahre gesammelt wurde, möglich geworden, die Produktion von $t\bar{t}Z$-Ereignissen zu vermessen.

Dazu wird eine Datenmenge von $139\,\mathrm{fb}^{-1}$ an aufgezeichneten Proton-Proton-Kollisionen, die in den Jahren 2015 bis 2018 vom ATLAS-Detektor am LHC mit einer Schwerpunktsenergie von $13\,\mathrm{TeV}$ aufgezeichnet wurden, verwendet. Die Anzahl zu erwartender Ereignisse des Signalprozesses und der durch das Standardmodell beschriebenen Untergrundprozesse wird mit Monte-Carlo-Simulationen abgeschätzt. Das Z-Boson wird anhand seiner Zerfallsprodukte identifiziert, die aus einem Elektron- oder Myonpaar mit entgegengesetzter elektrischer Ladung bestehen, dessen invariante Masse mit der des Z-Bosons kompatibel ist. Die Signatur der untersuchten $t\bar{t}Z$-Ereignisse wird somit durch die Zerfallsprodukte des assoziierten $t\bar{t}$-Systems bestimmt.

Der erste Teil dieser Disseration beinhaltet mehrere Studien, die im Rahmen einer Messung des inklusiven und differentiellen $t\bar{t}Z$-Wirkungsquerschnitts durchgeführt wurden. Hierzu werden die sensitivsten $t\bar{t}Z$-Zerfallskanäle mit drei oder vier isolierten Leptonen mit großem Tranversalimpuls im Endzustand verwendet. Mehrere simulierte Monte-Carlo-Datensätze zur Beschreibung $t\bar{t}Z$ bzw. der assoziierten Produktion einzelner Topquarks mit einem Z- sowie einem W-Boson werden validiert. Des Weiteren wird ein Versuch zur Minimierung der Beiträge dominanter Untergrundprozesse – der Produktion zweier Vektorbosonen sowie der assoziierten Produktion einzelner Topquarks mit Vektorbosonen – aufgeführt. Um dabei $t\bar{t}Z$ von den genannten Untergründen hinreichend abzugrenzen, wird das $t\bar{t}$-System teilweise rekonstruiert.

Im zweiten Teil der vorliegenden Disseration wird ein neuer Ansatz zur Messung des inklusiven Produktionswirkungsquerschnitts von $t\bar{t}Z$-Ereignissen mit zwei isolierten Leptonen im Endzustand, dem sogenannten $2\ell\mathrm{OS}$-Zerfallskanal, eingeführt. Im Gegensatz zu den Zerfallskanälen mit drei oder vier isolierten Leptonen im Endzustand ist der Anteil an Untergrundereignissen wesentlich größer. Zur Trennung von Signalereignissen von den beiden dominanten Untergrundprozessen – dileptonisch zerfallender $t\bar{t}$-Ereignissen und der assoziierten Produktion eines Z-Bosons mit Jets – werden zwei Boosted Decision Trees unabhängig voneinander auf jeweils einen dieser beiden Untergründe trainiert. Mithilfe des Unterscheidungsvermögens beider Boosted Decision Trees können Phasenraumbereiche mit einer hohen Reinheit Signal- bzw. Untergrundereignissen ausgezeichnet werden. Der Produktionswirkungsquerschnitt wird mittels eines profile-likelihood-Fits der Anzahl von simuliertem Monte-Carlo-Ereignissen an die Daten in diesen Regionen bestimmt, wobei ein Wert von

$$\sigma_{t\bar{t}Z} = 0,91 \pm 0,08\,(\text{stat.}) \pm 0,14\,(\text{syst.})\,\text{pb}$$

mit einer statistischen Signifikanz von $5,9\,\sigma$ gemessen wird. Der gemessene Wert des Produktionswirkungsquerschnitts stimmt mit den genauesten theoretischen Vorhersagen überein.

Abstract

The coupling of the top quark to the Z boson is precisely predicted within the Standard model of particle physics via the electroweak interaction. However, experimentally it is not yet well constrained and its value can vary significantly in many models featuring physics beyond the Standard Model. A process that is particularly sensitive to this coupling is the associated production of a top-antitop quark pair with a Z boson. The large centre-of-mass energy of the LHC and the tremendous amount of data collected in recent years have opened up the possibility to study this rare process which was previously inaccessible due to its small production cross section.

The production cross section of the $t\bar{t}Z$ process is measured, using $139\,\mathrm{fb}^{-1}$ of proton-proton collision data at a centre-of-mass energy of $13\,\mathrm{TeV}$, recorded by the ATLAS detector at the LHC in the years from 2015 to 2018. In order to estimate the expected number of events for both signal and the associated Standard Model background, Monte Carlo simulations are employed. The Z boson is identified by targeting events featuring a pair of electrons or muons with opposite electric charge and an invariant mass consistent with the parent particle. The final targeted signature is then characterised by the number of leptons from the decay of the associated $t\bar{t}$ system.

The first part of this thesis presents several studies conducted in the context of a $t\bar{t}Z$ cross-section measurement that targets the most sensitive decay modes of the $t\bar{t}Z$ process with three or four isolated leptons with high transverse momentum in the final state and reports measurements of both the inclusive and differential cross sections. A series of validation studies are performed for sets of simulated Monte Carlo samples used in the analysis to describe the signal process as well as one of the dominant backgrounds featuring the associated production of single top quarks with both a W and a Z boson. Furthermore, an attempt to reduce the contributions from the dominant background processes in these $t\bar{t}Z$ decay channels – the production of pairs of vector bosons and single top quarks in association with vector bosons – is shown by exploiting the discrimination provided by an algorithm based on a partial reconstruction of the $t\bar{t}$ system.

In the second part of this thesis, a new approach for a measurement of the $t\bar{t}Z$ production cross section in the dilepton channel, labelled $2\ell\mathrm{OS}$, is introduced. In contrast to the $t\bar{t}Z$ decay modes with three or four isolated leptons in the final state, the $2\ell\mathrm{OS}$ channel suffers from significantly larger background rates. In order to isolate the signal from the two dominant background processes – dileptonically decaying $t\bar{t}$ events and the associated production of a single Z boson with jets – two Boosted Decision Trees are independently trained, with each dedicated to a unique one of the two major background contributions. The output of the two classifiers is then combined in order to select phase space regions highly enriched in events from either signal or either one of the two dominant background processes. A profile-likelihood fit of the Monte Carlo prediction to the data within those regions is employed to determine the inclusive $t\bar{t}Z$ production cross section, which is measured to be

$$\sigma_{t\bar{t}Z} = 0.91 \pm 0.08\,(\mathrm{stat.}) \pm 0.14\,(\mathrm{syst.})\,\mathrm{pb}$$

with an observed statistical significance of $5.9\,\sigma$. The result is found to be in agreement with the most precise theoretical prediction.

Contents

1. Introduction

The foundation of quantum physics at the beginning of the 20th century laid the cornerstone of our current understanding of the universe on its smallest scales. Most of the theoretical works and experimental findings since that time were based on the concepts of the quantised nature of physical observables, the wave-particle dualism and the indeterminism of physical processes. A first comprehensive description of matter at the atomic level was provided by Niels Bohr [1], which was inspired by the model developed by Ernest Rutherford based on his scattering experiments with α-particles on gold nuclei [2]. Over the next two decades, the Bohr atomic model became refined by the works of, for example, Arnold Sommerfeld, Erwin Schrödinger, Wolfgang Pauli and Paul Dirac, and remained a valid description until today in its final version. Two notable events which provided strong vindication of the quantum model were the experimental discovery of the proton by Ernest Rutherford in 1919 [3] and of the neutron by James Chadwick in 1932 [4], both of which contributing significantly to the our knowledge of the sub-atomic world. At the time, a series of experiments confirmed that atoms consist of a central massive nucleus comprised of positively charged protons and uncharged neutrons, and a negatively charged shell consisting of several electrons. The electron was known to be fundamental already since it had been experimentally verified by Joseph Thomson in 1897 [5]. However, both the proton and the neutron were assumed to be fundamental as well until Robert Hofstadter discovered the proton to feature a finite charge distribution in the mid-1950s [6]. About ten years later, Friedman, Kendall and Taylor experimentally proved that the proton itself is not a fundamental particle and were able to reveal its internal structure [7]. These findings then led to the development of the parton model in order to describe the inner structure of the proton [8].

In the 1950s and 1960s, a multitude of particles was discovered at collider and fixed-target experiments which at the time had not been predicted by any theory. In light of these findings, Murray Gell-Mann and George Zweig introduced the quark model in 1964 [9–11], describing all the newfound particles – collectively referred to as hadrons – as composite states of only three distinct fundamental building blocks: the up, down and strange quarks (and their respective antiparticles). Shortly thereafter in 1970, a fourth quark, denoted as charm, was postulated to exist based on experimental absence of so-called flavour-changing neutral currents [12, 13]. It would take a further four years for its discovery to be made at the Stanford Linear Accelerator (SLAC) [14, 15]. In order to explain the observed CP violation effects in weak interactions, a third generation of quarks was theoretically predicted by Makoto Kobayashi and Toshihide Maskawa, [16] of which the lighter of the pair, the bottom quark, was discovered in 1977 at SLAC [17]. Its weak-isospin partner, the top quark, however, was given experimental proof in 1995 by the CDF and DØ experiments at the Tevatron ring collider [18, 19]. Since then, characteristics of the top quark have been extensively studied, in particular by the ATLAS and CMS experiments at the Large Hadron Collider (LHC).

The Standard Model of particle physics describes the set of known fundamental particles matter is built from and their interactions. It was mainly developed in the 1960s and 1970s by the pioneering work on quantum chromodynamics by Harald Fritzsch, Heinrich Leutwyler and Murray Gell-Mann as well as David Gross, David Politzer and Frank Wilczek and many more [20–22], as well as the work on the electroweak theory by Sheldon Glashow, Abdus Salam and Steven Weinberg [23–25]. In addition, electromagnetic interactions have been described by quantum electrodynamics which was formulated by Richard Feynman, Julian Schwinger, Shinichiro Tomonaga

and Freeman Dyson [26–31] already in the 1940s. Since then, the Standard Model remained to be proven very successful in the description of nature at the smallest scales. After the observation of the tau-neutrino in 2000 [32], the Higgs mechanism, which is responsible for the particles being massive, was the last important piece of the Standard Model which had to be proven by experiments. This was achieved by the ATLAS and CMS experiment with the discovery of the Higgs boson in 2012 [33, 34]. Since the value of a given particles' mass depends on the coupling strength to the Higgs boson, the top quark plays an important role in the electroweak sector of the Standard Model. Both the determination of its mass and of its coupling to the Higgs boson therefore serve as a consistency check of the Standard Model. Of equal importance is the investigation of the coupling of the top quark to the Z boson, which is an uncharged mediator of electroweak interactions and the third heaviest particle in the Standard Model below the top quark and the Higgs boson. The strength of the top-Z coupling is precisely predicted by the Standard Model but experimentally not yet well constrained and its value may vary in many models including physics beyond the Standard Model. A process which is particularly sensitive to the coupling strength of the top quark to the Z boson is the associated production of a top-antitop quark pair with a Z boson ($t\bar{t}Z$). The large centre-of-mass energy of the LHC and the tremendous amount of data collected in recent years have opened up the possibility to study this rare process which was previously inaccessible due to its small production cross section. The $t\bar{t}Z$ process was observed in experiment by the CMS collaboration only in 2018 for the first time [35]. Since then a series of measurements of the $t\bar{t}Z$ production cross section have been performed by ATLAS [36–39] and CMS [40] with increasing precision. Any observed deviations in the coupling strength of the top quark to the Z boson from its Standard Model prediction might hint to new effects in the electroweak symmetry breaking mechanism. Furthermore, the $t\bar{t}Z$ process is an irreducible background to several searches for phenomena predicted by theories describing physics beyond the Standard Model [41, 42], as well as to measurements of important Standard Model processes such as the associated production of a top-antitop quark pair with a Higgs boson [43] or single top-quark production in association with a Z boson [44, 45]. Any additional insight into the characteristics of $t\bar{t}Z$ production gained through measurements can therefore lead to improved measurements of such processes. Due to the large amount of available data, first differential cross-section measurements were performed by ATLAS [46]and CMS [47] as a function of different variables which probe the Standard Model predictions for the kinematics of the $t\bar{t}Z$ system. Such differential measurements offer sensitivity to differences between the predictions from various Monte Carlo generators, which can therefore serve as an important input to the modelling of physics processes.

For the measurements of the $t\bar{t}Z$ production cross section presented in this thesis, the full LHC Run 2 dataset of $139\,\mathrm{fb}^{-1}$ of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$, collected by the ATLAS experiment during the years of 2015 to 2018, was used. The expected rate of the signal and of various Standard Model background processes were estimated with Monte Carlo simulations. The Z boson is identified from its decay products which are, in the context of this thesis, a pair of isolated electrons or muons with opposite electric charge and an invariant mass compatible with the parent particle. The $t\bar{t}Z$ signature is therefore characterised by the final states of the decay of the associated $t\bar{t}$ system.

This thesis is structured as follows: Chap. 2 provides an overview of Standard Model and its important theoretical concepts. Additionally, an introduction to the physics of top quarks and in particular of the $t\bar{t}Z$ process at the LHC is given. The experimental setup is described in Chap. 3 with an introduction of the European Organisation of Nuclear Research (CERN) and the LHC, followed by a general description of the functionality of the various detector modules of the ATLAS experiment. In Chap. 4 the concepts of Monte Carlo simulation are introduced along with a discussion of the simulated samples used to describe the observed data. Furthermore, the

collected dataset and the necessary steps of data preparation are explained. The reconstruction of physics objects from recorded and simulated data is discussed in Chap. 5 focusing on the type of objects which were explicitly considered within this thesis. In Chap. 6, several studies are presented which were conducted in the context of an ATLAS $t\bar{t}Z$ cross-section measurement that targets the most sensitive decay modes of the $t\bar{t}Z$ process with three or four isolated leptons with high transverse momentum in the final state, and reports measurements of both the inclusive and differential cross sections. A new approach for a measurement of the $t\bar{t}Z$ production cross section, using events which feature two leptons from the decay of the Z boson and six jets from the $t\bar{t}$ decay, is introduced in Chap. 7, which includes a discussion of the concepts of the multivariate analysis techniques employed for the separation of signal from background events as well as of the statistical tools used for the extraction of the measured cross section. Finally, a conclusive remark and discussion of the results is given in Chap. 8

By convention, natural units are used throughout this thesis, which fixes the values for the speed of light, the reduced Planck constant, the vacuum permittivity and the Boltzmann constant to unity:

$$c = \hbar = \varepsilon_0 = k_\mathrm{B} = 1 \tag{1.1}$$

Therefore, energies, momenta and masses are measured in base units of electronvolts (eV) where typical scales at the LHC are in MeV, GeV, or TeV.

2. Theory

2.1. The Standard Model of Particle Physics

The Standard Model of Particle Physics sets out to describe our world at its most fundamental level. Since its development in the 1960s and 1970s, it has been extremely successful in describing elementary particles and its interactions. Up to the present, the Standard Model has met most experimental tests within the field of particle physics with high precision and many unknown characteristics of nature predicted by the Standard Model have been experimentally verified. One of its most celebrated accomplishments was the experimental confirmation of a particle consistent with the Higgs boson in 2012 by the two LHC experiments ATLAS and CMS [33, 34] since the original theoretical prediction of its existence half a century ago. However, despite its outstanding success, the Standard Model suffers from a few key limitations and leaves present physicists with some open questions. The unification of all interaction has not yet been achieved within the Standard Model, moreover, it does not account for the force of gravity at all. The various particle masses are not fixed by the theory but have to be determined from experiment. Beyond that, neutrinos are assumed to be massless which has been proven to be wrong as they undergo flavour transitions while propagating over large distances. This might be an indication for CP violation also being present in the lepton section of the Standard Model whereas it has only been measured in the quark sector so far. As a result, the Standard Model is not able to explain the huge asymmetry between matter and antimatter in the universe [48, 49].

In the Standard Model of Particle Physics, the elementary particles can be grouped into two main categories: particles with half-integer spin, referred to as fermions, which serve as the fundamental building blocks of matter, and particles with integer spin, referred to as bosons, that mediate the fundamental forces between fermions. The three types of interactions described by the Standard Model are the electromagnetic (EM) force, the weak force and the strong force. Most often encountered in everyday life, electromagnetism encompasses both electricity and magnetism and is mediated by the photon (γ). The weak interaction has three massive force carriers, two $W^{\pm}$ bosons and one neutral Z boson, and is the only fundamental interaction that allows for the direct conversion of a particle into another, for example within radioactive β-decays. The strong force is mediated by eight massless gluons and, amongst others, prevents nuclei from breaking apart. The fermions can be further subdivided into two groups, the so-called quarks and leptons. Whereas both quarks and leptons interact via the weak force, only the quarks interact via the strong force. In order to couple to the photon a particle has to be electrically charged which is true for quarks and half of the leptons. One notable characteristic of the Standard Model is its periodic structure of how particles are sorted into generations. For both quarks and leptons a total of three generations exist, each consisting of an up- and a down-type quark and of a charged lepton and the corresponding neutrino, respectively. A summary of all fermions and bosons of the Standard Model is given in Tabs. 2.1 and 2.2, respectively [48–50].

The Standard Model of Particle Physics is formulated in terms of a quantum field theory. By combining the principles of classical field theories on the one hand and of relativistic quantum mechanics on the other hand, quantum field theories are comprehensive techniques to describe elementary particles and their interactions with both the physical observables and the fields responsible for the interaction are quantised. In classical mechanics, the dynamics of a system can described by its Lagrangian $L(q_i, \dot{q}_i)$ which is a function of a set of generalised coordinates q_i and

Generation	Name	Symbol	Colour	Charge $[e]$	Mass
1	Up quark	u	yes	$+2/3$	$2.16^{+0.49}_{-0.26}$ MeV
	Down quark	d	yes	$-1/3$	$4.67^{+0.48}_{-0.17}$ MeV
	Electron neutrino	ν_e	no	0	< 2 eV
	Electron	e	no	-1	0.511 MeV
2	Charm quark	c	yes	$+2/3$	1.27 ± 0.02 GeV
	Strange quark	s	yes	$-1/3$	93^{+11}_{-5} MeV
	Muon neutrino	ν_μ	no	0	< 0.19 MeV
	Muon	μ	no	-1	105.66 MeV
3	Top quark	t	yes	$+2/3$	172.9 ± 0.4 GeV
	Bottom quark	b	yes	$-1/3$	$4.18^{+0.03}_{-0.02}$ GeV
	Tau neutrino	ν_τ	no	0	< 18.2 MeV
	Tau	τ	no	-1	1776.86 ± 0.12 MeV

Tab. 2.1.: Overview of the fermions in the Standard Model of Particle Physics. For each particle an antiparticle with opposite charge-like quantum numbers exists. In cases where no uncertainty on the measured mass is quoted, the experimental uncertainty is more precise than the number of significant digits states in the table. Upper limits for the neutrino masses from direct observations are given at 95 % CL for the electron and the tau neutrino while the limit for the muon neutrino mass is given at 90 % CL [51]

their time derivatives $\dot{q}_i$. The equations of motion are determined by solving the Euler-Lagrange equations,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \quad . \tag{2.1}$$

In quantum field theories, the system no longer comprises discrete particles but rather continuous fields. The Lagrangian will thus be replaced by Lagrangian density $\mathscr{L}(\phi_i, \partial_\mu \phi_i)$ which is described by the fields ϕ_i and their derivatives $\partial_\mu \phi_i$ and is related to the Lagrangian of Eq. (2.1) in the following way:

$$L = \int \mathscr{L}(\phi_i, \partial_\mu \phi_i)\, \mathrm{d}^3 x \quad . \tag{2.2}$$

The Euler-Lagrange equation (2.1) can then be rewritten as

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)}\right) - \frac{\partial \mathscr{L}}{\partial \phi_i} = 0 \quad . \tag{2.3}$$

The overall Lagrangian density of the Standard Model consists of three parts,

$$\mathscr{L}_{\mathrm{SM}} = \mathscr{L}_{\mathrm{QCD}} + \mathscr{L}_{\mathrm{EW}} + \mathscr{L}_{\mathrm{Higgs}} \quad , \tag{2.4}$$

with $\mathscr{L}_{\mathrm{QCD}}$ describing quantum chromodynamics (QCD), the theory of strong interactions, and $\mathscr{L}_{\mathrm{EW}}$ representing the electroweak (EW) theory, which is the theoretical foundation of the combination of electromagnetism and the weak force. The final term, $\mathscr{L}_{\mathrm{Higgs}}$, stands for the Higgs mechanism which is responsible for electroweak symmetry breaking thereby giving the elemen-

Name	Symbol	Mediated force	Colour	Charge $[e]$	Mass
Photon	γ	electromagnetic	no	0	$< 1 \cdot 10^{-18}$ eV
Gluon	g	strong	yes	0	0 (theory)
$W^{\pm}$ boson	$W^{\pm}$	weak	no	± 1	80.379 ± 0.012 GeV
Z boson	Z	weak	no	0	91.1876 ± 0.0021 GeV
Higgs boson	H	Higgs field	no	0	125.10 ± 0.14 GeV

Tab. 2.2.: Overview of the bosons in the Standard Model of Particle Physics. In contrast to the fermions, the gauge bosons in the Standard Models do not have an antiparticle associated except for the $W^{\pm}$ boson which can be seen as their respective antiparticles. In cases where no uncertainty on the measured mass is quoted, the experimental uncertainty is more precise than the number of significant digits states in the table [51].

tary particles their mass[(1)].

2.1.1. Gauge invariance and renormalisation

Two fundamental principles of quantum field theories are gauge invariance and renormalisability. If a physics process is denoted gauge invariant, its underlying equations and thus the process itself should not change under a phase transformation. In the following these two characteristics of quantum field theories will be illustrated by means of quantum electrodynamics (QED) [26–31], the quantum field theory of electromagnetic interactions.

For this purpose, it is convenient to start with the Lagrangian for a free spin-half particle,

$$\mathscr{L}_{\text{Dirac}} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \quad , \tag{2.5}$$

acting on the Dirac spinor fields $\psi(x)$ with the Einstein sum convention applied. The Dirac matrices γ^{μ} are a set of conventional 4×4 matrices generating a Clifford Algebra with the following properties:

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu} \tag{2.6a}$$

$$\left(\gamma^{0}\right)^{2} = 1 \tag{2.6b}$$

$$\left(\gamma^{i}\right)^{2} = -1 \qquad (i = 1, 2, 3) \tag{2.6c}$$

with the Minkowski metric

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad . \tag{2.7}$$

The γ^{μ} matrices are defined in the so-called Dirac representation by the Pauli spin matrices σ_i[(2)]

[(1)] In addition to the stated references, this section is based on Ref. [48, 49, 52, 53].

[(2)] The 2×2 unitary Pauli spin matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and the unit matrix in two dimensions[3],

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad i \in \{1,2,3\} \quad . \tag{2.8}$$

Inserting Eq. (2.5)into the Euler-Lagrange equation (2.3) yields the famous Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\,\psi(x) = 0 \tag{2.9}$$

which describes the free motion of a spin-half particle with mass m.

In quantum electrodynamics, gauge transformations are described by $U(1)_e$ which is part of the unitary group $U(N)$ representing the set of $N \times N$ matrices that satisfy the unitary condition $U^\dagger U = \mathbb{1}$. The subscript e denotes the electric charge which is the conserved quantity in QED. A possible representation is

$$U(\theta) = e^{-i\theta G} \quad , \tag{2.10}$$

in which G is called the generator of the group and θ is the rotation angle or parameter of the group. The Dirac equation (2.9) is invariant under a global $U(1)$ phase transformation

$$\psi(x) \to \psi'(x) = e^{ie\rho}\psi(x) \tag{2.11}$$

where ρ can be any real number, and, most importantly, is constant. This means that the physical result remains unchanged. However, the Dirac equation (2.9) is not invariant anymore under a local gauge transformation

$$\phi(x) \to \psi'(x) e^{ie\rho(x)}\psi(x) \quad , \tag{2.12}$$

with ρ now being a function of the space-time coordinates. Hence, inserting Eq.(2.12) into the Dirac equation (2.9) yields

$$\begin{aligned} (i\gamma^\mu \partial_\mu - m)\,\psi'(x) &= (i\gamma^\mu \partial_\mu - m)\, e^{ie\rho(x)}\psi(x) \\ &= e^{ie\rho(x)}\left[(i\gamma^\mu \partial_\mu - m)\,\psi(x) - e\partial_\mu \rho(x)\gamma^\mu \psi(x)\right] \quad . \end{aligned} \tag{2.13}$$

The additional factor $e\partial_\mu \rho(x)\gamma^\mu\psi(x) \neq 0$ spoils the invariance of Eq. (2.9) under a local gauge transformation. In order to maintain the invariance of the Dirac equation, the additional factor has to be counteracted. This can be achieved by replacing the partial derivative with the so-called covariant derivative D_μ which is defined as,

$$\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu, \quad , \tag{2.14}$$

thereby naturally introducing an interaction between the particle $\psi(x)$ and the vector gauge field $A_\mu(x)$.This newly created gauge field corresponds to the photon which is the mediator of electromagnetic interactions and couples to electrically charged particles with a coupling strength proportional to the electric charge e. The photon field has to transform similarly to the Dirac spinors

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \rho(x) \tag{2.15}$$

with the imaginary unit i within σ_2.

[3]The unit matrix in two dimensions is defined as

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad .$$

in order to satisfy local gauge invariance. The redefined Dirac equation is thus invariant under local gauge transformations and preserves its form,

$$(i\gamma_\mu D_\mu - m)\,\psi(x) = 0 \quad . \tag{2.16}$$

The full Lagrangian of quantum electrodynamics is then

$$\mathscr{L}_{\mathrm{QED}} = \bar{\psi}\,(i\gamma^\mu\partial_\mu - m)\,\psi + e\bar{\psi}A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad , \tag{2.17}$$

where the first term describes the free motion of a spin-half particle with mass m, followed by the second term representing the interaction of such a particle with a photon. The last summand in Eq. (2.17) accounts for the kinetic term of the photon field and has to be added to QED Lagrangian for consistency. The field strength tensor $F_{\mu\nu}$ is constructed from the electromagnetic four-vector potential A_μ,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu F_\mu \quad , \tag{2.18}$$

which makes up the electromagnetic Lagrangian for a free photon,

$$\mathscr{L}_{\mathrm{EM}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad , \tag{2.19}$$

and is also part of the Proca Lagrangian

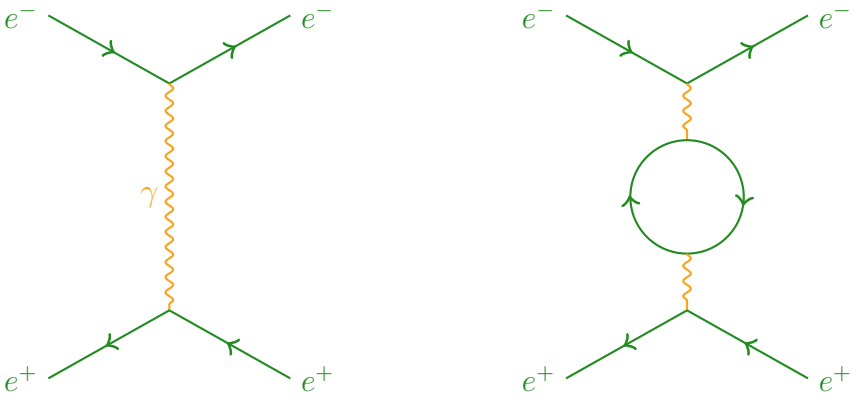
$$\mathscr{L}_{\mathrm{Proca}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2A_\mu A^\mu \quad , \tag{2.20}$$

which describes a massive vector field. However, any additional (mass) term such as that in Eq. (2.20) would spoil the local gauge invariance of the Dirac equation (2.17). On these grounds, the condition of QED being gauge invariant under local $U(1)$ transformations forces the photon to be massless and vice versa, which is in good agreement with the experimental results setting an upper limit of $1 \cdot 10^{-18}\,\mathrm{eV}$ on the photon mass [51]. The demand for massless force carriers in general a crucial prerequisite for local gauge invariance which ultimately leads to the Brout-Englert-Higgs mechanism [54–56] that causes particles to have a mass as well as the existence of a Higgs boson itself.

An important feature of the Standard Model is renormalisability, for which gauge invariance is an essential precondition [57]. Many theoretical calculations of physical observables as the cross section of particular physics processes or decay rates are performed in terms of perturbation theory since they are not analytically solvable anymore. In order for a theory to be classified as renormalisable, all divergent terms of a calculation can be absorbed within finite measurable quantities. The basic concept will be illustrated by means of electron-positron scattering, a process in which an electron and its antiparticle scatter off each other via the exchange of a photon – a process described within the framework of QED. The calculation of the so-called leading-order process shown in Fig. 2.1a can be done analytically without difficulty, however, it only yields an approximation. In quantum field theories, interactions between particles are used to be calculated perturbatively. That is, the overall interaction is quantified by a perturbative expansion in the coupling strength of the respective interaction force, taking into account all possible Feynman diagrams including in principle an arbitrary number of vertices[(4)]. Since each additional vertex contributes a factor proportional to the vertex coupling strength – a value less than unity – the greater the number of vertices, generally the smaller the impact on the overall result. Starting

[(4)] Point of intersection of two or more particles.

from the ground state – the vacuum – it is often sufficient to only consider the lowest possible number of virtual particles that are exchanged, referred to as leading order. In order to obtain the exact result, all higher-order corrections including so-called vacuum polarisation terms associated with the diagram depicted in Fig. 2.1b, have to be considered as well. In the calculation of the



(a) (b)

Fig. 2.1.: Feynman diagrams of electron-positron scattering via the exchange of a virtual photon. The leading-order process is shown on the left side in Fig. (a). A possible higher-order correction in which a virtual photon splits up into an electron-positron pair that annihilates again, is depicted on the right side in Fig. (b).

cross section of a given process, a divergent term in inevitably picked up which has the form

$$\int_{m^2}^{\infty} \frac{1}{z}\,\mathrm{d}z \quad , \tag{2.21}$$

with z denoting the momentum of the virtual particle and m its mass.

The concept of virtual particles has its origin deep in the structure of quantum field theories where interactions between particles are perturbatively described by the exchange of virtual particles. Therefore, in Feynman diagrams they only appear as internal lines and loop corrections, and they cannot be observed without entirely changing a physics process they are involved in. It is not necessary that a virtual particle has the same mass as the corresponding real particle thus allowing for virtual massive photons with three polarisation states, for instance. However, energy and momentum has to be conserved at the points of interaction. Since the vacuum polarisation effect corresponds to a quantum fluctuation, the range of forces carried by a virtual particle is limited by the uncertainty principle, which dictates an absolute upper limit on the product of the two conjugated quantities, mass and time. Thus, virtual particles with a high mass have a shorter lifetime [48, 49, 53].

The integral in Eq. (2.21) is logarithmically divergent for large z and thus the calculated cross section would ultimately be infinite which obviously contradicts experiment. In order to avoid the singularity, the integral is regularised by introducing a suitable upper cut-off mass M which, at a later step, is taken to infinity. The integral then reads

$$\int_{m^2}^{M^2} \frac{1}{z}\,\mathrm{d}z = \ln\left(\frac{M^2}{m^2}\right) \tag{2.22}$$

and can be split into a analytically calculable, finite term which is independent of M, and a term which blows up as $M \to \infty$. Crucially, all divergent terms that might appear in the calculation

will be absorbed into the physical quantities and are added in the calculation of the physical observables as corrections to the final result. Hence, the measured masses m' and the couplings g' are renormalised and are thus not the same as the bare quantities appearing in the Lagrangians (m, g) as they contain extra corrections δm and δg, respectively,

$$m' = m + \delta m \tag{2.23a}$$

$$g' = g + \delta g \quad . \tag{2.23b}$$

Since the corrections depend on the masses and the couplings of the particles involved in certain physical process, the final, measured masses and couplings are effectively energy-dependent, such that they are referred to as running masses and running couplings, respectively. For the given example of electron-positron scattering, it thus has to been distinguished between the bare electric charge e_0 used in theory calculations and the renormalised charge e that is measured in experiment. Both are related to each other by

$$e = e_0 \left(1 - \frac{e_0^2}{12\pi^2} \ln\left(\frac{M^2}{m^2}\right)\right)^{\frac{1}{2}} \tag{2.24}$$

were the infinity contained in the cut-off M are absorbed in the definition of the renormalised coupling. The electromagnetic charge is measured to be finite, though not to be constant, so the divergence of the logarithm must be cancelled by a counterpart. As a direct consequence, the coupling strength is no longer constant but depends on the energy scale of a particular physics process. In leading order, the QED coupling constant can be written as

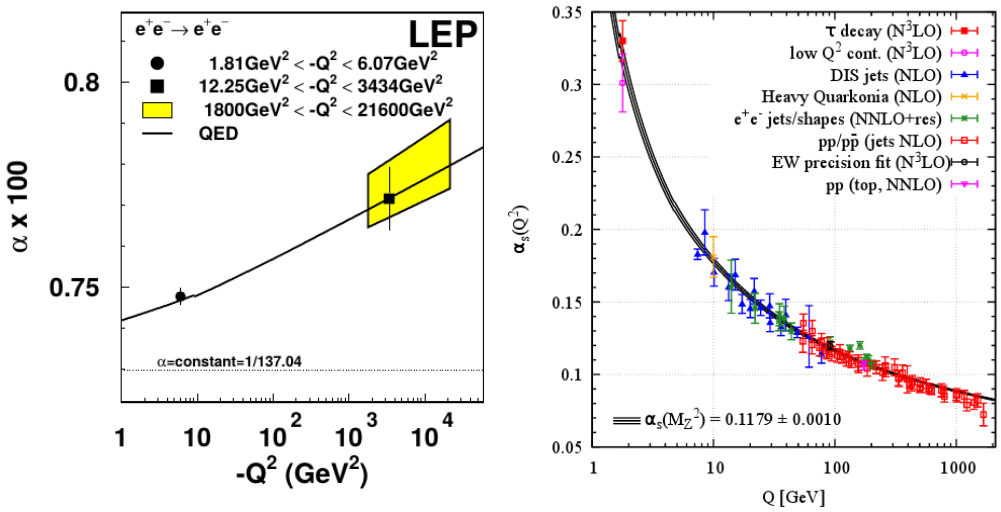
$$\alpha(Q^2) = \frac{e^2(Q^2)}{4\pi} = \frac{\alpha(\mu^2)}{1 - \frac{\alpha(\mu^2)}{3\pi} \ln\left(\frac{Q^2}{\mu^2}\right)} \quad , \tag{2.25}$$

where Q^2 denotes the momentum scale of the process being investigated and μ^2 is a reference renormalisation scale closely related to the cut-off value M^2. Fig. 2.2a shows experimental results proving that the value of the electromagnetic coupling constant increases at a larger momentum transfer and vice versa, as implied by Eq. (2.25)[(5)].

2.1.2. Quantum chromodynamics

Quantum chromodynamics is the theory of the strong force and describes the interactions between particles carrying a colour charge[(6)], namely quarks and gluon. In total three different colours charges exist: red, green and blue. According to QCD, colour charge is a conserved quantity and the net colour of any bound state is required to be zero. The three colours can be considered as axes spanning an abstract space with no axis in colour space being preferred to the others. Transformations between different colours correspond to rotations in colour space based on the $SU(3)_C$ symmetry group. In general, the special unitary group $SU(N)$ is a subgroup of $U(N)$ and has to fulfil the requirement that the determinant of U be equal to one, $\det U = 1$. Since any group described by $SU(N)$ has $N^2 - 1$ generators, QCD has eight generators, the 3×3 Gell-Mann

[(5)] In addition to the stated references, for this section Ref. [48, 49, 53, 58–60] have been used.
[(6)] Often referred to simply as colour.



(a) (b)

Fig. 2.2.: The running of coupling constants for the electromagnetic and strong force, α and α_s, respectively, as determined from experiment. Fig. (a) shows the energy dependence of α as measured at the LEP collider [61], while Fig. (b) depicts the experimental results for α_s [51]. The three data points in Fig. (a) measured different momentum transfer scales depicted as black and yellow dots perfectly match the theoretical expectation of QED. A horizontal line indicates the constant unscreened value of α. The coloured data points in Fig. (b) have been taken in various measurements. The quoted accuracy of NLO, NNLO and N^{3}LO associated with the measurements correspond to the consideration of the next-to-leading-order, next-to-next-to-leading-order and next-to-next-to-next-to-leading-order when comparing to the value calculated from QCD prediction at a reference scale equal to the mass of the Z boson, respectively.

matrices λ_a[7] that perform unitary transformations in colour space in the form of

$$U = \exp\left(i \sum_{a=1}^{8} \rho_a \lambda_a\right) \quad , \tag{2.26}$$

with ρ_a being real coefficients. Accordingly, QCD comprises eight different gauge fields G^a_μ, the so-called gluons, which are the force mediators of quantum chromodynamics.The superscript $a \in [1, \ldots, 8]$ refers to the index of a given gluon or generator.

The Lagrangian of quantum chromodynamics is given by

$$\mathscr{L}_{\mathrm{QCD}} = -\frac{1}{4} \sum_{a=1}^{8} G_{a,\mu\nu} G_a^{\mu\nu} + \sum_k \bar{q}_k (i\gamma^\mu D_\mu - m_k) q_k \quad , \tag{2.27}$$

with $G_{a,\mu\nu}$ as the gluon field strength tensor and q_k representing the spinors for different quark

[7]The Gell-Mann matrices are a set of generalised Pauli spin matrices for $SU(3)$.

(a) (b)

Fig. 2.3.: Two Feynman diagrams illustrating the self-coupling of gluons as the consequence of the non-Abelian structure of quantum chromodynamics. Fig (a) depicts a vertex with three gluons involved, whereas Fig. (b) shows a four-gluon vertex.

flavours k. The covariant derivative for QCD is defined as

$$\begin{aligned} D^\mu &= \partial^\mu + ig_s \sum_{a=1}^{8} \frac{\lambda_a}{2} G_a^\mu \\ &\equiv \partial^\mu + ig_s \frac{\boldsymbol{\lambda}}{2} \mathbf{G}^\mu \end{aligned} \tag{2.28}$$

where g_s is the strong coupling constant and $\boldsymbol{\lambda}$ and $\mathbf{G}_\mu$ are vectors of eight Gell-Mann matrices and eight gluon gauge fields, respectively. Similar to as in QED, the field strength tensor is defined as

$$G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f_{abc} G_b^\mu G_c^\nu \tag{2.29a}$$

where the factor f_{abc} represents the $SU(3)$ structure constants[8]with the indices $a, b, c \in [1, \ldots, 8]$. This can be expressed in a more condensed notation,

$$\mathbf{G}_{\mu\nu} = \partial_\mu \mathbf{G}_\nu - \partial_\nu \mathbf{G}_\mu - g_s \mathbf{G}_\mu \times \mathbf{G}_\nu \quad , \tag{2.29b}$$

which looks rather similar to the field strength tensor $F_{\mu\nu}$ of QED given by Eq. (2.18), except for the vector product at the end of the expression. This last term is a characteristic of quantum field theories that are based on non-Abelian groups which has the effect that the gauge bosons themselves carry charge and therefore may interact with each other. In case of QCD the gluons are in fact bicoloured carrying both a colour and an anticolour. Two representative Feynman diagrams of the self-interaction of gluons are shown in Fig. 2.3.

The QCD coupling constant is actually not constant but its value is dependent on the momentum transfer scale similar to the electromagnetic coupling constant described at the end of Sec. 2.1.1. At leading order, the coupling constant of quantum chromodynamics is given by

$$\alpha_s = \frac{g_s^2}{4\pi} = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi}(33 - 2n_f) \ln\left(\frac{Q^2}{\mu^2}\right)} \quad , \tag{2.30}$$

with n_f being the number of quark flavours participating in a particular interaction. Unlike the electromagnetic coupling constant in Eq. (2.25) which has a minus sign in front of the logarithm in

[8]In $SU(2)$ the structure constants are given by the totally antisymmetric tensor ϵ_{ijk}.

the denominator, α_s features a plus sign at this position which has its roots in the self-interaction of the gluons. Therefore, the strong coupling constant increases with decreasing momentum transfer and vice versa thus showing the very opposite behaviour of α of QED. An illustration of this feature can be found in Fig. 2.2b. This behaviour of the strong coupling constant is also referred to as anti screening whereas, in contrast, the behaviour of the coupling constant in quantum electrodynamics is called screening. If an electromagnetic charge is put into the vacuum, there will be still electron-positron pairs around arising from quantum fluctuations. In the vicinity of a charge the vacuum will be polarised, meaning that virtual particles of opposite charge will be attracted to the probed charge, whereas virtual particles of like charge are repelled. This leads to an overall cancellation of the field of the probed charge. However, as one approaches it, the vacuum effects diminish and the effective charge of the probe increases. The same applies for colour charges in QCD, with the crucial difference that the force-carrying particles themselves carry charges, unlike in QED. As each gluon carries both a colour and an anticolour, the polarisation of the vacuum is augmented, leading to an increase of the net coupling strength at larger distances. On the contrary, approaching the primary colour charge reduces the antiscreening effect of the surrounding virtual particles. The QCD coupling constant would feature a screening behaviour if the logarithm term of the denominator in Eq. (2.30) was greater than one, which requires more than 16 different quark flavours. As only six quark flavours are known, quantum chromodynamics remains an asymptotically free theory.

Two important phenomena of QCD, asymptotic freedom [21, 22] and confinement [62, 63], are closely related to the coupling constant's course. The first, asymptotic freedom, describes the behaviour of quarks at high energies or, equivalently, at small distances. In this regime, quarks behave as if there were quasi-free particles since the coupling constant α_s becomes increasingly small. However, at small energies, i.e. large distances, the strong coupling constant grows continually thereby increasing the attraction among the quarks[(9)]. Quarks can consequently not exist as free particles but are confined into bound colourless states called hadrons.

Depending on the quark content, hadrons can be further subdivided. Combinations of a quark and an antiquark with its anticolour cancelling the colour of the quark are referred to as mesons, whereas so-called baryons consist of either three quarks or three antiquarks with each quark carrying another colour and each antiquark carrying another anticolour, respectively[(10)]. Recently, the existence of so-called pentaquarks, bound colourless states of four quarks and an antiquark, has been revealed [64].

The formation of hadrons, referred to as hadronisation, is theoretically challenging to describe due to the low energy scales that are involved in this process. Because of confinement, instead of single quarks, bunches of colourless hadrons can be observed in experiment. At collider experiments, these bunches typically appear to be cone-shaped due to momentum conservation and are referred to as jets. The process of hadronisation and the construction of a jet will be discussed in more detail later on in Chap. 4 and Sec. 5.4, respectively[(11)].

[(9)] The breakdown of perturbation theory in the regime of large values of α_s as well as its experimental consequences will be discussed in Sec. 4.1.1 later on.

[(10)] Colour is an additive quantum number, so colours of the quarks of a baryon red, green and blue add up to "white", and for a meson the colour of the quark and the corresponding anticolour of the antiquark annihilate to "black". The picture that colours combine to "white" and cancel each other leaving "black" has been taken from chromatics in order to illustrate the idea of colourless hadrons. Of course, neither white nor black hadrons exist in nature.

[(11)] In addition to the stated reference, this section is based on Ref. [48–50, 53, 58, 60, 65, 66].

2.1.3. The electroweak theory

Although formulated separately, the electromagnetic and the weak interaction are both assumed to be a low-energy representation of the unified electroweak interaction based on a $SU(2)_L \otimes U(1)_Y$ gauge group. The electroweak theory contains four different gauge fields of which three $(W_\mu^1, W_\mu^2, W_\mu^3)$ are associated to the $SU(2)_L$ group and couple to the so-called weak isospin, whereas the B_μ belongs to the $U(1)_Y$ group and couples to the weak hypercharge of particle, which is given by

$$Y = 2(Q - I_3) \quad , \tag{2.31}$$

with Q being the electric charge and I_3 the third component of the weak isospin. Since parity is not conserved within the weak interaction, left-handed and right-handed particles are treated differently by the weak component of the electroweak force. The gauge bosons of $SU(2)_L$ couple only to left-handed particles and right-handed antiparticles which is known as parity violation [67, 68]. Thus, only left-handed particles and right-handed antiparticles carry a weak isospin other than zero and are arranged in so-called weak isospin doublets

$$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix} \quad \begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix} \quad \begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix} \quad \begin{pmatrix}u\\ d'\end{pmatrix} \quad \begin{pmatrix}c\\ s'\end{pmatrix} \quad \begin{pmatrix}t\\ b'\end{pmatrix} \quad . \tag{2.32}$$

For particles appearing in the upper portion of the doublets the third component of the weak isospin is $+1/2$, whereas particles of the lower portion have a third component of the weak isospin equal to $-1/2$ (for simplicity only particles and not antiparticles are given). In contrast, right-handed particles and left-handed antiparticles are weak isospin singlets with $I_3 = 0$. Commonly, particles on top are called up-type, and particles below are referred to as down-type. The meaning of the prime in the case of down-type quarks will be explained in Sec.2.1.4.

The electroweak Lagrangian,

$$\mathscr{L}_{\mathrm{EW}} = \sum_k i\bar{\psi}_k \gamma^\mu D_\mu \psi_k - \frac{1}{4}\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.33}$$

has a similar structure than that of QCD given by Eq. (2.27) except for absence of any mass term.

As the electroweak coupling to left-handed and right-handed particles is different, the covariant derivative in Eq. (2.33)

$$D_\mu = \partial_\mu + \frac{1}{2}ig\boldsymbol{\tau}\cdot\mathbf{W}_\mu + \frac{1}{2}ig'YB_\mu \tag{2.34a}$$

$$D_\mu = \partial_\mu + \frac{1}{2}ig'YB_\mu \tag{2.34b}$$

reads differently for the left-handed and right-handed component in Eqs. (2.34a) and (2.34b), respectively. Here, g represent the coupling strength of the $SU(2)_L$ portion and g' denotes the coupling strength of the $U(1)_Y$ portion of electroweak theory. Furthermore, $\boldsymbol{\tau}$ and $\mathbf{W}_\mu$ are vectors of the Pauli spin matrices, which are the generators of $SU(2)_L$ and the associated gauge fields, respectively. The field strength tensors of the electroweak interaction given in Eq. (2.33) are defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.35a}$$

$$\mathbf{W}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - g\mathbf{W}_\mu \times \mathbf{W}_\nu \quad . \tag{2.35b}$$

As was the case for $SU(3)_C$ in QCD, the $SU(2)_L$ group is a non-Abelian group which gives rise

to a self-interaction of its gauge bosons indicated by the vector product in Eq. (2.35b). Similar to QED the $U(1)_Y$ part does not provide a self-interaction of its mediator particles.

In Sec. 2.1.1 it has been shown that in order to preserve local gauge invariance of the theory, the gauge bosons have to be massless. However, experiments have shown that only the photon and the gluons are massless but the weak force mediators are quite massive. As a consequence, the electroweak symmetry has to be broken in order to both meet the experimental results and to satisfy the theoretical principles. In the Standard Model, electroweak symmetry breaking is described by the Brout-Englert-Higgs mechanism [54–56]. By adding a new complex scalar $SU(2)_L$ doublet field

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.36}$$

to the electroweak Lagrangian of Eq. (2.33) results into an additional term

$$\mathscr{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \tag{2.37}$$

where the covariant derivative is simply the one used in electroweak theory. The potential is defined by

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{2.38}$$

with λ begin a positive real number. Crucially, the minima of the given potential depend on the value of the parameter μ^2. If $\mu^2 > 0$, there is only one minimum at $\Phi = 0$. However, in case of $\mu^2 < 0$, the field Φ has a ground state different from zero,

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad , \tag{2.39}$$

for which the potential is minimised. The two minima of the potential are then $\pm v$ where v is defined as

$$v = \sqrt{\frac{|\mu^2|}{\lambda}} \tag{2.40}$$

and referred to as the vacuum expectation value (VEV) of the Higgs field. The procedure described is referred to as spontaneous symmetry breaking as the underlying symmetry is spoiled by the selection of a particular ground state. Due to the freedom of gauge transformation, ϕ can be written as an expansion around the VEV with a scalar field $\eta(x)$ which will be later identified with the Higgs field,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix} \quad . \tag{2.41}$$

With this representation of ϕ, the Higgs Lagrangian of Eq. (2.37) can be rewritten as

$$\mathscr{L}_{\text{Higgs}} = \left[\frac{1}{2} (\partial_\mu \eta)(\partial^\mu \eta) - \left|\mu^2\right| \eta^2 \right] + \frac{1}{2} \frac{g^2 v^2}{4} \left[\left|W_\mu^+\right|^2 + \left|W_\mu^-\right|^2 \right] + \frac{1}{2} \frac{v^2}{4} \left| g' B_\mu - g W_\mu^3 \right|^2 \ , \tag{2.42}$$

incorporating the physical mass eigenstates of the gauge bosons. This Lagrangian is given at leading order and comprises the free propagation of the Higgs field $\eta(x)$ with a mass $m_{\text{Higgs}} = \sqrt{2}\mu$. The mass terms for the electroweak gauge bosons arise from spontaneous symmetry breaking together with the subsequent expansion of Φ around the Higgs field VEV and the demand that the Lagrangian be locally gauge invariant under $SU(2)_L \otimes U(1)_Y$ symmetry. There are in general higher-order terms occurring in Eq. (2.42) including cubic and quartic gauge boson couplings to the Higgs field and self-interactions of the Higgs field which are not shown here. The free prop-

agation of the electroweak gauge boson is not included in the Higgs Lagrangian as well as it has been already given by Eq. (2.33).

The electroweak mixing angle, also referred to as Weinberg angle, is defined through the coupling strengths of the electroweak symmetry groups,

$$\theta_W = \tan^{-1}\left(\frac{g'}{g}\right) \quad \Leftrightarrow \quad \cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \tag{2.43}$$

and relates all couplings involved in the electroweak interaction by

$$e = g'\cos\theta_W = g\sin\theta_W \quad . \tag{2.44}$$

The electroweak gauge bosons and their masses can be expressed in terms of the electroweak gauge fields and the Weinberg angle as

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \qquad \text{with} \qquad m_\gamma = 0 \tag{2.45a}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \qquad \text{with} \qquad m_Z = \frac{m_{W^\pm}}{\cos\theta_W} \tag{2.45b}$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2) \qquad \text{with} \qquad m_{W^\pm} = \frac{gv}{2} \tag{2.45c}$$

Two $SU(2)_L$ gauge fields, W_μ^1 and W_μ^2, are absorbed into the $W^\pm$ bosons. The remaining W_μ^3 and B_μ, the gauge boson of the $U(1)_Y$ group, form two orthogonal superpositions, which are the photon known from electromagnetism and the electrically neutral but massive Z boson. The Weinberg angle that relates the masses of the $W^\pm$ and Z boson has been measured to have a value such that $\sin^2\theta_W = 0.2313$ [51].

The Higgs mechanism does not only give a mass to the electroweak gauge bosons but provides a general mechanism within the Standard Model to give elementary particles a mass. Fermion masses can are taken into account by expanding the Higgs Lagrangian with an additional mass term of the form

$$\tilde{g}_f \left(\bar{\Psi}_L^f \Phi \Psi_R^f + \bar{\Psi}_R^f \Phi^\dagger \Psi_L^f\right) \tag{2.46}$$

with Ψ denoting the left-handed and right-handed Dirac spinors of a fermion f, respectively, and $\tilde{g}_f$ being the corresponding Yukawa coupling strength of the fermion to the Higgs field. Hence, the particular value of the given fermion mass,

$$m_f = \tilde{g}\frac{v}{\sqrt{2}} \quad , \tag{2.47}$$

is governed by its respective Yukawa coupling, which is a free parameter in the Standard Model and thus has to be experimentally determined, and the Higgs VEV, which can be fixed by the Fermi constant G_F,

$$v = \left(\sqrt{2}G_F\right)^{-1/2} \quad , \tag{2.48}$$

and has been found to have a value of 246 GeV [69–71][(12)].

[(12)] In addition to the stated references, this section has been adapted from Ref. [48, 50, 53, 58, 59, 65].

2.1.4. The weak decay

The breaking of electroweak symmetry discussed in the previous section causes the mass eigenstates and weak eigenstates of massive strongly interacting particles not to be identical. They are related to each other by the unitary 3×3 CKM[13] flavour-mixing matrix that maps the mass eigenstates to the weak eigenstates and vice versa. Historically it can be seen as an extension of the GIM[14] mechanism which only includes the first and second generation of quarks and relies on the 2×2 Cabibbo flavour-mixing matrix [12, 72]. The mass eigenstates are described as a superposition of several weak eigenstates which allows quarks to change their flavour in weak processes mediated by a $W^{\pm}$ boson. The flavour of a particle is not a conserved quantity in terms of the weak interaction which gives rise to processes as radioactive β-decay or quark/meson mixing[15]. The relation between weak eigenstates, which are furnished with a prime, and mass eigenstates is given by

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad . \tag{2.49}$$

The transition of a quark of flavour i to flavour j is defined by the CKM matrix elements, and the probability of a transition is given by $|V_{ij}|^2$. The values of the CKM matrix elements are not predicted by the theory. They have been measured to be [51]:

$$V_{CKM} = \begin{pmatrix} 0.97420 \pm 0.00021 & 0.2243 \pm 0.0005 & 0.00394 \pm 0.00036 \\ 0.218 \pm 0.004 & 0.997 \pm 0.017 & 0.00422 \pm 0.00008 \\ 0.0081 \pm 0.0005 & 0.0394 \pm 0.0023 & 1.019 \pm 0.025 \end{pmatrix} \tag{2.50}$$

From the measured values it can be immediately seen that transitions within the same generation are most preferred since the diagonal elements are close to unity. Less likely are transitions between the first and the second generation, followed by transitions between the second and the third generation. The strongest suppression can be noted for transitions from the third to the first generation and vice versa.

From the structure of the matrix it can be deduced that only transitions between up-type and down-type quarks are permitted. In the Standard Model transitions of up-/down-type quarks into other up-/down-type quarks of another generation, referred to as flavour-changing neutral currents (FCNC), are forbidden at tree-level and moreover greatly suppressed at higher orders. In fact, the suppression of FCNC led to the prediction of charm quarks via the GIM mechanism in 1970 [12]. Two quark generations had not been sufficient to explain the CP violation observed in weak interactions on a theoretical basis, which is why a third generation of quarks had been postulated and could be experimentally established with the discovery of the bottom quark in 1977 [16, 17]. Described and conceptually introduced by the complex phase of the CKM matrix, CP violation is an explanation for the asymmetry observed between matter and antimatter in the universe. Prior to the introduction of a third fermion generation it was not possible to include the observed CP violation into the Standard Model since the unitary mixing matrix requires at least three dimensions in order to feature a non-vanishing complex phase.

However, CP violation in the quark sector is far too small to explain the amount of matter-antimatter asymmetry in the universe by itself. Contrary to the Standard Model presumption, neutrinos do have a small mass but non-zero mass [73, 74], which is the reason why they are ex-

[13] Named after Nicola Cabibbo, Makoto Kobayashi and Toshihide Maskawa for their outstanding pioneering work in this field [16, 72].

[14] Developed by Sheldon Lee Glashow, John Iliopoulos and Luciano Maiani.

[15] In electroweak interactions the weak $SU(2)_L$ component of the theory that not conserve flavour and violates parity.

posed to a similar dissimilarity of weak and mass eigenstates as the quarks[16] are. Thus, neutrinos may change their flavour during free propagation since a mass eigenstate can be expressed as the linear combination of various flavour eigenstates and vice versa. The relation between the two kinds of eigenstates is given by the so-called PMNS[17] matrix [75],

$$\begin{pmatrix} \nu_e \\ \nu_\nu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{2.51}$$

which is 3×3 unitary mixing matrix and equivalent to the CKM matrix from the quark sector. The ν_f with $f = e, \mu, \tau$ denote the three flavour eigenstates whereas $\nu_i,\ i = (1, 2, 3)$ refer to the neutrino mass eigenstates. Similar to the CKM matrix, an irreducible complex phase of the PMNS matrix also gives rise to CP violation in the lepton sector[18].

2.2. Parton distribution functions

A proton consists of three quarks, two up quarks and one down quark. Inside the proton, however, the quarks will interact with each other through the exchange of gluons. A powerful explanation for the dynamics of this interacting system is provided by the parton model.

The partonic content of a proton can be classified into three main categories. The main constituents are the three quarks previously mentioned, referred to as valence quarks, accounting for the basic structure of the proton. They are embedded in a sea composed of many ephemeral quark-antiquark pairs originating from quantum fluctuations, therefore referred to as sea quarks. The third category is a multiplicity of gluons which mediate the interaction between both the valence quarks and the sea quarks. The parton model was originally introduced by Richard Feynman in order to describe collisions of hadrons at high energies [8, 76]. Even after the development and experimental confirmation of quantum chromodynamics the parton model remains a justifiable approximation at high energies. In the parton model, partons are treated as quasi-free particles which in QCD is a direct consequence of asymptotic freedom. For small distances between coloured objects the strong coupling constant is typically small thus allowing for perturbative calculations of collision cross sections. At large distances, however, the strong coupling constant blows up which therefore confines quarks into colourless hadrons. As the value of α_s is typically larger than one in this regime, perturbative approaches cannot be applied below a certain energy scale. Instead, probability density functions of quark momenta within the proton can be used for the calculations. They are referred to as parton density or parton distribution functions (PDF) and give the probability to find a particular parton with a specific fraction x of the entire hadron's momentum at the resolution scale Q. Due to the inherent non-perturbative nature of partons which cannot be observed as free particles due to colour confinement, it is not possible to calculate the parton densities analytically for a given momentum transfer scale. For this reason they have to be extracted from fits to QCD-sensitive observables from deep-inelastic scattering experiments and hadron collider data. Nevertheless, the evolution in momentum transfer of the parton densities from a high scale to a lower scale, down to a cut-off scale of a few MeV, can be described by the DGLAP evolution equations [77–80].

At hadron colliders such as the LHC, the incoming protons at a collision provide broad beams

[16] By convention, in the quark sector the down-type particles have been chosen to be rotated. In the lepton sector, it appeared to be more convenient to rotate the up-type particles since it was the confirmation of them being massive that provides the opportunity to explore the field of flavour oscillations in the lepton sector.

[17] Named after Bruno Pontecorvo, Ziro Maki, Masami Nakagawa and Shoichi Sakata.

[18] In addition to the stated references, for this section Ref. [48, 49, 58, 65] have been used.

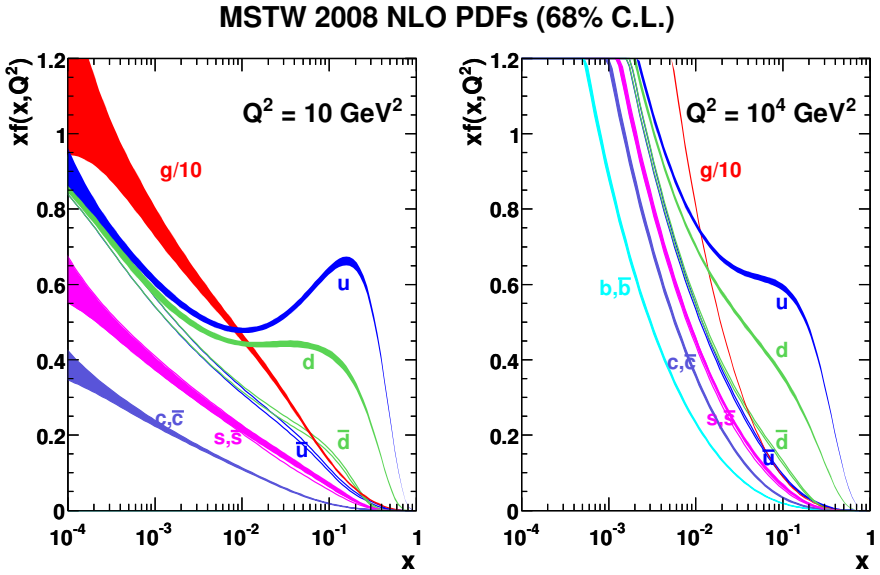


Fig. 2.4.: Values of the proton PDFs determined by the MSTW group [81] in 2008 at next-to-leading-order for two different energy scales, $Q^2 = 10\ \mathrm{GeV}^2$ (left) and $Q^2 = 10^4\ \mathrm{GeV}^2$ (right). The valence quarks carry a high momentum fraction of roughly $1/3$ with the up quark component exceeding the down quark component by a factor of approximately two. For smaller x, gluons and other quark flavours become more likely. Approaching towards higher energies, which is similar to an on average higher momentum transfer, more partons are involved each with a smaller x. For both figures the gluon contribution has been divided by 10 in order to reasonably fit into the plotting range.

of partons carrying a fraction x of the original proton momentum. This is the reason why parton distribution functions are commonly used at collider experiments to describe the substructure of protons and to quantify the outcome of high-energy proton-proton collisions. Fig. 2.4 shows parton density functions of the proton at two different momentum transfer scales. For a small momentum transfer, the valence quarks carry on average rather high momentum fractions, and sea quarks and gluons carry much smaller fractions of the proton's momentum. On the other hand, at higher momentum transfer, sea quarks and gluons contribute on average more and are thus more likely involved in proton-proton collisions. A correct calculation of the total cross section of a particular physics process requires all final-state hadrons to be considered. Difficulties in the calculation arising due to the low energies involved can be overcome by making use of the factorisation theorem [82] which states that a total inclusive cross section of a hadronic process $A + B \to X$ can be written as

$$\sigma_{AB} = \int \mathrm{d}x_a \mathrm{d}x_b \, f_{a/A}(x_a, \mu_\mathrm{F}^2) f_{b/B}(x_b, \mu_\mathrm{F}^2) \times \left[\hat{\sigma}_0 + \alpha_s(\mu_\mathrm{R}^2)\hat{\sigma}_1 + \cdots\right]_{ab\to X} \quad . \tag{2.52}$$

In this way, the calculation of the overall cross section can be factorised into the calculation of the leading-order partonic process $\hat{\sigma}_0$ with higher-order corrections $\hat{\sigma}_i$, $i = 1, 2, \ldots$, and the appropriate parton distribution functions $f_{a/A}$ and $f_{b/B}$ of the initial-state partons a and b that belong to the hadrons A and B, respectively. With a suitable factorisation scale μ_F^2 and renormalisation scale μ_R^2, the momentum fraction x_i of the involved partons and other phase space variables associated with the final-state X can be numerically integrated over. At the LHC, the partons a and b

would correspond to partons inside of two colliding protons. The final state could in principle be any resonance produced in the collision, for instance a Z boson produced by the annihilation of a quark-antiquark pair, $q\bar{q} \to Z$. (Furthermore, this section employs Ref. [49, 53])

2.3. The top quark

With a measured mass of approximately $173.3\,\mathrm{GeV}$[(19)] the top quark is by far the heaviest elementary particle known to date[(20)]. It has only been discovered in 1995 at the Tevatron experiments CDF and DØ [18, 19] since the Tevatron was the first collider in history with a centre-of-mass energy high enough to produce top quark pairs at an appropriate rate. Since then, its properties have been extensively probed by various experiments. As with all fermions, the mass of the top quark is a free parameter in the Standard Model and has thus to be determined by experiment. The mass, m, of a particle can be related to its lifetime τ and decay width Γ, by

$$\tau = \frac{1}{\Gamma} \sim \frac{1}{m^3} \quad , \tag{2.53}$$

where τ is the inverse of Γ [48]. Due to the exceptionally high mass the top quark has a lifetime of approximately $5 \cdot 10^{-25}\,\mathrm{s}$ which is smaller than the hadronisation time scale of roughly $10^{-24}\,\mathrm{s}$. The extremely short lifetime of the top quark allows the measurement of its bare properties. Furthermore, the top quark may perhaps play an important role in electroweak symmetry breaking as it has, according Eq. (2.47), a Yukawa coupling very close to unity, whereas all other fermions feature considerably smaller values. In addition, the exact value of its mass could have implications in terms of the overall stability of the universe [84]. Aside from the mass of the top quark, a growing effort in recent years has been placed on performing precision both theoretical calculations as well as measurements of its properties, for example its coupling to vector bosons, the correlation of spins, or differential cross sections. The hope is to gain deeper insights into phenomena which at present cannot be explained satisfactorily like flavour changing neutral currents or the coupling between fermions and the Higgs boson. Experimentally a detailed understanding of top quark processes is of great importance since they represent major backgrounds to searches for new physics beyond the Standard Model, including the predictions of many Supersymmetry [85] (SUSY) models.

2.3.1. Top quark production

Top quarks can either be produced individually in an electroweak process, referred to as single top quark production, or as pairs of top and antitop quarks via the strong interaction.

2.3.1.1. Single top quark production

The electroweak production of single top quarks can be categorised into three different channels, based on the lowest-order production diagrams at hadron colliders such as the LHC. An overview of the corresponding Feynman diagrams of single top quark production is depicted in Fig. 2.5. In the s-channel the annihilation of an up-type quark and a down-type antiquark (and vice versa) creates a virtual W boson which then splits up into an (anti)top quark and an (anti)bottom quark. The t-channel describes a flavour excitation process, whereby an up-type quark radiates a virtual W boson that couples to a bottom quark. The latter originates either from the quark sea of the

[(19)] The quoted value corresponds to a combination of measurements from the LHC experiments ATLAS and CMS and the Tevatron experiments CDF and DØ: $m_{\mathrm{top}} = 173.34 \pm 0.36\,(\mathrm{stat.}) \pm 0.67\,(\mathrm{syst.})\,\mathrm{GeV}$ [83]

[(20)] The top quark mass has a value similar to that of an entire gold atom.

proton or from gluon splitting, thereby being converted to a top quark. In the associated production of a W boson and top quark in the so-called Wt-channel either a down-type quark is increased in virtuality by a gluon and thus has sufficient virtuality to become a top quark by the radiation of a W boson, or a down-type quark, most likely a bottom quark, splits up into a W boson and a virtual top quark which subsequently becomes real through the coupling to a gluon.

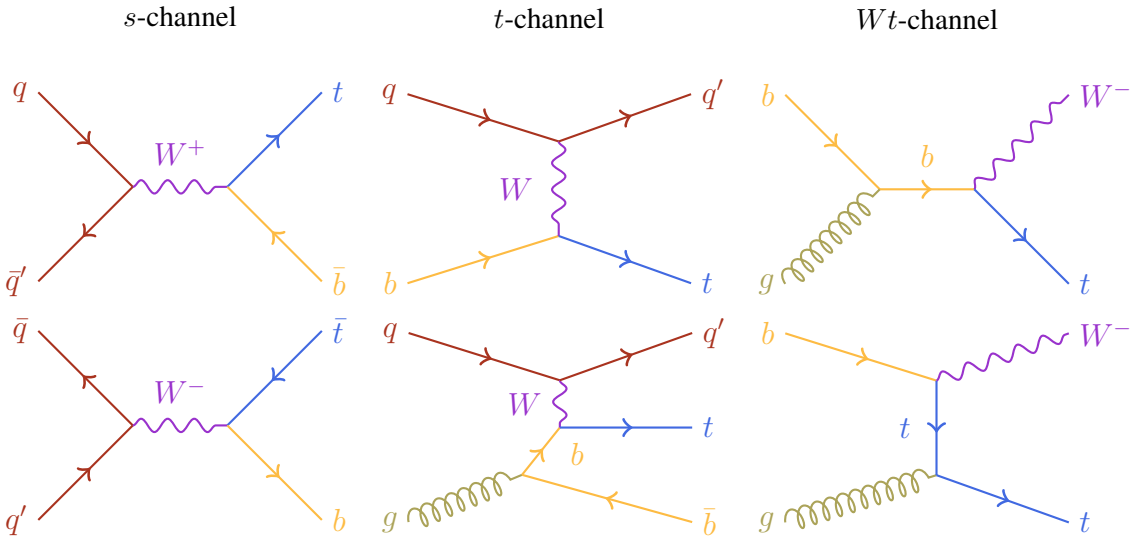


Fig. 2.5.: Representative Feynman diagrams for the electroweak production of single top quarks. The bottom and antibottom quarks could be replaced by any down-type quarks and antiquarks, respectively. Similarly, the quarks indicated by "q" depicted in red can be in principle bottom and top quarks, respectively. Given the structure of the CKM matrix, the cases depicted here are mostly favoured.

At proton-proton colliders such as the LHC, the cross section of single top quark production are different for producing a top quark or an antitop quark. Charge conservation forces up-type quarks to emit an W^+ boson and down-type quarks to radiate W^- bosons only. Furthermore, a W^+ boson will convert a bottom quark into a top quark, whereas a W^- boson produces an antitop quark from coupling to an antibottom quark. Since the proton has two up-type valence quarks but only one down-type valence quark, in the s-channel and in the t-channel the number of up-type quarks in the initial state of single top quark production is twice as large as the number of down-type quarks. Consequently, more top quarks are produced via the electroweak interaction than antitop quarks at proton-proton colliders, whereas at proton-antiproton colliders as the Tevatron this asymmetry does not appear. However, the Wt-channel does not contribute to the asymmetry between the number of produced top quarks and antitop quarks, respectively. In contrast to the s-channel and the t-channel, only down-type quarks (or up-type antiquarks) appear in the initial state of the Wt-channel which occur at the same rate for both proton-proton and proton-antiproton colliders, irrespective of whether the initial quark was a valence quark or has been taken from the quark sea of the proton.

2.3.1.2. Top quark pair production

At the LHC, top-antitop quark pairs ($t\bar{t}$) are predominantly produced via the strong interaction. In Fig. 2.6 the leading Feynman diagrams for the three categories of top quark pair production are depicted. As for the electroweak single top quark production, the s-channel describes the

annihilation of a quark and an antiquark thereby producing a gluon which subsequently splits up into a top-antitop quark pair. As gluons carry colour as well, they may couple to each other as discussed in Sec. 2.1.2 giving rise to a complementary process in the s-channel. In the so-called gluon-gluon fusion, two incoming gluons combine to a single gluon which is in case of top quark pair production a virtual particle and finally splits up into a top-antitop quark pair. As elaborated in more detail in App. A, the dominant production mechanism of top quark pairs at the LHC is gluon-gluon fusion to which not only the s-channel but also the t-channel contributes. The two gluons in the initial state do not annihilate but exchange a virtual top quark that gets converted into real particles at the coupling vertices with the gluons. The only difference to the u-channel is the interchange of the outgoing quarks with respect to the the incoming gluons.

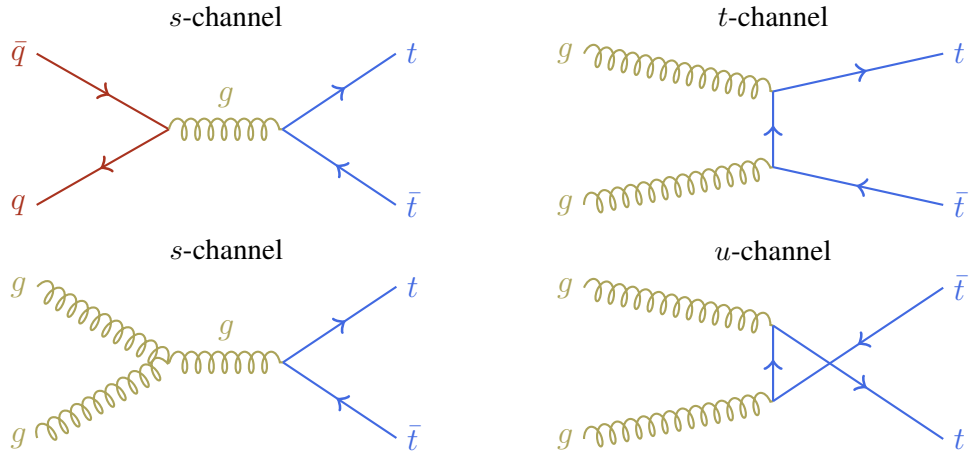


Fig. 2.6.: Representative Feynman diagrams for the strong top quark pair production.

Aside from the leading-order processes depicted in Fig. 2.6, higher-order processes such as those shown in Fig. 2.7 have to be taken into account. In strong processes additional gluons are radiated in either the initial state or the final state, referred to as initial-state radiation (ISR) and final-state radiation (FSR), respectively. Other corrections due to the exchange of purely virtual particles might occur in form of the emission and absorption of the same gluon, so-called loops.

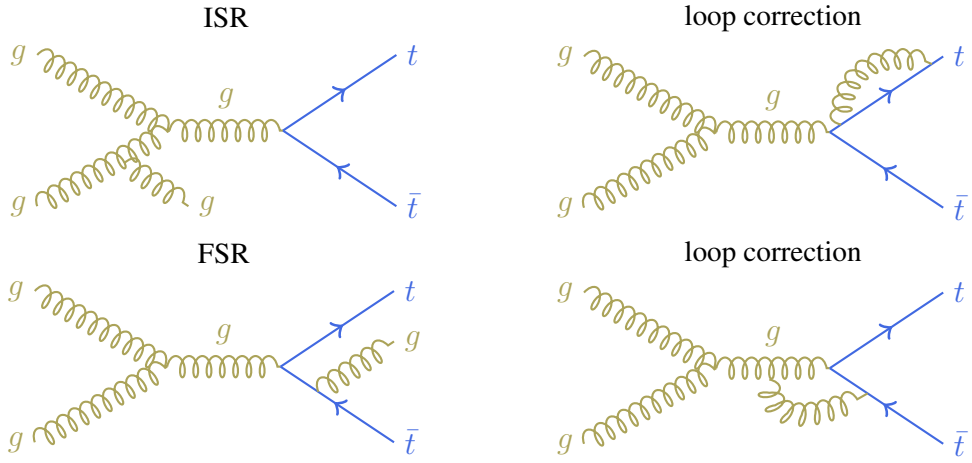


Fig. 2.7.: Representative Feynman diagrams for the strong top quark pair production including higher-order corrections.

To date, the cross section for top quark pair production at a proton-proton collider running with a centre-of-mass energy of 13 TeV has been calculated up to the next-to-next-to-leading-order (NNLO) in QCD including soft-gluon resummation at next-to-next-to-logarithmic order (NNLL) yielding a theoretical value of

$$\sigma_{t\bar{t}}^{\text{NNLO+NNLL}} = 832^{+20}_{-29}\,(\text{scale}) \pm 35\,(\text{PDF} + \alpha_s)\,\text{pb} \tag{2.54}$$

assuming a top quark mass of 172.5 GeV [86]. The scale uncertainty reflects uncertainties on the factorisation scale μ_F and renormalisation scale μ_R, while the second source of uncertainty stated is associated to possible choices of the PDF and the strong coupling constant α_s. Fig. 2.15 shows an overview of recent measurements of the $t\bar{t}$ production cross section performed by the ATLAS and CMS collaborations. The comparison with the theoretical prediction yields a good agreement for all measured values.

2.3.2. Top quark decay

In the Standard Model, top quarks can only decay via the electroweak interaction into a W boson and a down-type quark in which the latter almost exclusively appears to be a bottom quark. In the case of the top quark, the probability to decay to each of the particular three down-type quarks are, as discussed in Sec. 2.1.4, governed by the square of the CKM matrix elements, with the measured values [51] as given by Eq. (2.50):

$$|V_{td}|^2 = 6.6 \cdot 10^{-5} \tag{2.55a}$$
$$|V_{ts}|^2 = 1.6 \cdot 10^{-3} \tag{2.55b}$$
$$|V_{tb}|^2 = 1.04 \tag{2.55c}$$

The W boson decays further into either a charged lepton and its corresponding neutrino referred to as leptonic decay, or into an up-type and a down-type quark pair which is referred to as hadronic decay. Although technically allowed, the hadronic decay of the W boson into an (anti)top quark is kinematically forbidden. However, the other five quark flavours have to be considered as decay products of the W bosons. In about one third of all cases the W boson decays leptonically and in two thirds of all cases it decays hadronically. More details about the branching fractions[(21)] of the W boson can be found in Tab. 2.3.

$W \to e_n u_e$	$W \to \mu\nu_\mu$	$W \to \tau\nu_\tau$	$W \to q\bar{q}'$
$(10.71 \pm 0.16)\,\%$	$(10.63 \pm 0.15)\,\%$	$(11.38 \pm 0.21)\,\%$	$(67.41 \pm 0.27)\,\%$

Tab. 2.3.: Branching fractions of the W boson [51]. The leptonic branching fractions are given separately for each lepton flavour whereas the hadronic branching fraction is inclusive in terms of possible quark flavours.

The decay of top quark pairs is thus categorised according the subsequent W boson decays, resulting into three common decay channels. In the dileptonic channel, both W bosons decay leptonically giving rise to two neutrinos in the final state. Contrarily, the fully-hadronic channel features two hadronically decaying W bosons. The third so-called lepton+jets channel is a mixture of the two previous channels, namely where one W boson decays leptonically and the other into

[(21)] If a particle can decay into more than one final state, the probability for each final state is given by to so-called branching fraction which is defined as the fraction relative to the sum of all possible decays.

a quark-antiquark pair. Extra care must be taken within an analysis in the way τ-lepton decays are treated, since they may further decay both hadronically and leptonically due to their mass of roughly $1.8\,\mathrm{GeV}$. A leptonic decay of a τ-lepton via a W boson creates additional neutrinos in the final state of the $t\bar{t}$ decay which may cause problems at many hadron colliders as discussed later in Secs. 3.3.1 and 5.7. For this reason two numbers are stated for the branching fraction in case of the dileptonic and lepton+jets channel, respectively.

2.3.2.1. The $t\bar{t}$ dileptonic channel

With a predicted branching ratio of approximately $10\,\%$ considering τ-leptons and $5\,\%$ if not, the dileptonic channel is the least probable decay channel of the three main $t\bar{t}$ decay channels. Its final state comprises two bottom quarks from top quark decay and a total of two leptons and two neutrinos as can be seen in Fig. 2.8. Due to the two neutrinos in the final state, highly elaborated techniques have to be used in order to perform a kinematic reconstruction of the top quark pair system. However, the dileptonic channel offers a very clean event topology since it has a very low background contamination from strong processes since it typically features two isolated leptons in the final state.

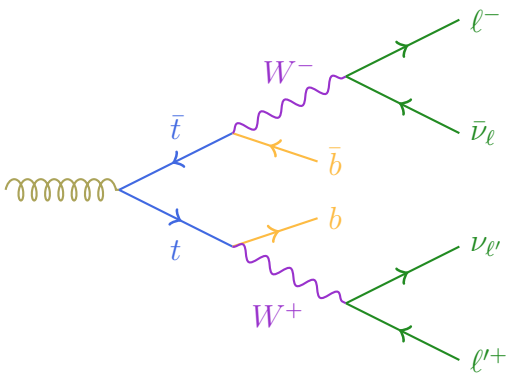


Fig. 2.8.: Dileptonic $t\bar{t}$ decay

2.3.2.2. The $t\bar{t}$ lepton+jets channel

The lepton+jets channel has a branching fraction of about $44\,\%$ with and roughly $34\,\%$ without explicitly taking τ-leptons into account. The final state is characterised by two bottom quarks, two light quarks from a W boson, one high energetic charged lepton and one neutrino (cf. Fig. 2.9). Despite the presence of a single neutrino, the kinematic reconstruction of the $t\bar{t}$ system can be performed much more straightforward compared to the dileptonic channel. The main background of this channel is the associated production of a vector boson with hadrons, while minor contributions originate from strong processes involving numerous hadrons. From an experimental perspective, the lepton+jets channel offers a good balance between a clean event signature and large branching fraction.

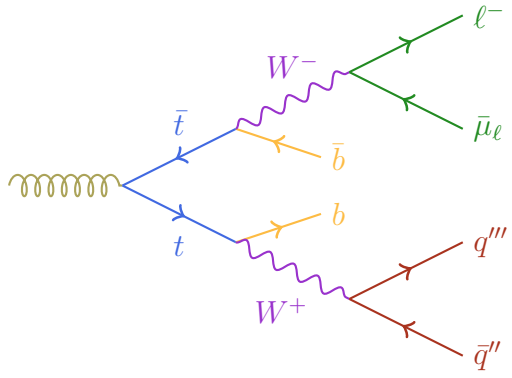


Fig. 2.9.: Lepton+jets $t\bar{t}$ decay

2.3.2.3. The $t\bar{t}$ fully-hadronic channel

In the fully-hadronic channel no τ-leptons are expected in the signal signature as both W bosons decay hadronically resulting in a branching fraction of about $45\,\%$. It thus is the most probable channel which top quark pairs decay in but also an experimentally very challenging one. Due to its fully-hadronic final state consisting of four light quarks from the W bosons and two bottom quarks as depicted in Fig. 2.10, not a single neutrino from either W boson has to be considered within the kinematic reconstruction as it was the case in dileptonic channel and partly in the lepton+jets channel as well. However, combinatoric effects caused by fully-hadronic backgrounds becomes a main background to this channel.

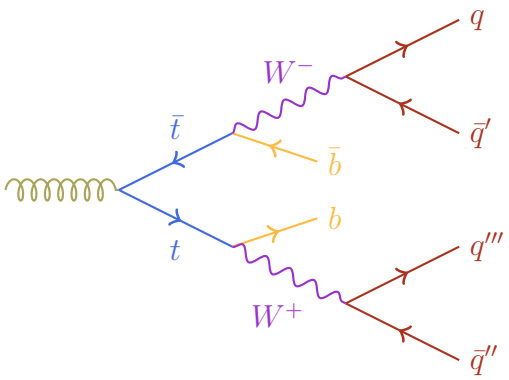


Fig. 2.10.: Fully hadronic $t\bar{t}$ decay

2.4. Top quarks in association with Z bosons

In addition to higher-order corrections mediated by the strong force as discussed in Sec. 2.3.1.2, electroweak corrections may also appear but their relative impact on the total $t\bar{t}$ production cross section is much smaller due to the lower coupling strength. A common case of such electroweak corrections is the associated production of top quark pairs and a Z boson ($t\bar{t}Z$) where the Z can radiated either from the incoming quarks or from both real and virtual top quarks in the final state. For real Z bosons the quark from which it was radiated had to be virtual as this process would have been kinematically forbidden otherwise.

The investigation of the associated production of top-antitop quark pairs and a Z boson allows one to directly measure the coupling strength between the top quark and the Z boson, which are the heaviest and third heaviest particle in the Standard Model, respectively. In addition, single top quark production in association with a Z boson has been studied as well, but higher background rates and the electroweakly dominated production noticeably limits the signal purity [44, 45]. Yet t-Z coupling is experimentally not well-constrained which is why its value may considerably vary in some physics models beyond the Standard Model. The $t\bar{t}$ process serves furthermore as an irreducible background to measurements of the coupling of the top quark to the Higgs boson as well as to Supersymmetry searches because of a similar final-state signature.

The production cross section of $t\bar{t}Z$ at $13\,\mathrm{TeV}$ has been calculated to be

$$\sigma_{t\bar{t}Z}^{\mathrm{NLO+NNLL}} = 0.863^{+8.5\,\%}_{-9.9\,\%}\,(\mathrm{scale}) \pm 3.2\,\%\,(\mathrm{PDF} + \alpha_s)\,\mathrm{pb} \tag{2.56}$$

at next-to-leading (NLO) and NNLL accuracy in the calculation considering both electroweak and strong corrections [87]. The most recent measurements of the cross section by the ATLAS and the

CMS collaborations,

$$\sigma_{t\bar{t}Z}^{\text{ATLAS}} = 1.05 \pm 0.05\,(\text{stat.}) \pm 0.09\,(\text{syst.})\,\text{pb} \quad (139\,\text{fb}^{-1}) \tag{2.57a}$$

$$\sigma_{t\bar{t}Z}^{\text{ATLAS}} = 0.95 \pm 0.08\,(\text{stat.}) \pm 0.10\,(\text{syst.})\,\text{pb} \quad (36.1\,\text{fb}^{-1}) \tag{2.57b}$$

$$\sigma_{t\bar{t}Z}^{\text{CMS}} = 0.95 \pm 0.05\,(\text{stat.}) \pm 0.06\,(\text{syst.})\,\text{pb} \quad (77.5\,\text{fb}^{-1}) \tag{2.57c}$$

are in reasonable agreement with the theoretical prediction [39, 46, 47]. The numbers in parentheses denote the amount of data used to derive the respective measured values in terms of the integrated luminosity, whose concept will be introduced and further explained in Sec. 3.2. However, the cross section is by almost a factor of 1000 smaller than the cross section of $t\bar{t}$ production. The main reason for this is the production of the Z boson which can only be accounted for by an additional electroweak interaction which has a much smaller coupling strength than the strong interaction. Furthermore, since the Z boson is a real particle, more energy per collision is needed to produce the $t\bar{t}Z$ system in a hard-scatter process. In Fig. 2.11 two possible Feynman diagrams are shown for $t\bar{t}Z$ production at the LHC. As for the case of $t\bar{t}$ production, also the decay of the $t\bar{t}Z$

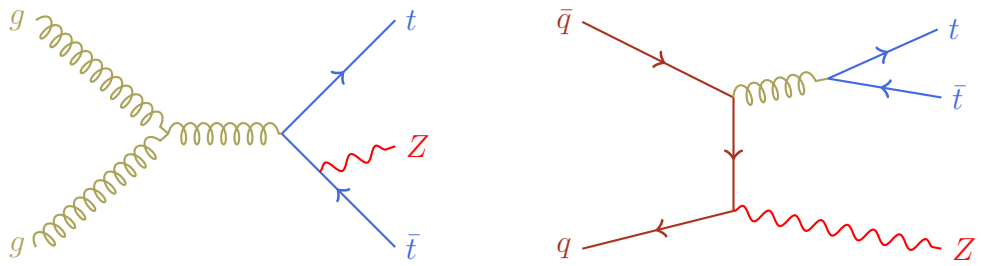


Fig. 2.11.: Representative Feynman diagrams for the strong top quark pair production in association with a Z boson.

process can be categorised which will be done in the following sections. Because of the additional Z boson more than just three decay channels are defined based on the number of leptons from the decays of the three vector bosons. Tab. 2.4 provides an overview of the dominant branching fractions for the decay of the Z boson. Since for the studies presented later on only three particular $t\bar{t}Z$ decay channels are dealt with, they will be looked at in greater detail. Single top quark production with a Z boson associated mentioned above, however, is explicitly not taken into account as desired processes.

$Z \to l^+l^-$	$Z \to \nu\bar{\nu}$	$Z \to q\bar{q}$
$(3.3658 \pm 0.0023)\,\%$	$(20.000 \pm 0.055)\,\%$	$(69.911 \pm 0.056)\,\%$

Tab. 2.4.: Branching fractions of the Z boson [51]. The leptonic branching fractions are averaged for all three lepton flavours.

2.4.1. The ttZ dilepton channel

The $t\bar{t}Z$ dilepton (2ℓ) decay channel features a fully-hadronic $t\bar{t}$ signature which has a fairly high branching fraction. Thus, as can be seen in Fig. 2.12, there is no neutrino present in the signature which allows for a full kinematic reconstruction. Furthermore, the Z boson decays into a pair of leptons with opposite electric charge and same flavour which can be experimentally well identified at the LHC. However, the small branching fraction of the Z boson to decay into a pair of leptons further limits the low statistics for an already rare process. In total the dilepton channel has a branching fraction of approximately 5 % (3 % without τ-leptons) which is the highest value for $t\bar{t}Z$ with a leptonic Z boson decay.

This channel has high background contributions from the production of a Z boson associated with multiple hadrons as well as from top quark pair production. Due to opposite-sign requirement on the lepton pair this channel is commonly referred to as 2ℓOS channel.

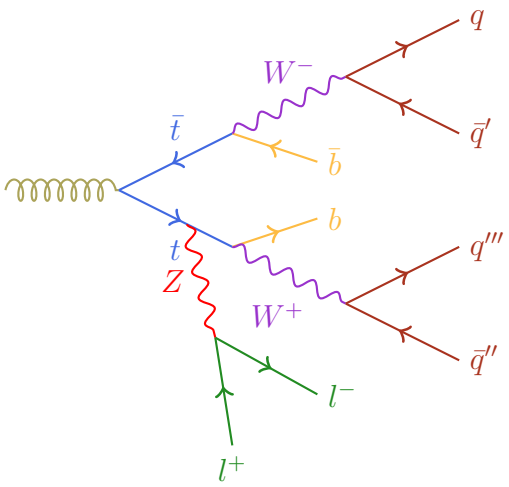


Fig. 2.12.: $t\bar{t}Z$ dilepton channel

2.4.2. The ttZ trilepton channel

The so-called trilepton (3ℓ) channel offers the best sensitivity at present due a good balance between a clean signal structure with a relatively low background contamination and a sufficiently high branching fraction with respect to the tetralepton channel. As it features a lepton+jets $t\bar{t}$ signature, a straightforward kinematic reconstruction of the signal signature can be performed. As for all decay channels described in this thesis, the Z boson decays to a pair of leptons with opposite sign and same flavour (OSSF). Due to the third lepton in the final state, the signal-to-background ratio is higher than for the dilepton channel. However, the branching fraction decreases to 4.5 % taking into account τ-leptons and 2 % if not. Due to the fact that the trilepton channel currently offers the greatest sensitivity, the first observation of $t\bar{t}Z$ at the LHC was made in the trilepton channel [35]. A representative leading-order Feynman diagram for the trilepton final state is shown in Fig. 2.13.

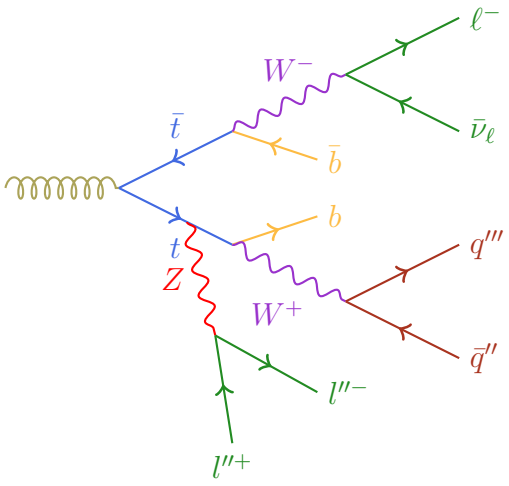


Fig. 2.13.: $t\bar{t}Z$ trilepton channel

2.4.3. The ttZ tetralepton channel

The tetralepton (4ℓ) channel, depicted in Fig. 2.14, offers by far the cleanest signature within the hadronic environment of the LHC. Both W bosons from the $t\bar{t}$ system decay leptonically giving rise to two final-state neutrinos. Thus the kinematic reconstruction of this signature is very involved. In total four leptons provide a good separation from backgrounds, although the very small branching fraction of $1\,\%$ limits signal statistics. Neglecting the contributions of τ-leptons the value of the branching fraction further drops down to $0.3\,\%$. However, previous measurements have exploited both the trilepton channel or the tetralepton channel [35–38, 40].

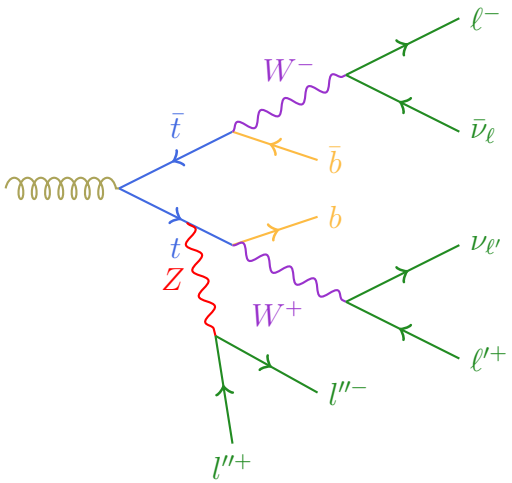


Fig. 2.14.: $t\bar{t}Z$ tetralepton channel

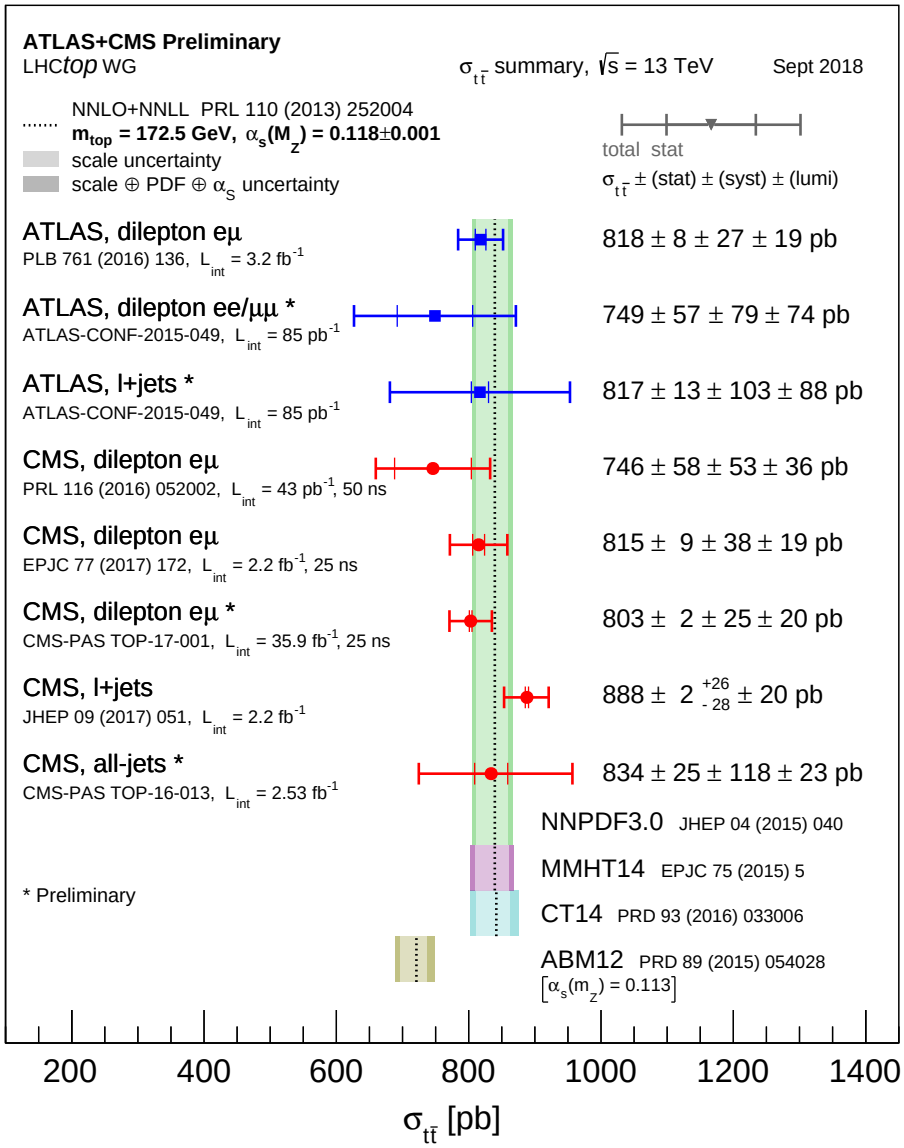


Fig. 2.15.: Summary of measurements of the top quark pair production cross section at the LHC experiments ATLAS and CMS for 13 TeV compared to the exact NNLO QCD calculation complemented with NNLL resummation [86]. The theory bands represent uncertainties due to renormalisation and factorisation scale, parton density functions and the strong coupling. The quoted results from measurement and theory calculation are based on an assumed top quark mass value of $m_{\mathrm{top}} = 172.5\,\mathrm{GeV}$ [88].

3. Experiment

3.1. Experiments at CERN

CERN[(1)] is the European Organisation of Nuclear Research founded in 1954 with its headquarters and accelerator complex located near Geneva, Switzerland. The largest particle physics laboratory in the world is operated at CERN, pursuing a variety of experiments in different areas of particle physics. CERN is one of the most important institutes of physics research in the world nowadays and was originally founded to provide a place for scientists of peaceful collaboration in Europe after many scientists had left Europe in the first half of the twentieth century. Since the foundation of CERN, the contributions of scientists from all over the world have led to remarkable progress in particle physics of which the most recent incident that received worldwide attention was the discovery of a new particle consistent with the Higgs boson predicted by the Standard Model [33, 34]. The main task of CERN is to probe the fundamental structure of the universe and has always been accompanied by revolutionary innovations helping day-to-day life. While trying to exceed the limit of knowledge, often entirely new working materials had to be developed in order to match with the technical challenges, thereby pushing the boundaries of technology at the time. This led to a number of technical advancements such as the invention of an early touch screen in 1970s [89], new possibilities of medical treatment, and the World Wide Web in 1989 [90]. Work at CERN must not have concerns with military requirements and, furthermore, all results have to be published or made generally available [91]. During more than 60 years of research at CERN, data have been measured by means of various experiments housed at accelerators of different types. The first accelerator, the Synchrocyclotron (SC), was already put into action only three years after CERN's foundation. Two years afterwards, the Proton Synchrotron (PS) was commissioned, which was the first circular accelerator at CERN. Since then, ring colliders have proven to be the dominant type of particles accelerators at CERN, most notably the Super Proton Synchrotron (SPS) where the W and Z bosons had been discovered in 1983 [92–94], and the Large Electron Position Collider (LEP) which served as the highest-energetic accelerator for leptons with a maximum centre-of-mass energy of $\sqrt{s} = 209\,\text{GeV}$ yet to date, enabling high-precision measurements of the electroweak interaction and an excellent confirmation of the Standard Model [95].

Fig. 3.1 shows a schematic overview of the CERN accelerator complex. Aside from the LHC accelerator complex which will be discussed in more detail in the following section, several other experiments are situated at CERN, for example the Antiproton Decelerator (AD) and its smaller equivalent ELENA[(2)], where antiprotons from fixed-target collisions of protons from the PS are decelerated in order to form bound states of antimatter. The antimatter is then investigated in various experiments located inside of the AD ring. In addition, the AMS[(3)] experiment, a particle detector mounted to the ISS[(4)] has a ground control centre on the CERN compound.

[(1)] The acronym follows the French name *Conseil européen pour la recherche nucléaire*.
[(2)] Extra Low ENergy Antiproton
[(3)] Alpha Magnetic Spectrometer
[(4)] International Space Station

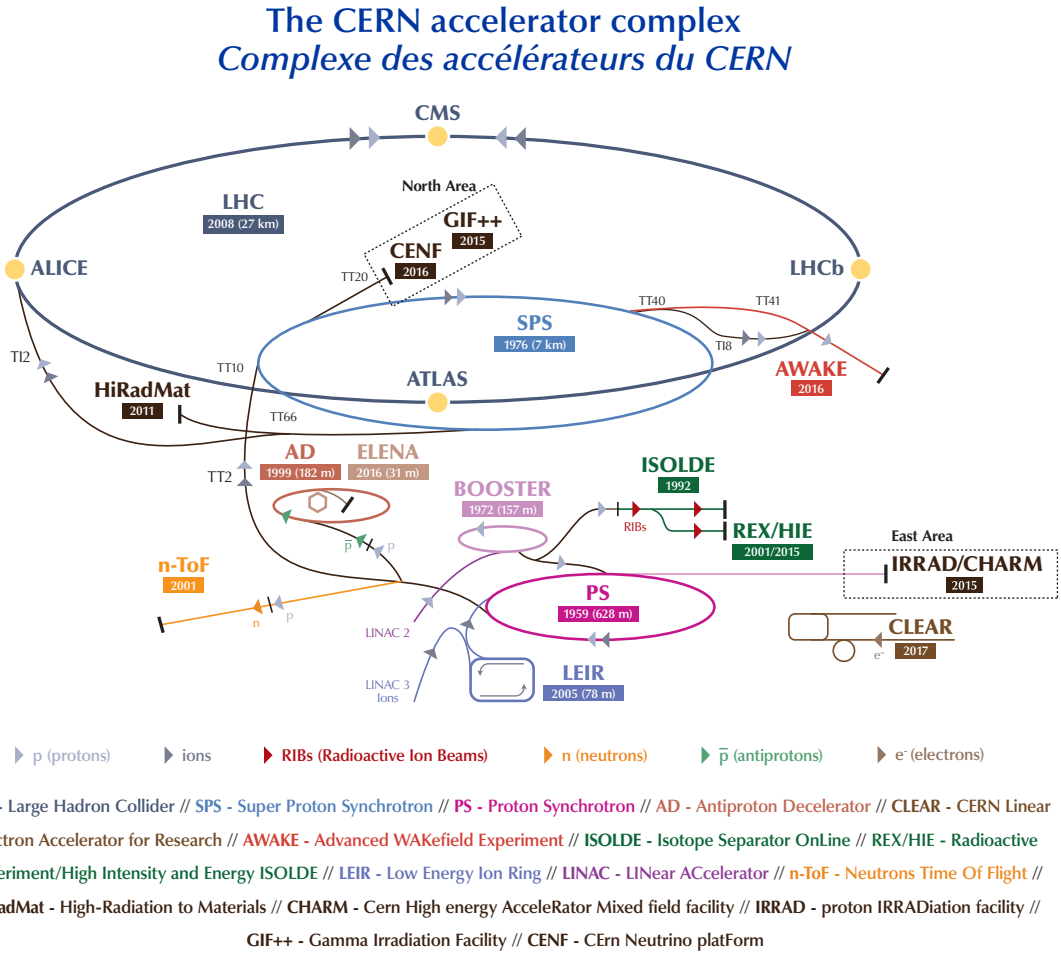


Fig. 3.1.: A schematic depiction of the CERN accelerator complex. The LHC ring is shown in dark blue with its four major experiments indicated by yellow dots. The pre-accelerator chains for protons and heavy ions are shown below including the LINAC 2, PS Booster, PS and SPS accelerators, as well as LINAC 3 and LEIR, respectively. Furthermore, the antimatter decelerator facilities as well as a number of smaller experiments are also included [96].

3.2. The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest particle accelerator and is located at CERN. It has been built to probe the Standard Model of Particle Physics and to search for new physics[(5)] phenomena in order to shed light on our understanding of physics at the high-energy scale. One of the significant limitations of a lepton collider such as LEP, where electrons and positrons were accelerated, is the high energy loss of the accelerated particles due to synchrotron radiation [50]. For this reason, a hadron collider was favoured instead of an upgraded electron-positron collider. The energy loss as a result of synchrotron radiation has to be compensated for

[(5)] New physics or physics beyond the Standard Model (BSM) comprises basically everything that cannot be described by the Standard Model, including massive neutrinos and thus neutrino oscillations.

in order to maintain a constant beam energy. The amount of energy a particle with electric charge Ze and mass m loses per turn, ΔE, is

$$\Delta E \propto \frac{(Ze)^2}{R} \cdot \gamma^4 \propto \frac{(Ze)^2}{R} \cdot \left(\frac{E}{m}\right)^4 \tag{3.1}$$

with beam energy E, radius R and the Lorentz factor $\gamma = E/m$. Due to the mass dependence to the power of four in the denominator of Eq. (3.1), protons ($m_{\text{proton}} = 939.272\,\text{MeV}$ [51]) lose a factor of about $10 \cdot 10^{13}$ less energy than electrons ($m_{\text{electron}} = 0.511\,\text{MeV}$ [51]) at the same energy. Thus, hadron colliders operated with (anti-)protons typically reach higher beam energies due to the reduced energy loss due to synchrotron radiation, and are therefore often called discovery machines since new hypothetical particles are often expected to have high masses [85]. However, a notable disadvantage of hadron colliders are the hadronic remnants in the events[(6)] that likely impede to perform measurements at very high precision. Operation of the LHC started in 2008, and the first period of data taking at the LHC started on 20th November 2009 with a centre-of-mass energy of 7 TeV and continued through the end of 2013. Within this period, widely called LHC Run 1, the centre-of-mass energy has been slightly increased to 8 TeV. Already on 30th November 2009, the LHC became the world's highest-energy particle accelerator beating the previous record of 1.96 TeV held by the Tevatron accelerator at Fermilab [97]. After two years of technical shutdown, the LHC came back online in April 2015 to perform a most successful data taking period until early December 2018 at a constant centre-of-mass energy of 13 TeV, referred to as Run 2. Being in technical shutdown since then, the LHC is foreseen to restart operation in 2021 for another few years, denoted as Run 3, before being upgraded to the high-luminosity LHC project which is supposed to be ready in 2026.

The LHC is circular machine composed of alternating straight and curved segments with a circumference of 27 km. It is located approximately 100 m underground in the same tunnel as its predecessor LEP. Despite being able to accelerate protons to energies of up to 7 TeV it is also designed to be filled with lead nuclei reaching a centre-of-mass energy of 2.76 TeV per nucleon. In order to reach those high values, it proved to be beneficial to increase the energies of the particles gradually by passing them through series of pre-accelerators. For this purpose, the existing infrastructure of former accelerators and experiments was integrated into the LHC acceleration chain.

Hydrogen atoms from a conventional gas bottle lose the electrons in an electric field and are then passed to the LINAC 2 linear accelerator where they are accelerated to an energy of up to 50 MeV. Then, the protons are injected into the Proton Synchrotron Booster (PSB) and the PS itself to be further accelerated to 1.4 GeV and 25 GeV, respectively. The proton beam is split into four parts of equal length before being injected into the Booster in which they are conducted in four separate, superimposed rings. In this way, within the same acceleration cycle a higher number of particles from the LINAC 2 can be filled into the PS and ultimately into the LHC. Furthermore, the decrease of the beam quality due to repulsive forces between the particles at such low energy intensities can be minimised. However, in the PS the beam parts are combined again. Besides, lead ions begin their journey at the LINAC 3 accelerator and speed up within the LEIR[(7)] before getting injected into the PS. From the PS the particles are, irrespective of their kind, transferred to the SPS where the energy of protons is further increased to 450 GeV. The SPS is the first accelerator in the chain being located underground as up to the PS all acceleration devices have been situated at the surface. In addition, a couple of fixed-target experiments are associated to the SPS ring, for

[(6)] An event is defined as the crossing of two particle beams which a hard scatter interactions occurs in.
[(7)] Low Energy Ion Ring, the equivalent for heavy ions to the PSB.

example the NA61/SHINE[8] and NA62 detectors or the COMPASS[9] experiment.

The particles are filled into the LHC where they circulate in two parallel vacuum beam pipes rotating clockwise and counter-clockwise. In order to accelerate the protons to their collision energy of $6.5\,\mathrm{TeV}$, eight radiofrequency (RF) cavities operating at a frequency of $408\,\mathrm{MHz}$ are employed per beam. Particles arriving at the RF cavities undergo either a local acceleration or deceleration from the oscillating electric field depending on whether the particles are travelling too fast or too slow, respectively. As a consequence, the beams at the LHC are actually not continuous but are comprised of more than 2000 small bunches each consisting of roughly $6 \cdot 10^{11}$ particles. The same technique as described above is used at all stages of the LHC complex to accelerate the particles which is the reason why the beam has its bunch structure from the very beginning on. In total 1232 dipole electromagnets, each $15\,\mathrm{m}$ in length, are employed to guide the particles around the LHC. They generate a peak magnetic field of $8.33\,\mathrm{T}$ which requires an electric current of $12\,500\,\mathrm{A}$. A multitude of multipole magnets is used to steer and adjust the beams. For instance, quadrupole magnets focus the particles in a beam closer together thereby increasing the number of colliding particles in a collision. Those magnets have to be installed in pairs since they focus only in one plane perpendicular to the beam direction. In fact, units of successive bending and focusing magnets are installed throughout the whole LHC ring, while the RF cavities are situated at one place only. Moreover, the particle beams are injected and ejected by a combination of horizontally deflecting fast-pulsed magnets and vertically-deflecting magnets, referred to as kicker and septum magnets, respectively. In order to control the particle beams at energies as high as those at the LHC, both the RF cavities and most of the magnets including the dipole and quadrupole elements have to be operated superconductively. The cables of the magnets are therefore made of niobium-titanium alloy so superconductivity below a temperature of $10\,\mathrm{K}$ can be reached. It is sufficient to operate the RF cavities at a temperature of down to $4.5\,\mathrm{K}$, though the magnets are even further cooled down to $1.9\,\mathrm{K}$ using gigantic amounts of liquid helium. In contrast, all the pre-accelerators are operated at normal room temperature.

The protons tend to drift apart due to the electromagnetic repulsive force they expose to each other inside the bunches. Several measures have been installed at the LHC to maintain focused beams thus protecting the fragile inner layers of the experiments from damage and preventing from unintentional interactions with the beam pipes. In addition to quadrupole magnets constantly acting on the beam, dedicated multipole magnets were installed in order to reduce the amplitude of the oscillation of the particle trajectories around the nominal which is induced by the repetitive focusing and defocusing of the quadrupole magnets in a specific plane perpendicular to the beam direction. Rather robust devices are the so-called beam collimators which are primarily installed at the LHC in front of a detector and consist of two tight metal jaws which leave a narrow slit in between. Hence, all particles at a beam radius greater than the width of the slit are scattered off thereby preventing the outer particles of the beam from passing the collimator device.

The two beams cross each other at various points of the LHC accelerator ring for two different purposes. On the one hand, the two beams cross without colliding in order to balance small spatial displacements arising due to the slightly different radii of the inner and outer beam line. On the other hand, with the help of dedicated multipole magnets the two beams are made to cross each other at designated points corresponding to the centres of the detectors. The total rate of proton-proton collisions, R, is given by

$$R = \sigma \cdot L \quad . \tag{3.2}$$

where the instantaneous luminosity L describes the number of particles crossing a unit area per

[8] SPS Heavy Ion and Neutrino Experiment
[9] Common Muon and Proton Apparatus for Structure and Spectroscopy

unit of time and σ denotes the interaction cross section. For particle colliders, Eq. (3.2) reads

$$L = \frac{nN_1N_2}{4\pi\sigma_x\sigma_y} \cdot f \quad , \tag{3.3}$$

with f denoting the revolution frequency of n colliding proton bunches. The LHC is designed to have bunch collisions at a rate of $40\,\mathrm{MHz}$ corresponding to a spacing of $25\,\mathrm{ns}$ in time between adjacent bunches. The number of protons within the colliding bunches is given by N_1 and N_2, respectively. The size of the beams at the collision point, σ_x and σ_y, are determined by so-called van-der-Meer scans [98].The total interaction cross section σ serves as a measure for the probability of an interaction to happen and has been originally interpreted as an effective area transverse to the relative motion of two particles within which they must meet in order to scatter. It consists of the sum of the inclusive cross sections σ_i of all processes that take place at the collision. The total number of events of a certain process i is then

$$N_i = \sigma_i \cdot \mathcal{L} \tag{3.4}$$

where the time-integrated luminosity

$$\mathcal{L} = \int L \,\mathrm{d}t \tag{3.5}$$

is a reflective of the amount of data collected within a certain time interval.

Additional corrections can be introduced to refine the approximate expression for the instantaneous luminosity given in Eq. (3.3). The beam quality is accounted for by the so-called normalised emittance ε considering particles in the beams with transverse motion. In addition, an amplitude function β^* quantifies the squeezing of the beam optics at the interaction point, where the star indicates the value at the collision point. As the beams collide not exactly head-on but at an angle $\theta_c \neq 0$, a geometric reduction factor

$$F = \left[\sqrt{1 + \left(\frac{\theta_c\sigma_z}{2\sigma^*}\right)^2}\right]^{-1} \tag{3.6}$$

is used to correct the luminosity with respect to head-on collisions. Here, σ_z denotes the average longitudinal spread of the bunches and σ^* represents the transverse beam size at the interaction point. With these corrections, Eq. (3.3) for the instantaneous luminosity at particle colliders can be rewritten as

$$L = \gamma \cdot \frac{nN_1N_2}{4\pi\varepsilon\sqrt{\beta_x^*\beta_y^*}} \cdot f \cdot F \quad , \tag{3.7}$$

where the beam width is expressed in terms of the emittance and amplitude function,

$$\sigma_{x,y} = \sqrt{\varepsilon \cdot \beta_{x,y}} \quad . \tag{3.8}$$

Moreover, the relativistic proton velocities are accounted for with the relativistic Lorentz factor γ.

In total seven experiments are located at the LHC whereof the biggest is the ATLAS[(10)] detector which will be explained in detail in the next section. It is a general-purpose detector designed to cover a broad field of topics in particle physics. The CMS[(11)] detector is located at the LHC ring diametrically opposite to ATLAS. Both experiments have similar physics goals allowing complementarity and share their areas of research and benefit from cross-checking each other. The

[(10)] A Toroidal LHC ApparatuS
[(11)] Compact Muon Solenoid

largest special-purpose experiment of the LHC involves the ALICE[12] detector. It was specifically designed to investigate the state of matter dominant shortly after the big bang, referred to as quark-gluon plasma, which is produced in heavy-ion collisions. The fourth major LHC experiment is LHCb[13] which, in contrast to the ATLAS, CMS and ALICE detectors, features an asymmetric setup. The LHCb detector was built to explore the origin of the matter-antimatter imbalance by studying the B-hadron produced in proton-proton collisions. Next to CMS the TOTEM[14] detector measures remnants from deep-inelastic particle collisions scattered along the beam pipe and, in addition, performs studies on the measurement of the proton size. The smallest detector at the LHC is LHCf[15] which was designed to study particles generated almost directly in line with the colliding proton beams but also to look for cosmic rays. It therefore consists of two detectors on either side of the interaction point of ATLAS[16].

3.3. The ATLAS detector

The ATLAS experiment [101, 102] is multi-purpose particle detector housed underground in the experimental cavern at LHC point 1. It was designed to cover a wide range of research topics in particle physics such as the top quark, the Higgs boson and the Standard Model in general, but also physics beyond such as SUSY or Dark Matter searches most comprehensively. Thus it supposed to provide high-grade particle detection within the high-luminosity environment of the LHC peaking at a value of the instantaneous luminosity equal to $10^{-34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. For this purpose, the information from several distinct layers, each featuring various subdetectors with particular technologies, is combined. In summary, the overall concept of the ATLAS detector is to prove identification and reconstruction of all fundamental particles of the Standard Model.

The ATLAS detector is situated around the LHC beam pipe with the nominal interaction point at the geometric centre of the experiment. As depicted in Fig. 3.2, it has a cylindrical shape and features a forward-backward symmetry. The cylindrical portion is divided into a central part, commonly referred to as barrel, with most of the components concentrically aligned around the beam pipe, and two disk-shaped outer parts at both ends of the cylinder, therefore called end-caps. ATLAS consists of three main detector categories which are installed in concentric layers around the interaction point: the inner detector, the calorimeters, and the muon spectrometer. The momenta of charged particles are measured with very high precision by particular tracking detectors. The ATLAS detector features an excellent electromagnetic calorimeter for electron and photon identification as well as a hadronic calorimeter for the measurement of strongly interacting particles producing jets. Special emphasis is placed on the detection and measurement of muons by a dedicated detection system, and a reasonable identification power of τ-leptons and heavy-flavour particles is provided. In addition, a system of superconducting electromagnets providing magnetic fields is employed for the measurement of the particles' momenta. The vast amount data recorded by the various detector components is handled by a custom-designed trigger and data acquisition system. Particles can be identified by reconstructing their way through the detector using the combined information of all subdetectors. With an overall length of about $46\,\mathrm{m}$ and a diameter of $25\,\mathrm{m}$ ATLAS is by far the largest detector of all LHC experiments. However, with a weight of roughly $7000\,\mathrm{t}$ it is a fairly low-weight construction with respect to its dimensions, as for example the CMS experiment is almost twice as heavy but significantly smaller in size. The

[12] A Large Ion Collider Experiment
[13] Large Hadron Collider beauty
[14] TOTal Elastic and diffractive cross section Measurement
[15] Large Hadron Collider forward
[16] In addition to the stated references, this section is based on Ref. [49, 50, 98–100].

ATLAS detector would be able to swim on the ocean if it was made watertight.

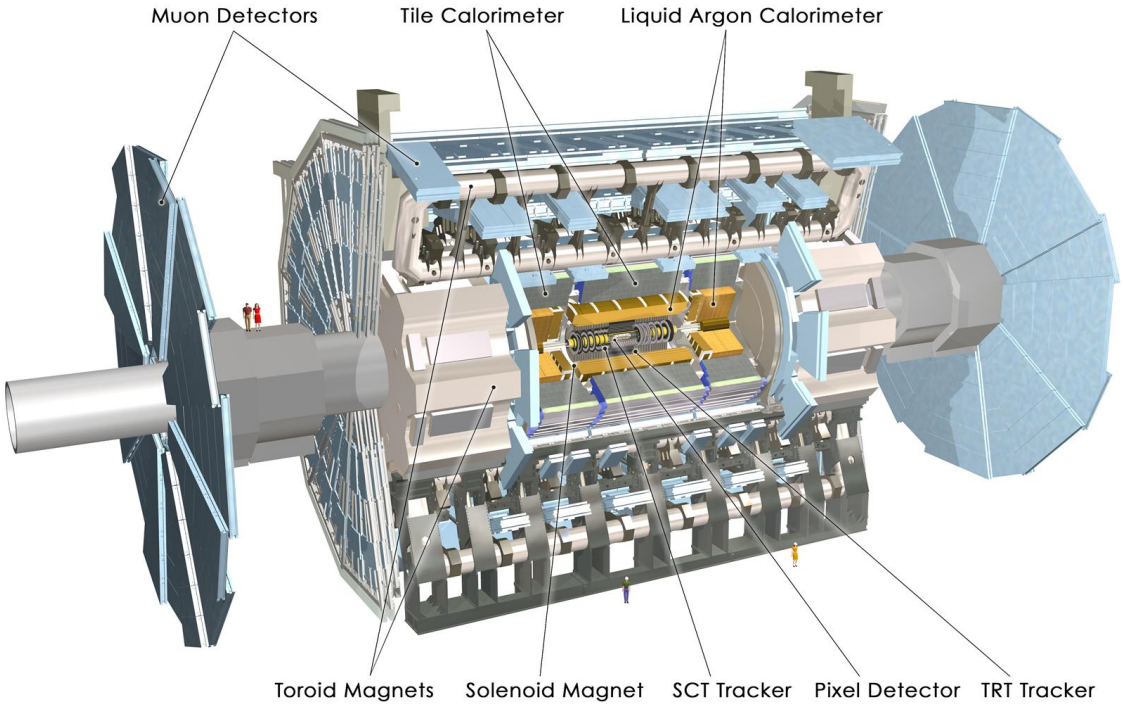


Fig. 3.2.: Computer-generated image of the ATLAS detector showing the three layers of main detector components. From the nominal interaction point at the geometric centre going outwards, the inner detector, the calorimeters and the muons system are concentrically aligned. The big toroidal coils of the magnet system are clearly visible [103].

3.3.1. The coordinate system

The ATLAS experiment employs a right-handed coordinate system which has its origin at the nominal interaction point in the centre of the detector. The z-axis is defined by the beam direction with the positive side pointing eastwards and the negative side pointing westwards. In the transverse plane perpendicular to the beam axis (x-y), the positive x-axis points towards the centre of the LHC ring, while the positive y-axis points upwards to the surface. Due to the cylindrical shape of the detector, a different coordinate system than Cartesian coordinates are conventionally used to describe the positions of objects. The azimuthal angle ϕ is measured around the beam line in the transverse plane with $\phi = 0$ corresponding to the positive x-axis. As viewed from the positive z-axis, ϕ describes the angle between the positive x-axis ($\phi = 0$) and another point in the x-y-plane, increasing and decreasing in the counter-clockwise and clockwise direction, respectively, such that is covers the range $\phi \in [-\pi, +\pi]$ in radians. Therefore, the positive y-axis corresponds to $\phi = +\pi/2$, and the negative y-axis features a value of the azimuthal angle equal to $-\pi/2$. The polar angle θ defines the separation of an object from the beam axis with the origin defined on the positive z-axis. For example, with a clockwise rotation around the x-axis, θ increases until it reaches the x-y-plane ($\theta = \pi/2$) from where it continues to increase towards the negative z axis ($\theta \in [0, \pi]$). However, the polar angle θ is not invariant under Lorentz boosts which is the reason why the rapidity y is instead used to describe the movement of particles in z-direction. The

rapidity is defined as

$$y = \frac{1}{2} \ln \left[\frac{E + p_z}{E - p_z} \right] \tag{3.9}$$

relative to the beam axis and is Lorentz-invariant under boosts along the longitudinal axis for both massive and massless objects. In the relativistic limit where the particles are travelling close to the speed of light, or equivalently in the approximation of negligible particle masses ($p \gg m$), rapidity becomes equal to the pseudorapidity η, defined as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad . \tag{3.10}$$

As the value of η is exactly determined by the polar angle θ, the pseudorapidity is predominantly used to describe the angle of a particle relative to the beam axis. An object in the transverse plane ($\theta = \pi/2$) features $\eta = 0$, whereas coordinates parallel to the beam at either $\theta = 0$ and $\theta = \pi$ correspond to $\eta = \pm\infty$. Due to the symmetry of the detector, coordinates are typically stated in terms of the modulus, $|\eta|$. Coordinates with high values of $|\eta|$ are said to be in the so-called forward region. Differences in angular space (η-ϕ-plane) between two objects are defined by

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.11}$$

where $\Delta\eta$ and $\Delta\phi$ is the difference in rapidity and the angular separation in the transverse plane, respectively [51, 101].

The LHC collides protons which are no fundamental particles but are rather comprised of so-called partons (quarks and gluons), each carrying a certain fraction of the proton's momentum in z-direction. In a collision of two protons, the hard-scatter interactions of interest take place between one of the partonic constituents from each respective proton and not between the protons themselves. The momentum fraction of a parton is not a priori known but can only be roughly estimated with the help of parton distribution functions, as described in Sec. 2.2. However, since the colliding particles propagate in z-direction at almost speed of light, the total momentum in the transverse plane is roughly zero before the collision. It is therefore useful to make use of the projections of a given object's total momentum into the transverse plane only:

$$p_\mathrm{T} = \sqrt{p_x^2 + p_y^2} \tag{3.12a}$$

$$E_\mathrm{T} = \sqrt{m^2 + p_\mathrm{T}^2} \tag{3.12b}$$

If a physics observable depends on the momentum, it will be restricted to the transverse plane, too, as shown above for the case of the transverse energy.

3.3.2. The magnet system

The ATLAS magnet system, shown schematically in Fig. 3.3, consists of two parts: the central solenoid magnet (CS) enclosing the inner detector in the barrel, and three toroid magnets within the muon spectrometer which are further subdivided into a barrel toroid (BT) and two end-cap toroids (ECT) [101, 102]. In total, the magnet system has a length of 26 m and a diameter of 22 m.Made of aluminium-stabilised niobium-titanium alloy, it is cryogenically cooled down to 4.5 K and has a stored energy amount of 1.6 GJ during operation. With more than 8000 m^3 of enclosed volume and a total weight of approximately 1300 t it is the largest superconducting magnet in the world [104].

With a length of 5.3 m and a bore diameter of 2.44 m, the CS is designed to provide a homoge-

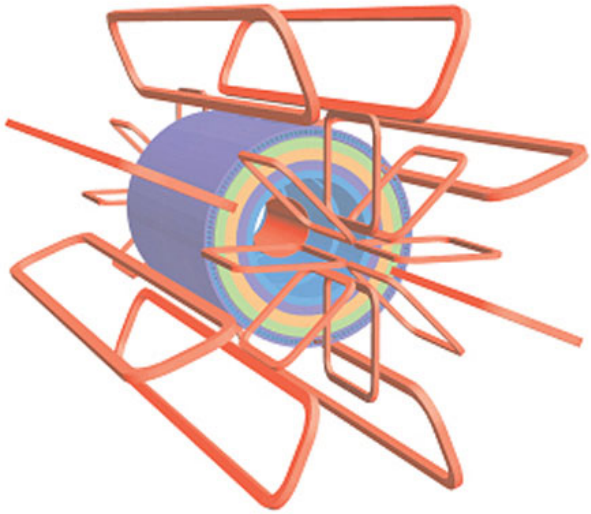


Fig. 3.3.: Schematic view of the ATLAS magnet system coloured in red [105]. The barrel solenoid surrounds the inner detector inside of the calorimeters which are depicted in blue, purple, green and orange. The toroid magnet comprises eight big coils in the barrel regions and eight smaller coils in each end-cap being symmetrically arranged in circles.

neous magnetic field of $2\,\mathrm{T}$ in the central tracking volume of the inner detector with the magnetic field lines parallel to the beam line. The central solenoid it is capable of generating a peak field of $2.6\,\mathrm{T}$. Despite being only $4.5\,\mathrm{cm}$ thick it has an overall weight of about $5\,\mathrm{t}$ since an iron yoke is employed in order to contain the stray field and to amplify the magnetic field strength [106].

The three toroid magnets are comprised of eight large air-core coils each which radially aligned around the calorimeters with at equal distances to each other. Extending over a length of $25.3\,\mathrm{m}$ and with an outer diameter of $20.1\,\mathrm{m}$ and inner diameter of $9.4\,\mathrm{m}$ the barrel toroid is by far the largest of all of the magnetic subsystems in ATLAS. The end-cap toroids are much more compact in contrast, each with a length of $5\,\mathrm{m}$ and outer and inner diameters of $10.7\,\mathrm{m}$ and $1.65\,\mathrm{m}$, respectively. The toroid magnets deliver a magnetic field perpendicular to the beam line of approximately $0.5\,\mathrm{T}$ in the barrel region and $1.0\,\mathrm{T}$ in the end-caps for the muon detectors, yet much higher peak values of $3.9\,\mathrm{T}$ and $4.1\,\mathrm{T}$ can be reached by the BT and the ECT, respectively. In contrast to the barrel solenoid, for the ATLAS toroid magnets no iron yoke is employed. A single toroid coil has weight of about $100\,\mathrm{t}$ in the barrel region and $30\,\mathrm{t}$ in the end-cap parts due to the large amounts of so-called cold mass needed to sustain the cryogenic cooling. The end-cap magnets are inserted into the barrel toroid at each end such that they close with the barrel outer boundary which can be seen in Figs. 3.2 and 3.3. Furthermore, the ECT coils are rotated with respect to the BT coils by $22.5°$ in order to provide a radial overlap of the magnetic fields thereby optimising the bending power in the interface region between barrel and end-caps. A slight decrease in bending power has to be taken into account in the interface region corresponding to a pseudorapidity range of $1.3 < |\eta| < 1.6$ [107, 108].

According to Lorentz's law, a charge q moving with a velocity $\mathbf{v}$ within a magnetic field with flux density $\mathbf{B}$ experiences a force

$$\mathbf{F}_\mathrm{L} = q\,(\mathbf{v} \times \mathbf{B}) \tag{3.13}$$

which deflects the charge's path perpendicular to both the velocity vector $\mathbf{v}$ and the magnetic field $\mathbf{B}$ [109]. Hence, the trajectories of charged particles appear as curved paths inside the ATLAS

detector. This phenomenon is exploited in the measurement of the momentum of charged particles in the inner detector and the muon spectrometer as will be described in Secs. 3.3.3 and 3.3.5, respectively. From the direction of the curvature with respect to the straight path the sign of the electric charge of the reconstructed particle associated to the track can be deduced. The angular deflection of a particle in a magnetic field is proportional to its momentum. Employing the relation

$$p[\mathrm{GeV}] = 0.3\, B[\mathrm{T}]\, R[\mathrm{m}] \tag{3.14}$$

between the momentum p, the magnetic field strength B and the radius R of the deflected trajectory, the latter provides a measure for a particle's momentum [50]. The relative resolution of the momentum measurement was found to be directly proportional to the transverse momentum of a particle and to be anti-proportional to the magnetic field [110]. The use of superconducting magnets to provide sufficiently high magnetic field strengths is necessary in order to attain high-precision momentum measurements in the LHC environment.

3.3.3. The inner detector

The inner detector (ID) is the innermost detector component of ATLAS assembled at a distance to the beam of only $3\,\mathrm{cm}$. The ATLAS ID is shown in Fig. 3.4 enabling a look inside of the inner detector in both the barrel and the end-cap regions. From the beam line outwards it comprises three independent parts: the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT). In the barrel region the subdetector modules are arranged as concentric cylinders enclosing the beam line, whereas in the end-caps the detector elements are mounted on disks perpendicular to the beam axis. Separate images of the ID barrel and end-cap parts are also depicted by Figs. 3.5a and 3.5b, respectively. With an overall length of $6.2\,\mathrm{m}$ and an outer radius of $155\,\mathrm{cm}$ the ATLAS inner detector covers a region in space of $|\eta| < 2.5$. Precision tracking elements are installed in the ID centre within a radius of $56\,\mathrm{cm}$, followed by the continuous tracking portion.

The ATLAS inner detector was designed to provide a precise vertex and particle identification as well as high-precision momentum measurements for charged particles. In order to meet the high demands imposed by the benchmark physics processes and precision measurements within the large track density at the LHC, a hybrid system of semiconducting layers and gas-filled tubes is employed which is immersed in the $2\,\mathrm{T}$ magnetic field generated by the central solenoid magnet as described previously in Sec. 3.3.2. By convention, a track is defined as the reconstructed trajectory based on the hits of the charged particle within several layers of the detector. Thus, from the curvature of a track the momentum and charge of the corresponding particle can be inferred. In addition, primary[17] and secondary[18] vertices, positions corresponding to intersections of two or more reconstructed tracks, are determined through the retracement of tracks. This offers the possibility to separate particles originating from different parton collisions, and additionally supports the identification of charged particles composed of heavy-flavour quarks as well as of τ-leptons [112, 113].

[17] Vertices are identified from an extrapolation of the measured tracks back to their origin. The point of the parton-parton interaction of interest of a proton beam crossing is referred to as primary vertex (PV), defined as the interaction with the highest sum of squared transverse momenta of the associated tracks.

[18] The crossings of tracks of either particles from secondary decays of particles associated to the PV, for example B-hadrons, or from particles from parton-parton collisions not associated to the PV are referred to as secondary vertices (SV).

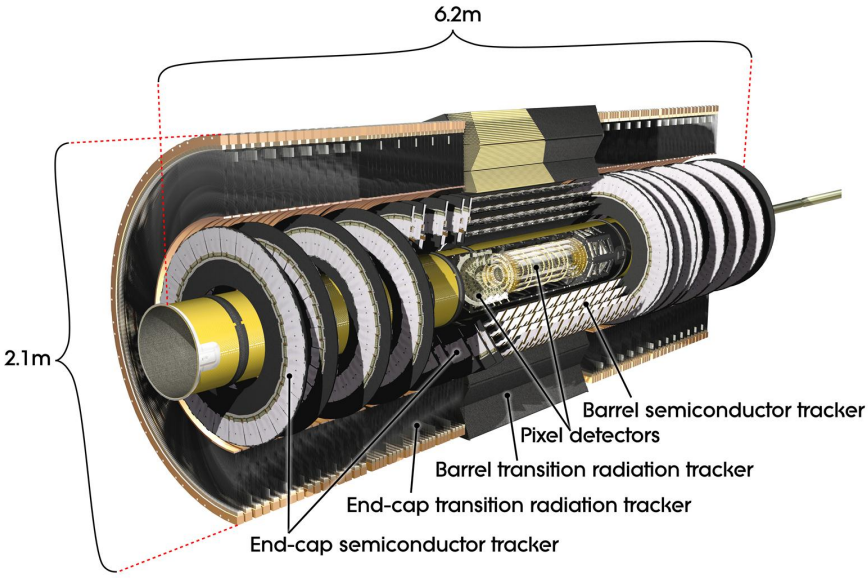
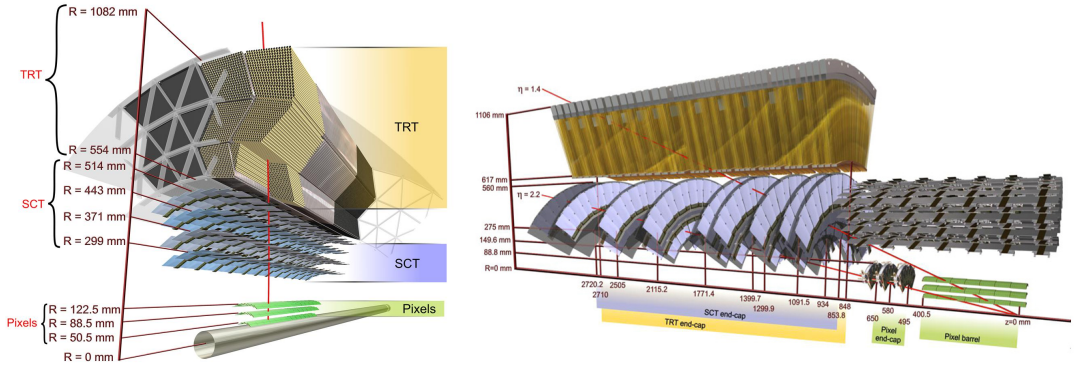


Fig. 3.4.: Computer-generated cut-away view of the ATLAS inner detector. The concentric cylinders in the barrel region and the end-cap disks are partly removed so that the innermost parts can be seen [111].

3.3.3.1. The pixel detector

Both the precise reconstruction of vertices and the accurate momentum measurement very close to the interaction point are the tasks of the ATLAS pixel detector. In the barrel region, the pixel detector consists of four layers of silicon pixel detectors concentrically aligned around the beam axis at a distance of approximately $3\,\mathrm{cm}$, $5\,\mathrm{cm}$, $9\,\mathrm{cm}$ and $13\,\mathrm{cm}$, respectively, which results in a coverage of $|\eta| < 1.7$. Each layer is segmented into pixels with a minimum size in R-$\phi \times z$ of $50 \times 400\,\mu\mathrm{m}^2$, where R-ϕ corresponds to the transverse plane and z to the beam direction. In order to cover the volume in the pseudorapidity range of $1.7 < |\eta| < 2.5$, five disks of silicon pixels are mounted perpendicularly to the beam line at distances between $11\,\mathrm{cm}$ and $20\,\mathrm{cm}$ in the end-caps. The granularity of approximately $10\,\mu\mathrm{m}$ in the R-ϕ plane and $115\,\mu\mathrm{m}$ in the z-direction is equal for both the barrel and end-cap regions [114].

In the original design of the ATLAS ID only the three outer barrel pixel layers were in place which proved to be sufficient for the centre-of-mass energies of $7\,\mathrm{TeV}$ and $8\,\mathrm{TeV}$ in the first years of LHC operation. However, before the centre-of-mass energy was increased to $13\,\mathrm{TeV}$, an additional pixel layer was inserted between the innermost layer of the barrel and the beam pipe during the LHC shutdown lasting from February 2013 to April 2015. With the addition of the new pixel layer the radius of the beam pipe was decreased by a few centimetres. The new pixel layer, referred to as insertable B-layer (IBL), provides the necessary increase in the resolution of both track and vertex measurement given the higher charged-track multiplicity caused by the elevated centre-of-mass energy. The pixels of the IBL have a size of 50×250 in R-$\phi \times z$, and are thus of a higher granularity than for the other three layers. With a full coverage of $|\eta| < 2.5$ the IBL provides a spatial precision of $8\,\mu\mathrm{m}$ and $40\,\mu\mathrm{m}$ in the transverse plane and in beam direction, respectively [115, 116].



(a) (b)

Fig. 3.5.: Computer-generated schematic views of the ATLAS inner detector. Fig. (a) shows the components of the barrel region, while Fig. (b) depicts a cut-away image of one of the two end-cap modules. The gas-filled tubes are aligned parallel to the beam line in the barrel and perpendicular to the beam axis in the end-caps [111].

3.3.3.2. The semiconductor tracker

The ATLAS semiconductor tracker, which completely encompasses the pixel detector, consists of eight concentric semiconducting silicon strip layers in the barrel located at radial separations of between $30\,\mathrm{cm}$ and $51\,\mathrm{cm}$ running parallel to the beam line. In either end-cap nine disks with radially aligned semiconducting silicon strips are mounted at distances between $85\,\mathrm{cm}$ and $273\,\mathrm{cm}$ from the interaction point. In addition, in both the barrel and the end-cap regions so-called stereo strips rotated by an angle of $40\,\mathrm{mrad}$ are employed in order to measure both the longitudinal and transverse coordinates in each region of the detector. The barrel section of the SCT covers $|\eta| < 1.4$ in pseudorapidity range, and the end-cap modules cover the remaining $1.4 < |\eta| < 2.5$. The intrinsic resolution per SCT module is $17\,\mu\mathrm{m}$ in R-ϕ-plane for both barrel and end-caps and $580\,\mu\mathrm{m}$ in z-direction in the barrel and in $580\,\mu\mathrm{m}$ in R-direction for the end-caps. The SCT modules are well visible in Fig. 3.4 and Fig. 3.5.

The pixel detector and the semiconductor tracker are both made of silicon which makes it prone to irradiation damage over the course of the operation time. The detector modules are therefore cooled below $-5\,^{\circ}\mathrm{C}$ in order to suppress electronic noise from radiation damage [112, 113].

3.3.3.3. The transition radiation tracker

In contrast to the pixel detector or the SCT, the ATLAS transition radiation tracker uses thin-walled proportional drift-tubes, referred to as straws or straw tubes. The straws are filled with xenon gas and are equipped with a thin gold-plated tungsten wire measuring $30\,\mu\mathrm{m}$ in diameter, which features a fast response due to the good mechanical and electrical properties. The walls are only $4\,\mathrm{mm}$ thick and kept at a voltage of $1.5\,\mathrm{kV}$, while the wire is held at ground potential. When a charged particle traverses a straw thereby ionising the gas inside the tube, the free electron inside the tube will drift towards the wire being amplified and read out.

As depicted in Fig. 3.5, the straw tubes are aligned parallel to the beam axis in the barrel region and have a length of $144\,\mathrm{cm}$. The wires inside of those tubes are divided into half at around $\eta = 0$ to form two separate detectors. In the end-caps, the straws are shortened to $37\,\mathrm{cm}$ and are radially

arranged on 18 circular layers. The TRT provides information about the particles' tracks only in transverse plane with a precision of 130 µm per straw. However, due to the large number of hits a particle leaves on average, the reconstruction of continuous tracks is enabled within pseudorapidity ranges of $|\eta| < 2.0$.

Along with continuous tracking, the main purpose of the TRT is to discriminate between electrons and heavier charged particles such as pions or kaons. The space between the straw tubes are filled with polymeres in order to create so-called transition radiation which is emitted by highly relativistic charged particles traversing a material boundary with different refractive indices. As this effect depends on the relativistic factor $\gamma = E/m$, it is strongest for electrons and less pronounced for particles with a higher mass. The photons from transition radiation are absorbed by the gas atoms inside the straw tubes therefore amplifying the measured signal [112, 113].

3.3.4. The calorimeter system

In order to measure the energy of particles, ATLAS employs a system comprised of two separate subsystems, which are distinguished based on the differing type of dominant signatures being targeted. The electromagnetic (EM) calorimeter was specifically designed to measure the energy of electromagnetically interacting particles such as photons and electrons.In contrast, the hadronic calorimeters were constructed to measure the energy of particles by means of the strong force, for example of protons and neutrons. Both subdetectors are so-called sampling calorimeters which means that they are not homogeneous blocks of a single material. Instead, the calorimeters consist of alternating layers of high-density active material, where particles interact with the matter and therefore lose energy, and read-out layers that measure the energy lost by the particles in the adjacent active layers. The ATLAS calorimeter system covers a pseudorapidity range of $|\eta| < 4.9$ employing different techniques which are suited to the differing requirements of the physics processes of interest and the different radiation environments depending on the respective regions of space. The capability of covering such a large $|\eta|$-range is particularly important for the determination of the so-called missing transverse energy which will be described in more detail in Sec. 5.7. A crucial function of the calorimeters is to prevent all particles except for muons to reach the muon spectrometer which fully encloses the calorimeters.

An overview of the various ATLAS calorimeters is given in Fig. 3.6. The innermost part is the electromagnetic calorimeter which entirely surrounds the inner detector and the central solenoid magnet. It is predominantly located in the barrel but extends somewhat into the end-cap regions. In contrast, the end-caps host most parts of the hadronic end-cap and forward calorimeters. The outermost subdetector is the hadronic tile calorimeter which extends over the whole length of the calorimeter system [101, 102].

3.3.4.1. The electromagnetic calorimeter

For the ATLAS electromagnetic calorimeter liquid argon (LAr) is employed as the active medium and lead serves as the absorber. A full coverage of the azimuthal angle and for a pseudorapidity of $|\eta| < 3.2$ is provided due to the accordion-shaped kapton electrodes and lead plates. Particles originating from the inner detector will produce cascades of secondary charged particles and photons, referred to as electromagnetic showers, due to electromagnetic interactions within the absorber. The shower constituents then ionise the liquid argon atoms in the active layers. The energy of an incoming particle is proportional to the number of charged shower constituents and can thus be measured by counting the number of ionised atoms of the active material. Each time a particle in the shower converts or radiates a photon, it loses a certain amount of its energy. This process

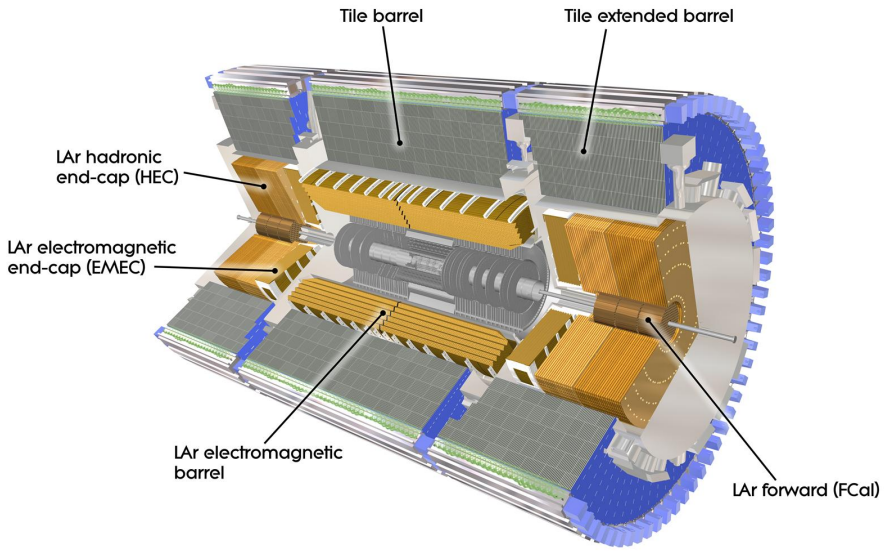


Fig. 3.6.: Computer-generated cut-away image of the ATLAS calorimeter system. The electromagnetic calorimeter fully encloses the inner detector with the central solenoid magnet and is itself encompassed by the hadronic calorimeter. The end-cap regions consist of the forward calorimeters which are surrounded by the electromagnetic and hadronic end-cap calorimeters. The outermost parts are occupied by the hadronic tile calorimeter in both barrel and end-cap regions [117].

proceeds until the particles are stopped. The exact shape of a shower and its position inside the calorimeter is determined from the hits in the individual calorimeter cells.

The ATLAS EM calorimeter is divided into a barrel portion capable of covering regions of $|\eta| < 1.475$, and two end-cap modules that cover a pseudorapidity of $1.375 < |\eta| < 3.2$. The overlap between barrel and end-caps has been deliberately introduced in order to ensure a continuous coverage in $|\eta|$. The barrel EM calorimeter itself consists of two identical halves separated by a small gap of $4\,\mathrm{mm}$ in the middle ($z = 0$). The end-cap calorimeters can be further subdivided into two coaxial wheels, with the inner wheel covering a region in space of $2.5 < |\eta| < 3.2$ and the outer wheel covering a pseudorapidity range of $1.375 < |\eta| < 2.5$, respectively. Over the pseudorapidity range of $|\eta| < 1.8$, the so-called presampler, an instrumented layer of liquid argon, precedes the actual calorimeter correcting for energy losses that may appear prior to the EM calorimeter.

In total the electromagnetic calorimeter has a thickness of at least 22 radiation lengths[(19)] in the barrel and of at least 24 radiation length in the end-caps. The barrel EM calorimeters is embedded in the vacuum of the barrel cryostat which surrounds the inner detector and therefore shares its vacuum vessel with the central solenoid. The two end-cap cryostats, however, are separate and contain the electromagnetic and hadronic end-cap calorimeters as well as the integrated forward calorimeters [118].

[(19)] The radiation length is the characteristic length for the energy loss of high energy particles within matter electromagnetically interacting with it.

3.3.4.2. The hadronic calorimeter system

The ATLAS hadronic calorimeter is comprised of three different components: the tile calorimeter, the hadronic end-cap calorimeters (HEC) and the forward calorimeters (FCal). All parts together cover a pseudorapidity range of $|\eta| < 4.9$. Unlike in the EM calorimeter, impinging particles deposit their energy mostly due to inelastic nuclear interactions with the absorber material, mediated by the strong force. The particle cascades produced in the hadronic calorimeters are therefore referred to as hadronic showers [50, 101, 102].

The tile calorimeter

As depicted in Fig. 3.6, the main part of the ATLAS hadronic calorimeter system is the tile calorimeter which completely surrounds all other parts of the calorimeter system. The absorber medium consists of steel whereas scintillating plastic tiles are employed as active material. Attached to the tiles are wavelength-shifting fibres converting the light from ionisation into visible light which is then able to be detected by photomultiplier tubes. The energy of incoming particles is proportional to the amount of collected light.

The tile calorimeter is composed of a barrel portion and, instead of an end-cap design, two extended barrel regions featuring the barrel architecture. With the barrel cylinder a pseudorapidity range of $|\eta| < 1$ is covered, overlapping with the coverage of the extended barrels of $0.8 < |\eta| < 1.7$ in order to prevent gaps in $|\eta|$. The tile calorimeter extends radially from an inner radius of $2.28\,\mathrm{m}$ to an outer radius of $4.25\,\mathrm{m}$. It is further segmented perpendicular to the beam axis into three layers, approximately 1.5, 4.1 and 1.8 interaction lengths[(20)] thick in the barrel and 1.5, 2.6 and 3.3 interaction lengths thick in the extended barrels. At the outer rim of the instrumented region of the tile calorimeter, the detector has a total thickness of 9.7 interaction lengths in the barrel and of 10 interaction lengths in the end-caps [119].

The liquid-argon hadronic end-cap calorimeters

The hadronic end-cap calorimeter is comprised of two independent wheels with an outer radius of $2.03\,\mathrm{m}$ in each end-cap. It is placed directly next to the EM calorimeter end-caps with which it shares the LAr cryostats. In contrast to the electromagnetic calorimeter, copper is used as the absorber material in the HEC while liquid argon serves as the active medium. As the hadronic end-cap calorimeter covers a pseudorapidity range of $1.5 < |\eta| < 3.2$, it overlaps with the tile and forward calorimeters [118].

The liquid-argon forward calorimeters

The liquid argon forward calorimeter encompasses both the beam pipe outside of the inner detector with its front face at roughly $4.7\,\mathrm{m}$ distance from the interaction point. It shares the end-cap cryostat with the end-cap electromagnetic and hadronic calorimeters as this significantly reduces the radiation background levels and provides uniformity of the calorimetric coverage at the same time. In order to avoid back-scattering into the electromagnetic calorimeter, it is separated from the FCal by approximately $1.2\,\mathrm{m}$. The gap severely limits the depth of the forward calorimeter which is why it has an extremely high-density design. With a length corresponding to about ten interaction lengths it covers a pseudorapidity range of $3.1 < |\eta| < 4.9$ Each end-cap consists of three modules which each employ liquid argon as the active medium but which differ in the

[(20)] The nuclear interaction length is defined as the mean distance a hadronic particle travels before undergoing an inelastic nuclear interaction [49].

choice of the absorber material. The first part closest to the interaction point employs copper as the absorber being optimised to predominantly measure electromagnetically interacting particles, whereas the other two parts have much denser absorber elements made of tungsten in order to fully contain hadronic showers along the beam line [118].

3.3.4.3. The calorimeter resolution

The ATLAS calorimeter is segmented in so-called calorimeter cells, which are distinct regions within the calorimeter that can be read out separately. The granularity of the calorimeter is thus defined by the size of the cells. The cell size of the EM calorimeter ranges from $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ to 0.025×0.1 and 0.075×0.025 in the barrel, and from $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ to 0.1×0.1 in the end-caps, respectively. This fine granularity of the electromagnetic calorimeter is ideally suited for precision measurements of electrons and photons. However, the hadronic calorimeter system features larger cells ranging from $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ to 0.2×0.2. More precisely, in the central and extended barrel regions of the tile calorimeter the cells have a size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ up to 0.2×0.1, while the forward calorimeter features a granularity between $\Delta x \times \Delta y = 3.0\,\mathrm{cm} \times 2.6\,\mathrm{cm}$ and $5.4\,\mathrm{cm} \times 4.7\,\mathrm{cm}$.

3.3.5. The muon spectrometer

As so-called minimally ionising particles muons lose little energy within the calorimeters and are thus the only detectable Standard model particles that may travel beyond the calorimeter system. Therefore, a giant detector system dedicated to the identification and reconstruction of muons has been built around the calorimeter. With a mean lifetime of $2.2\,\mu\mathrm{s}$ [51] muons are even able to leave the detector. The muon spectrometer (MS) makes up the largest and outermost portion of the ATLAS detector. It consists of several subdetector modules specifically designed to trigger on and to measure the tracks and thus the momenta of muons within a pseudorapidity range of $|\eta| < 2.7$. The measurement of muons is based on the deflection of their tracks in the magnetic field provided by the large superconducting air-core coils of the toroid magnets. In the central region $|\eta| < 1.4$ the muon tracks are bent by the large barrel toroid magnet, whereas for $1.6 < |\eta| < 2.7$ the bending is exclusively provided by the two smaller end-cap toroids. In the transition region of $1.4 < |\eta| < 1.6$ the deflection is provided by a combination of the barrel and the end-cap magnetic fields.

Depicted in Fig. 3.7, the muon spectrometer is divided into a barrel and two end-cap portions which differ in their design. In the barrel region three concentric cylinders consisting of two types of tracking chambers, so-called monitored drift tubes (MDT) and resistive plate chambers (RPC), surround the beam axis at distances of $5\,\mathrm{m}$, $7.5\,\mathrm{m}$ and $10\,\mathrm{m}$, respectively, and provide a tracking coverage of $|\eta| < 1.05$. Each end-cap portion consists of three circular layers of tracking chambers covering $1.0 < |\eta| < 2.7$ which are mounted perpendicularly to the beam line at distances of approximately $7\,\mathrm{m}$, $14\,\mathrm{m}$ and $21\,\mathrm{m}$ from the interaction point, respectively. A fourth ring-shaped layer in the end-cap is mounted at roughly $11\,\mathrm{m}$ between the innermost and the central wheel. In addition to the typically employed monitored drift tubes and thin gap chambers (TGC), the innermost wheel features a layer of cathode strip chambers (CSC) in order to cover large pseudorapidities close to the interaction point. The individual modules of the wheels are symmetrically arranged in sixteen overlapping sections to prevent a particle from escaping untracked.

Asides from allowing for muon identification and reconstruction, the MS features its own trigger system for regions in space at $|\eta| < 2.4$. The MS trigger consists of resistive plate chambers in the barrel and thin gap chambers in the end-caps. Serving a threefold purpose, its task is to

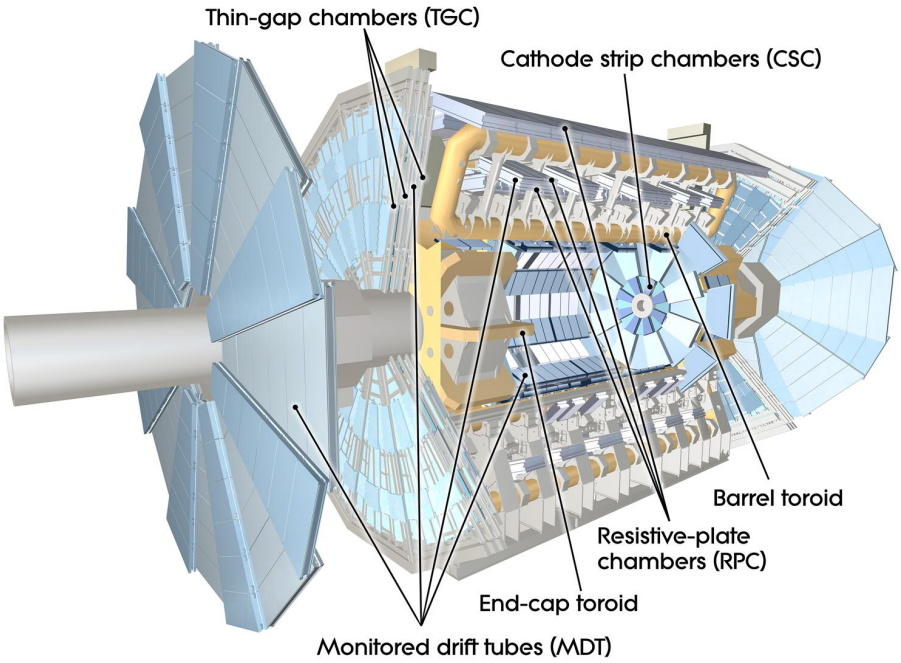


Fig. 3.7.: Computer-generated cut-away view of the ATLAS muon spectrometer. It completely encloses the calorimeter system and ensures no particle being untracked due to its distinctive design. The MS is comprised of four different types of detectors capable of measuring and triggering on muon tracks [120].

identify and distinguish between bunch crossings which requires that it provide a time resolution of better than $25\,\mathrm{ns}$, to provide well-defined thresholds on the transverse momentum of muon tracks, and to measure the muon coordinate orthogonal to the coordinates determined by the tracking chambers [101, 102, 121].

3.3.5.1. The monitored drift tubes

The monitored drift tubes provide high-precision tracking measurements in the forward direction over the full η-acceptance of the muon spectrometer. However, no information about the transverse coordinates is provided. The MDT chambers consist of six to eight layers of aluminium drift tubes of lengths between $70\,\mathrm{cm}$ and $630\,\mathrm{cm}$. The tubes are filled by a gaseous mixture of argon and carbon dioxide kept a high pressure such that impinging muons ionise the gas by creating electron-ion pairs. The electrons drift towards the central gold-plated tungsten rhenium wire which is kept at high voltage, where they create an avalanche of electrons that can be measured. Per tube an average resolution in z-direction of $80\,\mu\mathrm{m}$ is achieved with the previously described setup which results in a z-resolution of approximately $35\,\mu\mathrm{m}$ for each MDT chamber [101, 102, 121].

3.3.5.2. The cathode strip chambers

On the innermost disk of the end-cap muon spectrometer cathode strip chambers, capable to withstand the high rate of impinging particles close to the interaction point, are installed instead of monitored drift tube chambers. The CSC modules are multi-wire proportional chambers covering a pseudorapidity range of $2.0 < |\eta| < 2.7$. In contrast to the monitored drift tubes which are segmented into tubes each with a single wire, the cathode strip chambers employ multiple wires

pointing radially outwards within a single module. In order to provide two-dimensional position information, the read-out cathodes are segmented both parallelly and perpendicularly to the wires. Each chamber has a longitudinal resolution of $40\,\mu\mathrm{m}$ and a transverse resolution of $40\,\mathrm{mm}$ [101, 102, 121].

3.3.5.3. The resistive plate chambers

Resistive plate chambers are high-precision tracking chambers used as triggers in the barrel region of the muon spectrometer. Fast read-out modules were added to cope with the high rate of bunch crossings at the LHC. In total three layers are employed in the barrel region covering $|\eta| < 1.05$. The resistive plate chambers consist of parallel electrodes with a gas-filled gap of $2\,\mathrm{mm}$ between the two plates and no wires. The read-out of the electric signals is segmented into strips capacitively coupled to the electrodes. Due to the small size of the RPC modules, a time resolution of less than $2\,\mathrm{ns}$ is achieved [101, 102, 121].

3.3.5.4. The thin gap chambers

In order to trigger on muons, thin gap chambers are employed in the end-caps of the muon spectrometer for regions in the pseudorapidity region of $1.05 < |\eta| < 2.7$. Thin gap chambers are small multi-wire proportional chambers with a fast response and thus are well-suited as trigger detectors. In addition, as they are aligned perpendicularly to the monitored drift tubes, a rough position measurement of the ϕ-coordinate can be performed with a resolution of about $5\,\mathrm{mm}$ to $10\,\mathrm{mm}$ [101, 102, 121].

3.3.6. The forward detectors

In addition to the main ATLAS detector, three smaller detector systems cover the far-forward region of the ATLAS underground cavern. The main purpose for LUCID[21] and ALFA[22], two such detectors, is the exact determination of the luminosity delivered to the ATLAS experiment. At a distance of $\pm 17\,\mathrm{m}$ from the interaction point LUCID detects inelastic proton-proton scattering in order to both measure the integrated luminosity and to provide an online monitoring of the instantaneous luminosity and beam conditions. It consists of 15 thin gas-filled aluminium tubes outside of each ATLAS end-cap which are arranged parallel to the beam pipe at a radial distance of $10\,\mathrm{cm}$ to the beam pipe [122]. The ALFA detector is located at a distance of approximately $\pm 240\,\mathrm{m}$ at either side of ATLAS interaction point. It employs elastic scattering to measure the absolute luminosity and was designed to approach as close as $1\,\mathrm{mm}$ to the beam line. By means of the optical theorem, the elastic-scattering amplitude within the forward region can be connected to the total scattering cross section and can then be used to extract the luminosity [123].

The third such detector, ZDC[23] plays a key role in the determination of the centrality of heavy-ion collisions which is strongly correlated to the number of far-forward neutrons. It is placed at approximately $\pm 140\,\mathrm{m}$ from the ATLAS interaction point and detects neutrons within a pseudorapidity range of $|\eta| > 8.2$. For this purpose, it consists of layers of alternating quartz rods and tungsten plates [101, 102].

[21] LUminosity measurement using Cherenkov Integrating Detector
[22] Absolute Luminosity For ATLAS
[23] Zero-Degree Calorimeter

3.3.7. The trigger system and data acquisition

The 25 ns bunch spacing at the LHC translates into a bunch-crossing rate of 40 million times per second. Such a collision rate requires incredibly fast read-out electronics and a sturdy detector design, though the available disk space would be filled within a minute. In order to reduce the rate of the data recorded by ATLAS, so-called triggers are used to select events that are potentially of interest based on distinctive features and signatures.

The ATLAS experiment employs a two-stage trigger system comprised of a hardware-based level-1 trigger (L1) and a software-based high-level trigger (HLT). Back in 2012, the LHC accelerated the proton beams such that a centre-of-mass energy of 8 TeV was reached, with a bunch spacing of 50 ns. However, in 2015 the centre-of-mass energy was elevated to 13 TeV leading to an increase of the total proton-proton cross section of 20 % and to an inclusive hard interaction cross section twice as big as before. Furthermore, the halved bunch spacing of 25 ns caused an increase in both luminosity and detector pile-up. In total, the overall interaction rate rose by a factor of six. In order to cope with the more challenging conditions, the ATLAS trigger system was upgraded between 2013 and 2015. At the beginning of the LHC operation, the HLT was divided into two parts, the level-2 trigger and an event filter, The combination of the two modules in the scope of the trigger upgrade significantly reduces the complexity of the software-based trigger [124, 125].

The L1 trigger uses preprocessed information from the calorimeters, the muon spectrometer and the forward detectors applying custom electronics within the respective detector components. At this stage the data rate is reduced from the LHC bunch cross rate of 40 MHz down to 100 kHz thereby decreasing the amount of data to be stored from 100 PiB s^{-1} without any filter applied to approximately 160 GiB s^{-1}. The decision as to whether or not to record an event has to be made within 2 µs since the L1 trigger must be able to decide for each bunch crossing. The L1 trigger consists of several parts, as illustrated schematically in Fig. 3.8. Coarse-grained information from the liquid argon and tile calorimeters is used by the level-1 calorimeter trigger (L1Calo) in order to identify energy clusters and particle types providing a count of objects above certain energy thresholds. It also delivers a preliminary estimate for the missing transverse energy as well as particular information for τ-lepton identification. The level-1 muon trigger (L1Muon) employs resistive plate chambers and thin gap chambers with fast read-out electronics in order to identify muon candidates, as described in Sec. 3.3.5. The information from the several MS subsystems is combined in the Muon-to-CTP interface (MUCTPI) which transmits the information to the L1 centrepiece, the central trigger processor (CTP). Topological criteria based on information from the calorimeters and the muon spectrometer, for example angles between or invariant masses of trigger objects, can be evaluated with the level-1 topological trigger (L1Topo). The CTP then combines the information of all subdetectors and the interface systems and decides which events to keep and to eventually pass to the HLT based on so-called regions of interest (ROI) [126–129].

The high-level trigger consists of computing farms employing a large number of rack-based computers and further reduces the L1 output rate of 100 kHz to 1.5 kHz. With a typical event size of 1 MiB at the ATLAS experiment approximately 1.5 GiB per second of data are finally recorded. The events selected by the L1 trigger are further processed by many different sophisticated selection algorithms using full granularity detector information. Due to the complex event structure, the high-level decision process is the very consuming in computing power and particularly in time. In order speed up the decision, the regions of interest formed by the L1 CTP are used as input for the HLT instead of whole events, thereby reducing the amount of information to be processed. Thus, the HLT is able to decide within 200 ms whether an event should be saved or not. Typically about more than 90 % of all collision data is rejected by the use of triggers. This is one of the reasons why the LHC delivers such a tremendous amount of data given the often tiny cross sections of the

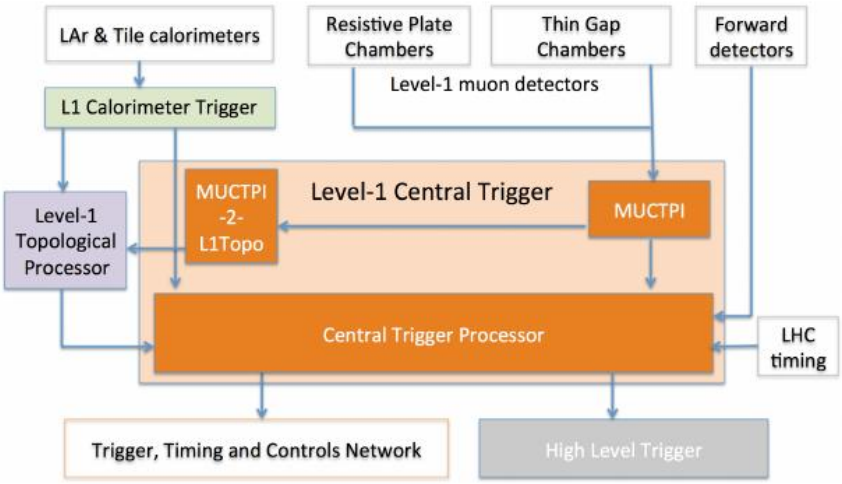


Fig. 3.8.: Flowchart of the ATLAS level-1 trigger system depicting its subcomponents and in the way in which are connected. The information from the various subdetectors is gathered by interface modules before being transmitted to the central trigger processor. The final decision as to which events to keep and thus to pass to the high-level trigger is performed by the central trigger processor based on the regions of interest developed based on the combined information from all detector components [126].

physics processes of interest [129, 130].

The events selected by the trigger system are fully read out by the detector and permanently stored at the CERN Tier-0 computing centre. This so-called raw data undergoes several steps of preprocessing, including the reconstruction of physics objects, calibration and the preparation of more user-friendly data formats. In order to handle the incredibly large amounts of data in the range of several PiB per year produced by the LHC experiments, the Worldwide LHC Computing Grid was developed. It is comprised of more than 170 computing centres in over 40 countries maintained by universities and scientific research institutes affiliated to an LHC experiment. Copies of both raw and preprocessed data are transferred to thirteen large computing centres, referred to as Tier-1, each with sufficient storage, computing power and data-transmission bandwidth. The Tier-1 centres are responsible for larger areas that have typically been defined according to geographical and political conditions. Several smaller so-called Tier-2 computing centres are associated to a Tier-1 centre where the data is further distributed from. Besides handling big parts of simulated data generation and reconstruction, a total of more than 160 Tier-2 centres provide the access for scientists to the grid infrastructure from local workstations and personal computers [131, 132].

3.4. Pile-up and underlying event

For a given bunch crossing, typically more than one scattering process takes place at high-luminosity colliders. For each scattering process, a distinct vertex can be reconstructed by extrapolating the measured tracks, leading to a multitude of vertices within a single event. A typical event recorded by the ATLAS detector is shown in Fig. 3.9 which features 65 distinct vertices along the beam line. The presence of additional vertices in the addition to that from the hard-scatter process of interest makes it difficult to identify the primary vertex of interest and leads to

additional signals in the various detector components. Such additional interactions thus degrade the resolution of the measurement of objects from the interaction of interest and are therefore considered as background to the hard-scatter interaction.

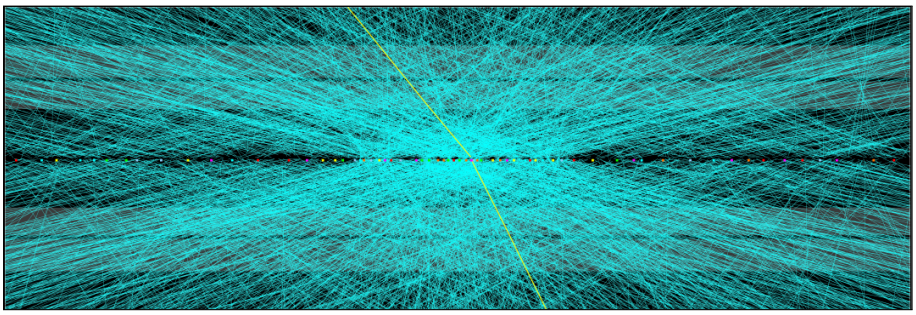


Fig. 3.9.: Display of a $Z \rightarrow \mu^+\mu^-$ candidate event recorded by the ATLAS experiment on 29th September 2017 with a collision energy of 13 TeV. The event has a total of 66 reconstructed vertices that are depicted as coloured dots along the beam line. The interaction of interest is the one associated to the two muon tracks highlighted in yellow [133].

A common source of such undesired contributions is the presence of additional proton-proton collisions within the same bunch crossing, referred to as in-time pile-up (PU). The number of interaction per bunch crossing is continuously monitored and can be parametrised by a Poisson distribution. The average number of particle interactions per bunch crossing $\langle\mu\rangle$ then provides a measure of the hard-scatter activity not associated to the primary vertex. The measured distributions for the ATLAS experiment are presented in Fig. 3.10 for the years of 2015 to 2018. As can be seen in the image, the mean number of interactions increased continuously over the years.The rise of $\langle\mu\rangle$ goes along with the increase of the instantaneous luminosity in the respective years. In 2017 a second peak towards the right side can be observed. Due to a technical shortcoming the bunch spacing had to be increased which implies loss in statistic. In order to compensate, the number of particles within a bunch had been raised leading to a slightly higher average value of $\langle\mu\rangle$ than in 2018 [134, 135].

In contrast, so-called out-of-time pile-up is a detector-related effect caused by remnants from the previous bunch crossing. In case the time certain detector components need to be prepared for the recording of the subsequent collision is longer than the bunch spacing of the LHC, the retarded detector signals contribute to the noise originating from in-time pile-up.

At hadron colliders an additional effect, commonly referred to underlying event (UE), has to be taken into account. Since the colliding particles are not elementary but consist of smaller building blocks, the parton-parton interaction of interest is likely polluted with a variety of interactions occurring in the realm of the same proton-proton collision. The underlying event is mainly composed of soft and semi-hard multiple parton scatterings, but it may also include gluon and photon radiation as well, depending on the particular definition of the UE [136].

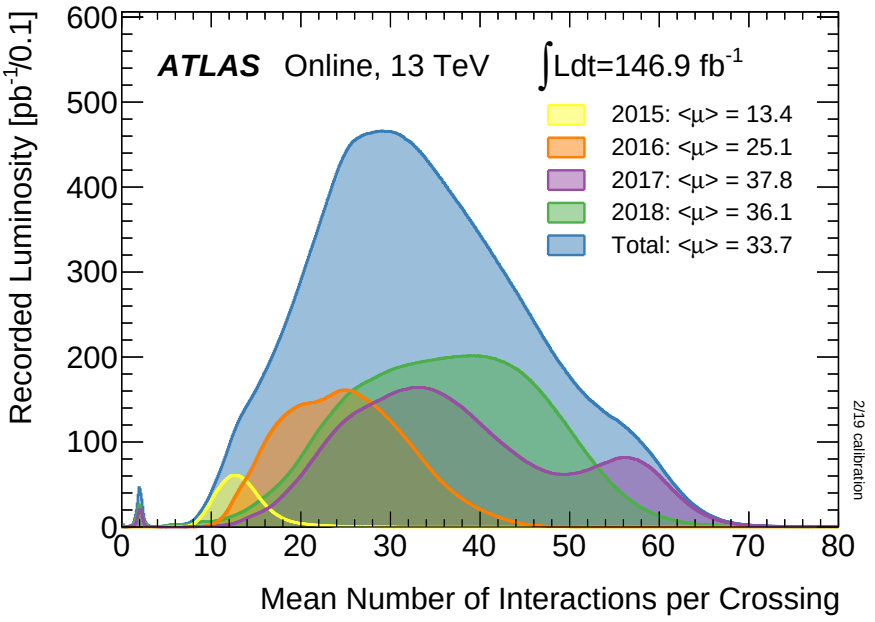


Fig. 3.10.: Mean numbers of interactions per bunch crossing for the proton-proton collision data recorded by ATLAS in Run 2 at a centre-of-mass energy of 13 TeV shown separately for each year and in total. The Poisson distributions are weighted by the luminosity of the respective year of data-taking. The total integrated luminosity stated in the plot is the total amount of data that has been collected by the ATLAS detector within the four years. The average values stated in the legend correspond to the mean of the respective Poisson distribution and have been calculated for each bunch [137].

4. Data and Monte Carlo samples

4.1. Simulated data

Monte Carlo (MC) simulations are widely used tools in many fields in physics to make predictions of physical observables and verify models with experiment. In particular, they are important for high-energy physics, where the aim is to generate large numbers of simulated collision events which can then be compared to experimental data. The simulated data are produced with dedicated computer programs, referred to as MC generators, which are passed to algorithms specifically designed to simulate final states of high-energy physics in full detail down to the level of individual particles.

4.1.1. Event generation

In quantum mechanics, the exact outcome of an experiment cannot be a priori predicted. Calculations only provide a probability to observe a certain measured value. Within the allowed region of phase space, a particle is described by a complex wave function which contains the probability information for the set of possible outcomes of measurements. Before the interaction with the particle its state is undetermined, however, the probability distribution that the measured data follows is uniquely determined by the state of the quantum mechanical system. Therefore, single measurements of a particular physics process feature different outcomes, though a large number of measurements reproduce the corresponding probability distribution. As a natural consequence, this principle is applied within quantum field theories to describe fundamental particles and their interactions. In Monte Carlo generators, the concept of random numbers are employed to make choices on the outcome of an experiment and to reproduce probability distributions for physical processes [138].

The term Monte Carlo method refers to a numerical technique for calculating probabilities and related quantities employing sequences of random numbers. In case of a single random variable, a set of random values $r_1, r_2, \ldots$ uniformly distributed in $[0, 1]$ is generated. In practice, this task is typically accomplished by a computer algorithm called a random number generator. The set $r_1, r_2, \ldots$ is then used to determine another sequence $x_1, x_2, \ldots$ such that they are distributed according to a desired probability density function (p.d.f.) $f(x)$. The values of x can be thought of as a set of simulated measurements from which the probabilities for x to take on values in certain regions of the phase space can be estimated. This is effectively similar to integrating over $f(x)$ in the regions of interest, however, the usefulness of this technique becomes apparent within multidimensional problems where the integration of a joint p.d.f. $f(x, y, z, \ldots)$ over a complicated regions cannot be analytically done anymore [139].

Two important concepts relevant to Monte Carlo generators are perturbation theory and the factorisation theorem. If the result for a given problem cannot be computed exactly, it can still be approximated employing a perturbative ansatz, which is a common approach within quantum field theories. This approach is justified if the problem can be expressed in terms of a power series,

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots \quad , \tag{4.1}$$

which safely converges as long the expansion parameter, x, is strictly smaller than one, $x < 1$. If x is sufficiently small, the solution to a given problem is adequately approximated by taking

into account the first few orders only. As one considers additional higher-order terms, the calculations become more involved. At the LHC, the underlying theory used to describe proton-proton collisions is predominantly QCD. Thus, cross section calculations feature an expansion in powers of the strong coupling constant, α_s. The theoretical description of the hard-scatter interaction of two partons is typically performed by deriving the matrix elements of the corresponding Feynman diagrams. Similar to the wave function in quantum mechanics, the matrix element provides the information about the probability of an interaction. The Feynman diagram with the lowest number of vertices is referred to as leading-order process or tree-level, and in case of $t\bar{t}$ production a few examples were shown in Fig. 2.6. Higher-order corrections, as depicted in Fig. 2.7, involve a higher number of contributing fermion and boson lines which increases the accuracy of the calculation but also its complexity. The matrix elements are then used for the calculation of the differential cross section of the process of interest which again serves as the probability distribution for the fundamental particles associated to that process. A multitude of candidate events are created by scanning through the phase space of the generated particles spanned by their degrees of freedom, each defined by the choice of parameters at a certain point in phase space. As the differential cross section of such a candidate event is directly related to its probability to exist with the given set of parameters, the overall pattern reproduces the underlying probability distribution of the process. Therefore, the simulated events should in principle have the same particle content and kinematics as the corresponding scatter process in a real experiment, given a sufficient theoretical understanding of the process and its correct implementation in the Monte Carlo generator. From a technical point of view, the differential cross sections derived for the candidate events are re-used as scale factors, referred to as MC generator weights, in order to reflect the likelihoods of the events.

In order to reflect the conditions at a hadron collider, a couple of additional steps have to be performed on top of the modelling of the hard-scatter interaction. The factorisation theorem allows the event-generation step to be divided into several almost independent steps. The sets of rules are iteratively applied to construct a physical state that becomes more complex with each step, as schematically depicted in Fig. 4.1. Since the incoming hadrons are composite particles, the initial state of the collision has to be determined by means of the appropriate parton distribution functions. The hard-scatter interaction itself and the calculation of its cross section is perturbatively described, including both the initial and final state partons and short-lived resonances, for example top quarks or Higgs bosons, as well as the decay products.

Partons from the hard-scatter interaction emit gluon radiation somewhat analogously to the radiated photons from electrically charged particles in the context of QED. The gluons can either emit further gluons or split into an antiquark-quark pair, each of which can radiate gluons again. The resulting cascade of partons, referred to as parton shower, represent higher-order corrections to the hard subprocess and are convenient to model the effect of all higher-order processes. However, since it is not feasible to calculate these corrections exactly, an approximate, perturbative treatment of QCD is adopted instead which is typically formulated as an evolution from a high scale in momentum transfer of the hard process down to a low scale of the order of $1\,\mathrm{GeV}$ associated with the confinement of the partons into hadrons. An iterative approach is employed, starting with the generation of one parton splitting subsequent to the hard process. The final state of that splitting is then treated as a new hard process which undergoes the same procedure. The term splitting refers to both the emission of a gluon and the decay of a gluon into a pair of quarks or gluons. The probability of such a splitting is given by the DGLAP splitting functions, describing the evolution from a high momentum transfer scale to a lower one. If a parton shower is developed from an outgoing parton of the hard subprocess, it is referred to as final-state shower. In contrast, so-called initial-state showers are simulated for the incoming partons of the hard-scatter interaction.

The various event-generation steps have to be merged in a consistent manner. In order not to

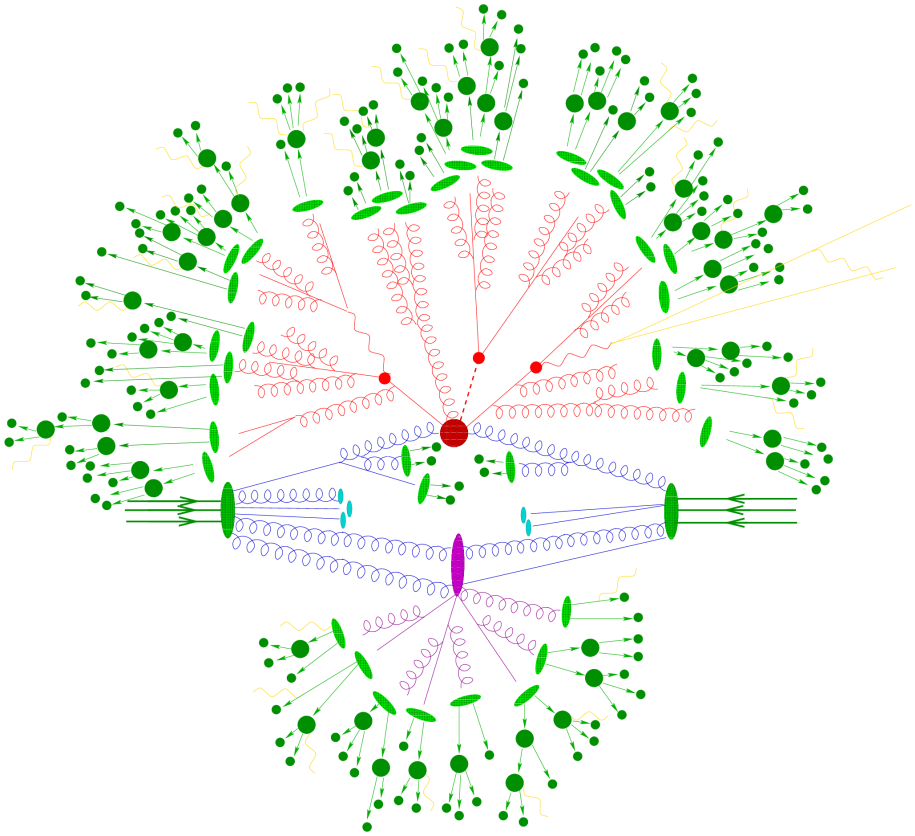


Fig. 4.1.: Sketch of a proton-proton collision simulated by a Monte Carlo event generator. The hard-scatter interaction of interest is illustrated as a big red spot in the middle centre of the picture. Outgoing particles and subsequent decays are also depicted in red, whereas the incoming gluons are shown in blue. Accompanying soft proton remnants are represented by light blue blobs. A secondary hard-scatter event and its associated decay products are indicated by the purple blob. The light green blobs represent parton-to-hadron transitions, followed by the decay of the hadrons indicated by dark green spots. The yellow lines depict soft photon radiation [140].

double-count a certain feature and not to perform mismatching of particles belonging to different stages of the event generation, the matrix elements and parton showers are typically simulated separately from each other and connected afterwards by employing dedicated matching algorithms. In the so-called CKKW matching [141, 142], matrix elements are merged with the parton shower by means of reweighting the matrix elements with the so-called Sudakov form factors [143], which describe the probability for a parton to evolve from a harder momentum scale to a softer scale without emitting another parton either in the initial or final state. Additionally, the emission of showers is vetoed in phase space regions that have already been covered by the parton-level configurations. In contrast, the MLM matching algorithm [144] matches partons from the matrix element calculation to parton showers. In case the showers are either missing or fail to match the partons created at matrix-element level, an event will be rejected and has to be re-generated. For either matching algorithm, as this procedure might get computationally very intensive, the number of additional partons considered in the hard-scatter process is usually limited.

In the regime of asymptotic freedom, the partons emerging from parton showers behave as if they were free particles. Due to the non-Abelian nature of QCD, the strong coupling constant

increases its value towards low momentum scales[1]. This leads to a breakdown of perturbation theory in the evolution of parton showers, and the dynamics enter a non-perturbative phase. Thus, parton shower can only be used down to a cut-off scale of approximately $1\,\mathrm{GeV}$. Furthermore, if the coupling strength increases, the quarks are confined into colourless bound states, referred to as hadrons. Since the formation of colour-neutral hadrons, referred to as hadronisation, cannot be perturbatively described, phenomenological QCD-inspired models are instead employed to characterise the process of hadronisation.

Within the so-called string model [145, 146] a separate gluonic confinement field is assumed to be stretched between each colour charge and its respective anticolour. Its name is derived from the fact that at large distances the potential energy of colour sources induced by the gluon self-attraction increases linearly with their separation which corresponds to an attractive force independent of the distance. Therefore, the colour field is compressed into tube-like regions, referred to as strings, which can be thought of a field lines similar to those in the case of classical electric fields, for example. As two high-energy quarks from a colour-anticolour pair move apart from each other, the potential energy of the gluonic string stretched between the two grows until it reaches the typical scale of hadron masses. At this point it becomes energetically more favourable for the string to break thereby creating a new quark-antiquark pair. The new quark is then connected to the original antiquark by one of the two string segments, and the new antiquark is connected to the original quark by the other string segment. As the two colour-anticolour pairs move further apart, the intermediate strings stretch and break again, which leads to another two new quark-antiquark pairs. This series of reactions continues until all the available energy has been converted into partons which are connected by very short string segments which are then identified as hadrons. The hadronisation within a system of numerous partons, as it is typically the case for hard-scatter events at hadron colliders, is more complicated than that a for a single quark-antiquark pair. Each parton in the system has a unique colour partner to which it is connected by a string that stretches and breaks as previously described. Since gluons carry both a colour and an anticolour, they are connected to two other partons which leads to kinks in the string evolution. In order to properly model these soft, non-perturbative QCD effects, various parameters of the Monte Carlo generators have to be adjusted such that the outcome is as data-like as possible, referred to as tuning. The values of the tuning parameters are obtained from the comparison of Monte Carlo generated events to measured data.

A complementary approach, the so-called cluster model [147], is based on the evidence that at evolution scales much less than the hard subprocess, the partons in a shower are clustered in colourless groups with an invariant mass distribution only depending on the momentum transfer scale. This phenomenon, referred to as colour pre-confinement, implies that the colour flow is local so that each colour line at the end of the shower is connected to an anticolour partner at the same scale. The clusters are then identified as temporary hadronic states that decay into the observed final-state hadrons according to the available phase space.

Aside from the hadrons originating from the hard-scatter interaction, extra hadrons are likely produced from interactions between partons not taking part in the hard subprocess. Since protons are not fundamental but consist of three valence quarks embedded in sea of quarks and gluons, a high-energetic collision of two protons typically features a multitude of additional parton-parton interactions. The interference between those multiple parton interactions and the hard process, as well as with remnants of the collided protons further impede the modelling. The additional QCD-induced activity within the same collision of two initial-state protons is referred to as underlying event, which differs from pile-up since the latter refers to multiple proton-proton collisions within the same bunch crossing. However, in experiment the combined and overlapping effects from both

[1] The concept of the running QCD coupling strength was introduced in Sec. 2.1.2.

phenomena impact the measured signals in the detector. Many effects of the underlying event cannot be adequately described from theory and are therefore controlled by tunable parameters whose values have been determined from experimental data.

The final stage of the event generation is to model the subsequent decay of the unstable hadrons produced during the hadronisation process. Some are deemed stable[(2)], or are sufficiently long-lived, such that their decays are visible in the detector, so a large variety of final-state hadrons and their decay modes has to be simulated.

4.1.2. Data preparation

After the event generation step, particles are propagated through a simulation of the interaction with all components of the ATLAS detector using the GEANT4 toolkit [148]. The detector configuration reflects the full detector geometry including known misalignments and distortions. Subsequently, the ATLAS digitisation software converts the hits produced by the detector simulation into digitised detector responses, referred to as digits. Energy deposits are converted into voltages and currents similar to the read-out signals created by real data events, thereby simulating the read-out electronics. A digit is typically produced when the voltage or a current of a particular read-out channel rises above a certain pre-configured threshold within a particular time window. At this stage, the detector noise and other environmental backgrounds from the detector cavern and the LHC are added. In order to save computing time for the detector simulation, the information of several interactions from the same bunch crossing is overlayed to include pile-up effects during digitisation as well [149].

Whereas the L1-trigger is implemented with hardware in the real detector, it is simulated for Monte Carlo events using the digits constructed from artificial hits as input. The high-level trigger then takes the L1 inputs in to order to make subsequent decisions related to the trigger selection. The object reconstruction applied to the events selected by the HLT is identical for simulation and data, with the exception that truth[(3)] information is available only in MC datasets. Instrumental calibrations and detector alignment are adapted from real data-taking conditions. In addition, dedicated scale factors are applied to simulated events in order to match the object identification, reconstruction and trigger efficiencies, energy scales and resolutions determined from real data. In general, simulated events are treated equally to those from real data as much as possible and undergo the same reprocessing procedure [132, 150].

Due to the complicated detector geometry and detailed physics description incorporated in the GEANT4 framework, the full simulation of the ATLAS detector is very intensive both in computing power and time. Therefore it is impracticable to achieve the required simulated statistics for many physics studies without an alternative and faster solution. One such alternative, referred to as ATLFAST-II, was established as the standard fast simulation, providing large statistics to supplement full-simulation samples. With a simplified description of both the detector geometry and the shower parametrisation within the calorimeters, events are simulated as fast as possible, though keeping the ability to run the standard ATLAS reconstruction algorithms. However, in order to maintain a reasonable level of accuracy, a full simulation with GEANT4 of only a few essential subdetectors accompanies the fast simulation [149, 151].

The statistics of a Monte Carlo sample can be expressed by the ratio of the number of generated events N and the cross section cross section σ_{MC} a certain process has been produced with, assuming each generated event corresponds to an event recorded in the experiment. The so-called

[(2)] A particle is typically considered to be stable if $c\tau \geq 1\,\mathrm{cm}$.

[(3)] The so-called truth information is a history of the interaction from the generator, including incoming and outgoing particles, and consists of a detailed listing of all involved particles and their interactions at the various steps of the event generation.

integrated luminosity of a MC sample,

$$\mathcal{L}_{\mathrm{MC}} = \frac{N}{\varepsilon_{\mathrm{filter}} \cdot \sigma_{\mathrm{MC}} \cdot k} \quad , \tag{4.2}$$

features two additional terms in the denominator. The first, $\varepsilon_{\mathrm{filter}}$, describes to so-called filter efficiency, which is defined as the fraction of produced events compared to the case without any filter applied. Since the generation of events is very time-consuming, dedicated filters are applied during Monte Carlo generation in order to both enhance the speed and to increase the amount of generated events in certain region of phase space. The generation of events can therefore be limited to those reflecting particular features of the physics processes of interest, for example certain decay modes. The filter efficiency $\varepsilon_{\mathrm{filter}}$ corrects for the reduced phase space that was available during the event generation. The second factor, k, accounts for higher-order corrections to the production cross section. In the numerator, the number of generated weights, N, incorporates both the experimental scale factors and the so-called generator weights. The generator weights are specific for each event and account for the fact that some points in the available phase space are physically more probable than others, as described previously in Sec. 4.1.1. In addition, generator weights are employed to correct for double-counting of the same process, for example a leading-order process which emits radiation at parton-shower level may reproduce a process that has already been accounted for by a next-to-leading order calculation. In order to properly compare the simulated events passing all selection criteria to data, they have to be scaled to the corresponding integrated luminosity of the data to which they are compared. Therefore, a scale factor,

$$f = \frac{\mathcal{L}_{\mathrm{data}}}{\mathcal{L}_{\mathrm{MC}}} \quad , \tag{4.3}$$

is applied to each individual simulated event. Simulated samples with a luminosity higher than the integrated luminosity of the data are scaled down, meaning that one recorded data event is represented by more than one generated event. In contrast, in the opposite case the simulated samples are scaled up and a MC event represents more than one data event. It is therefore important to keep that ratio below one in order not to be affected by statistical effects arising from the limited number of Monte Carlo events.

4.1.3. List of simulated samples

Monte Carlo simulated samples are used to model both the signal and background processes considered in the studies that are presented in this thesis. They have been either passed through a full simulation of the ATLAS detector based on GEANT4 or through a fast ATLFAST-II simulation. In order to model the different physics processes studied at the LHC, a variety of Monte Carlo generators is available, given the possibility to combine them with different sets of parton distribution functions. Every MC generator shows a different level of accuracy within the various steps of Monte Carlo generation as explained in Sec. 4.1.1, although some programs are capable of covering several or even all steps of Monte Carlo generation. Thus, the simulated samples are typically produced with a combination of different generators each best suited to the requirements of the respective MC generator step and physics process it is responsible for.

For the generation of the hard-scatter interaction three main programs are typically employed in ATLAS. POWHEG-BOX [152, 153] calculates the matrix elements (ME) at NLO which are matched to the parton shower (PS) simulation according to the POWHEG method [154, 155]. The MADGRAPH5_AMC@NLO [156] event generator is a combination of MADGRAPH5 [157], which allows for the generation of any physics model that can be expressed with a Lagrangian,

and the aMc@NLO [158] tool with which matrix elements calculated at NLO can be merged with parton shower simulations. Multi-parton matrix elements can be calculated at NLO with the Sherpa [159, 160] general-purpose event generator which is also suited to simulate parton showers and to merge them with NLO matrix elements. In addition, the modelling of the underlying event and hadronisation can be performed by Sherpa. A broad range of collider physics processes can be simulated with the Pythia [161–163] generator. Widely used as an interface to hard-scatter generators except for Sherpa, the Pythia program models parton showers, the underlying event and hadronisation. As an alternative to Pythia, the Herwig [164, 165] generator is employed. Heavy flavour decays, which are particularly important for the studies presented in this thesis, are modelled using the EvtGen [166] program, except for processes generated with Sherpa. In all samples, the top quark mass value used in the generation of the hard-scatter process is set $172.5\,\mathrm{GeV}$, and the mass value of the Higgs boson is assumed to be $125\,\mathrm{GeV}$. If the generation of a sample involves the explicit modelling of bottom quarks, a mass value of $4.75\,\mathrm{GeV}$ is adopted for the bottom quarks. A common approach to estimate the impact of systematic uncertainties on a given measurements as a result of the choice of a certain MC generator is to compare simulated samples of the same physical process which have been produced with a different set of MC generators each. For example, if the uncertainty associated to the theoretical prediction of the parton shower is investigated, the computation of the matrix element is made with the same program.

The associated production of top-antitop quark pair with a heavy vector boson (W or Z) is simulated with the MadGraph5_aMc@NLO 2.3.3 which provides matrix element calculations at NLO in the strong coupling constant, α_s, using the NNPDF3.0NLO PDF set [167]. The decay of top quarks is modelled at LO with the MadSpin [168, 169] package in order to preserve the correlation between the spins of the top and antitop quarks, which is a property that can impact the kinematics of the final-state particles from the hard-scatter process. For the simulation of the parton shower and for the hadronisation process, the generation of the hard-scatter interaction is interfaced with Pythia 8.210 using the A14 set of tuned parameters [170] and the NNPDF2.3LO PDF set [171]. Decays of charm and bottom quarks are simulated with the EvtGen 1.2.0 program. The theoretical $t\bar{t}W$ and $t\bar{t}Z$ cross sections are calculated with MadGraph5_aMc@NLO at NLO accuracy both in QCD and in EW theory according to Ref. [172], including possible contributions to the cross section from virtual photons, γ^*, and from the interference between the Z boson and γ^* [173]. Contributions from off-shell Z bosons with an invariant mass of below $5\,\mathrm{GeV}$ are not taken into account. Theoretical uncertainties on the signal prediction are evaluated with several alternative simulated samples. The Sherpa 2.2.1 generator at NLO accuracy with up to one additional parton was employed to generate $t\bar{t}Z$ samples with the Z boson decaying only leptonically ($Z \to \ell\ell$), which were used to estimate uncertainties associated to the calculation of the matrix element. The default Sherpa parton shower is used along with the NNPDF3.0NNLO PDF set [167]. A similar set of Sherpa samples, providing the most precise prediction of the $t\bar{t}Z$ process so far, was simulated with the MEPS@LO [174, 175] setup, featuring up to one additional parton at NLO and two additional partons at leading order. In order to estimate uncertainties related to the parton shower simulation, an alternative signal sample has been generated with the MadGraph5_aMc@NLO generator interfaced to Herwig 7 and HE-UE for the underlying event tune. The evaluation of uncertainties associated to the modelling of ISR is performed with two sets of simulated samples produced using the same settings as for the nominal signal samples, but with an up and down variation in the A14 tune, respectively.

The production of two heavy vector bosons (W and Z) with additional partons, featuring two charged leptons and two neutrinos ($\ell\ell\nu\nu$), three charged leptons and one neutrino ($\ell\ell\ell\nu$), and four charged leptons ($\ell\ell\ell\ell$) in the final state, is simulated using the Sherpa 2.2.2 generator. Such

samples are referred to as WZ /ZZ+jets in the course of this thesis. Multiple matrix elements are matched and merged with the SHERPA parton shower based on the Catani-Seymour dipole factorisation scheme [176, 177] using the MEPS@NLO [174, 175, 178, 179] prescription. The virtual QCD corrections for matrix elements at NLO are provided by the OPENLOOPS [180, 181] library. The samples are generated using the NNPDF3.0NNLO PDF set, along with a dedicated set of tuned parton shower parameters developed by the SHERPA authors. Events with not more than one additional parton are simulated at NLO precision, whereas events with two or three additional partons are simulated at leading order. An additional set of diboson samples, which are each normalised to a respective cross section calculated at NNLO precision, has been generated with SHERPA 2.2.1 using the NNPDF3.0NLO PDF set. The matrix elements are calculated for up to one additional parton at NLO and two or three additional partons at LO precision and are merged afterwards with the SHERPA parton shower using the MEPS@NLO prescription.

The production of three massive vector bosons with up to six charged leptons in the final state is simulated with SHERPA 2.2.2 using the NNPDF3.0NNLO PDF set. By means of the COMIX [176] and OPENLOOPS matrix element generators, final states with no additional partons are calculated at NLO, whereas final states with up to three additional partons are calculated at leading order.

For the production of single top quarks in association with heavy vector bosons sets of samples are employed for final states with two leptons and final states with three or four leptons, respectively. For final states with three or four charged leptons, single top quarks in association with a W and a Z boson (tWZ) are modelled using MADGRAPH5_AMC@NLO 2.2.2 at NLO with NNPDF3.0NNLO PDF set. The simulated events are interfaced with PYTHIA 8.235 using the A14 tune and the NNPDF2.3LO PDF set. Modelling uncertainties of the tWZ process have been evaluated by the comparison to a sample produced with the same generator but employing a different treatment of the interference between $t\bar{t}Z$ and tWZ, referred to as diagram removal[4] [182] (DR) [183]. The production of single top quarks in association with a Z boson and one extra parton (tZq) is generated using MADGRAPH5_AMC@NLO 2.3.3 at NLO with the NNPDF3.0NNLO PDF set. The events are interfaced with the PYTHIA 8.230 generator along with the A14 tune and the NNPDF2.3LO PDF set. As for $t\bar{t}Z$, the samples include contributions from off-shell Z bosons, and samples with different variations of the A14 tune are compared. However, alternative samples are used to simulate events with two leptons in the final state. The tWZ process is simulated with the MadGraph5_aMc@NLO generator with the NNPDF3.0NLO PDF set. The parton shower is modelled using PYTHIA 8.212 along with the A14 tune and the NNPDF3.0NLO PDF set. tZq events are simulated with the MADGRAPH5 generator at leading order, while the parton shower is modelled using PYTHIA 8 in conjunction with the A14 tune and the NNPDF2.3LO PDF set.

Both top quark pair production and the associated production of a single top quark with an additional W boson (Wt) are generated at NLO using POWHEG and the NNPDF3.0NLO PDF set. The parton shower, hadronisation and underlying event are modelled with PYTHIA 8.230, using the A14 tune and the NNPDF2.3LO PDF set. Heavy-flavour decays are simulated with EVTGEN-1.6.0. The events are normalised to their NNLO cross section including the resummation of soft gluon emission at NNLL accuracy derived with Top++ [184]. In case of single top production, diagram removal[5] is applied in order to account for the interference with $t\bar{t}$.

The simultaneous production of three or four top quarks ($t\bar{t}t$ and $t\bar{t}t\bar{t}$, respectively) as well as the

[4] All NLO Feynman diagrams for tWZ overlapping with the the $t\bar{t}Z$ definition are removed from the calculation of the tWZ matrix element.

[5] All NLO Feynman diagrams for Wt overlapping with the the $t\bar{t}$ definition are removed from the calculation of the Wt matrix element.

production of a $t\bar{t}$ pair with two W bosons ($t\bar{t}WW$) are simulated at LO using MADGRAPH 2.2.2 interfaced to PYTHIA 8.186 with the A14 tune and the NNPDF2.3LO PDF set.

The simulated samples are showered with PYTHIA 8 using the A14 tune. Samples featuring the production of a Higgs boson in association with a W or Z boson are generated with PYTHIA 8.186 using the A14 tune and the NNPDF2.3LO PDF set.

The production of a Z boson with jets (Z+jets) is simulated with SHERPA 2.2.1 using the NNPDF3.0NLO PDF set. The matrix elements are calculated up to two additional partons at NLO and for three and four additional partons at leading order, respectively.

Top-antitop quark pairs with an associated photon ($t\bar{t}\gamma$) are modelled using MADGRAPH5_AMC@NLO 2.3.3 along with the NNPDF3.0NLO PDF set. The showering is simulated with PYTHIA 8.212 using the A14 tune and the NNPDF2.3LO PDF set. The production of single photons in association with a Z boson (γZ) is produced with SHERPA 2.2.4 along with the NNPDF3.0NLO PDF set.

In order to account for multiple parton-parton interactions in the same and neighbouring bunch crossings, the pile-up is modelled by overlaying the original hard-scatter event with a set of minimum-bias events simulated with PYTHIA 8.186 using the NNPDF2.3LO PDF set and the A3 tune for each hard-scatter interaction. As the Monte Carlo samples were simulated in some cases before the data were collected, the distribution of the number of additional proton-proton interactions for the simulated samples is reweighted in order to match the pile-up level present in the experimental data.

During the LHC Run 2, the data-taking conditions, most notably the number of pile-up interactions within a collision, were different in the years of 2015 and 2016, 2017 and 2018, respectively. Therefore, different sets of MC samples, referred to campaigns, were produced for these time periods each reflecting the respective data-taking conditions. The list of the ATLAS MC campaigns for Run 2 and the corresponding data-taking periods is given in Tab 4.1.

Data-taking period	2015	2016	2017	2018
MC campaign		mc16a	mc16d	mc16e

Tab. 4.1.: List of Monte Carlo production campaigns with the corresponding data-taking periods.

4.2. ATLAS data

The data used for the studies presented in this thesis reflects the full dataset of proton-proton collisions at a centre-of-mass energy of 13 TeV recorded by the ATLAS detector during the LHC Run 2, which corresponds to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The number of parton-parton collisions per beam crossing was steadily increased from 2015 to 2018. Whereas in the first year of data-taking the maximum value of the instantaneous luminosity was $5 \cdot 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, the peak value in 2018 reached $21 \cdot 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, thereby exceeding the design value of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [99] by more than a factor of 2[(6)]. In Tab 4.2 both the integrated luminosity and the peak value of the instantaneous luminosity per year are presented. The increase in instantaneous luminosity led to a growing dataset for each single year of Run 2 but also to higher pile-up rates as discussed in Sec. 3.4.

As the LHC is put into operation after a shutdown, a pilot beam consisting of a single proton bunch is injected into either beam pipe in order to test the entire machine. If neither peculiar

[(6)] The design value had already been reached in July 2016.

Year	Integrated luminosity [fb^{-1}] good for physics	Peak instantaneous luminosity [$\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$]
2015	3.2	$5.0 \cdot 10^{33}$
2016	33.0	$13.8 \cdot 10^{33}$
2017	44.3	$20.9 \cdot 10^{33}$
2018	58.5	$21.0 \cdot 10^{33}$
Full Run 2	139.0	

Tab. 4.2.: Overview of the integrated luminosity recorded by ATLAS after applying data quality requirements and of the instantaneous luminosity delivered to ATLAS in the years of 2015 to 2018. For the integrated luminosity the values per year are summed to show the full amount of ATLAS data available for physics analysis of Run 2. The instantaneous luminosity can be seen to increase from 2015 to 2018. All values are given at an accuracy of one decimal place [137].

features of the beam nor a malfunction of the machine can be observed, the full beam is injected and accelerated to the collision energy. Once stable-beam conditions, a state in which the LHC is running as desired, are reached, the experiments at the LHC commence to take data. After a time span of typically 10 h to 20 h of stable beams, the proton bunches have thinned out too much in order to maintain a certain number of collisions per bunch crossing. Therefore, in such a case the beams are dumped onto a solid target located in an underground cavern alongside the LHC tunnel, and the LHC gets refilled with two new beams of proton bunches. It takes approximately half an hour until the beams reach their maximum energy of 6.5 TeV.

Due to a slight inefficiency of the data acquisition system and various detector components, not the entire data provided by the LHC can be recorded by the ATLAS detector. Once the LHC is in stable-beam mode after a refill, the high voltage and the pre-amplifiers for the tracking system have to ramp up, for example. Thus it takes some time until they work at full efficiency. Besides, a few permille of the recorded data is employed for luminosity calibrations. As can be seen in Fig. 4.2, this leads to a slightly smaller amount of data of $147\,\mathrm{fb}^{-1}$ ATLAS was able to record with respect to the $156\,\mathrm{fb}^{-1}$ delivered by the LHC. A subset of the data recorded with stable beams and optimal functional conditions of the ATLAS detector, corresponding to approximately $139\,\mathrm{fb}^{-1}$, passed all quality requirements on the reconstructed physics objects and can thus be used for physics analysis, therefore referred to as good for physics.

The basic time unit for storing ATLAS data in the so-called luminosity block (LB), a period of approximately one minute of data taking over which the experimental conditions are assumed to be constant. In particular, it is assumed that the instantaneous luminosity stays almost constant over the duration of the LB. Luminosity blocks of data passing all requirements on the beam stability, the operational status of all ATLAS detector components and the reconstruction performance are summarised in so-called Good Run Lists (GRL). However, if a LB has been identified to contain data that has been recorded under suboptimal conditions, it will be excluded from physics analysis. Several luminosity blocks which share roughly the same stable running conditions, for example the filling scheme of the bunches or the pursued instantaneous luminosity, form a so-called period. Depending on the year of Run 2, the ATLAS data consists of 15 to 20 periods per year [98].

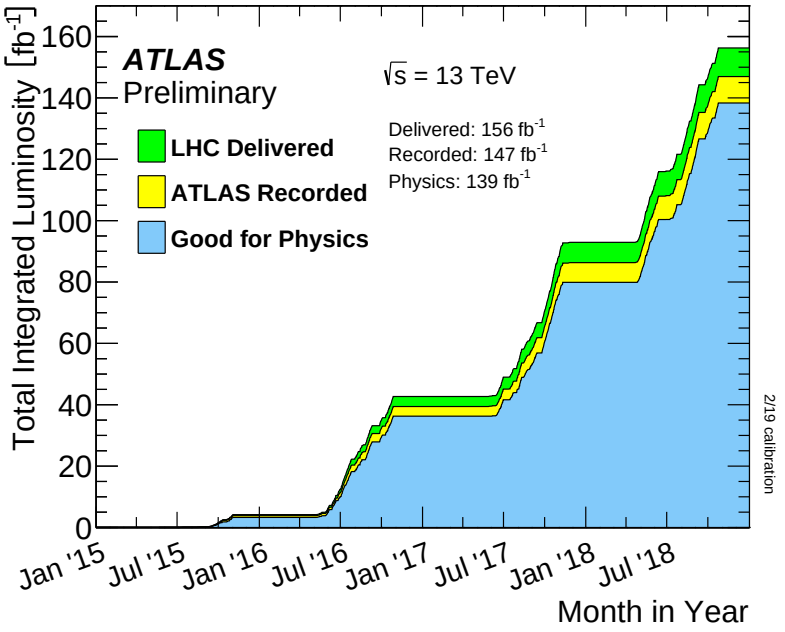


Fig. 4.2.: Evolution of the luminosity per time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be of good quality to be used for physics (blue) of proton-proton collisions at a centre-of-mass energy of $13\,\mathrm{TeV}$. The recorded luminosity reflects the inefficiency of the data acquisition as well as the inefficiency of various detector components during the data taking. For data to be of good quality all quality criteria for the reconstructed physics objects have to be fulfilled [137].

5. Physics objects

In order to make use of the recorded data, the detector signals have to be converted into physical objects. Fundamental particles produced in a collision can be reconstructed with specialised algorithms using the information from the various subdetector systems. The reconstruction procedure is the same for real data and for Monte Carlo simulated events, providing preliminary candidates for physics objects which are refined with additional identification requirements in order to become the final physics objects used within analyses. After the actual reconstruction procedure potential overlaps between the objects have to be removed.

The algorithms used to reconstruct and identify the physics objects were developed using knowledge of the distinct detector signatures produced by different particles and particle types. Selections of events featuring well-known processes are used to continually improve those algorithms. Fig. 5.1 shows a sketch of the signatures of all detectable particle types within a cut-out of the ATLAS detector in the barrel region. Particles which predominantly interact electromagnetically, for example electrons and photons, deposit their energy in the electromagnetic calorimeter, whereas hadrons such as protons, pions and kaons lose their energy primarily in the hadronic calorimeter. The curved trajectories of electrically charged particles, such as electrons and muons, as a result of the Lorentz force are depicted as solid lines in the figure. The trajectories can be reconstructed as track segments in both the tracking region of the detector and, in case of muons, in the outermost muon spectrometer. Muons are the only detectable particles described by the Standard Model that travel beyond the calorimeters and therefore induce tracks in the muon system. Neutrinos interact exclusively via the weak force, which makes it very hard to detect them. Thus, they typically leave the detector unseen, indicated by a dashed line.

The studies presented in Chaps. 6 and 7 make use of reconstructed objects as electrons and muons, jets, and the so-called missing transverse energy. Reconstructed photons or hadronically decaying tau leptons are objects commonly employed in ATLAS analyses. but are not relevant to the signal process of interest in this thesis and are therefore not further described.

5.1. Tracks and vertices

The trajectories of charged particles are reconstructed as tracks in the inner detector and in the case of muons also in the muon spectrometer, as discussed in Secs. 3.3.3 and 3.3.5, respectively. The momentum and the sign of the electric charge of the particles can be inferred from the bending of the reconstructed tracks within the magnetic fields of solenoid and toroid magnets [135, 186]. For the reconstruction of tracks an iterative algorithm is employed which is based on fitting a set of measurements of the inner detector to a trajectory model. Triplets of hits in the pixel and semiconductor tracker systems pointing towards the interaction point serve as seeds which are identified using an inside-out approach. With a combinatorial Kalman filter technique [187], additional hits in the pixel and SCT systems further away from the interaction point are associated to the reconstructed track in an iterative fashion. Each time another hit is added to the track candidates, a χ^2-fit is re-evaluated using the ATLAS global χ^2 track Fitter [188]. The ambiguity due to hits associated with multiple track candidates is resolved with a dedicated algorithm that employs track quality requirements in order to solve such ambiguities. However, these ambiguous hits are not taken into account within the fit procedure of track segments. Only track candidates of sufficient quality are passed onto the transition radiation tracker where its hits are included into the

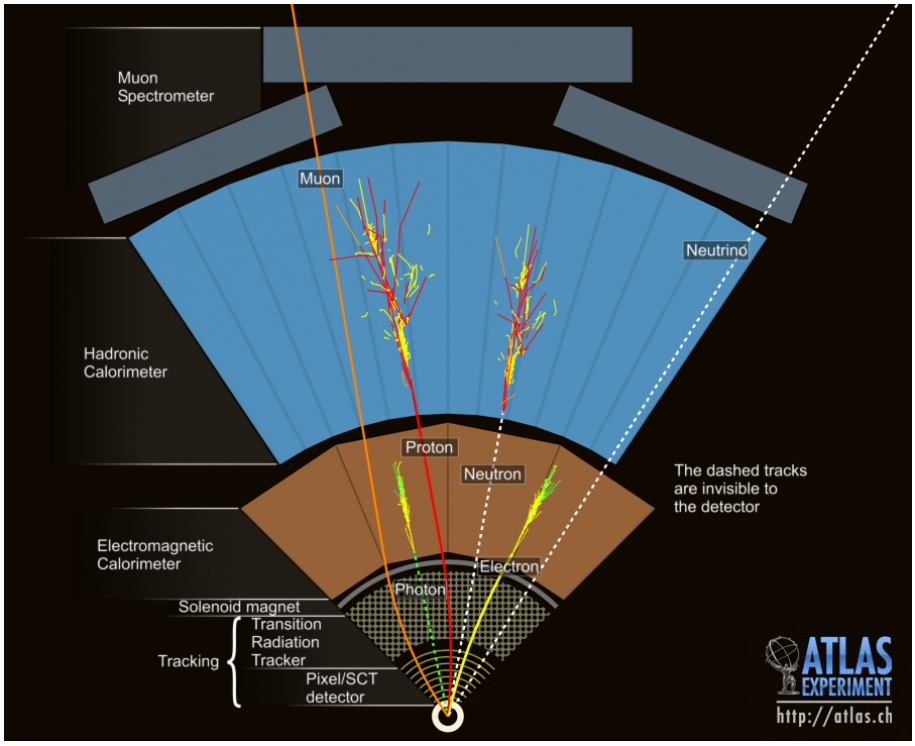


Fig. 5.1.: Schematic cross-sectional view of particle signatures in the ATLAS detector. From the different patterns that are left in the detector components the particles produced in a collision can be reconstructed. Trajectories of electrically neutral particles are indicated by a dashed line [185].

track reconstruction procedure. Track segments within the TRT which were so far not included in the reconstruction procedure are extrapolated into the pixel detector and semiconductor tracker in order to form tracks of particles which were produced within interactions of particles from the hard-scatter interaction. In addition, track candidates are reconstructed from segments in the TRT that have no extension into the silicon detectors. In order to be considered, a track candidate has to have a minimum transverse momentum of $400\,\mathrm{MeV}$ and must be within a pseudorapidity range of $|\eta| < 2.5$. Furthermore, at least seven hits in the silicon detectors are required. Characteristic parameters are the ratio of the track charge to its momentum, the azimuthal and polar angles denoting its position within the detector, and the transverse and longitudinal impact parameters which are defined as the minimum distance to the interaction point in the transverse plane and in z-direction, respectively. These parameters are used to set quality requirements on the tracks but are not used for the fitting procedure [189–191].

The points in space corresponding to intersections or projected origins of two or more tracks are referred to as vertices and are reconstructed with an iterative procedure which comprises two main stages. In the first stage, for all tracks that originate from a point compatible with the interaction region and survive certain quality requirements based on the track parameters, a vertex seed is determined. The x- and y-coordinates of the starting point are taken from the centre of the beam spot in the transverse plane. For the z-coordinate of the starting point the maximum of the z-coordinates of all tracks, calculated at the point of the closest approach to the reconstructed interaction point of the respective track, is used. The optimal vertex position is determined with the so-called adaptive vertex fitting algorithm [192] which performs an iterative χ^2-minimisation with the position of the seed as the starting point and the parameters of reconstructed tracks as

input measurements. Each input track is assigned a weight which reflects the compatibility of the track with the vertex estimate, and the vertex position is recomputed with the weighted tracks. This procedure is repeated afterwards and the track weights are recalculated with respect to the new vertex position. After the last iteration, the final track weights are evaluated and all tracks found to be incompatible with the vertex position by more than seven Gaussian standard deviations are removed from the vertex candidate and are considered as inputs for additional vertex candidates. The above described procedure is repeated until all tracks in an event have been associated to a vertex or no additional vertex can be found in the remaining set of tracks. Vertices with at least two associated tracks with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ are considered as valid primary vertex (PV) candidates. The final primary vertex considered for a given events and deemed to be associated with the hard-scatter interaction of interest is defined as the PV candidate with the largest sum of squared transverse momenta, $\sum_i (p_{\mathrm{T},i})^2$, of its associated tracks and serves as origin for the physical objects of interest in an event [193, 194].

5.2. Electrons

An electron is defined as an object consisting of groups of isolated, neighbouring energy deposits in the calorimeter, referred to as clusters, and at least one matched track of a charged particle in the inner detector. For the reconstruction of electron and photon candidates, clusters of energy deposits measured in topologically connected cells of the electromagnetic and the hadronic calorimeter, referred to as topo-clusters, are formed and matched to tracks in the inner detector. The matched topo-clusters serve as an input to the reconstruction of so-called superclusters of variable-size which are used to build the physical electron and photon objects. Depending on whether ID tracks, photon conversion vertices or neither of them can be matched to such a supercluster, the reconstructed object will be classified as an electron, converted photon or unconverted photon, respectively [195].

The topo-cluster reconstruction algorithm [196, 197] looks for local maxima of the cell significance,

$$\varsigma_{\mathrm{cell}} = \frac{E_{\mathrm{cell}}}{\sigma_{\mathrm{cell}}^{\mathrm{noise}}} \quad , \tag{5.1}$$

which is defined as the ratio of the deposited energy E_{cell} over the average noise $\sigma_{\mathrm{cell}}^{\mathrm{noise}}$ within the respective calorimeter cells. Energy deposits are considered as seeds for the subsequent cluster formation if the significance of the cell that initiated the cluster has a value of at least four. Neighbouring clusters are sequentially scanned and added to the initial seed clusters based on their cell significance which has to be greater than two. If two given clusters contain the same neighbouring cell, these two are merged. In contrast, a cluster is split if it contains two or more local maxima each with an energy deposit of above $500\,\mathrm{MeV}$ and at least four neighbouring cells with smaller signals. The reconstruction of electron and photon candidates then starts from those topo-clusters with a transverse energy of at least $400\,\mathrm{MeV}$, using the energy of the corresponding cells in the EM calorimeter only except in the crack-region [195].

Tracks from the inner detector are then associated with the topo-cluster candidates based on the agreement of the respective η- and ϕ-coordinates. Energy clusters which have a longitudinal and lateral shower profile compatible with that of an EM shower are employed to create so-called regions of interest (ROI). They are derived by sequentially scanning the EM calorimeter similarly to the previous description in the context of the topo-clusters except for using a fixed-size window. In order to discriminate electron candidates from pions, a pattern recognition algorithm is applied which models the energy loss due to interactions of the particles with the detector material under the hypothesis that the particles are pions by default. If a track candidate fails the pion hypothesis

within an ROI due to large energy losses, it is fit again with the electron hypothesis from a modified pattern recognition based on a Kalman filter formalism [187] which allows for energy losses of up to 30 % at each material intersection, for example due to bremsstrahlung. The track candidates are then fit with the global χ^2-fitter [188] allowing for additional energy losses when the standard track fit fails. In addition, track candidates with hits in the silicon tracking detectors that can be matched to the fixed-size clusters are refit by means of a Gaussian-sum filter method [198, 199], which takes non-linear energy loss effects into account, to either confirm or reject the pion hypothesis. These tracks are subsequently matched to the EM topo-clusters by extrapolation from the inner detector into the calorimeter system. In the case that more than one track matches a cluster, a primary track is chosen based on criteria such as the distance between extrapolated tracks and clusters, the number of hits in the pixel detector, and the presence of a hit in the innermost silicon layer of the tracking detectors [195, 200].

In the first stage of the supercluster reconstruction, the track-matched EM topo-clusters are tested for the use as seed cluster candidates. The initial list of EM topo-clusters is sorted in descending order according to their transverse energy which is calculated using the EM energy only. Each cluster is separately tested and considered as seed cluster if it features a minimum transverse energy of 1 GeV and if it matched to a track with at least four hits within the silicon tracking detectors. In the second stage, EM topo-clusters near the seed candidates are identified as satellite cluster candidates and added to the seed candidates if they satisfy the necessary selection criteria, in order to form the final superclusters. A given cluster cannot be used as a seed cluster anymore if it has already been considered as a satellite cluster and added to another seed cluster [195].

Afterwards, initial energy and position calibrations are applied to the superclusters. Tracks of charged particles within the inner detector are matched to the superclusters similarly to the procedure applied for the EM topo-clusters. Since electron and photon superclusters are build independently, both an electron and a photon candidate can be built from a given seed cluster. In the case that a particular object was reconstructed from a supercluster which has a track of good quality attached and no photon conversion vertices, the object is identified as an electron candidate and only this electron object is created. In contrast, if the cluster has neither any good tracks attached nor photon conversion vertices, the only object built from the respective cluster is a photon candidate. If neither of the two cases applies, both an electron and a photon object are created and explicitly marked ambiguous. The final classification is based on a series of requirements specific to each analysis, including both the existence and the number of silicon hits as well as of conversion vertices [195].

Inaccuracies in the energy measurement due to the material in front of the calorimeter and the choice of a predefined cluster size are accounted for by calibrating the cluster energy based on simulated $Z \to e^+e^-$ events. Uniformity corrections are applied to data in order to equalise the response of the longitudinal EM calorimeter layers between data and Monte Carlo simulations. Residual disagreements are corrected for with in-situ measurements of the energy scale within $Z \to e^+e^-$ events. Simulated events are calibrated by efficiency scale factors which depend on the cluster energies and pseudorapidities derived from reference measurements of $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ events. The electron four-momentum is constructed from the fully calibrated cluster energy as well as from the momentum and the angular information of the associated tracks [195, 201–203].

The tracks of so-called prompt electron candidates are required to be compatible with the primary vertex in order to ensure their association with the hard-scatter interaction of interest. Prompt electrons are typically referred to as signal within the object reconstruction and identification procedure, whereas contributions from non-prompt electrons, such as the conversion of a photon in an electron-positron pair or electrons from secondary vertices, are deemed to be background. In

order to minimise contributions of non-prompt electron candidates, the longitudinal distance of the electron candidate with respect to the PV, with polar angle θ and with transverse impact parameter d_0, is required to satisfy both $|z_0 \cdot \sin\theta| < 0.5\,\text{mm}$ and $|d_0|/\sigma(d_0) < 5$. The quantity $|d_0|/\sigma(d_0)$ is referred to as the transverse impact parameter significance, where $\sigma(d_0)$ is the uncertainty in the measured value in the numerator [195, 201].

The candidates for prompt electrons have to pass various identification criteria which are designed to discriminate from non-prompt electron candidates not associated with the primary vertex. In addition, jets that left signatures in the detector similar to those of an electron and which were thus misidentified as electrons, referred to as fake electrons, add to the fraction of non-prompt electron candidates. For the identification of prompt electron candidates a likelihood (LH) method is employed which takes properties from both the tracks and calorimeter clusters as inputs. Information about the reconstructed electrons related to the shapes of the electromagnetic shower in the electromagnetic calorimeter, the tracking and the track-to-cluster matching are simultaneously evaluated and combined into a single discriminant,

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B} \quad . \tag{5.2}$$

Signal and background probability density functions, $P_{S/B}$, of the discriminating variables are used to calculate an overall probability for an object to be identified as signal or background, parametrised by the likelihood functions $\mathcal{L}_{S/B}$ which are defined as

$$\mathcal{L}_{S/B} = \prod_{i=1}^{N} P_{S/B,i}(x_i) \quad . \tag{5.3}$$

The x_i represents the value of the ith input variable out of in total N input quantities. Four well-defined working points (WP) for electron identification are provided in ATLAS which are referred to, in order of increasing background rejection, as VeryLooseLH, LooseLH, MediumLH, and TightLH[(1)]. The working points are constructed such that the set of objects selected by a certain WP is a strict subset of the set of objects passing the proximate WP with looser criteria. The LH discriminant is defined identically for the various working points which differ only in terms of the cut applied to this discriminant. A looser WP allows for a higher efficiency for selecting true signal electrons, but allows a higher number of fake or non-prompt electrons to be selected as well, and vice versa. The choice of a particular working point therefore corresponds to a desired trade-off between these contrasting effects. Depending on the choice of the working point, the signal efficiencies for selecting true electron candidates with a given value of the transverse energy of 25 GeV are in the range of 78 % to 90 % As the shower shapes are dependent on the energy and pseudorapidity values of the electron candidates, the identification working points have slightly different efficiencies depending on the respective location of an electron candidate in the detector. Electrons considered in the studies presented in this thesis have to pass the MediumLH quality requirements. However, for the overlap removal discussed in Sec. 5.6, electron objects are required to fulfil the looser criteria imposed by the LooseAndBLayerLH WP. This working point is identical to the Loose WP except for the fact that in addition at least one hit in the innermost layer of the pixel detector of the primary track is required [200, 203].

The prompt production of an electron from a semi-leptonic decay of a hadron, hadrons misidentified as leptons or photons, as well as the conversion of a photon into an electron-position pair can degrade the resolution of the electron reconstruction. Therefore, in addition to identification

(1) "LH" stands for "likelihood".

criteria, electrons are required to be sufficiently isolated from other reconstructed objects.For this purpose, two isolation variables were used of which one, referred to as $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$, uses tracking data, whereas the other, $E_{\mathrm{T}}^{\mathrm{varcone0.2}}$, uses information from the calorimeter. The track-based $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ is defined as the scalar sum of the transverse momenta of all associated tracks within a cone of

$$\Delta R = \min\left(0.2, 10\,\mathrm{GeV}/p_{\mathrm{T}}\right) \tag{5.4}$$

around the electron candidate, excluding the primary track of the electron candidate and tracks from background effects such as converted bremsstrahlung photons. Similarly, $E_{\mathrm{T}}^{\mathrm{varcone0.2}}$ consists of the sum of transverse energies of all clusters associated to the electron candidate within a cone of $\Delta R = 0.2$ around the seed cluster. Both the relative amount of the transverse momentum and the transverse energy within the angular cone, $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$, with respect to the total transverse momentum is required to be below $6\,\%$,

$$p_{\mathrm{T}}^{\mathrm{varcone0.2}}/p_{\mathrm{T}} < 0.06 \tag{5.5a}$$

$$E_{\mathrm{T}}^{\mathrm{varcone0.2}}/p_{\mathrm{T}} < 0.06 \quad , \tag{5.5b}$$

for the electrons considered in this thesis in order to be selected after the overlap removal, which will be described in Sec. 5.6 [195, 201]. Furthermore, electron candidates must feature a minimum p_{T} of $7\,\mathrm{GeV}$. The corresponding calorimeter clusters have to be within a range of $|\eta^{\mathrm{cluster}}| < 2.47$, though electron candidates situated in the transition region between the barrel and end-cap calorimeters of $1.37 < |\eta^{\mathrm{cluster}}| < 1.52$, referred to as crack-region, are excluded. A summary of the requirements posed on electron objects used for this thesis is given in Tab. 5.1.

		Electrons	Muons
Acceptance		$p_{\mathrm{T}} > 7\,\mathrm{GeV}$ $\lvert\eta^{\mathrm{cluster}}\rvert < 2.47$ except $1.37 < \lvert\eta^{\mathrm{cluster}}\rvert < 1.52$	$p_{\mathrm{T}} > 7\,\mathrm{GeV}$ $\lvert\eta\rvert < 2.5$
Quality	(before OR)	LooseAndBLayerLH	Medium
	(after OR)	MediumLH	Medium
Isolation	(track-based)	$p_{\mathrm{T}}^{\mathrm{varcone0.2}}/p_{\mathrm{T}} < 0.06$	$p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.06$
	(calorimeter-based)	$E_{\mathrm{T}}^{\mathrm{varcone0.2}}/p_{\mathrm{T}} < 0.06$	–
Impact parameters		$\lvert d_0\rvert/\sigma(d_0) < 5$ $\lvert z_0 \cdot \sin\theta\rvert < 0.5\,\mathrm{mm}$	$\lvert d_0\rvert/\sigma(d_0) < 3$ $\lvert z_0 \cdot \sin\theta\rvert < 0.5\,\mathrm{mm}$

Tab. 5.1.: Summary of the object definitions for electrons and muons. The electron identification criteria are looser before the overlap removal (OR) procedure than those afterwards. In contrast, for muons a single working point is applied.

5.3. Muons

Muons candidates are reconstructed independently in the inner detector and the muon spectrometer and are combined afterwards to form a muon track. The reconstruction in the inner detector is identical to the procedure described for electrons in Sec. 5.2. Within the muon spectrometer

however, track segments are formed from hit patterns identified in each single muon chamber separately [204].

Inside the monitored drift chambers, tracks segments are formed by fitting straight lines to the measured hits. This yields particle trajectories in the bending plane of the muon spectrometer, whereas the trajectories in the orthogonal plane are determined within the resistive plate chambers and the thin gap chambers, respectively. Separate measurements of track segments are performed within the cathode strip chambers. After all segments that were identified separately in the various subdetectors, an overlap removal procedure prevents the repeated use of the same segments from different subsystems and layers of the muon spectrometer for more than one muon track candidate [204].

In order to combine the information from the separate reconstruction procedures in the inner detector and the muon spectrometer, four different algorithms are applied and, accordingly, four types of muon candidates can be distinguished, based on the available information from both detectors. If a track can be fully reconstructed in both the inner detector and the muon spectrometer, so-called combined muons are built by extrapolating tracks from the muon system to the inner detector and matching to an ID track. However, in a small fraction of cases the extrapolation is executed reversely, starting in the ID and trying to find matching track segments in the MS. A global fit of the track is then performed to hits in both detector systems. If the signatures left in the muon spectrometer by a muon candidate are insufficient for a full reconstruction, muon candidates denoted as segment-tagged muons are reconstructed from a track in the inner detector which matches at least one track segment inside the MDT or CSC chambers. Such cases occur when a muon candidate passes only one layer of the muon system because of its low p_{T} or if it is measured in a phase space region with a reduced acceptance. Calorimeter-tagged muons consist of tracks in the inner detector which can be matched to energy deposits in the calorimeter system that are compatible with signatures from minimally ionising particles. Such muon candidates have thus a lower purity with respect to track-based reconstructed muons. However, due to the big pseudorapidity range covered by the calorimeter system, calorimeter-tagged muons provide the possibility to compensate for regions with a lower acceptance in the muon spectrometer. The so-called extrapolated muons are built with the help of information from the MS only. The tracks have to originate from a point compatible with the collision point. Furthermore, hits in at least two layers of the muon spectrometer in the barrel region and at least three layers in the end-caps are required. The muon acceptance is thereby extended into the pseudorapidity region of $2.5 < |\eta| < 2.7$ which exceeds the acceptance of the inner detector. In order to resolve possible overlaps between the different types of muon candidates, ID tracks matchable to more than one type are assigned based on a hierarchy. Preference is given to combined muons first, then to segment-tagged muons and finally to calorimeter-tagged muons. The overlap removal for extrapolated muons is executed separately as it relies on additional criteria such as the track fit quality and the number of hits [204].

The muon transverse momentum is calibrated in order to correct for observed differences between data and Monte Carlo simulation. As in the case of electrons, the calibration weights for the momentum scale and resolution in both the inner detector and the muon spectrometer are derived from reference measurements in $Z \to \mu^+\mu^-$ and $J/\phi \to \mu^+\mu^-$ events, and are functions of the transverse momentum itself and the pseudorapidity of the muon candidate. The transverse muon momentum is then evaluated from the combination of the weighted average of the corrected momenta from both detector systems [204].

The decay of hadrons such as pions and kaons is an important background for the identification of muons since those decays causes kinked tracks and induce significant differences between the transverse momentum measurements in the inner detector and in the muon spectrometer, respectively. A set of quality criteria is therefore applied to the muon candidates in order to target only

prompt muons from the hard-scatter interaction. Based on the ratio of the charge and the transverse momentum of a muon candidate, Q/p_{T}, the absolute value of the difference between the momentum measurements inside the inner detector and the muon system, the number of hits in the respective detector layers and the fit quality of the combined track fit, five muon identification working points are defined. These working points are denoted as Loose, Medium, High, Low-p_{T} and High-p_{T}. Similarly to the case of electron identification, the probability to select background decreases with tighter identification criteria, and the signal efficiency slightly diminishes at the same time. For the Loose, Medium and Tight working points, a muon passing tighter criteria must also have passed the looser requirements. Low-p_{T} and High-p_{T} working points have been explicitly developed to increase the identification efficiency for the respective momentum regimes. The muon candidates considered in this thesis are identified using the Medium WP which has an efficiency of $97\,\%$ for muons with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$. The muon efficiencies of the various working points are evaluated with simulated $t\bar{t}$ events and defined as the fraction of ID tracks associated with a reconstructed muon passing the Medium selection criteria [204, 205].

Muon isolation requirements are derived similarly to those for electrons based on the additional activity within a cone around the muon candidate. However, for muons considered in this thesis only track-based isolation criteria have been applied, using the isolation variable $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$ which is defined as the scalar sum of transverse momenta of all tracks not associated to the muon candidate under investigation within a cone size ΔR which depends on the transverse momentum

$$\Delta R = \min\left(0.3, 10\,\mathrm{GeV}/p_{\mathrm{T}}\right) \quad . \tag{5.6}$$

Due to the distinct signature muons leave in the detector, the cone size has been chosen to be slightly larger with respect to the electron case. In order to be sufficiently isolated, the p_{T}-weighted track-based isolation variable for a muon candidate has to be smaller than 0.06,

$$p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.06 \quad , \tag{5.7}$$

after the overlap removal to be described in Sec. 5.6 [204].

Muons candidates are required to have a minimum transverse momentum of $7\,\mathrm{GeV}$ and absolute pseudorapidity value satisfying $|\eta^{\mathrm{cluster}}| < 2.5$. Moreover, as the muon tracks need to be compatible with the primary vertex, the significance of the transverse impact parameter, defined in the same way as for the electron case, has to be $|d_0|/\sigma(d_0) < 3$ and the longitudinal distance of the muon candidate with respect to the PV is required to be $|z_0 \cdot \sin\theta| < 0.5\,\mathrm{mm}$, respectively. In addition, background from cosmic muons[(2)] is suppressed by vetoing muon candidates with $|z_0| > 1\,\mathrm{mm}$ and $|d_0| > 0.22\,\mathrm{mm}$ [204, 205]. A summary of the object definitions for muons is shown in Tab 5.1.

5.4. Jets

Due to colour confinement in the context of QCD, quarks and gluons hadronise and form cone-shaped showers of colourless particles, referred to as jets which serve as a proxy to the original partons for experimentalists. Each hadron induces a cascade of hadronic particles in the hadronic calorimeter which are used for the measurements of the jet energies.

[(2)] When high-energy cosmic rays hit the Earth's atmosphere, they create numerous secondary particles such as pions or kaons which further decay into muons. Those muons travel towards the Earth's surface without further decaying, due to the relativistic effect of time dilation. Since muons are minimally ionising particles, they penetrate the ground and can therefore be detected by the underground experiments at the LHC such as ATLAS. Cosmic muons are typically not compatible with particles that originate from the interaction point.

Various methods have been developed to reconstruct a jet and to separate the set of corresponding hadron showers out of thousands of calorimeter signals which can be categorised into two basic classes. The cone algorithms [206–208] perform an iterative top-down approach, starting with a cone of a defined radius around the energy deposit with highest measured value in an event. For each iteration the vector sum of the momenta of all particles enclosed by that cone is calculated. If the resulting direction is not compatible with the cone axis of the seed particle, it is taken as the new seed for the next iteration. In the other case, the cone is taken as a jet and all its constituents are removed from the iteration. Then, the procedure is repeated with the next most-energetic particle as seed for a cone. Such algorithms are typically infrared and collinear unsafe, which means that the emission of partons with low momentum and collinear splitting changes the outcome of the jet reconstruction noticeable. On top of that, the resulting jets are substantially dependent on accuracy of MC generators and resolution of the detector.

In contrast, sequential recombination algorithms [209] follow a bottom-up approach. Their defining parameter is the measure of distance between two entities in the η-ϕ-plane,

$$d_{ij}^2 = \min_{i,j} \left(k_{\mathrm{T},i}^{2p}, k_{\mathrm{T},j}^{2p} \right) \cdot \frac{\Delta R_{ij}^2}{R^2} \quad , \tag{5.8}$$

where $k_{\mathrm{T},i}$ and $k_{\mathrm{T},j}$ denote the transverse momentum of the jet constituents i and j, respectively, being considered for the particular step in the iteration. The distance in angular space between the two partons, ΔR_{ij}, is according to Eq. (3.11) defined as

$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \quad . \tag{5.9}$$

The radius parameter R defines the size of the cone which encloses all partons j that can be found within that cone around the entity i, and has to be given to the algorithm by the user. The algorithms iteratively combines all possible jet constituents and recalculates the distances between the remaining partons for each iteration, starting from the two entities with the smallest d_{ij}. This procedure stops if the distance d_{iB},

$$d_{iB} = k_{\mathrm{T},i}^{2p} \quad , \tag{5.10}$$

between the ith parton and the beam axis is smaller than its distance to any other entity j, calling i an jet an removing it from the list of entities. Depending on the choice of the parameter p, different variations of the family of sequential recombination algorithms have been defined. The k_{T}-algorithm [210, 211] corresponds to $p = 1$ and is hierarchical in the relative transverse momentum of the constituents, whereas the Cambridge-Aachen algorithm [212, 213] has $p = 0$ is therefore momentum-independent. The current standard choice for the majority of ATLAS analyses, including those described in this theses, features $p = -1$ and is referred to as anti-k_{T} algorithm [214]. It has the advantage of being both collinear and infrared safe [215] and is efficient in computing time. Due to the minimum requirement in Eq. (5.8), the jets grow around hard seeds and are grown with an almost circular profile.

In ATLAS jets can be built from different types of constituents. The jets used for the studies in the 3ℓ and 4ℓ channels of $t\bar{t}Z$ are reconstructed from topo-clusters of calorimeter cells and are therefore called topological jets [196]. Seeds are localised in the hadronic calorimeter as single calorimeter cells if the cell significance exceeds a threshold of four. Neighbouring cells from the same calorimeter layer and in adjacent layers are added to the cluster if their cell significance exceeds half of the threshold. If multiple local maxima in the energy distribution of the previously formed cluster are found, the seeds are merged to a single cluster. The energy of the cells is measured as if the deposits were induced electromagnetic showers without taking into account appropriate corrections for hadronic shower. The topological clusters constructed from all these

input cell energies are then used to build jet candidates using the sequential anti-k_{T} algorithm with a radius parameter of $R = 0.4$. Tracks measured in the inner detector are matched with the jet candidates employing a method referred to as ghost-association [216]. A set of so-called ghost-tracks with the same angular information as the measured tracks but with infinitesimal momentum is added to the list of topological clusters to be combined to jets. In this way, the list of tracks that are associated with a jet is obtained without altering the result of the anti-k_{T} reconstruction algorithm. Contributions to the jet energy from additional pile-up activity are subtracted both on a per-jet and per-event level by means of so-called jet areas [217], which are a measure for the susceptibility of a given jet to radiation such as pile-up or underlying event, and are determined via ghost-association. An evenly-distributed sea of ghost particles is added to the set of jet constituents to be clustered by the jet reconstruction algorithm. From the number of ghosts that have been included in the jet its volume can be determined.

An alternative approach to topological jets has been developed, referred to as particle-flow [218], which picks up the idea of the ghost-association of ID tracks to clusters within the calorimeter system. For the particle-flow jets that are used within the studies in the 2ℓOS channel, the measurements from both the inner detector and the calorimeter are combined to form signals representing individual particles. The energy deposited in the calorimeter by all charged particles is subtracted and the jet reconstruction is then performed on an ensemble of particle-flow objects consisting of the remaining calorimeter energy and tracks which are matched to the hard-scatter interaction. The advantage of this method is that by combining the two subdetector systems, the strength of the tracker in reconstructing charged particles are complemented by the ability of the calorimeter to reconstruct both charged and neutral particles in order to obtain an optimal reconstruction. At low energies, the tracker has the better resolution, whereas at high energies the calorimeter resolution is superior. Furthermore, outside the geometrical acceptance of the ID only calorimeter information is available, so it is necessary for the algorithm to carefully decide which energy measurement to be used in order to avoid double counting. In advance of the actual formation of particle-flow objects, a cell-based subtraction algorithm removes possible overlaps between momentum and energy measurements in the inner detector and the calorimeters, respectively. Well-measured tracks are matched to single topological clusters based on their compatibility in angular space and the expected energy deposited by the particle which also created the track is computed based on the cluster position and the track momentum. For each track-cluster system, the probability that the particle has deposited its energy in more than one cluster is evaluated and, if necessary, additional clusters are added to the track-cluster system in order to recover the full shower energy. Then, the expected energy is subtracted from the set of matched clusters cell-by-cell. If the remaining energy in the track-cluster system is consistent with the expected shower fluctuations of the signal created by a single particle, the remnant clusters are removed, and vice versa. After this final step, the set of selected tracks and the remaining topo-clusters in the calorimeter should ideally represent the reconstructed objects of an event without double counting of energy between the subdetectors. Hence, they are used as input to the anti-k_{T} algorithm to build jets with a radius parameter of $R = 0.4$ [218].

In order to account for differences in the detector performance for electromagnetically and hadronically interacting particles, the energy measurements within each single cell are calibrated to the energy scale of the electromagnetic calorimeter. Jet candidates are calibrated to correct for contributions from the underlying event and for residual detector effects. The jet energy scale (JES) calibration is a multi-stage procedure involving a correction of the origin of jet candidate to point into the direction of the PV instead of the centre of the detector, which does not affect the jet energy. Additional corrections are performed accounting for pile-up, based on the density of tracks and the number of simultaneous parton-parton scatter interactions. The jet four-momenta are cor-

rected such that their values agree with those from jets from MC simulated reference samples without the detector simulation applied. These corrections are dedicated to transitions between different calorimeter technologies, a varying detector response with respect to different pseudorapidity regions, and changes in the calorimeter granularity. The corresponding calibration factors are functions of both p_{T} and $|\eta|$. The particle composition as well as the shape of the hadronic showers are dependent on whether a jet was initiated by a quark or by a gluon [219–221]. The global sequential correction method [222] has been designed to reduce the impact of the parton flavour on detector response and jet reconstruction by using global jet properties such as the fraction of the measured energy in particular regions of the calorimeters or the track multiplicity. Finally, in-situ calibrations account for differences in the jet p_{T} measurement between real and simulated data using well-calibrated objects as a reference [223–225]. Such differences arise from imperfections in the description of the detector response and material within the simulation, electromagnetic and hadronic interactions with the detector, as well as from limitations in the modelling of the underlying event, jet formation and pile-up. Similarly, the jet energy resolution (JER) is calibrated using reference measurements of the p_{T} asymmetry in dijet events [224–226].

Jets which are not associated with the primary vertex of an event but arise from pile-up activity can be identified with a likelihood-based discriminant, referred to as jet-vertex-tagger (JVT). The JVT discriminant is constructed as a multivariate combination of two variables which include both tracking and calorimeter information and are designed to separate jets associated to the hard-scatter interaction from pile-up jets [227, 228].

The jets considered in this thesis have a minimum transverse momentum of $25\,\mathrm{GeV}$ and are required to be within $|\eta| < 2.5$. In addition, topological jets with a transverse momentum below $120\,\mathrm{GeV}$ must satisfy JVT > 0.59 in order to suppress pile-up contributions. Jets built with the particle-flow algorithm, however, are required to feature a value of the JVT discriminant greater than 0.2 if $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$. Tab. 5.2 provides an overview of the jet definitions including flavour tagging which is discussed in the next section, Sec. 5.5.

	Topological jets	Particle-flow jets
Acceptance	$p_{\mathrm{T}} > 25\,\mathrm{GeV}$ $\lvert\eta\rvert < 2.5$	$p_{\mathrm{T}} > 25\,\mathrm{GeV}$ $\lvert\eta\rvert < 2.5$
Jet-Vertex-Tagger	JVT > 0.59 if $p_{\mathrm{T}} < 120\,\mathrm{GeV}$ and $\lvert\eta\rvert < 2.5$	JVT > 0.2 if $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $\lvert\eta\rvert < 2.4$
b-tagging	MV2c10 algorithm Fixed $85\,\%$ WP and PCBT	DL1 algorithm Fixed $85\,\%$ WP

Tab. 5.2.: Summary of the object definitions for jets. For topological jets, the MV2c10 b-tagging algorithm is used employing both a fixed WP and a pseudo-continuous technique, wheres particle-flow jets are tagged with the DL1 algorithm at a fixed WP only.

5.5. Flavour tagging

Jets which have been initiated by a bottom quark, referred to as b-tagged jets or simply b-jets, can be identified and separated from jets originating from lighter quarks, gluons or hadronically decaying taus, based on the slightly different signatures in the detector of the respective jets. Bottom

quarks typically hadronise to B-hadrons which have a significant lifetime of $c\tau > 450\,\mu\text{m}$ [51] since the decay of a b-quark into a top quark is kinematically forbidden and the decay into a quark of a lighter generation is strongly suppressed due to the CKM matrix elements as explained in Sec. 2.1.4. Therefore, they travel a visible distance through the detector before further decaying at a secondary vertex which is displaced from the PV.

In order to identify b-jets, the so-called MV2c10 algorithm [229, 230] is applied within the studies concerning the $t\bar{t}Z$ decay channels with three or four leptons in the final state. This algorithm combines the outputs of three independent algorithms in a boosted decision tree which is trained to discriminate b-jets from jets initiated by charm quarks, referred to as c-jets, and other jet activity. Due to the large lifetime of B-hadrons, tracks generated by them tend to feature a large transverse and longitudinal impact parameter, which are defined as the closest approach of a track to the PV in the transverse plane and z-direction, respectively. This feature is exploited by the IP3D [229, 230] algorithm which combines both impact parameters into a likelihood-based discriminant. The likelihood-based SV1 [229, 230] algorithm aims to identify secondary vertices by building two-track vertices from all tracks matched to a jet and calculating the probability that those vertices originate from the decay of long-lived particles, photon conversion or hadronic interactions with the detector material. JetFitter [229, 230] is a decay chain multi-vertex reconstruction algorithm which exploits the topological structure of weak decays of charmed hadrons and B-hadrons inside jets and tries to reconstruct the full B-hadron decay chain, taking into account subsequent hadron decays.

A complementary method, referred to as the DL1 tagger [230, 231], was used for b-jet identification in the analyses of $t\bar{t}Z$ final states with two leptons. The DL1 tagger uses inputs from various flavour tagging algorithms The DL1 tagger incorporates all inputs used by the MV2c10 tagger as well as a number of additional inputs and combines them in a deep neutral network [232]. Several variables created by the SMu algorithm [230], which exploits the topology of muons coming from semi-leptonic decays of heavy-flavoured hadrons, and by RNNIP [230], an algorithm based on a recurrent neural network [233] which relies on properties of tracks associated to jets similarly IP3D, are employed as discriminating variables.

A jet is said to be deemed to have originated from a b-quark if the output values of the MV2c10 and DL1 algorithms are above a threshold of 0.11 and 0.41, respectively. Those values correspond to an $85\,\%$ efficiency for the MV2c10 tagger and an efficiency of $77\,\%$ for the DL1 tagger in order to select true b-jets, and were evaluated from simulated $t\bar{t}$ events and validated in data using a $t\bar{t}$-enriched selection [234]. The given b-tagging efficiencies correspond to a mistag[(3)] rate of $1/2$ for c-jets and $1/28$ for light-flavoured jets for the MV2c10 tagger applied to jets built from topological clusters only, whereas the mistag rates for the DL1 tagger in conjunction with jets built from particle-flow objects are $1/4$ and $1/141$ for c-jets and light-flavoured jets, respectively. Hadronically decaying taus can be rejected with an efficiency of $84\,\%$ by the MV2c10 and with an rejection efficiency of $94\,\%$ by the DL1 tagger, based on the b-tagging efficiency working points which were employed for the respective b-taggers.

In addition, the analyses in the 3ℓ and 4ℓ channels employ an alternative method, referred to as pseudo-continuous b-tagging (PCBT), in order to optimise the b-jet identification. Rather than using a single fixed working point for the categorisation of a whole event, different thresholds on the b-tagging discriminant out of a set of four working points with b-tagging efficiencies of $60\,\%$, $70\,\%$, $77\,\%$ and $85\,\%$ are applied to individual jets. A more detailed prescription of the choices of combinations b-tagging working points will be given later in Sec. 6.1.

Similar to other physics objects, simulated events are calibrated with scale factors dependent on the jet transverse momentum and pseudorapidity which are employed to account for discrepancies

[(3)] If a jet other than a b-jet was tagged as such, it is denoted mistagged.

in the b-tagging between data and Monte Carlo simulation. The scale factors are determined from different reference processes best suited for the respective flavour to tag on [234, 235].

5.6. Overlap removal

The object candidates are reconstructed by specialised algorithms tailored to identify the respective type of objects. Since the algorithms share information among each other from various subdetectors, the reconstruction and identification of particles not always gives an unambiguous result, for example an electron typically being reconstructed as a candidate for both an electron and a jet. In order to resolve ambiguities between objects that were reconstructed twice and to avoid double counting of the same physics object, a procedure referred to as overlap removal is employed after the reconstruction. If two object candidates are too close to each other, one will be removed in favour of the other and thus will no longer be considered within any further step.

Electron candidates which share a track in the inner detector with a muon candidate are discarded. In the case that two electron candidates are reconstructed with a common associated track in the inner detector, the electron candidate with lower transverse momentum is dropped. Furthermore, the closest jet to an electron candidate within an angular distance of $\Delta R < 0.2$ is removed in order to avoid double-counting of energy deposits in the calorimeters. However, if the distance between this jet and the electron candidate is $0.2 < \Delta R < 0.4$, the jet is rather kept and the electron candidate rejected instead. Background contributions from muons originating from meson or heavy-flavour hadron decays inside jets are suppressed by removing muon candidates which are separated from the nearest jet by $\Delta R < 0.4$. In case this jet has fewer than three associated tracks, the muon candidate is rather kept and the jet is dropped instead in order to preserve the high-energy muon candidate [236, 237].

5.7. Missing transverse energy

Neutral weakly-interacting particles, such as the neutrino, rarely interact with matter. Hence, such particles leave the detector without leaving any noticeable signatures and thus cannot be reconstructed. However, their kinematic properties an sometimes be inferred by employing momentum conservation. Since the incoming protons are composite particles with their net momentum exclusively aligned in the z-direction, the exact momentum fraction carried by the colliding partons in a particular event are not known. Therefore, at the LHC momentum conservation applies only in transverse plane in which the momenta of all visible and invisible particles have to add up to zero. Consequently, a significant imbalance in the sum of the transverse momenta of all visible particles implies that a particle must have left the detector unseen. This momentum imbalance, referred to as missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, is defined as the negative vector sum of all reconstructed objects surviving the overlap removal,

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} := -\sum_{i=1}^{N} \vec{p}_{\mathrm{T},i} \quad , \tag{5.11}$$

where N denotes the number reconstructed objects.

The reconstruction of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ can be divided into the sum over all hard objects and a so-called soft term. The hard component comprises all calibrated hard objects, which in the studies presented in this thesis are electrons, muons and jets. If one considered additional physics objects such as taus or photons, those would have to be taken into account for the calculation as well. Otherwise they enter the calculation either as electrons, jets, or via the soft term which consists of contributions

from tracks from the primary vertex which could not be associated to any reconstructed object. The soft term, in contrast, can derive from energy deposits in the calorimeters which includes contributions from electrically neutral particles, but this method is by far less robust with respect to pile-up. The precision of the determination of the missing transverse momentum can be increased by employing jet vertex tagging technique which suppresses the events to be polluted with jets originating from pile-up, as discussed in Sec. 5.4. In order to achieve the best possible $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ resolution, the objects used for the calculation have to be calibrated. However, the final resolution is limited due to detector inefficiencies and a finite resolution of the reconstructed objects [238–240].

In addition, ISR disrupts the momentum balance in the transverse plane. Instead of the vector quantity $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, its magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$, referred to missing transverse energy, and the angular information of the missing momentum vector in the transverse plane, denoted by the azimuthal angle $\phi_{\mathrm{T}}^{\mathrm{miss}}$, are commonly used in physics analyses.

With the method described the number of undetectable particles in an event cannot be determined as the calculation will always result in a single vector. This poses a challenge to analyses with more than one expected invisible particle, for example measurements of $t\bar{t}$ in the dileptonic channel, or in particular searches for new hypothetical particles predicted by BSM theories which are typically constructed to leave the detector unnoticed. In such cases, other techniques, which usually employ parts of or the entire event topology, have to be applied.

6. Studies for the $t\bar{t}Z$ cross-section measurement in the trilepton channel

The studies presented in Secs. 6.2 and 6.3 were performed in the context of the most recent measurement of the $t\bar{t}Z$ cross section at the ATLAS experiment [46], using the full Run 2 set of data collected by the ATLAS experiment in proton-proton collisions from 2015 to 2018 at a centre-of-mass energy of $13\,\mathrm{TeV}$ which corresponds to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. In advance of the sections dedicated to the studies, an outline of the analysis will be given in Sec. 6.1, highlighting the parts which are necessary in order to provide the relevant context to the above-mentioned studies.

6.1. Analysis overview

The production of a top-antitop quark pair is a relatively rare SM process compared to $t\bar{t}$ production and a source of different multilepton final states, as discussed in Sec. 2.4. Precise measurements of the $t\bar{t}Z$ cross section are therefore of particular interest. Aside from the inclusive measurement to determine the total $t\bar{t}Z$ production cross section, differential cross-sections measurements were performed as functions of several variables which are sensitive to the underlying kinematic properties as well as to subtleties in the production mechanism of the $t\bar{t}Z$ system, using the Iterative Bayesian Unfolding technique [241]. These differential cross-section measurements are sensitive to differences between the settings of various Monte Carlo generators, and thus could potentially provide an important input to the tuning of MC parameters.

In addition to the criteria specified in Chap. 5, the reconstructed objects used for the $t\bar{t}Z$ cross-section measurement have to fulfil certain quality requirements in order for the corresponding events not to be rejected. Events for which a certain subdetector was observed not to work at full functionality are removed. A given event must contain at least one electron or one muon that caused a trigger not to reject the respective event. The analysis focuses on the two most sensitive signatures with a high sensitivity to the $t\bar{t}Z$ process, namely on events with three or four light leptons, comprising electrons and muons[(1)], in the final state, referred to as trilepton and tetralepton channels, respectively. Based on these signatures, several phase-space regions were selected, intended to achieve the best sensitivity to $t\bar{t}Z$ production with either one or both top quarks featuring a leptonically decaying W boson, referred to as signal regions. They were optimised to feature a high signal purity and a low contributions from other processes. In order to restrict the uncertainties on the measurement due to limited statistics, a sufficiently high number of events has to be selected while keeping the signal-to-background ratio high. In total three signal regions targeting the trilepton or 3ℓ channel were defined for which the selection criteria are listed in Tab 6.1. Since the studies presented in the two subsequent sections focus on the 3ℓ channel, the definitions of regions corresponding to this channel are given here, whereas, the selection criteria chosen for the 4ℓ channel can be found in Ref. [46]. The minimum requirement on the transverse momentum on the leading, subleading and third lepton[(2)] is $27\,\mathrm{GeV}$, $20\,\mathrm{GeV}$

[(1)] In the frame of this thesis, only electrons and muons are considered as leptonic final states. Since their masses are small with respect to taus, they are grouped together as light leptons. Leptonically decaying taus are implicitly considered either as electron or muon.

[(2)] The physics objects of a certain type in a given event, for example electrons or jets, are for convenience arranged in descending order according to their transverse momentum. For instance the electron with the highest or second-

and 20 GeV, respectively, and the sum of the three lepton charges is required to be ± 1. Since the Z boson is expected to decay into a pair of electrons or muons, the opposite-sign-same-flavour (OSSF) lepton pair with the invariant mass closest to the Z boson mass is considered to originate from the Z boson. Additionally, its invariant mass, labelled as $m_{\ell\ell}$, has to be compatible with the Z boson mass within a range of 10 GeV. Furthermore, all OSSF lepton combinations that can be found in an event must have an invariant mass of greater than 10 GeV in order to remove potential contributions from events featuring low-mass resonances. Only those jets satisfying $p_{\mathrm{T}} > 25$ GeV are considered for a given event. The three signal regions differ only by their values of the requirements imposed on the multiplicity of reconstructed jets and b-jets which are given in Tab 6.1. For the differential measurement, the b-jet requirement is fixed to an 85 % efficiency WP, requiring at least two b-tagged jets for the 3ℓ-Z-2b3j signal region, which in the following will be denoted as differential signal region. In contrast, the inclusive measurement is performed with a combination of two orthogonal regions employing the pseudo-continuous b-tagging technique: 3ℓ-Z-1b4j-PCBT with a tight WP of 60 % to identify a single b-jet, and 3ℓ-Z-2b3j which features two b-jets at 70 % efficiency. In order to ensure these two regions to be orthogonal, a veto is applied to events which pass the 60 % requirement featuring an additional b-jet tagged with 70 % efficiency. The different approaches are intended on the one hand to reduce the contributions of the WZ+jets background, which becomes dominant within the signal regions if one did not apply the PCBT technique, for the inclusive measurement, and on the other hand to achieve sufficient statistics for the differential measurement which motivates the choice of a looser b-tagging WP. In addition to the definitions of the specialised signal regions, the so-called preselection level is defined by the requirements of exactly three light leptons and at least two jets with a minimum transverse momentum of 25 GeV imposed on top of the quality requirements described above.

Variable	3ℓ-Z-1b4j-PCBT	3ℓ-Z-2b3j-PCBT	3ℓ-Z-2b3j	3ℓ-WZ-CR
$N_\ell (\ell = e, \mu)$	3	3	3	3
	≥ 1 OSSF pair with $\lvert m_{\ell\ell} - m_Z \rvert < 10$ GeV			
	for all OSSF combinations: $m_{\ell\ell} > 10$ GeV			
$p_{\mathrm{T}}(\ell_1,\ \ell_2,\ \ell_3)$	> 27 GeV , > 20 GeV , > 20 GeV			
N_{jets}	≥ 4	≥ 3	≥ 3	≥ 3
$N_{b\text{-jets}}$	1@60 %	≥ 2@70 %	≥ 2@85 %	0@85 %
	veto add. b-jets @70 %			

Tab. 6.1.: Definitions of the 3ℓ signal regions and the WZ control region to constrain the WZ+jets background. For the inclusive cross-section measurement, a combination of the two signal regions with PCBT and the WZ control region is employed, whereas for the differential measurement only the 3ℓ-Z-2b3j signal region with a fixed b-tagging working point is used [46].

Several physics processes described by the Standard Model other than $t\bar{t}Z$ can lead to final states very similar to $t\bar{t}Z$ and thereby contaminate the signal regions. Such processes are referred to as backgrounds and can be subdivided into so-called prompt and non-prompt contributions. In addition to the signal regions, the definition for a region intended to be dominated by the prompt

highest transverse momentum in an event is referred to as the leading or subleading electron, respectively. The analogous p_{T}-based ordering is used for jets.

WZ+jets background with only a small amount of contamination from signal events, referred to as WZ+jets control region, is given in Tab 6.1. Since WZ+jets production suffers from large theoretical uncertainties, a dedicated WZ control region was introduced in order to help use the measured data to constrain the production rates. The definition of the WZ control region is equal to those for the signal regions except for the requirements on the jet and b-jet multiplicities. The same ATLAS analysis employs a similar control region to constrain the ZZ+jets background in the case of the tetralepton channel. In addition, various background processes that comply with the requirements from Tab. 6.1, such as the associated production of single top quarks with one or two heavy vector boson (tWZ, tZq), are included. Other processes such as the associated production of a $t\bar{t}$ pair with a Higgs or W boson ($t\bar{t}H$ or $t\bar{t}W$, respectively), ZZ+jets and those listed in Sec. 4.1.3 only contribute to a limited extent and are therefore estimated with Monte Carlo simulated samples, just as tWZ and tZq.

Background contributions from non-prompt and so-called fake[(3)] leptons, however, are estimated employing a data-driven, so-called matrix method [242, 243]. The matrix method relies on a differing response of the prompt and fake leptons to the identification, isolation and impact parameter requirements imposed during the object reconstruction procedure. Thus, two sets of data events with the same selection criteria as for the signal regions but with modified lepton identification and isolation requirements are employed. Leptons passing the object criteria required in the signal regions are referred to as tight leptons, whereas leptons which fail those requirements are denoted as loose leptons. For loose leptons, no isolation requirements are applied, and the identification criterion for loose electrons is relaxed to the working point that is used before the overlap removal procedure. The simplified case of a single lepton can be used to illustrate the general concept of the matrix method mathematically: The total number of fake leptons that enter the signal regions, N_F, can be evaluated by inverting the matrix of a linear system of equations,

$$\begin{pmatrix} N_T \\ N_L \end{pmatrix} = \begin{pmatrix} \varepsilon & \chi \\ (1-\varepsilon) & (1-\chi) \end{pmatrix} \cdot \begin{pmatrix} N_R \\ N_F \end{pmatrix} \quad , \tag{6.1}$$

where N_T and N_L denote the observed number of events with a lepton passing or failing the tight object definitions, respectively. They are related to the number of events with a real or fake lepton, N_R and N_F, by the so-called real lepton efficiency ε and fake lepton efficiency χ. These two values have to be determined separately. In case of the $t\bar{t}Z$ 3ℓ channel it is obtained from data events with exactly two leptons with the same electric charge and at least one b-jet tagged with the $85\,\%$ WP, indicating the probabilities of a real lepton and a fake lepton satisfying the tight lepton requirements, respectively. The matrix is inverted via the maximisation of a likelihood function [244] which consists of a product of Poisson probability functions. The observed number of leptons selected with looser criteria, N_L, and of prompt leptons, N_T, which satisfy the nominal lepton requirements are fixed, while the expectation values of the Poisson functions, the number of real and fake leptons in the signal region, N_R and N_F, are obtained from likelihood maximisation. The single-lepton case shown in Eq. (6.1) can be extended to an arbitrary number of leptons. For more than one lepton, the estimate for fake leptons then comprises all possibilities for a fake lepton

[(3)] Objects which were unintentionally misidentified as light leptons are typically referred to as fakes or fake leptons. They can originate from meson decays, photon conversion, or light jets that accidentally create detector signatures similar to those from a light lepton. In contrast, non-prompt leptons typically originate from heavy-flavour hadron decays. If not stated otherwise, fake and non-prompt leptons will not be distinguished anymore in the following but commonly denoted as fake leptons or fakes.

to occur in an event, and reads in case of the $t\bar{t}Z$ 3ℓ channel:

$$\begin{aligned} N_F = & N_{FFF} \\ & + N_{FFT} + N_{FTF} + N_{TFF} \\ & + N_{FTT} + N_{TFT} + N_{TTF} \quad . \end{aligned} \tag{6.2}$$

The formula explicitly takes into account the order of the leptons, such that for example the number of events with exactly one fake lepton is given by the sum of all summands in the third line of Eq. (6.2). The expected contributions from fake leptons are subtracted from the data sample and considered as a distinct background process. Uncertainties related to the evaluation of real and fake rates are quantified and translated into an associated systematic uncertainty which has to be taken into account in the cross-section measurement.

The production cross section of the $t\bar{t}Z$ process is extracted by performing a simultaneous maximum-likelihood fit [139, 245] to the number of events in the trilepton and tetralepton signal regions, as well as the WZ+jets and ZZ+jets background control regions. For events from the 3ℓ channel, only the two PCBT signal regions dedicated to the inclusive measurement are included in the fit. The ratio between the measured value of the inclusive $t\bar{t}Z$ production cross section and its corresponding Standard Model prediction, referred to as signal strength $\mu_{t\bar{t}Z}$, as well as the normalisation factors of the WZ+jets/ZZ+jets backgrounds used to extrapolate the corresponding event yields into the signal regions, are free parameters of the fit. In contrast, the contributions from the other backgrounds are fixed at their respective expected value obtained from Monte Carlo simulation and are allowed to only vary within the associated uncertainties. Since the fit is based on the so-called profile-likelihood technique, the corresponding likelihood function is composed of the product of Poisson probability functions given by the observed event yields in the signal and control regions. The sources of experimental, theoretical and statistical uncertainties related to the considered objects and physics processes are accounted for in the fit by nuisance parameters which are constrained by Gaussian functions. The measured values of $\mu_{t\bar{t}Z}$ from the individual fits in the trilepton and tetralepton regions, as well as in the combination of all regions, are converted to an inclusive measurement of the $t\bar{t}Z$ cross section using the theoretically predicted value of $0.88\,\text{pb}$ [172, 173] from the nominal MC simulated signal sample, calculated for a centre-of-mass energy of $13\,\text{TeV}$. The uncertainties on this theoretical value originate from variations of the factorisation and renormalisation scale, as well as of the PDF and α_s. The inclusive $t\bar{t}Z$ production cross section in proton-proton collisions was measured to be

$$\sigma(pp \to t\bar{t}Z) = 1.05 \pm 0.05\,(\text{stat.}) \pm 0.09\,(\text{syst.})\,\text{pb} = 1.05 \pm 0.10\,\text{pb} \quad , \tag{6.3}$$

which is within the uncertainties in good agreement with the Standard Model prediction given in Eq. (2.56),

$$\sigma_{t\bar{t}Z}^{\text{NLO+NNLL}} = 0.864^{+0.07}_{-0.09}\,(\text{scale}) \pm 0.03\,(\text{PDF} + \alpha_s)\,\text{pb} \quad , \tag{6.4}$$

calculated for a centre-of-mass energy of $13\,\text{TeV}$ at NLO and NNLL accuracy in the calculation considering both electroweak and strong corrections [87].

Further details on this measurement as well as of the differential cross-section measurements can be found in Ref. [46].

6.2. Comparison of Monte Carlo generators

Systematic uncertainties on a given measurement from the choice of the Monte Carlo generator are typically estimated by assessing the impact on the measurement of simulated samples of the same

physical process, each generated with a different set of Monte Carlo algorithms. As discussed in Sec. 4.1.3, a so-called nominal sample is employed to perform the various measurements. In order to derive the systematic uncertainties for the description of the hard-scatter process, a different Monte Carlo generator is used for the modelling while the parton shower is simulated using the same program for both samples, and vice versa. In this section such systematic uncertainties for the signal process and for the important tWZ background are investigated.

6.2.1. Comparison of matrix element generators for simulated ttZ production

The nominal sample employed to model the $t\bar{t}Z$ process was produced with the MADGRAPH5_AMC@NLO generator interfaced to with PYTHIA 8 to simulate both the hadronisation process and the parton shower, as discussed in Sec. 4.1.3. Two additional sets of simulated samples were generated with SHERPA 2.2.1 at NLO accuracy in the description of the hard-scatter interaction, where one of which features the radiation of multiple partons within the calculation of the matrix element, referred to as the multileg[(4)] sample. Although this technique is computationally intensive, it provides the most precise description of the $t\bar{t}Z$ process so far available. However, due to differences in the implementation of the MADGRAPH5_AMC@NLO and SHERPA generators the content of the samples produced with either of them were seen to slightly differ from each other in terms of the number and kinematics of generated particles. A validation was therefore necessary in order to facilitate the use of the SHERPA samples for the analysis.

The comparison of nominal and alternative samples was performed both in terms of yields within the defined phase-space regions as well as graphically. For convenience only the signal region designed to perform the differential cross-section measurement and the WZ control region were used. The weighted yields within the selected regions before any fit was performed are shown in Tab. 6.2, scaled to the theoretical cross sections of the respective samples and to the integrated luminosity of the full Run 2 ATLAS dataset. As the SHERPA samples only feature $t\bar{t}Z$ final states with a leptonically decaying Z boson, events with a Z boson decaying to a pair of quarks or neutrinos from the inclusive nominal dataset were not considered. The numbers from

	3ℓ-Z-2b3j	3ℓ-WZ-CR
MADGRAPH5_AMC@NLO+PYTHIA 8	354.87 ± 1.41	43.29 ± 0.51
SHERPA 2.2.1 NLO	380.42 ± 0.84 (7.2 %)	37.11 ± 0.23 (14.2 %)
SHERPA 2.2.1 NLO multileg	402.84 ± 1.20 (13.5 %)	39.12 ± 0.41 (9.6 %)

Tab. 6.2.: Expected yields in the 3ℓ differential signal region and WZ control region, stated for the nominal MADGRAPH5_AMC@NLO+PYTHIA 8 and SHERPA samples separately. The values quoted in parentheses correspond to the relative difference with respect to the MADGRAPH5_AMC@NLO+PYTHIA 8 sample. The uncertainties on the yields correspond to the Monte Carlo statistical error.

both SHERPA samples exceed those from the nominal sample by 7 % to 14 % in the signal region, whereas in the WZ control region the yields of both SHERPA samples are 10 % to 14 % lower.

A comparison of the shapes of some kinematic variables was performed by means of the distributions of the various signal samples which have been normalised to the respective number of

[(4)] A line in a Feynman diagram is often called a leg. The term multileg refers to so-called external legs which correspond to the incoming and outgoing particles of an interaction. The greater the number of outgoing particles originating from higher-order corrections are considered in the calculation of the matrix element, the more precise the description of the hard-scatter process.

entries, as illustrated in Fig. 6.1. Each row corresponds to a particular variable: the jet multiplicity in the upper row and the leading lepton transverse momentum in the lower row. The different lines in the figures represent the three respective simulated signal samples. The subfigures on the left show the distributions for the differential signal region, and those on the right side the same variables for the WZ control region. In the lower panels below the main figures the ratio of the yields from the SHERPA samples with respect to the nominal sample is depicted. As for the case of the event yields summary, the error bars show the Monte Carlo statistical errors. The distributions

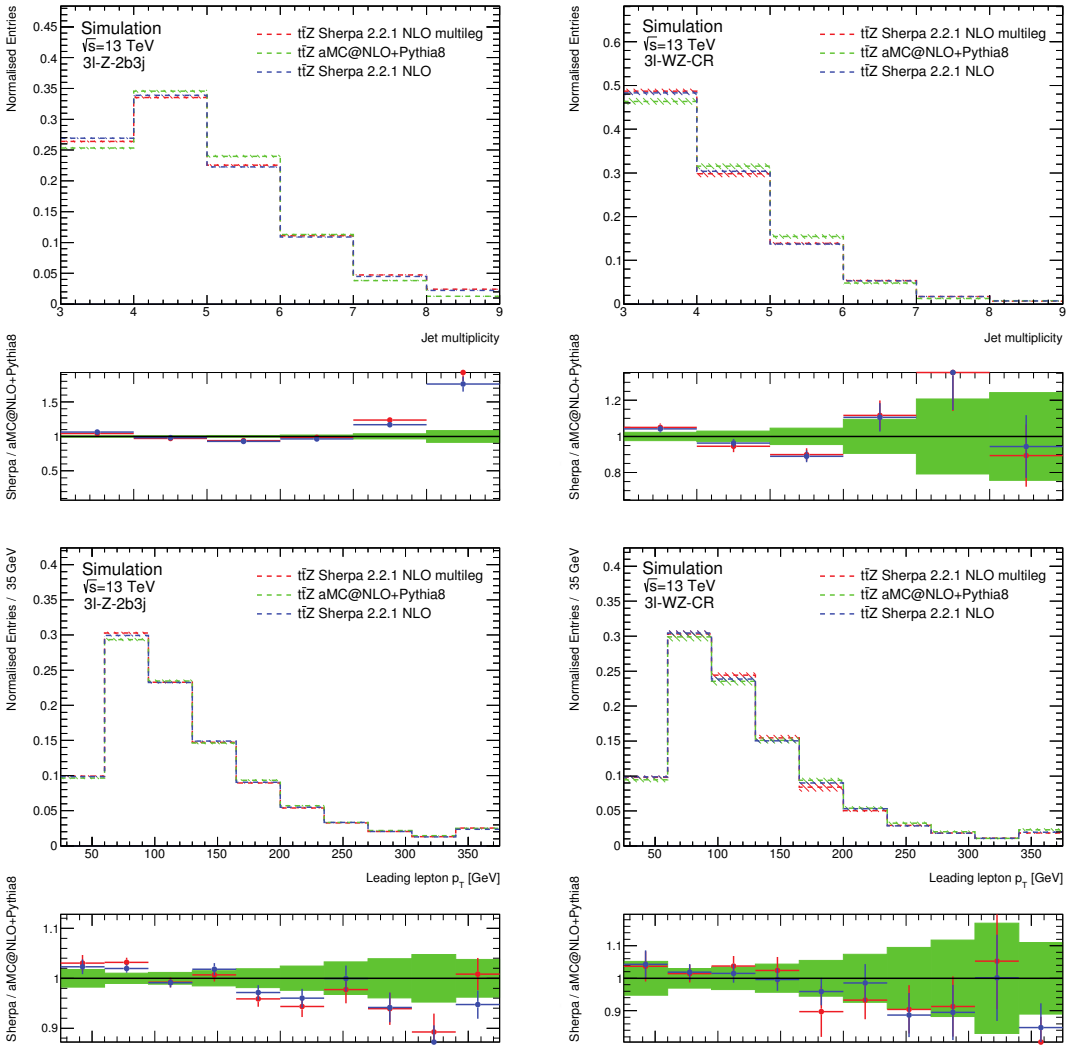


Fig. 6.1.: Normalised distributions of the jet multiplicity in the top row and the leading lepton transverse momentum in the bottom row, shown for the differential signal region in the left column and the WZ control region on the right side, respectively. The coloured lines correspond the various $t\bar{t}Z$ Monte Carlo samples. The pads below the normalised distributions show the ratios of either SHERPA sample to the nominal signal model. The error bars in all subfigures are statistical only.

show a good level of agreement between the two SHERPA samples. Differences between the nominal and the alternative samples are of the order of 10 % to 20 % and within statistical uncertainties most of the time. In the case of a few bins in more extreme phase-space regions of either a high

jet multiplicity or at high values of transverse momentum the differences are seen to be larger.

The checks showed the alternative $t\bar{t}Z$ samples to be reasonably consistent with the nominal signal model and the differences to be within the expected range. Both SHERPA samples were then used to show different Monte Carlo predictions for the differential results.

6.2.2. Comparison of diagram removal techniques for simulated tWZ productions

The tWZ background was modelled at NLO accuracy using the combination of the MADGRAPH5AMC@NLO+PYTHIA 8 Monte Carlo generators. During the generation of the simulated sample the diagram removal (DR) technique was employed to prevent overlaps between $t\bar{t}Z$ and tWZ. Although $t\bar{t}Z$ and tWZ are the processes of interest in the context of these studies, the DR techniques can be more clearly illustrated by the analogous processes of Wt and $t\bar{t}$ production, respectively. The production of a single top quark in association with a W boson, Wt may interfere with the production of $t\bar{t}$ and tWb, which describe NLO contributions to the Wt process with additional b-quarks. The relevant Feynman diagrams are shown in Fig. 6.2. The top-antitop

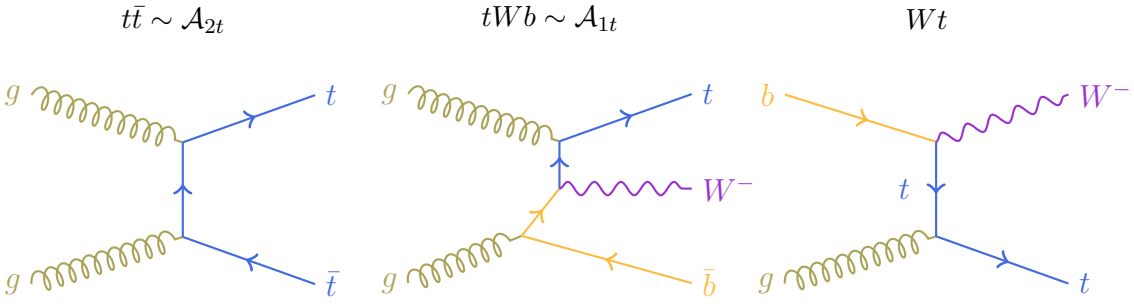


Fig. 6.2.: Feynman diagrams for the $t\bar{t}$ production on the left, the production of tWb in the middle, and for the electroweak production of a single top quark in association with a W boson on the right. Both the left and middle diagram represent a NLO contribution to the Wt process. The additional b-quark for Wt is likely produced within the LHC environment.

quark production involves the production of two on-shell quarks each decaying to a W boson and a b-quark with its matrix element denoted by $\mathcal{A}_{2t}$. In contrast, the tWb process features only one on-shell top quark and is therefore labelled with $\mathcal{A}_{1t}$. The squared matrix element for tWb production, which enters the calculation of the Wt cross section at NLO accuracy, can thus be written as

$$\begin{aligned} |\mathcal{A}_{tWb}|^2 &= |\mathcal{A}_{1t} + \mathcal{A}_{2t}|^2 \\ &= |\mathcal{A}_{1t}|^2 + 2\Re\left(\mathcal{A}_{1t}\mathcal{A}_{2t}^*\right) + |\mathcal{A}_{2t}|^2 \end{aligned} \tag{6.5}$$

where $\mathcal{A}_{2t}^*$ represents the complex conjugate of the $t\bar{t}$ amplitude. In order to subtract the contribution from the second top quark, the nominal sample employs the first version of diagram removal that was developed, referred to as DR1, where $\mathcal{A}_{2t}$ is simply set to zero such that both $t\bar{t}$ production and the interference between $t\bar{t}$ and tWb is cancelled in the production of the simulated samples. In order to estimate the impact of systematic uncertainties from the application of the diagram removal procedure, an alternative sample was employed which was generated with the same set of Monte Carlo generators as the nominal sample but features a different version of the diagram removal technique, referred to as DR2. In the case of the DR2 technique, rather than completely preventing the generation of $t\bar{t}$, only the square of the corresponding amplitude, $|\mathcal{A}_{2t}|^2$, is set to zero such that the interference term in Eq. (6.5), $2\Re\left(\mathcal{A}_{1t}\mathcal{A}_{2t}^*\right)$, is kept.

A comparison of the tWZ event yields between the nominal and the alternative sample was performed for the three 3ℓ signal regions before the Monte Carlo samples were fit to data. The numbers quoted in Tab 6.3 are scaled to both the theoretical cross sections of the respective Monte Carlo samples, which differ on the order of $10\,\%$, and to the integrated luminosity of the full Run 2 dataset. In addition, the yields are decomposed into the contributions from the various Monte Carlo production campaigns. It can be seen that the numbers from the DR2 sample are

	mc16a		mc16d		mc16e		All campaigns	
3ℓ-Z-1b4j-PCBT	5.84 / 5.21	(10.8 %)	5.98 / 5.55	(7.2 %)	10.18 / 6.51	(36.1 %)	22.00 / 17.27	(21.5 %)
3ℓ-Z-2b3j-PCBT	4.68 / 3.39	(27.6 %)	5.24 / 3.11	(40.6 %)	8.30 / 4.22	(49.2 %)	18.22 / 10.73	(41.1 %)
3ℓ-Z-2b3j	9.53 / 7.78	(18.4 %)	10.13 / 7.00	(30.9 %)	16.13 / 9.17	(43.2 %)	35.79 / 23.95	(33.1 %)
Preselection	88.45 / 81.52	(7.9 %)	89.33 / 83.95	(6.0 %)	148.89 / 112.67	(24.3 %)	326.67 / 278.14	(14.9 %)

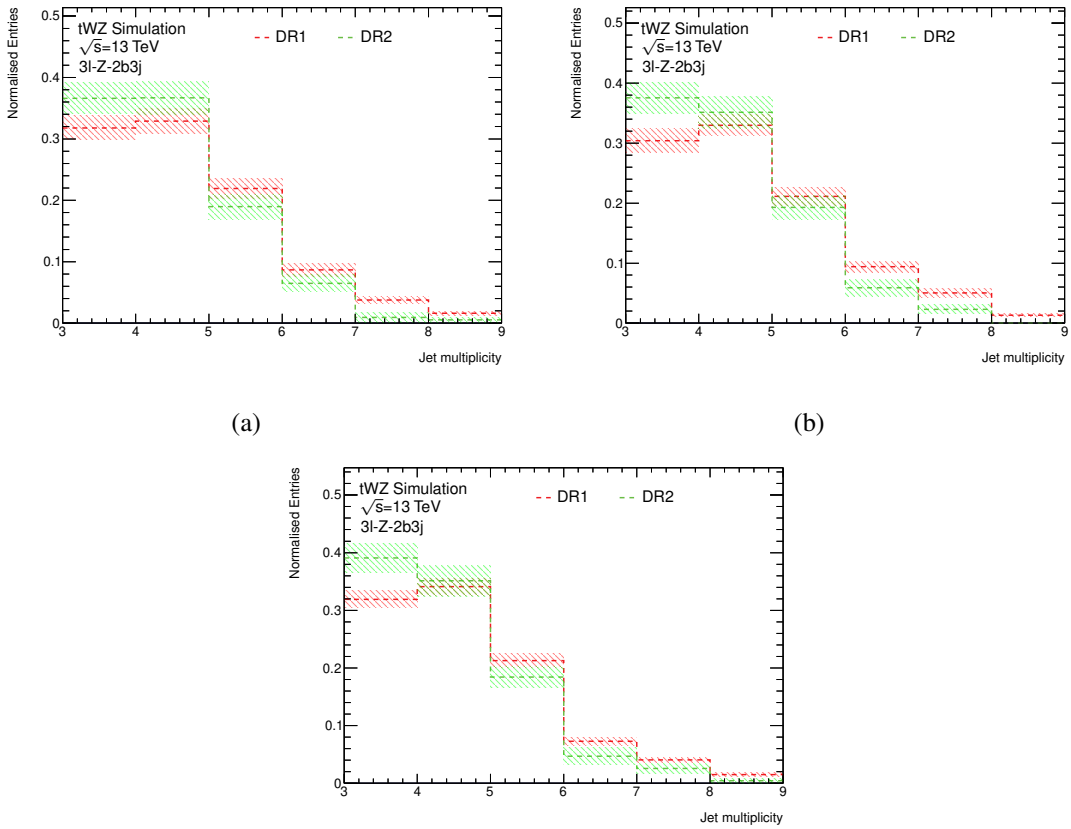
Tab. 6.3.: Expected yields in the various 3ℓ signal regions and the preselection decomposed into the various MC campaigns and stated separately for the nominal DR1 (left) and alternative DR2 (right) sample. The values quoted in parentheses correspond to the relative difference with respect to the nominal sample, calculated with Eq. (6.6). The numbers are given without any statement about the statistical error.

consistently lower that those from the DR1 sample. The relative yields differences with respect to the nominal sample range from $6\,\%$ to almost $50\,\%$, where the largest discrepancies can be observed for the mc16e campaign.

In order to check whether the observed differences in event yields are concentrated in particular regions of phase space, a set of distributions of several kinematic quantities, normalised to the respective number of entries, are shown in Fig. 6.3 as well as in Figs F.1 and F.2. Each panel of those figures shows a particular variable for a single MC campaign, and the different coloured lines correspond to the two samples either produced with DR1 or DR2. For convenience only the signal region used for the differential cross-section measurements is shown. The error bars are statistical only. The distributions of the two tWZ samples show a reasonable level of agreement with each other. Notable differences between the shapes are visible for the distributions of the jet multiplicity and the leading lepton transverse momentum. In case of jet multiplicity, the DR2 curve shows a different behaviour in the first two bins from the left for mc16a than for the other two MC campaigns. The course of the DR2 curve of the leading lepton p_{T} is different for the mc16e campaign than for mc16a and mc16d.

A similar set of distributions to those of Fig. 6.3 is shown in Fig. 6.4, but for which the simulated process is fixed to either DR1 or DR2 and the three coloured lines correspond rather to the various MC campaigns. The shapes of all MC campaigns seem to be consistent with each other within statistical uncertainties for all distributions of the DR1 sample and for both the jet and b-jet multiplicity of the DR2 sample. The distribution of the leading lepton transverse momentum for the mc16e campaign of the DR2 sample, however, can be seen to be clearly different from expectation.

The reason for the considerably high values of the relative difference between the two tWZ samples as well as the disagreements observed in both sets of normalised distributions will be investigated in the next section.



(a)

(b)

(c)

Fig. 6.3.: Normalised distributions of the jet multiplicity for mc16a, mc16d and mc16e in Figs. (a), (b) and (c), respectively, shown for the 3ℓ differential signal region. The red line corresponds to the DR1 sample, and the green line to the DR2 tWZ sample. The error bars in all panels include only MC statistical uncertainties.

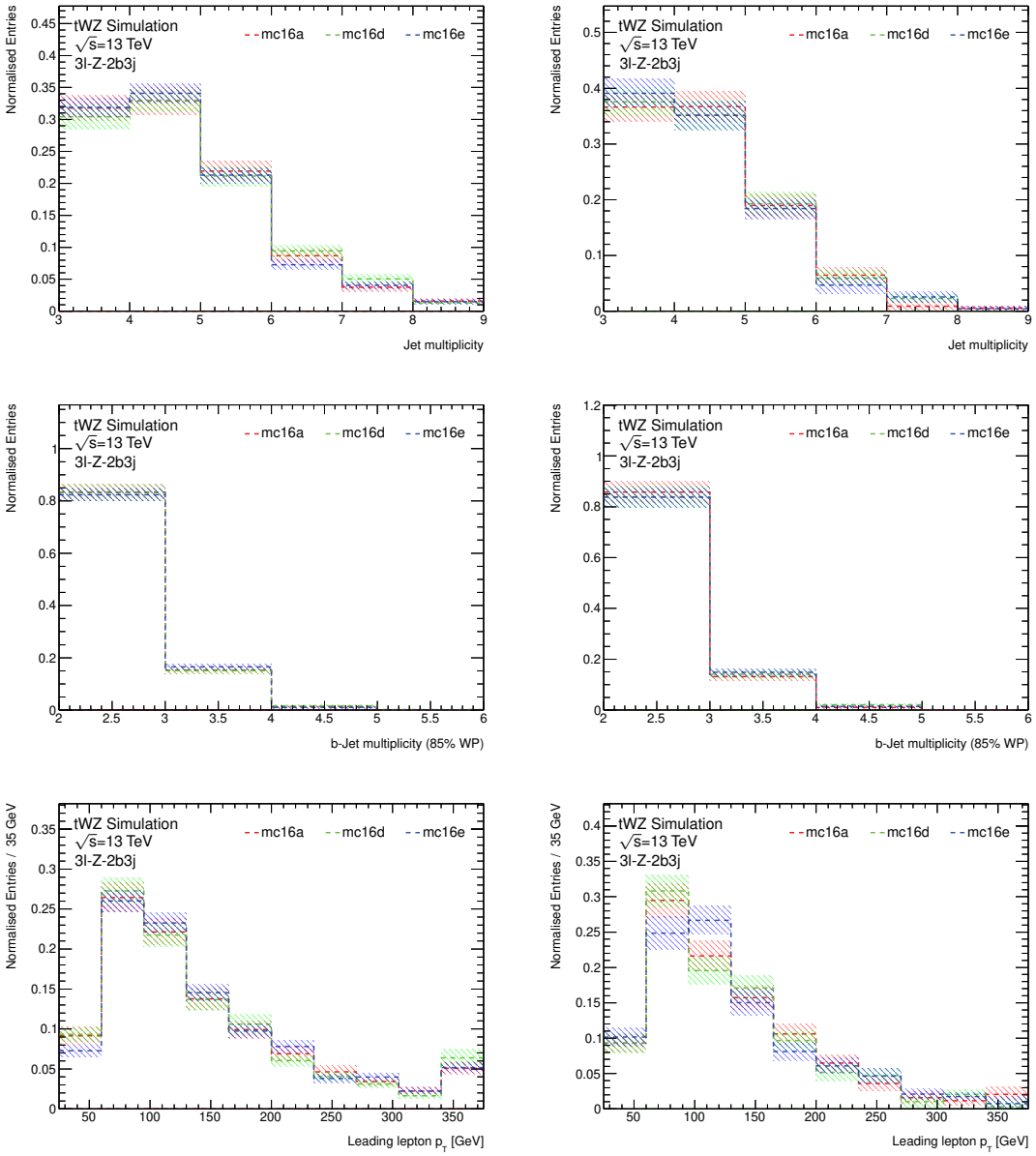


Fig. 6.4.: Normalised distributions of the jet multiplicity in the top row, the b-jet multiplicity in the middle row and the leading lepton transverse momentum in the bottom row, shown for the 3ℓ differential signal region. The coloured lines correspond to the three MC campaigns, on the left for the DR1 sample and on the right for the DR2 sample. The error bars in all subfigures are statistical only.

6.2.3. Investigation of large Monte Carlo generator weights on tWZ estimations

The tWZ samples were generated at next-to-leading order, as mentioned in Secs. 4.1.3 and 6.2.2 causing an interference effect of tWZ and $t\bar{t}Z$ which is accounted for by the diagram removal technique during the production of the simulated samples. In addition to that, the samples produced at NLO accuracy have an impact on the distributions of the generator weight which accounts for the fact that some points in the available phase space are physically more probable than others, as described in Sec. 4.1.1. In order to properly compare simulated events to data, the Monte Carlo samples have to be scaled to the respective cross section they were produced with and normalised to the integrated luminosity of the data. The various scale factors from the calibration and reconstruction procedures, as explained in Sec. 4.1.2 and Chap. 5, respectively, and the generator weight have to be applied as well.

As a measure for the differences between the nominal and alternative tWZ samples the relative in yields between the nominal and alternative tWZ samples with respect to the nominal sample, ΔN_{DR}, was chosen:

$$\Delta N_{\text{DR}} = \frac{N_{\text{DR1}} - N_{\text{DR2}}}{N_{\text{DR1}}} \tag{6.6}$$

The N_{DR1} and N_{DR2} denote the event yields of the DR1 and DR2 tWZ samples, respectively. The values calculated according to Eq. (6.6) in the 3ℓ signal regions can be found in Tab. 6.4. The left column presents the yields from unweighted samples, also referred to as raw yields. The MC generator weight was applied to the yields in the central column, and for those in the right column the whole set of weights, including the scale factors from object reconstruction, were taken into account. From the numbers quoted in Tab. 6.3, the generator weight can be seen to be one of the main causes of the large yield differences between the DR1 and the DR2.

	Unweighted	Generator weight only	All weights
3ℓ-Z-1b4j-PCBT	0.9 %	16.0 %	21.5 %
3ℓ-Z-2b3j-PCBT	0.2 %	36.8 %	41.1 %
3ℓ-Z-2b3j	1.8 %	28.1 %	33.1 %
Preselection	1.1 %	8.9 %	14.9 %

Tab. 6.4.: Relative yields differences between the tWZ DR1 and DR2 samples with respect to DR1, calculated with Eq. (6.6), given for the 3ℓ signal regions and the preselection. The numbers are shown for completely unweighted yields in the left columns, for partly weighted yields with only the MC generator weight applied in the middle column and for fully weighted yields in the right column.

The Monte Carlo generator weights were further investigated separately for the various MC campaigns. As can be seen in Tab. 6.5, the fractions of negative generator weights within the 3ℓ signal regions as well as for the preselection are consistently higher for the DR1 sample than for the DR2 sample. Since in the implementation of DR2 the interference term between $t\bar{t}$ and tWb is not forbidden, the available phase space is different with respect to that of the DR1 version and different fractions of negative and positive MC generator weights were expected. However, this does not account for the features observed in Fig. 6.5, which depicts the absolute value of the MC generator weight normalised to the respective number of entries in red for DR1 and in blue for DR2 at preselection level. Each subfigure corresponds to a particular MC campaign. In the Fig. 6.5a depicting the generator weight for mc16a, both the DR1 and DR2 sample feature one sharp peak

	DR1	DR2
3ℓ-Z-1b4j-PCBT	22.9 %	28.2 %
3ℓ-Z-2b3j-PCBT	18.1 %	30.5 %
3ℓ-Z-2b3j	19.8 %	28.8 %
Preselection	21.0 %	24.8 %

Tab. 6.5.: Fractions of negative Monte Carlo weights for both tWZ samples. The values for the DR1 sample in the left column and the DR2 sample in the right column are given for the 3ℓ signal regions and the preselection.

each. In the Fig. 6.5b presenting the generator weights for mc16d, however, an additional smaller peak can be observed at an approximately ten times higher value of the generator weight than for the large peak. The extra peak account for roughly 3900 raw events in the case of DR1, and 3700 raw events for DR2, respectively. In Fig. 6.5c, which shows the generator weight for mc16e, only the DR2 samples reveal such an additional peak which accounts for approximately 3800 raw events, whereas the DR1 sample features only one peak.

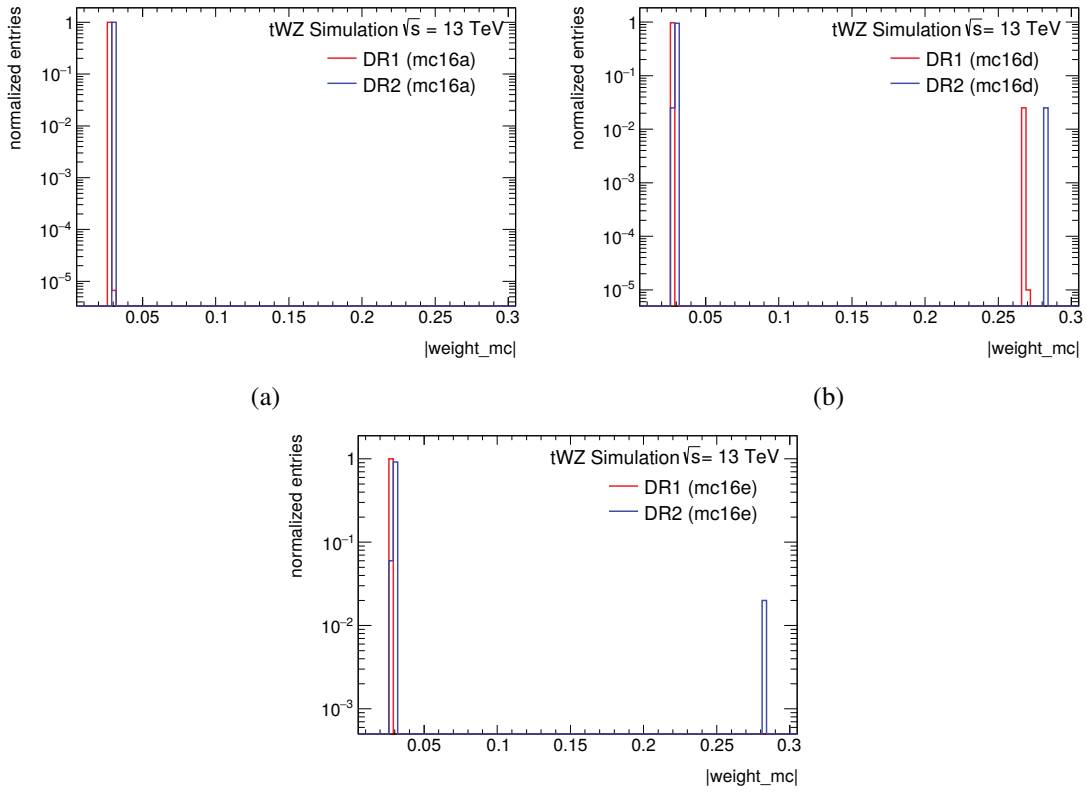


Fig. 6.5.: Normalised distributions of the MC generator weights for the mc16a, mc16d and mc16e production campaigns depicted in Figs. (a), (b) and (c), respectively. The distributions for DR1 in red and DR2 in blue are normalised to their respective number of entries and shown on a logarithmic scale. Only preselection selection criteria were applied.

The events with a generator weight ten times higher than for the rest, thus causing the three additional peaks in the distributions shown in Fig. 6.5, amount for only $2\,\%$ of all generated events. However, the overall impact of these is roughly an order of magnitude larger due to the large weight which makes them subject to large statistical fluctuations. A fault within the Monte Carlo generation procedure was finally identified as the reason of the unbalanced distribution of the MC generator weights, which is why those events are neither physically nor statistically meaningful and were consequently removed from the pool of simulated tWZ events. In order to preserve a proper scaling of the tWZ Monte Carlo samples according to Eq. (4.3), the initial number of weighted tWZ events had to recalculated since it is affected by the generator weights. All other quantities which enter Eq. (4.2) remained constant, though.

The effect of the removal of the events with a high generator weight modulus can be seen in Fig. 6.6 which depicts the relative yield differences between the DR1 and DR2 sample with respect to DR1 for the 3ℓ signal regions as well as at preselection level. The values are given separately for each MC campaign as well as for their combination. The left subfigure shows the initial situation with all values of the generator weight included. The stated values therefore correspond to those from Tab. 6.4. In the right subfigure, however, events with an absolute value of the Monte Carlo generator weight of greater than 0.1 are vetoed which is equal to simply ignoring all events within the lower peaks on the right of the various panels of Fig. 6.5. While relative differences of up to $40\,\%$ remain, these can be seen to be more consistent between the various Monte Carlo campaigns.

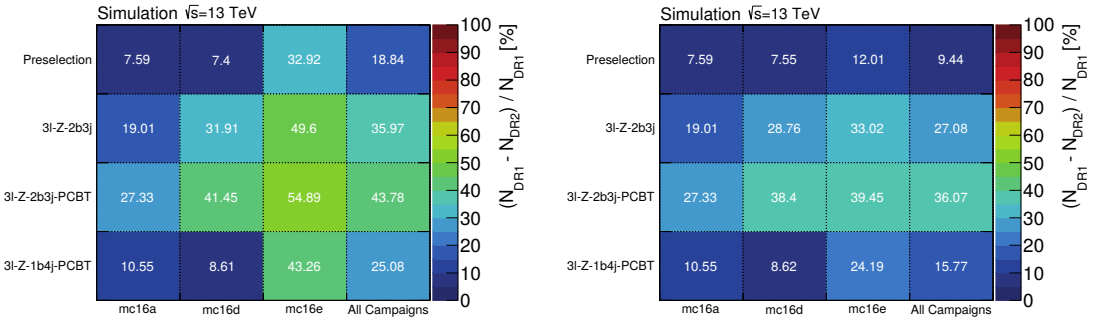


Fig. 6.6.: Relative yield differences in per cent between the DR1 and DR2 sample shown for the various 3ℓ signal regions as well as at preselection level, as written on the y-axis. The values are stated for each MC campaign separately as well as for the combination of all campaigns, indicated by the x-axis. In the left figure all selected events have been used to compute ΔN_{DR}, whereas in the subfigure on the right events with an absolute value of the MC generator weight of greater than 0.1 are vetoed which requires to reevaluate the initial number of weights tWZ events. All weights and scale factors have been applied.

In order to check the impact of the removal of events with high absolute values of the MC generator weight on the physics effects intended to model by the variation of the diagram removal scheme, several kinematic variables sensitive to the $t\bar{t}Z$ signature were investigated. For example, the distributions of the transverse momentum of the Z boson are shown in Fig. 6.7, whereas the rapidity of the Z boson and the azimuthal difference between the Z boson and the top quark with a subsequent decay into a charged lepton and the corresponding neutrino are depicted in Figs. F.3 and F.4, respectively. All those variables were built from the parton-level[(5)] information of the

[(5)] Parton-level objects are obtained from the Monte Carlo generation history of the $t\bar{t}Z$ system and can therefore be used

Monte Carlo truth[6] record in order to highlight the differences between DR1 and DR2 induced by the physics effects and avoid them being disturbed by any detector or reconstruction inefficiencies. For each figure, the various panels correspond to another Monte Carlo production campaign. All distributions are normalised to their respective number of entries. The DR1 sample is depicted as a red solid line, and the DR2 sample is represented by the blue dots. Both the normalised distributions and the ratios of DR2 to DR1 shown below each subfigure feature statistical errors only. Since no events were removed from the mc16a dataset, it serves as the reference for the two other MC campaigns. The distributions of the various MC campaigns are consistent with each other within the statistical uncertainties, and no significant deviations of either mc16d or mc16e from mc16a can be observed.

The tWZ samples were used in the 3ℓ and 4ℓ analysis with a veto applied on absolute values of the MC generator weight greater than 0.1 in order to avoid a potential overestimation of the systematic uncertainty related the diagram removal scheme due to single events with a large absolute value of the MC generator weight.

6.3. Optimisation studies based on top quark reconstruction

The targeted resonances of the $t\bar{t}Z$ process, the top quarks and the Z boson, cannot be directly detected due to their short lifetime, and only the signatures left by their decay products are able to be measured in the detector. However, from the detected objects the original particles can be reconstructed by employing an algorithm which associates the detector-level objects with the original partons of interest. In order to reduce the number of possible assignment, several kinematic requirements based on the final-state signature as well as information about the electric charge of the particles and, in case of quarks, about b-tagging are employed. Such a kinematic reconstruction provides the possibility to probe the $t\bar{t}Z$ system and its heavy resonances directly, and furthermore allows for background contributions to be suppressed as they feature different kinematics than the signal the reconstruction algorithm is optimised for. The aim of the studies presented in the subsequent sections is to minimise the uncertainties on the measured $t\bar{t}Z$ cross section by employing the kinematic reconstruction of the $t\bar{t}Z$ system.

6.3.1. Leptonic-side top quark reconstruction

The top or antitop quark of the $t\bar{t}$ system for which the W boson decays into an electron or a muon and the corresponding neutrino ($\ell\nu$), referred to as leptonic-side top quark, was reconstructed for events in the 3ℓ regions. This partial reconstruction of the $t\bar{t}$ system allows for the construction of several variables which are sensitive to the kinematics of the top quarks and the Z boson of the $t\bar{t}Z$ process. For the reconstruction it is assumed that the neutrino of the leptonically decaying W boson is the dominant source of missing transverse energy for a given event. Both the magnitude and the azimuthal angle of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are assigned to the neutrino directly. The four-vectors of the neutrino and the lepton not associated to the Z boson, referred to as non-Z lepton, are added together and the combined mass is set to $80.385\,\mathrm{GeV}$ [51]. With the application of such a W boson mass constraint, the neutrino momentum in the z-direction, p_z^ν, can be determined via a

only for simulated events. At parton level, the information about both initial and final state partons of a collision as well as about short-lived resonances such as top quarks and Z bosons and their decay products is available. In contrast to reconstructed objects, neither the hadronisation process of coloured objects nor the interaction of particles with the detector material is simulated at this stage.

[6] Common denomination for the information about the generated particles and their decay a Monte Carlo generator typically provides in addition to what is known for real data. In addition to parton-level information it includes information about the final-state objects such as jets but without any simulation of them with the detector.

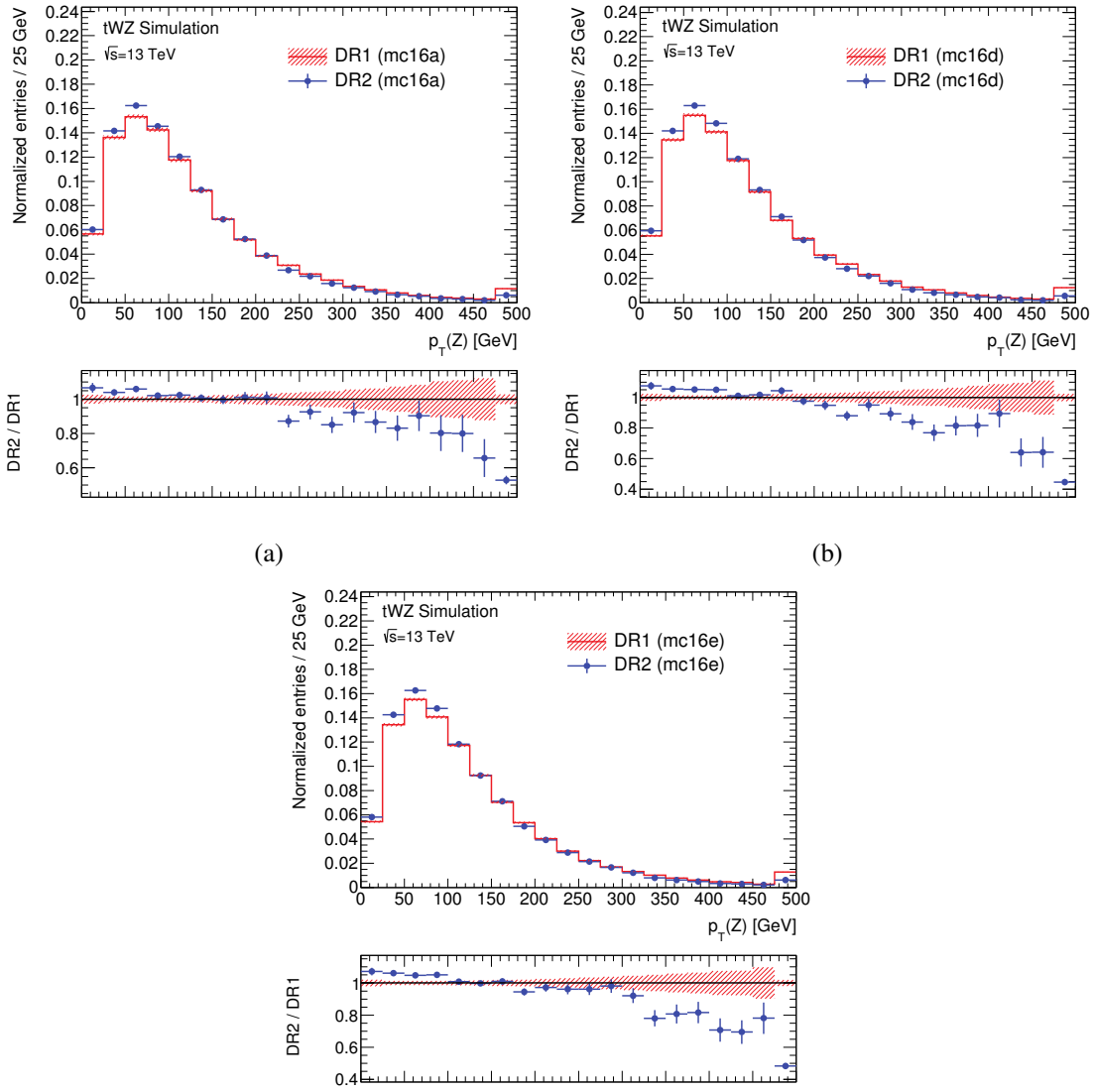


Fig. 6.7.: Normalised distributions of the Z boson transverse momentum at parton level. The various panels (a), (b) and (c) corresponds to the mc16a, mc16d and mc16e campaigns, respectively. All distributions for DR1 in red and DR2 in blue are normalised to their respective number of entries. The ratio pads below each plot show the ratio of DR2 to DR1. The error bars only statistical MC uncertainties.

quadratic equation,

$$A(p_z^\nu)^2 + Bp_z^\nu + C = 0 \quad , \tag{6.7}$$

where the terms A, B and C depend on the transverse momentum and the azimuthal angle of neutrino, the fixed value of the W boson mass, and the kinematics of the non-Z lepton[(7)].

For the case that the discriminant of Eq. (6.7) is positive, namely that $B^2 - 4AC > 0$, two possible solutions of p_z^ν are obtained, giving rise to two unique candidate four-vectors of the neutrino. Both values are considered for the subsequent step which is to combine a neutrino candidate with the non-Z lepton and a jet in order to reconstruct a top quark candidate. For a certain fraction of events, however, the values of A, B and C are such that the discriminant of the quadratic equation is negative and thus no real solution can be found. In such cases, the missing transverse energy, which is equivalent to the assigned neutrino transverse momentum, will be smeared until a single solution for p_z^ν can be found. This procedure is equivalent to solving analytically for the value of the neutrino transverse momentum for the case which yields a non-negative discriminant, $B^2 = 4AC$.

The single neutrino candidate, or, in the case of two solutions, the pair of neutrino candidates is then combined with a jet in order to form a leptonic-side top quark candidate. From the two jets with the highest b-tagging output score, the one with the smaller ΔR separation from the given $\ell\nu$ system is selected. It is possible that the selected jets differ for the two neutrino candidates in a given event since the ΔR criterion is applied separately.

An output weight from the leptonic-side top quark reconstruction is defined based on the invariant mass distribution ($m_{b\ell\nu}$) from correctly reconstructed top quarks in simulated $t\bar{t}Z$ events. The reconstructed top quarks used to fill the reference distribution are formed from parton-level neutrinos and reconstructed leptons and jets which have been matched to the corresponding parton-level objects. The weighting function therefore represents the idealised distribution that can be achieved with a perfect assignment of reconstructed objects to their corresponding parton-level objects aside from residual corrections to the object kinematics. The $m_{b\ell\nu}$ distribution which was employed in the context of this thesis can be found in Fig. 6.8. For events with a possible top quark candidates, the two invariant masses are assigned an output weight based on an interpolation of the $m_{b\ell\nu}$ reference distribution and the candidate with an invariant mass most consistent with a leptonic-side top quark is selected, whereas the other solution is not considered at any later point. The interpolated value of the output weight which was seen highest for a given event was stored in order to provide a discriminating variable suppressing contributions of background processes. The reference distribution of $m_{b\ell\nu}$ was normalised such that its maximum has a value of one. Consequently, the interpolated value is thus representative of the probability of how consistent a reconstructed top quark candidate is with a leptonic-side top quark.

In addition to the definition of the signal regions, the output weight from the leptonic-side reconstruction can be used to further suppress contributions of background processes by cutting away events with an output weight below a certain value. This cut has to be set in favour of a high signal efficiency while rejecting most of the background events. The studies presented in this thesis were designed such that while keeping the signal-to-background ratio sufficiently high, the impact of the systematic and statistical uncertainties associated with the signal and background contributions is to be minimised.

6.3.2. Minimisation of the relative uncertainty on the measured cross section

Monte Carlo simulated samples are widely used in high energy physics to make predictions about and verify or reject certain physics models based on the comparison with data. Under ideal cir-

[(7)] A detailed calculation of the neutrino kinematics be found in App. B.

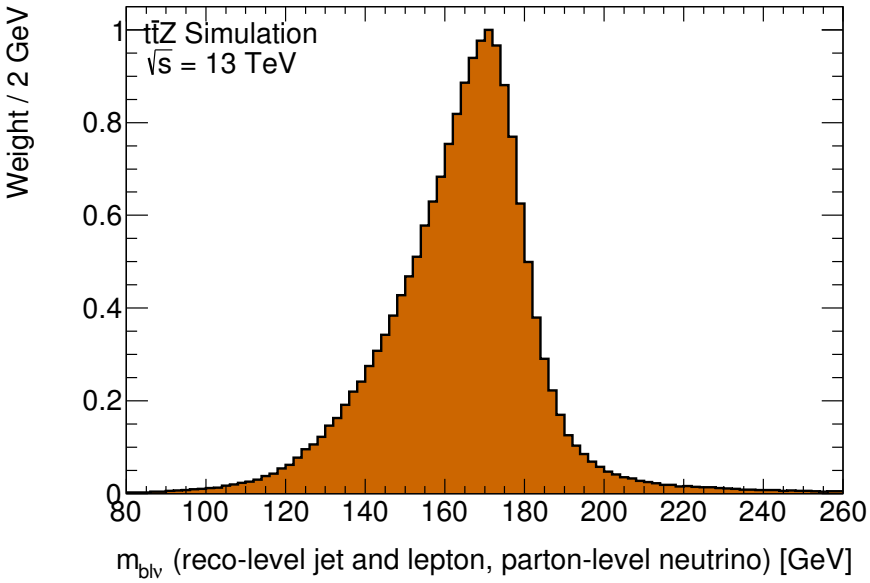


Fig. 6.8.: Invariant mass distribution from correctly reconstructed top quarks in simulated $t\bar{t}Z$ events ($m_{b\ell\nu}$). The reconstructed top quarks used to fill the distribution are formed from the parton-level neutrino and the reconstruction-level lepton and jet matched to the corresponding parton level objects.

cumstances, real data and Monte Carlo predictions would be in perfect agreement with each other, provided the knowledge of all necessary theoretical and experimental parameters. However, errors on the measured data are caused by experimental uncertainties such as a finite resolution of the detector or noise from the read-out electronics. Similarly, the production of simulated samples is affected by theoretical uncertainties related to the limited degree of precision of the theoretical description of both perturbative and non-perturbative processes such as hard processes at high-order accuracy and hadronisation, respectively. All these factors have to be corrected for with dedicated calibration methods which themselves entail certain uncertainties. Therefore, the data-MC agreement is typically not perfect within most physics analyses.

Many analyses of LHC data represent a form of counting experiment. That is, an analysis can be seen as consisting of $n \in \mathbb{N}$ measurements which follow an underlying p.d.f. $f(x)$ that is a priori unknown. The properties of $f(x)$ can be inferred, however, based on the observations $x_1, \ldots, x_n$ by constructing appropriate functions of the x_i in order to estimate the various properties of $f(x)$. Often there is a hypothesis for the p.d.f., $f(x;\theta)$ which depends on an unknown parameter θ or a set of unknown parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m), m \in \mathbb{N}$. The goal is then to construct a function of the observed x_i to estimate those parameters. This procedure of estimating parameter values given the data $x_1, \ldots, x_n$ is called parameter fitting and is widely used in particle physics. Both experimental and theoretical uncertainties, denoted as systematic uncertainties, as well as the statistical error are accounted for with a distinct parameter [139].

In order to reject or accept the hypothesis of the existence of some physics process, and in the latter case also to measure its production cross section, the compatibility of the observed data with the expectation needs to be determined. The probability distribution of the data is inferred with by a fit and the compatibility of data and expected events is tested. The parameter of interest is the so-called signal strength, $\mu_{t\bar{t}Z}$, which represents the ratio of the observed number of events for the given process in measured data with respect to the number predicted from simulation. As

an example, a value of $\mu_{t\bar{t}Z} = 0$ corresponds to the case of absent signal, whereas a value of $\mu_{t\bar{t}Z} = 1$ indicates that the amount of signal events predicted from Monte Carlo and measured in data is equal.

The methodology just described is employed for the ATLAS cross-section measurement in the 3ℓ and 4ℓ channels [46]. However, in order to get a sense of the sensitivity achievable by applying a kinematic reconstruction of the leptonic-side top quark, the impact of the combined uncertainties on the measured cross section was estimated using the yields and corresponding errors for a given cut on the reconstruction output weight. The rate of events[8] for some physical process produced within the proton-proton collisions inside the ATLAS detector can be described, according to Eq. (3.2), by the product of the corresponding cross section and the instantaneous luminosity:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \sigma \cdot L \tag{6.8a}$$

$$\mathrm{d}N = \sigma \cdot \int \mathrm{d}t\, L \tag{6.8b}$$

$$\sigma = \frac{N}{\mathcal{L}} \tag{6.8c}$$

The production cross section σ can thus be expressed in terms of the number of events of the respective process, N, and the total amount of data, represented by the integrated luminosity $\mathcal{L}$. The signal contribution is then estimated from the difference between data and simulated background events. The number of observed $t\bar{t}Z$ events for a specific process, N, is rewritten in terms of data and Monte Carlo events of all background processes:

$$N = N_{\text{Data}} - N_{\text{Bkg}} \tag{6.9}$$

The signal Monte Carlo samples were produced inclusively with respect to the $t\bar{t}$ and Z boson decay modes. As only electrons and muons were targeted as the final states of the Z boson decay, the number of events in Eq. (6.8c) has to be corrected for with a branching fraction $\mathcal{BR}$ of approximately $2.3\,\%$[9]. Instrumental inefficiencies of the detector and a small fraction of signal events that are found to be outside of the acceptance of the signal regions are accounted for by a correction factor $\varepsilon_{t\bar{t}Z}$. This factor is defined as the number of selected signal events $N_{t\bar{t}Z}$ divided by the total amount of generated $t\bar{t}Z$ events of the targeted final state,

$$\varepsilon_{t\bar{t}Z} = \frac{N_{t\bar{t}Z}}{\sigma_{\text{MC}} \cdot \mathcal{L} \cdot \mathcal{BR}} \quad , \tag{6.10}$$

where σ_{MC} denotes the cross section value employed at the production of the respective Monte Carlo samples. The measurable production cross section of $t\bar{t}Z$ including all those correction factors therefore reads

$$\sigma_{t\bar{t}Z} = \frac{N_{\text{Data}} - N_{\text{Bkg}}}{\varepsilon_{t\bar{t}Z} \cdot \mathcal{L} \cdot \mathcal{BR}} \tag{6.11}$$

which can be transformed into

$$\sigma_{t\bar{t}Z} = \frac{N_{\text{Data}} - N_{\text{Bkg}}}{N_{t\bar{t}Z}} \sigma_{\text{MC}} \tag{6.12}$$

[8] The event rate is the number of events per unit of time.

[9] The branching fraction of the $t\bar{t}$ system decaying in the lepton+jets channel equals roughly $34\,\%$ (cf. Sec. 2.3.2). Multiplied with the branching fraction of a Z boson to decay into a pair of electrons or muons of $6.73\,\%$ [51], the overall $t\bar{t}Z$ branching fraction for 3ℓ events has a value of $2.3\,\%$.

by inserting Eq. (6.10) into Eq. (6.11). The expression for the cross section is now dependent on the uncertainties of both the modelling and the data, but is also sensitive to potentially missing contributions of background processes which are not considered. The dependencies on the integrated luminosity $\mathcal{L}$ and on the correction factor $\varepsilon_{t\bar{t}Z}$, however, are absorbed. The fraction $(N_{\text{Data}} - N_{\text{Bkg}})/N_{t\bar{t}Z}$ in Eq. (6.12) corresponds to the signal strength,

$$\mu_{t\bar{t}Z} = \frac{N_{\text{Data}} - N_{\text{Bkg}}}{N_{t\bar{t}Z}} \tag{6.13}$$

scaling the amount of predicted signal events from Monte Carlo.

The errors on the luminosity and on $\varepsilon_{t\bar{t}Z}$ were assumed to be sufficiently small with respect to the systematic uncertainties related to the various simulated processes. For this reason, Eq. (6.12) is a function of three parameters, N_{Data}, N_{Bkg} and $N_{t\bar{t}Z}$. The uncertainty on the measured cross section can therefore be derived by means of Gaussian error propagation [245]. Given a function f of several variables x_i which have to be statistical independent from each other, the error on f is given by

$$\Delta f(x_i) = \sqrt{\sum_i \left(\frac{\partial f(x_i)}{\partial x_i} \Delta x_i \right)^2} \quad . \tag{6.14}$$

The error on the cross section in Eq. (6.12) thus reads

$$\Delta\sigma_{ttZ} = \frac{\sigma_{\text{MC}}}{N_{t\bar{t}Z}} \sqrt{(\Delta N_{\text{Data}})^2 + (\Delta N_{\text{Bkg}})^2 + \left(\frac{N_{\text{Data}} - N_{\text{Bkg}}}{N_{t\bar{t}Z}} \right)^2 (\Delta N_{t\bar{t}Z})^2} \quad , \tag{6.15}$$

and the relative uncertainty on the cross section can hence be written as

$$\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}} = \frac{1}{N_{\text{Data}} - N_{\text{Bkg}}} \sqrt{(\Delta N_{\text{Data}})^2 + (\Delta N_{\text{Bkg}})^2 + \left(\frac{N_{\text{Data}} - N_{\text{Bkg}}}{N_{t\bar{t}Z}} \right)^2 (\Delta N_{t\bar{t}Z})^2} \quad . \tag{6.16}$$

From Eq. (6.16) it can be seen that the overall relative cross-section uncertainty depends on the number of data and simulated events but also on their associated errors. As previously mentioned in Sec. 6.3.1, the output weight of the leptonic-side top quark reconstruction method was employed to perform an iterative search for a minimum of Eq. (6.16). The number of respective events and uncertainties given to the formula depends on the respective cut value on the output weight. For this purpose, the range in the signal efficiency of 70 % to 100 % was scanned in integer steps of 1 %. Values below 70 % were not considered in order to maintain a sufficiently high number of signal events. At each efficiency point the relative cross section uncertainty was calculated from data and simulated events and their uncertainties determined by the respective cut on the reconstruction output weight. The set of considered uncertainties contains various experimental uncertainties related to the measurement and reconstruction of the physics objects, electrons and muons, jets, and missing transverse energy, as well as from calibration and flavour tagging. In addition, theoretical uncertainties on the modelling of the signal and various background processes, as well as uncertainties induced by the choice of certain Monte Carlo generators, as described in Sec. 6.2, were taken into account. The statistical uncertainty of a certain number of events N was derived from the Poisson error, $\sqrt{N}$ [245], applied on the fully weighted yields.

The result of this procedure is depicted in Fig. 6.9 which shows the relative cross section uncertainty as a function of the signal efficiency in the range of 70 % to 100 %. The red dots indicate the found values of $\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z}$ for the respective signal efficiency values. As a reference value for the case without setting any cut on the output weight, a dashed blue line was added to the figure,

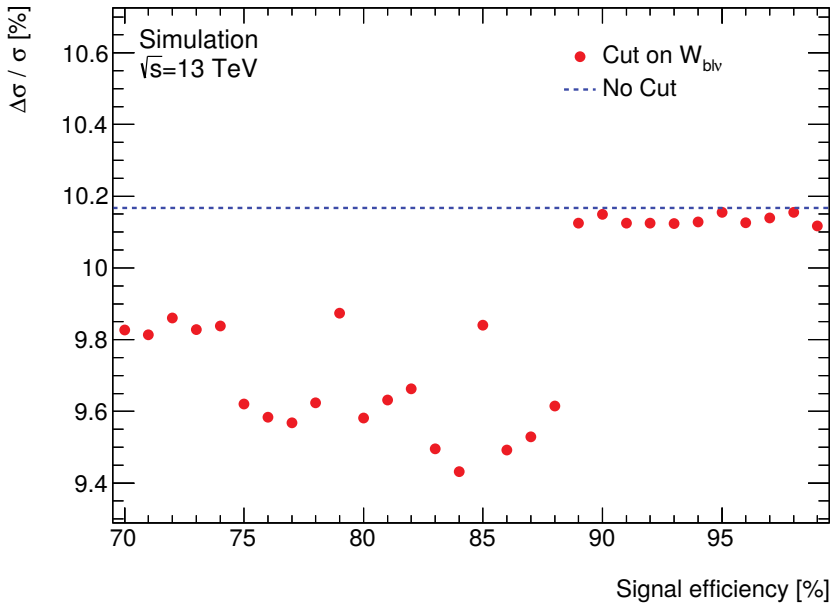


Fig. 6.9.: Distribution of the relative uncertainty on the cross section derived by means of Eq. (6.16), plotted against the signal efficiency of the respective cuts on the output weight distribution within the range of 70 % to 100 %. The red dots represent the found values of the relative uncertainty and the dashed blue line serves as reference for the case without any cut. The distribution of the red dots has a minimum at a signal efficiency value of 84 % and features a horizontal form close to the reference line at a signal efficiency of 89 % and higher.

corresponding to a value of approximately $\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z} = 10.15\,\%$. The resulting distribution of the values of the relative uncertainty has a spread of below 1 % and yields a minimum of about $\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z} = 9.4\,\%$ for a signal efficiency of 84 %. At a signal efficiency of 89 % a sharp cut-off can be observed from the red dots gathered close to the reference line in this range. This is on the one hand related to the finite binning of the histograms that were used to evaluated both the event yields and their uncertainties, and on the other hand above 89 % signal efficiency the distribution of signal and background events do not change anymore. The study was performed with events passing the selection criteria of the 3ℓ differential signal region only. In Fig. 6.10 the output weight, $W_{b\ell\nu}$ from the leptonic-side top quark reconstruction, used for the evaluation of the relative cross section uncertainty, is shown. Although the full background composition was employed to determine the values of $\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z}$ presented in Fig. 6.9, only the three dominant background processes, tWZ, tZq and WZ, are depicted in Fig. 6.10. The distributions of simulated signal and background events, depicted by the coloured lines, were scaled such that the most populated bin has a peak height of unity and individually normalised for each process. The vertical red dashed line on the left indicate the cut value which corresponds to a signal efficiency of 84 %. The direction of the red arrow on top of the vertical red line implies that the events left to the red line are cut away.

Ultimately for the final $t\bar{t}Z$ cross-section measurement analysis [46], it was decided not to apply a cut on the reconstruction output. This was primarily due to the fact that the modest gains in precision did not justify the additional time that would be required for a detailed systematic study. For example, the reduced number of events impacts the statistical error and a good balance between

systematic and statistical uncertainties has to be determined, in particular for the differential part of the cross-section measurement. It could be shown, however, that the method presented in this section offers, as intended, the prospect for improved precision in future versions of the analysis, when additional time would allow for a more detailed optimisation.

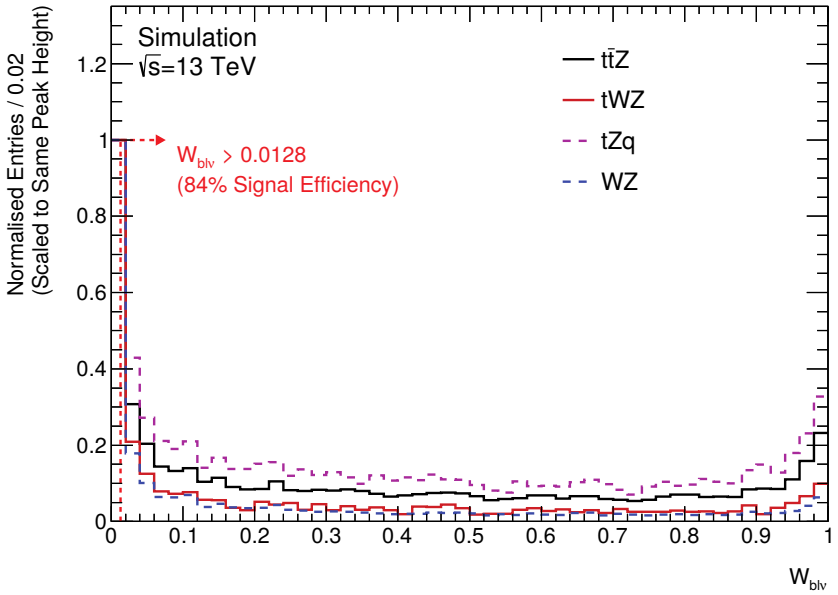


Fig. 6.10.: Normalised distributions of the output weight from the leptonic-side top quark reconstruction, depicted for the signal in black and the three dominant backgrounds in red, pink and blue, respectively. The distributions are normalised such that the maximum of either distribution is equal to one. The position for setting a cut corresponding to a signal efficiency of $84\,\%$ is indicated by the vertical red dashed line.

7. Prospects for a new $t\bar{t}Z$ cross-section measurement in the dilepton channel using multivariate techniques

Observation of the $t\bar{t}Z$ process with a statistical significance of more than five standard deviations was so far achieved for the 3ℓ and 4ℓ channels only, as they offer the highest sensitivity [35, 39, 46, 47]. However, the so-called 2ℓOS channel constitutes a promising extension to explore the $t\bar{t}Z$ process and thereby probe the Standard Model. The latest ATLAS measurement in the 2ℓOS channel was performed with data collected in 2015 and 2016, corresponding to an integrated luminosity of $36.1\,\mathrm{fb}^{-1}$ [39]. In the course of this chapter, a new attempt for the measurement of the $t\bar{t}Z$ production cross section in the 2ℓOS channel will be presented using a dataset of $139\,\mathrm{fb}^{-1}$. The analysis strategy is based on the approach adopted in Ref. [39], but employs a refined procedure and novel approaches in order to separate signal events from the various background processes with respect to Ref. [39].

7.1. Analysis strategy

The present section outlines the overall strategy of the 2ℓOS analysis. The two subsequent sections will then describe the multivariate analysis techniques employed and the statistical interpretation of the analysis, respectively.

7.1.1. Targeted analysis regions

Three phase-space regions were defined and optimised in order to feature both a high signal purity and low background contribution. The 2ℓOS channel targets events which feature a leptonically decaying Z boson and a fully-hadronically decaying $t\bar{t}$ system. Based on the final-state signatures of such targeted decays, the following selection criteria were defined.

In addition to the object definition criteria specified in Chap. 5, all detector components are required to have been operating according to detector-specified quality standards at the time the event was recorded. In addition, a given event must contain at least one electron or one muon which caused a trigger to fire and, specifically for real data events, have to recorded within a luminosity block where the detector was known to be functioning under proper running conditions. The selected events are required to contain exactly two light leptons[(1)] with opposite measured electric charge and the same flavour. The invariant mass of this opposite-sign-same-flavour (OSSF) lepton pair, $m_{\ell\ell}$, must be consistent with the Z boson mass within a $10\,\mathrm{GeV}$ window in order to accommodate detector response and resolution effects. Minimum requirements of $30\,\mathrm{GeV}$ and $15\,\mathrm{GeV}$ were placed on the leading and subleading lepton[(2)] transverse momentum, respectively. The targeted phase-space regions differ in the number of required jets and b-jets, respectively, where only jets with a minimum transverse momentum of $25\,\mathrm{GeV}$ and a fixed b-tagging WP with $77\,\%$ efficiency to tag a true b-jet, as defined in Sec. 5.5, are considered. An ideal candidate 2ℓOS event is expected to feature at least six hadronic jets, with two of these being identified ("b-tagged") as having originated from b-quarks. In order to maximise the sensitivity to the $t\bar{t}Z$ process due to limited statistics, two of the regions are defined for cases where either one jet or one b-jet of

[(1)] As in the case of the 3ℓ and 4ℓ selections described in Sec. 6.1 for the studies in the 2ℓOS channel, only electrons and muons were explicitly taken into account. Taus implicitly contribute if they decay leptonically.

[(2)] The physics objects are are listed in descending order according to their transverse momentum.

the 2ℓOS final-state signature is missing, labelled as 2ℓ-Z-1b6j and 2ℓ-Z-2b5j, respectively. The third region, 2ℓ-Z-2b6j, targets the full signature of the $t\bar{t}Z$ process. The selection criteria of the three targeted phase-space regions are presented in Tab. 7.1 A list of the expected yields within

Variable	2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
$N_\ell(\ell = e, \mu)$	2	2	2
	OSSF ($ee, \mu\mu$) pair with $\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$		
$p_\text{T}(\ell_1,\ \ell_2)$	$> 30\,\text{GeV}$, $> 15\,\text{GeV}$		
N_jets	≥ 6	5	≥ 6
$N_{b\text{-jets}}$	1@77 %	$\geq$ 2@77 %	$\geq$ 2@77 %

Tab. 7.1.: Definitions of the 2ℓOS target regions. The requirements on the leptons are specifically designed to select an on-shell Z boson. The targeted phase-space regions differ in the respective number of selected jets and b-jets. For the b-tagging the 77 % efficiency working point was employed.

the targeted regions is given in Tab. 7.2. The numbers for the $t\bar{t}Z$ signal and the considered background processes are taken from Monte Carlo simulation and are quoted with the corresponding statistical uncertainties based on the limited number of simulated events. The most dominant background is the associated production of a Z boson with additional jets for which the yields are given separately for the different flavour contributions of the final-state jets. The smallest relative contribution in terms of the overall Z+jets yields in the signal regions corresponds to events which do not contain any bottom- or charm-flavoured hadrons at parton-level. Such events are labelled as Z+light flavour and can be sufficiently suppressed due to the high rejection efficiency of light-flavoured jets of the DL1 b-tagging algorithm of above 99 % 5.5. The next higher fraction of Z+jets, denoted as Z+c, has at least one jet found to stem from a charmed hadron in the final state, and the majority of Z+jets events features at least one hadron originating from a bottom quark. The second-largest background contribution in the 2ℓOS channel is the production of $t\bar{t}$ events decaying in the dilepton channel. For the estimation of the contribution from dileptonic $t\bar{t}$ events a data-driven method was employed which will be elaborated on in Sec. 7.1.2 in further detail. The contributions from other SM background processes, such as the associated production of top quark pairs and single top quarks and either a vector or Higgs boson, as well the production of two bosons, are substantially smaller based on simulation, but are still significant and their predicted yields are incorporated into the analysis. With the exception of the 2ℓ-Z-1b6j region, where the amount of predicted events exceeds the data by a factor of 20 %, the agreement of the observed data events with the expected yields is very good. A significant difference in the background composition of the 2ℓ-Z-1b6j region with respect to the other target regions is the by construction the restriction to events featuring a single b-jet. The increased amount of Z+light flavour as well as of Z+c events was deemed to be the reason for the increased Z+light flavour and Z+c fractions. Due to the b-jet requirement in this region the suppression of those backgrounds is not as substantial as for the region with the requirement of at least two b-jets. For the Z+light flavour background fraction dedicated phase-space regions which feature a high purity of such events were able to be defined with the purpose to estimate the rate of this background from data. This procedure will be elaborated on later in Sec. 7.1.3.

The modelling of various kinematic variables was checked within all targeted phase-space regions. In Fig. 7.1 the transverse momentum of both the leading lepton and the leading jet are

Process	2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
$t\bar{t}Z$	302.3 ± 1.5	264.1 ± 1.3	443.2 ± 1.8
DD $t\bar{t}$	1209 ± 36	3857 ± 63	1875 ± 44
Z+light flavour	1209 ± 28	74 ± 17	18 ± 14
Z+c	3573 ± 30	670 ± 22	375 ± 11
Z+b	4774 ± 17	5221 ± 23	2697 ± 11
$t\bar{t}X$ $(X = W, H, \gamma)$	33.67 ± 0.93	93.8 ± 1.6	57.4 ± 1.1
VV, VH $(V = Z, W)$	491.9 ± 4.5	225.7 ± 3.8	150.4 ± 2.9
Other (tWZ, tZq, Wt)	105.7 ± 2.1	162.4 ± 3.2	118.4 ± 2.1
Total SM	$12\,553 \pm 58$	$10\,762 \pm 73$	5832 ± 49
Data	9929	10 661	5425
Data/SM	0.79	0.99	0.93
$t\bar{t}Z$ purity	2.4 %	2.5 %	7.6 %

Tab. 7.2.: Estimated and measured yields from simulation as well as those from measured data of the targeted analysis regions of the 2ℓOS analysis. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. Only the Monte Carlo statistical errors are shown.

presented. The distributions of the simulated processes were normalised to their respective cross section and scaled to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. In addition, the generator and calibration weights were applied. For each figure the rightmost bin contains, in addition to the events falling into it, the contributions from outside the depicted variable range, referred to as overflow. The data agree reasonably well with the Monte Carlo prediction within the uncertainties except for the 2ℓ-Z-1b6j target region where an excess of simulated events of roughly 20 % can be observed for all kinematic variables. Additional distributions of various kinematic variables can be found in Figs. F.5–F.8.

Although the target regions were constructed to target signal events specifically, the 2ℓOS channel is characterised by large background contributions, especially from dileptonically decaying $t\bar{t}$ events with additional jets and the associated production of a Z boson and jets. In order to sufficiently isolate the signal from the vast amount of selected background events, a multivariate analysis (MVA) technique referred to as a Boosted Decision Tree (BDT) was employed. Rather than training one BDT per target region in order to separate $t\bar{t}Z$ from the total background as was done in the previous ATLAS measurement in the 2ℓOS channel, two Boosted Decision Trees were trained separately against each of the two dominant backgrounds, dileptonic $t\bar{t}$ and Z+jets, within each of the targeted phase-space region. In this way, each BDT can focus entirely on the characteristics of the respective background process in order to increase its separation power. The two classifiers are then depicted in two-dimensional representation as illustrated in Fig. 7.2, such that for each of targeted phase-space regions (2ℓ-Z-1b6j, 2ℓ-Z-2b5j, 2ℓ-Z-2b6j) three dedicated subregions are selected in order to feature predominantly $t\bar{t}Z$, $t\bar{t}$ or Z+jets events, respectively. The detailed procedure of the multivariate analysis will be elaborated on in Sec. 7.2.

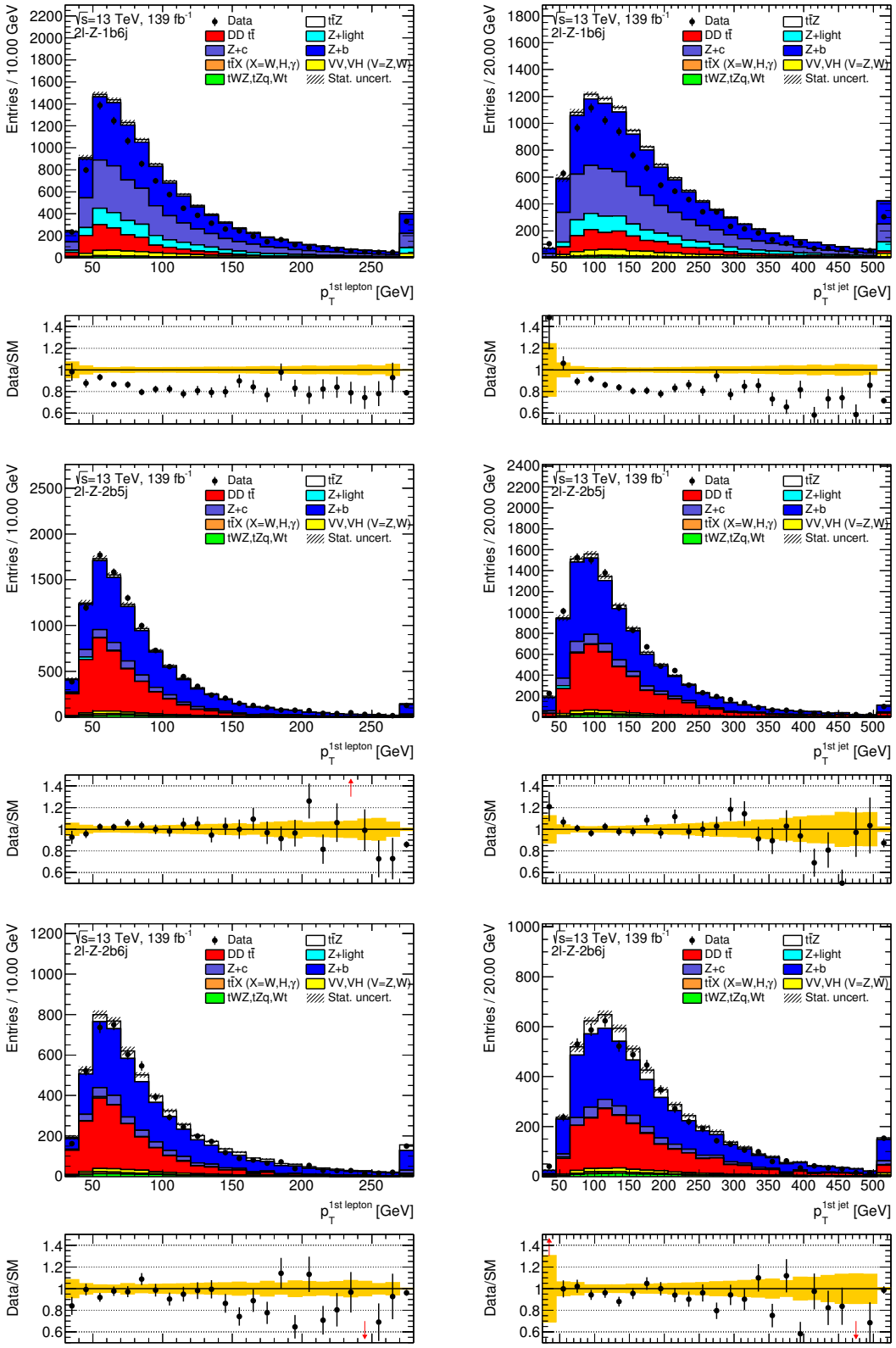


Fig. 7.1.: Distributions of the leading lepton p_{T} in the left column and of the leading jet p_{T} in the right column, shown for the three target regions. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

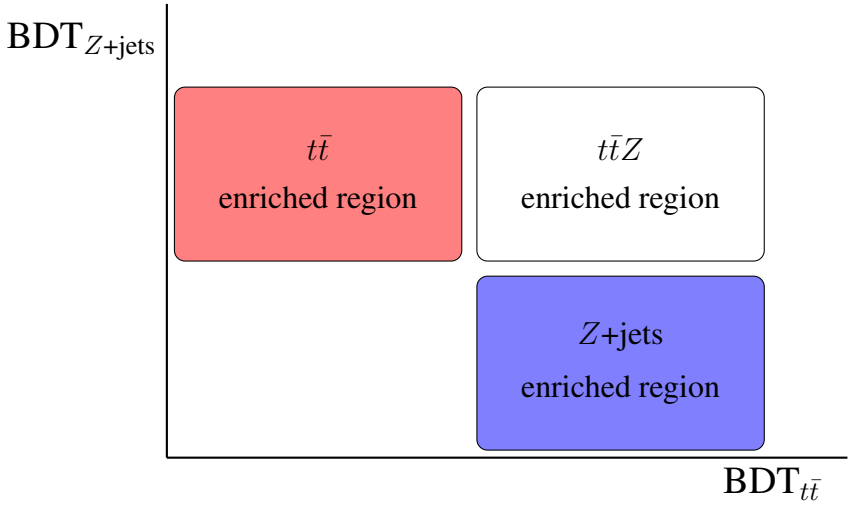


Fig. 7.2.: Sketch of the multivariate analysis strategy for the $t\bar{t}Z$ cross-section measurement in the 2ℓOS channel. In each of the target regions one BDT is trained against the $t\bar{t}$ background and a second BDT is trained against Z+jets events. By constructing a two-dimensional space based on the two BDT outputs, phase-space regions can be determined which are enriched predominantly in either, $t\bar{t}$ or Z+jets events.

7.1.2. Estimation of the tt̄ background

Dileptonically decaying $t\bar{t}$ events are a particularly important background to the $t\bar{t}Z$ process in the 2ℓOS channel. An ideal such event features only two b-jets from the decay of top quarks themselves into a W boson and a b-quark at tree-level. Therefore, at least three additional jets have to be produced within the dileptonic $t\bar{t}$ events, for example from pile-up, in order to be selected. As the modelling of this large number of additional jets suffers from large theoretical uncertainties, the $t\bar{t}$ background was estimated using a data-driven technique, following the approach employed in Ref. [39].

For each of the target regions a complementary region is constructed which is identical in all details to the original target region, but with the same-flavour ($ee, \mu\mu$) requirement replaced with a different-flavour ($e\mu$) requirement, referred to as $t\bar{t}$ validation regions. The corresponding selection criteria are summarised in Tab. 7.3 and differ to those of the target regions given in Tab. 7.1 in the requirement on the two leptons to be of different flavour (DF), which strongly suppresses at a very large extent all Standard Model processes featuring a real Z boson. As can be seen from the numbers given with the corresponding statistical uncertainties from Monte Carlo listed in Tab. 7.4, all selected phase-space regions feature a purity of dileptonic $t\bar{t}$ events of above $95\,\%$. Possible contributions from any other processes are highly suppressed. A generally good level of agreement between data and simulated events was observed.

In Fig. 7.3 the distributions of the leading lepton as well as the leading jet transverse momentum are presented for all $t\bar{t}$ validation. The simulated events were normalised to their respective theoretical cross sections as well as an integrated luminosity of $139\,\mathrm{fb}^{-1}$. Both generator and calibration weights were applied and the overflow bin was added. The data are well described by the Monte Carlo simulation within the statistical uncertainties for the two validation regions featuring at least two b-jets. Irrespective of the phase-space regions all distributions feature higher yields in the measured data with respect to the Monte Carlo simulation within the most populated range of the shown distributions. The observed deviations are at a level of roughly $20\,\%$ and are deemed

Variable	$e\mu$-1b6j	$e\mu$-2b5j	$e\mu$-2b6j
$N_\ell(\ell = e, \mu)$	2	2	2
	OSDF ($e\mu$) pair with $\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$		
$p_\text{T}(\ell_1,\ \ell_2)$	$> 30\,\text{GeV}$, $> 15\,\text{GeV}$		
N_jets	≥ 6	5	≥ 6
$N_{b\text{-jets}}$	1@77 %	$\geq$ 2@77 %	$\geq$ 2@77 %

Tab. 7.3.: Definitions of the $t\bar{t}$ validation regions for the data-driven $t\bar{t}$ estimation. The requirements on the leptons are specifically designed to select two leptons of opposite flavour, $e\mu$, as taus are not explicitly considered. The targeted phase-space regions differ in the respective number of selected jets and b-jets, respectively. For the b-tagging the 77 % efficiency working point was employed.

Process	$e\mu$-1b6j	$e\mu$-2b5j	$e\mu$-2b6j
$t\bar{t}Z$	6.69 ± 0.56	17.71 ± 0.74	14.11 ± 0.78
MC $t\bar{t}$	1071.8 ± 6.5	3695 ± 12	1697.8 ± 8.1
Z+light flavour	0.04 ± 0.03	–	–
Z+c	0.19 ± 0.07	0.19 ± 0.14	–
Z+b	0.50 ± 0.11	0.61 ± 0.14	0.30 ± 0.08
$t\bar{t}X$ $(X = W, H, \gamma)$	34.19 ± 0.99	98.8 ± 1.6	53.4 ± 1.1
VV, VH $(V = Z, W)$	6.05 ± 0.28	2.10 ± 0.34	1.13 ± 0.12
Other (tWZ, tZq, Wt)	25.5 ± 1.9	66.8 ± 3.0	31.3 ± 2.0
Total MC	1097.3 ± 6.5	3720 ± 12	1715.6 ± 8.1
Data	1251	3971	1925
Data/MC	1.14	1.07	1.12
$t\bar{t}$ purity	98 %	99 %	99 %

Tab. 7.4.: Estimated and measured yields from simulation as well as those from measured data of the $t\bar{t}$-enriched regions featuring the $e\mu$-selection criteria. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\text{fb}^{-1}$. All relevant weights were applied. Only the Monte Carlo statistical errors are shown.

to be related to the somewhat limited precision of the description of $t\bar{t}$ by Monte Carlo generators events in the high-jet-multiplicity regime as in the 2ℓOS channel [246]. Additional distributions of various kinematic variables can be found in Figs. F.10–F.13.

As a check whether the simulated $t\bar{t}$ events selected with the same-flavour (SF) requirement represent the same physics as those featuring an different-flavour lepton pair, the shapes of several kinematic variables were tested for consistency. For this purpose, distributions of either sample normalised to their respective number of entries were compared with each other, such as the examples depicted in Fig. 7.4. Before the distributions were normalised, all appropriate weights were

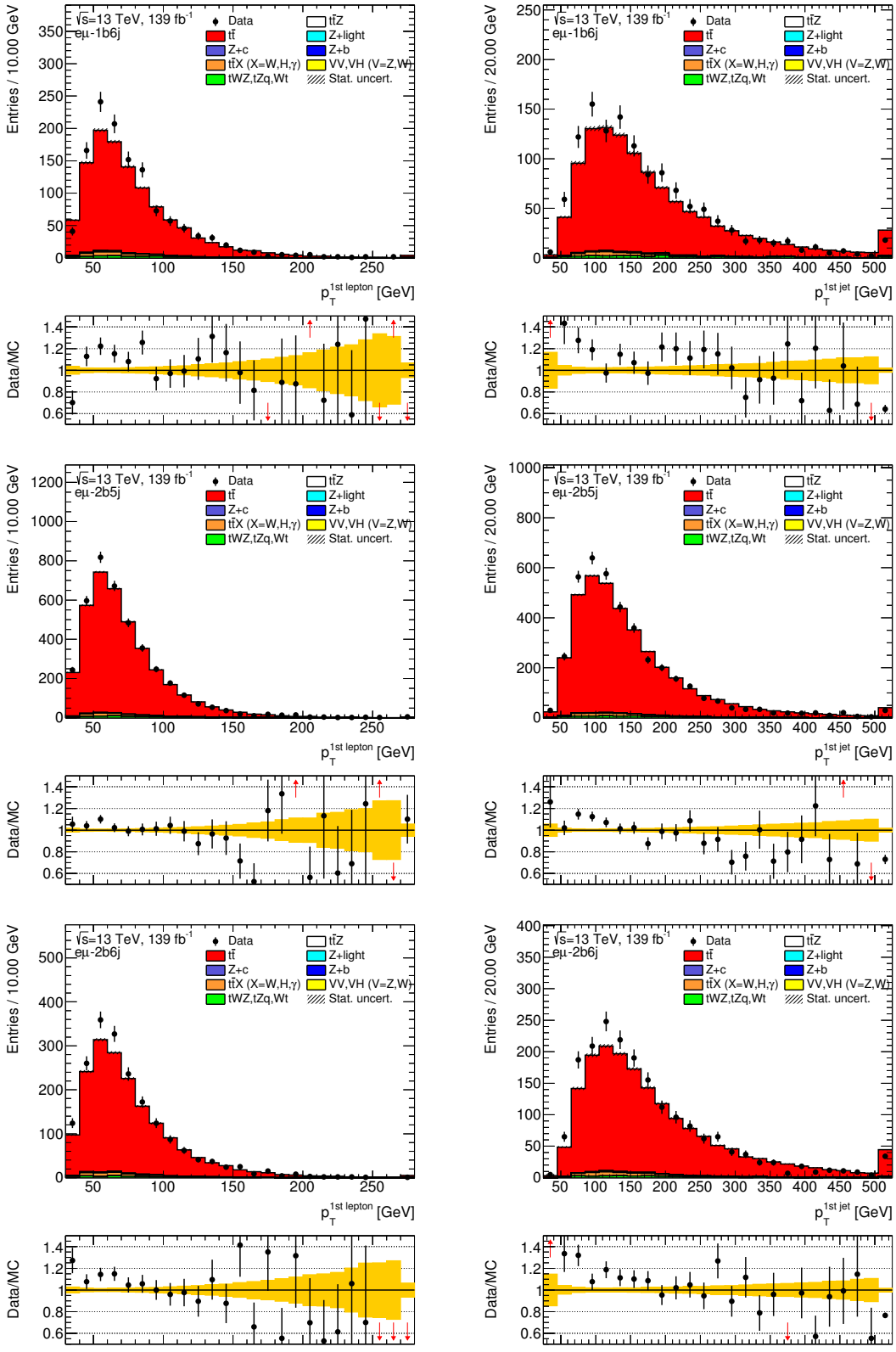


Fig. 7.3.: Distributions of the leading lepton p_{T} in the left column and of the leading jet p_{T} in the right column, shown for the three $t\bar{t}$ validation regions. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of 139 fb^{-1}. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

applied and contributions to the overflow bins were included. Both $t\bar{t}$ samples were found to be consistent with each other within the statistical uncertainties for all investigated variables. Further such control plots can be found in Figs. F.14–F.17.

For the data-driven $t\bar{t}$ estimation, the Monte Carlo prediction of all background processes except for $t\bar{t}$ are subtracted from the data within the DF-selected phase-space regions. The different detector acceptance and selection efficiencies for the SF and DF selections for dileptonic $t\bar{t}$ events are taken into account with a correction factor $C_{t\bar{t}}$, which is defined as the ratio between the Monte Carlo prediction of dileptonic $t\bar{t}$ for the same- and different-flavour selection, respectively:

$$C_{t\bar{t}} = \frac{N_{t\bar{t}}^{\ell\ell}}{N_{t\bar{t}}^{e\mu}} \tag{7.1}$$

The values of $C_{t\bar{t}}$ were calculated according to Eq. (7.1) separately for each target region with the results quoted in Tab. 7.5. The quoted errors consists of the statistical uncertainties of both set of $t\bar{t}$ events, selected with either an DF or SF criterion, and the amount of mismodelling within the respective validation region. Despite a certain mismodelling of the data in the $e\mu$-validation region the values for $C_{t\bar{t}}$ are very close to unity. However, a further investigation of this issue was not feasible on the time scale of this thesis.

2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
1.005 ± 0.140	1.000 ± 0.070	0.999 ± 0.120

Tab. 7.5.: Correction factors for the data-driven $t\bar{t}$ estimation to take the different detector acceptance and selection efficiencies of the dileptonic $t\bar{t}$ events into account. The factor was calculated for each target region separately according to Eq. (7.1). The stated errors are built from statistical uncertainties from the $t\bar{t}$ yields and the amount of mismodelling within the validation regions.

The data-driven $t\bar{t}$ estimate was employed within most steps of the analysis, such as the extraction of event yields and within the statistical treatment employed for the cross-section measurement. For the training of the various Boosted Decision Trees, however, the Monte Carlo estimation of the $t\bar{t}$ process, featuring the standard SF-selection, was used. The numbers and distributions of $t\bar{t}$ events in Tab. 7.4 and Fig. 7.3 are taken from MC $t\bar{t}$ simulation only as the data-driven estimate was based on those.

7.1.3. Z+light background normalisation

The associated production of a Z boson and additional jets is the major background to the $t\bar{t}Z$ signal within the 2ℓOS channel. Due to the limited precision of the modelling of this important background process in phase-space regions with particularly high jet and b-jet multiplicities, the simulated numbers of jets and b-jets suffer from large systematic uncertainties. Therefore, it is more favourable to obtain the normalisation of the Z+jets process not from Monte Carlo simulation but rather to measure it from data within dedicated phase-space regions highly enriched in Z+jets events.

From the yields in Tab. 7.2 and the distributions in Fig. 7.1 one can observe a rather uniformly distributed increased level of predicted events with respect to data within phase-space regions which feature an elevated fraction of Z+light flavour events. For the determination of the Z+light flavour normalisation two dedicated regions in phase space, referred to as Z+light flavour control regions, were defined with the selection criteria summarised in Tab. 7.6. The requirements

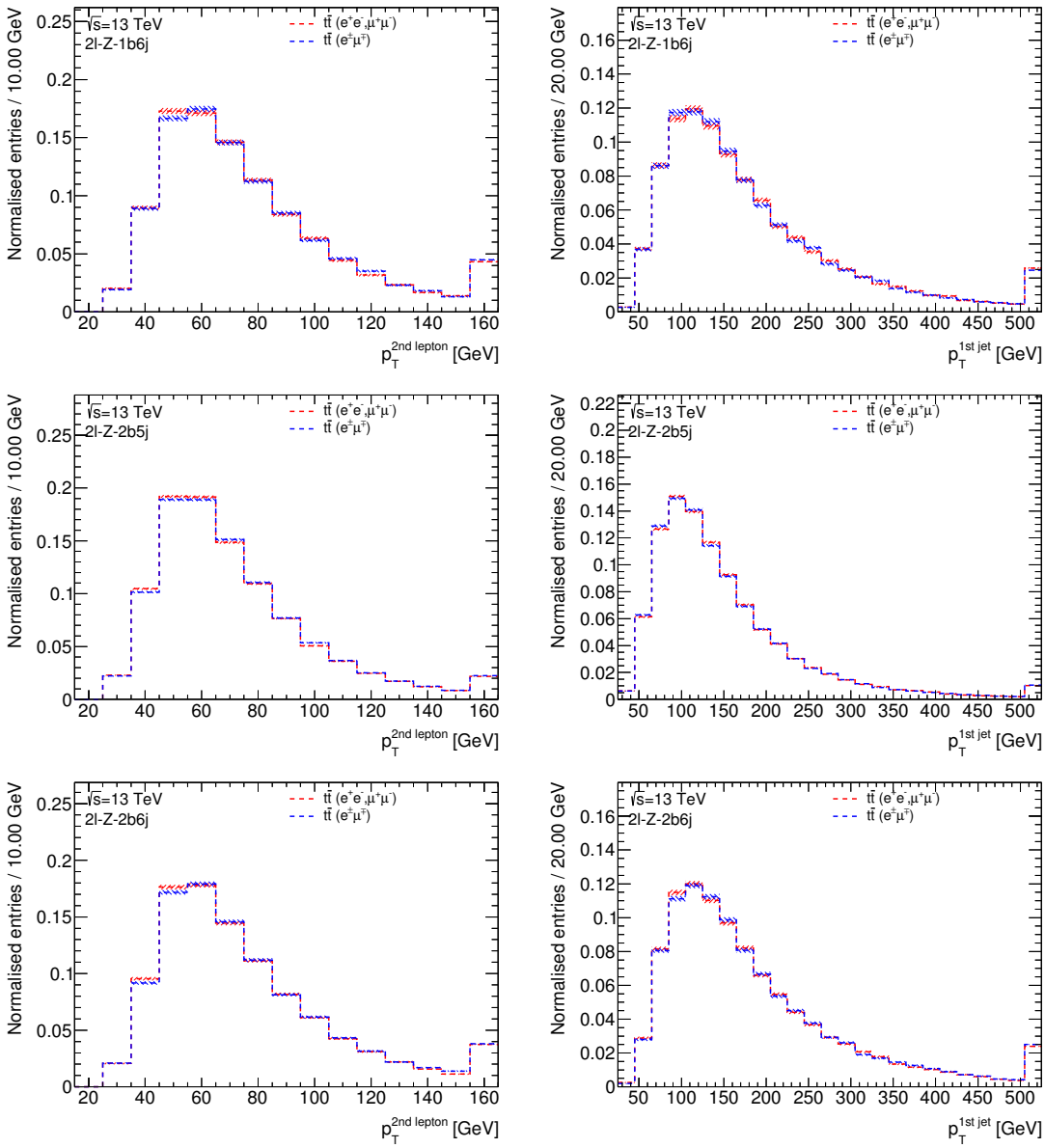


Fig. 7.4.: Distributions of the leading lepton p_{T} in the left column and of the jet p_{T} in the right column, shown for $t\bar{t}$ events selected with a SF-criterion in red and for $t\bar{t}$ events selected with an DF-criterion in blue. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. All distributions were normalised to their respective number of entries. The errors indicated by the striped bands represent the statistical uncertainties from Monte Carlo simulation.

on the lepton selection are equal to those in the target regions. In order to ensure a dominant contribution from Z+light flavour events as well as orthogonality to the target regions, a veto on b-tagged jets is applied using a b-tagging efficiency of 77 %. Due to the different kinematics with respect to the jet multiplicity within Z+jets events, one phase-space region featuring exactly five jets and another region with six or more jets was defined, respectively. The yields of data and sim-

Variable	2ℓ-Z-0b5j	2ℓ-Z-0b6j
$N_\ell(\ell = e, \mu)$	2 OSSF ($ee, \mu\mu$) pair with $\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$	2
$p_\text{T}(\ell_1,\ \ell_2)$	$> 30\,\text{GeV}\ , > 15\,\text{GeV}$	
N_jets	5	≥ 6
$N_{b\text{-jets}}$	0@77 %	0@77 %

Tab. 7.6.: Definitions of the Z+light flavour control regions. The requirements on the leptons are specifically designed to select an on-shell Z boson. The targeted phase-space regions differ in the respective number of selected jets while b-jets are vetoed. For the b-tagging the 77 % efficiency working point was employed.

ulated events within the Z+light flavour control regions are listed in Tab. 7.7 which also quotes the statistical uncertainties on each value based on the generated Monte Carlo statistics for the respective process. Both control regions were found to feature a purity of Z+light flavour events of higher than 60 % as well as an increased amount of predicted events over data which is deemed to be caused by the dominant fraction of the Z+light flavour background. The number of predicted $t\bar{t}$ events was taken from Monte Carlo simulated samples in the Z+light flavour control regions rather than from the data-driven estimate described in Sec. 7.1.2 since this method was not validated for phase spaces without any b-jet.

As it was done for the target regions, the modelling of the Z+light flavour control regions was investigated with the help of various kinematic variables. The transverse momentum of both the leading lepton and the leading jet as well as the missing transverse energy are depicted in Fig. 7.5. All simulated processes were normalised to their respective cross section and scaled to an integrated luminosity of $139\,\text{fb}^{-1}$. Both generator and calibration weights were applied and contributions to the overflow bins were included. In the range where most of the events are cumulated the agreement between data and simulated events is sufficiently flat, albeit a slight negative slope trend is visible. The predicted events generally exceed the data by a factor of approximately 20 % as it was similarly observed within the 2ℓ-Z-1b6j region which features the highest Z+light flavour fraction of all target regions. Therefore, the Z+light flavour was deemed to be causing this factor of 20 % also for the 2ℓ-Z-1b6j target region and the selected control regions were able to be proven a reasonable choice in order to determine of the Z+light flavour normalisation from data. Distributions similar to those depicted in Fig. 7.5 are depicted in Fig. F.9 but for different kinematic variables.

The determination of dedicated control regions for the Z+c and Z+b background contribution, however, deemed not to be feasible based on the first set of event-selection criteria, but will be revisited at a later stage of the analysis. Since both the Z+c and Z+b background contribution a signature which is quite similar to that of signal events, they represent the largest relative fraction of events to the Z+jets background within the target regions. As outlined at the end of Sec. 7.1.1, a dedicated BDT was trained for each target region in order to separate the signal from the Z+jets

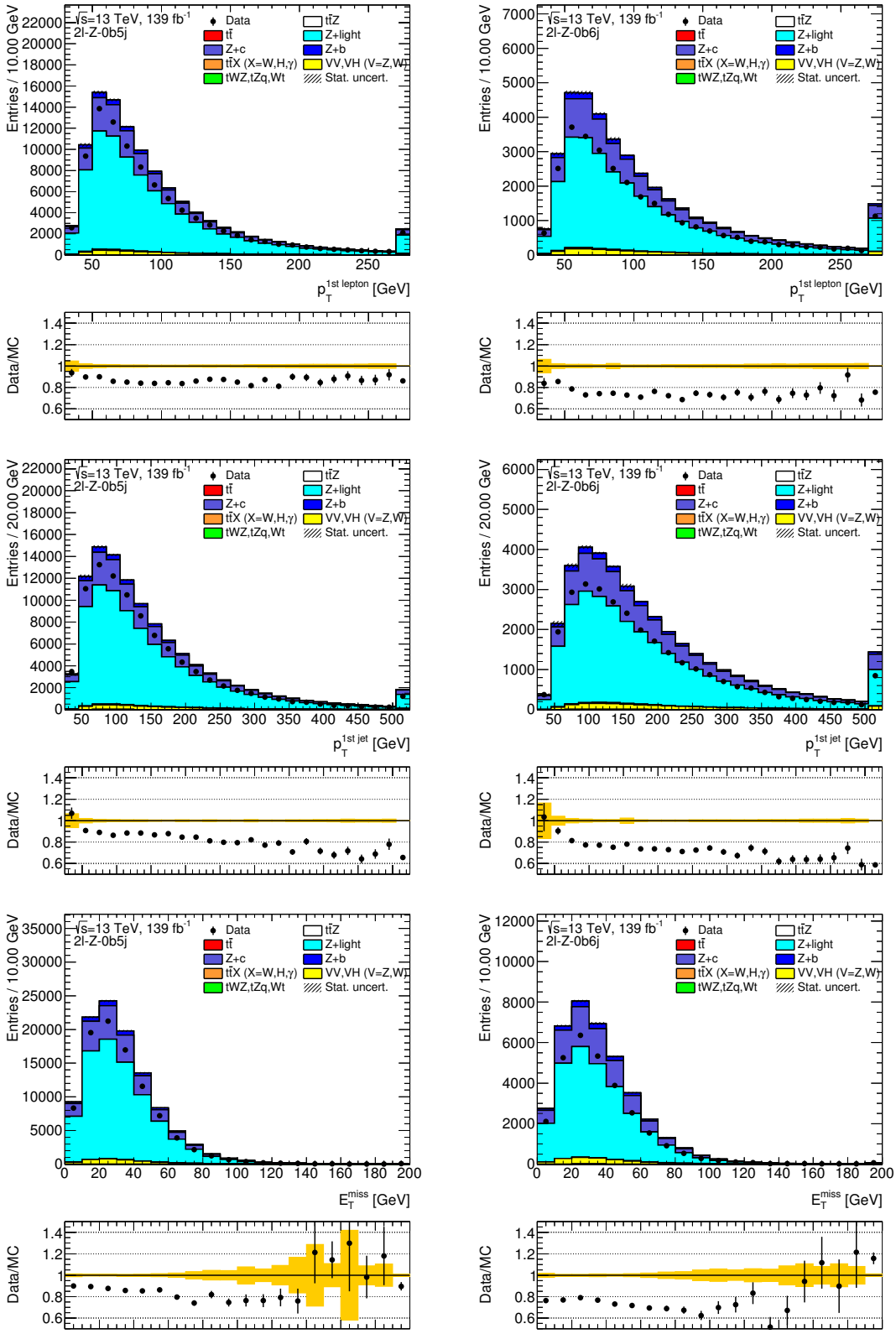


Fig. 7.5.: Distributions of the leading lepton p_{T} in the top row, the leading jet p_{T} in the middle row, and of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the bottom row. The distributions of the 2ℓ-Z-0b5j region are presented in the left column and those showing the 2ℓ-Z-0b6j region are on the right. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band are statistical only.

Process	2ℓ-Z-0b5j	2ℓ-Z-0b6j
$t\bar{t}Z$	67.43 ± 0.66	69.99 ± 0.74
MC $t\bar{t}$	522.9 ± 5.0	214.2 ± 2.9
Z+light flavour	$79\,260 \pm 300$	$26\,140 \pm 120$
Z+c	$22\,100 \pm 94$	9500 ± 41
Z+b	3300 ± 22	1381.8 ± 9.2
$t\bar{t}X$ $(X = W, H, \gamma)$	15.07 ± 0.69	7.00 ± 0.43
VV, VH $(V = Z, W)$	3500 ± 11	1658.4 ± 7.6
Other (tWZ, tZq, Wt)	61.8 ± 1.9	28.9 ± 1.2
Total SM	$115\,190 \pm 320$	$42\,240 \pm 130$
Data	94 024	29 323
Data/MC	0.81	0.69
Z+light flavour purity	69 %	62 %

Tab. 7.7.: Estimated and measured yields from simulation as well as those from measured data of the Z+light flavour control regions of the 2ℓOS analysis. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\text{fb}^{-1}$. All relevant weights were applied. Only the Monte Carlo statistical errors are shown.

contribution with the aim to be able to use the resulting Z+jets-enriched phase-space regions as control regions for at least one of the Z+c and Z+b background contributions.

7.2. Multivariate analysis

In the previous ATLAS measurement performed in the 2ℓOS channel [39], a single BDT was trained within each target region for the purpose of constructing a single output variable to be used to discriminate between candidate signal and background events. For the new approach, however, in each target region two Boosted Decision Trees were trained separately against a single one of the backgrounds in each case, either dileptonic $t\bar{t}$ or Z+jets production.

A commonly employed procedure for the separation of signal and background events is to set cuts on the individual variables for which the shapes of signal and background processes are sufficiently different from each other. The use of multivariate analysis techniques becomes favourable over such cut-based approaches if signal and background events cannot be sufficiently discriminated anymore by conventional means. In contrast, MVA methods are able to exploit the discriminating features of both the shapes of and the correlations between many variables in order to retain as much as information from the various inputs. In order to solve a classification problem, several variables with a certain amount of discrimination between different physical processes are combined within a classifier in order to build a single output quantity which should then be sensitive to the discriminating features provided by all input variables. In general, MVA techniques can be used for both classification and regression purposes. The aim of a regression procedure is, on the other hand, to find an underlying functional behaviour of a given distribution or at least to find an approximation for it. However, as the use case defined within this thesis is clearly a classification problem, the focus will be entirely on this type in what follows.

7.2.1. Boosted Decision Trees

The method chosen for the studies presented by this theses is referred to as Boosted Decision Tree (BDT), which is one of the most-employed MVA methods in high-energy physics. It has the clear benefits of high a stability with respect to both the handling of large input datasets and to statistical fluctuations of the input variables. The BDT belongs to the family of the so-called supervised learning algorithms [247], which means that it makes use of training events for which the desired output is already known in order to determine the mapping function from the input variables to the final decision. In the context of this thesis, the implementation of a BDT from the Toolkit for Multivariate Analysis [248] (TMVA) was employed, which allows for the processing, evaluation and application of a broad spectrum of different multivariate techniques using the data analysis framework ROOT [249].

A decision tree is a tree-structured classifier which evolves from a single root node which comprises all input events irrespective of whether they belong to the signal or background datasets. With a binary decision using the variable with the highest discrimination power the input dataset is cut into two subsets which are treated independently from now on and which serve as root nodes for the next iteration, as illustrated in Fig. 7.6. This procedure is repeated until all nodes fulfil certain stop criteria which have to be defined in advance of the BDT training. The final nodes should then feature a high purity of events that have been classified as either signal-like or background-like. Each event is assigned a specific value of the final output score which indicates how signal-like or background-like it was classified. The Boosted Decision Trees implemented in TMVA have a defined output score range of $[-1, +1]$, with a value of $+1$ representing an absolute signal-like classification and a value of -1 an absolute background-like classification [248].

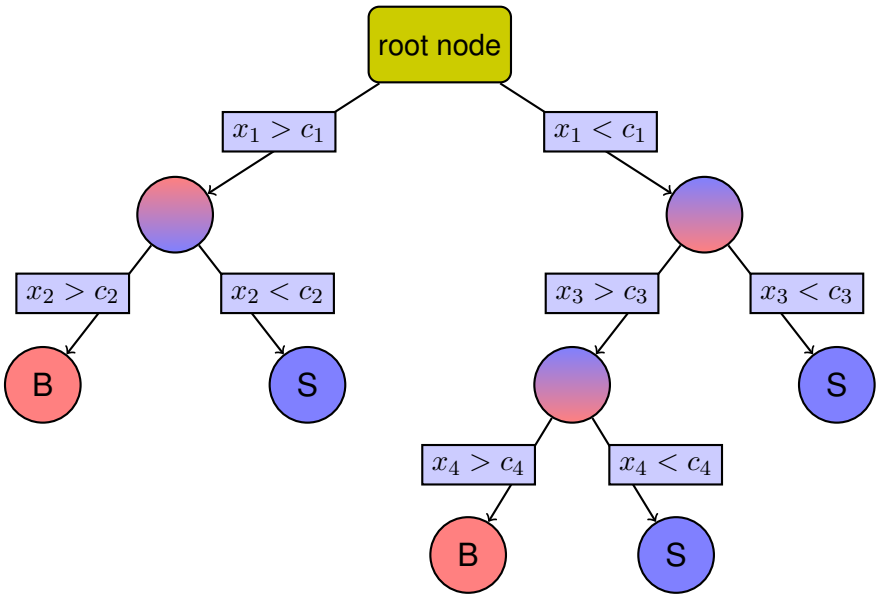


Fig. 7.6.: Schematic illustration of a decision tree. Beginning at the root node, a sequence of binary decisions is applied to a dataset by setting cuts at specific values c_i on the distribution of a variable x_i. The leaf nodes at the ends of the branches are labelled "S" and "B" depending on whether the majority of events ending up in the respective node was classified as signal or background, respectively. The image was adapted from [248].

In order to determine which variable is best to be cut on at a certain node and to find the optimal cut value, TMVA provides several methods which are listed in Tab. 7.8. Based on the purity p of

true signal events of a certain parent node, each separation method uses a different mathematical expression in order to optimise the separation between the two prospective daughter nodes. The purity is defined as the ratio of signal events to the sum of signal and background events,

$$p = \frac{S}{S+B} \quad , \tag{7.2}$$

such that nodes consisting exclusively of true background events have a purity of $p = 0$. Since cuts predominantly selecting background are as important as cuts with a high signal efficiency, the criteria are constructed such that they are symmetric with respect to the event classes. Therefore,

Separation type	Mathematical expression
Gini index	$p(1-p)$
Cross entropy	$p - \ln(p) - (1-p)\ln(1-p)$
Misclassification error	$1 - \max(p, 1-p)$
Statistical significance	$S/\sqrt{S+B}$

Tab. 7.8.: List of separation techniques as provided by TMVA with their mathematical expressions. They are used to determine both the variable with the highest discrimination at a certain node and the corresponding optimal value to cut on. All methods except the statistical significance feature a maximum for a purity of $p = 0.5$ and fall off to zero if the node consists of exclusively signal or background events.

when building a decision tree, the same variable might be employed several times whereas others might not be used at all. The more often a single variable was used to make a decision for the growth of a decision tree, the higher it is ranked. In addition, each count is weighted by the output value of the mathematical expression of the employed separation method and the number of events in the respective node [248].

A well-known shortcoming of decision trees is the instability with respect to statistical fluctuations as it is the case for most MVA techniques. In order to prevent training on statistical fluctuations, commonly referred to as overtraining, several countermeasures were developed, such as the limitation of both the depth of a tree up to a given maximum number of nodes as well as the minimum number of events within the final nodes. Another possibility to increase the stability with respect to statistical fluctuations of the input datasets is not to train a single decision tree but to construct a forest consisting of many of such trees. After the training of a single decision tree, misclassified events are assigned a weight such they will be treated with particular care during the training of the subsequent decision tree. In addition, the number of grown trees is typically restricted in order to prevent from overtraining. By applying such a reweighting, also referred to a boosting [250, 251], the classification performance greatly improves which is the reason why Boosted Decision Trees are such a widely-used MVA techniques. A small drawback of the boosting procedure, however, is the increased difficulty to keep track of individual decisions at the same time. The final classification of an event is made by taking the weighted average of all decision trees. The two boosting methods which are enabled by TMVA, AdaBoost and Gradient Boost, are both elaborated in App. C.1 in more detail [248].

The shapes of the input variables are sometimes in a form such that the entire information content cannot be extracted from them. For such cases TMVA provides a set of preprocessing methods in order to transform the shapes of the input variables into more appropriate forms as well as to reduce correlations between several variables, which accelerates the response time of the applied MVA technique. The available preprocessing techniques range from a simple normalisation of

the variables to more involved decorrelation and transformation methods which are discussed in App C.2 in more detail.provide Boosted Decision Trees are highly favoured MVA techniques in many fields of application including high-energy physics for many reasons, among them the most important being the robustness with respect to statistical fluctuations of the input variables as well as the ability handling large datasets. Boosted Decision Trees are mostly insensitive to variables with poor discrimination between signal and background since in each node the split decision is made with the help of the most discriminating variable only. In general, not much tuning and adjustment of the setting parameters is necessary in order to obtain reasonable results with a BDT which makes it one of the best "out-of-the-box" classifiers.

7.2.2. Definitions of discriminating variables

A number of key variables was identified which featured pronounced differences in the shapes for the signal and the two dominant background processes. Only those variables which were shown to provide the highest level of discrimination for a given phase-space region and for the discrimination of a particular background process were chosen for the training of the various Boosted Decision Trees. In the following, all variables used within a BDT training will be explained and a short physical motivation will be given.

A comprehensive list of all input variables can be found in Tab. 7.9 for the training against the $t\bar{t}$ background, and the variables for the training against the Z+jets background are summarised in Tab. 7.10. The lists are sorted according the importance of the respective variables within the training procedure as determined by the BDT machinery. In addition, similar listings of the variables ranked according to their so-called discriminating power are given in Tabs. G.1 and G.2, respectively. The separation power $\langle s^2 \rangle$ used to rank the variables based on their individual ability to discriminate between the signal and background before the training, as well as to decide whether to use a certain variable for the training of one of the various Boosted Decision Trees, was calculated according to

$$\langle s^2 \rangle = \frac{1}{2} \sum_{i=1}^{N} \frac{(n_{S,i} - n_{B,i})^2}{n_{S,i} + n_{B,i}} \tag{7.3}$$

where $n_{S,i}$ and $n_{B,i}$ represent the amount of signal and background events in bin i, respectively. For all variables employed for the various BDT training procedures the distributions of the normalised signal and dominant background processes as well as the level of agreement between data and total MC prediction were checked, in order to ensure that the input variables were well modelled. The normalised distributions of the variables were found not to exhibit any problematic features such as unexpected discontinuities in the shapes, and a reasonable agreement of the data and Monte Carlo simulations could be observed. A few examples of such distributions are depicted in Fig. 7.7, presenting a selection of highly-ranked variable per background process and target region, normalised to the respective number of entries. The same variables are shown in Fig. 7.8 with all processes scaled to their respective cross section and normalised to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. Data and simulated events agree quite well within the statistical uncertainties for the target regions, although within 2ℓ-Z-1b6j an excess of simulated events over the data is present which could already observed in Fig. 7.1, for example. An overview of all discriminating variables can be found in App. F.9, presenting the distributions normalised to their respective number of entries. The corresponding figures with all simulated processes scaled to their respective cross section and normalised to an integrated luminosity of $139\,\mathrm{fb}^{-1}$ are shown in App. F.10.

Despite the selection requirements on the lepton pair, the amount of selected signal events in the target regions feature a substantial fraction of cases where the $t\bar{t}$ system of $t\bar{t}Z$ decays is the

lepton+jets channel. In order to have a training model predominantly consisting of $t\bar{t}Z$ events with a fully-hadronic $t\bar{t}$ decay and an electron-positron or muon-antimuon pair compatible with a Z boson in the final state, the signal dataset was divided into two subsets comprised of events with the desired signature, which was used for training procedures, and the rest, respectively. The two subsets of signal events are presented separately in Fig. 7.7 and the corresponding distributions shown in App. F.9. The separation power of $t\bar{t}Z$ events with a fully-hadronic $t\bar{t}$ system against either of the dominant background processes were calculated according to Eq. (7.3) and are quoted in the figures. The ordering of the variables quoted in Tabs. G.1 and G.2 is derived using the training dataset of $t\bar{t}Z$ and the respective background datasets.

7.2.2.1. Discrimination of the $t\bar{t}$ background

In this section, the input variables which were employed for the BDT training procedures against the $t\bar{t}$ background are defined, together, in several cases, with some of the underlying reasons for the exhibited discriminating features.

Leading and subleading lepton transverse momenta $\boxed{p_{\mathrm{T}}^{\ell_i} \quad (i = 1, 2)}$

For signal events, the two leptons are assumed to originate from the decay of the Z boson which is either radiated from one of the top quarks before its decay receiving a considerable amount of recoil energy, or even from an high-energy initial-state quark. In either case the leptons are expected to feature an on average higher transverse momentum than the leptons from W boson decays of dileptonic $t\bar{t}$ events after the top quark decay. In addition, in the case of leptons originating from the $t\bar{t}$ system, a certain amount of the momentum the W received from the top quarks is taken away by the neutrino.

Separation in η-ϕ-space of the two leptons $\boxed{\Delta R(\ell, \ell)}$

Due to the higher momentum the two leptons feature a smaller separation on η-ϕ-space for signal events than for background.

Missing transverse energy $\boxed{E_{\mathrm{T}}^{\mathrm{miss}}}$

At tree-level, the $t\bar{t}Z$ signature features no source of missing transverse energy leading to small values of $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by reconstruction and identification inefficiencies. In contrast, the amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected to considerably higher for dileptonically decaying $t\bar{t}$ events due to the two neutrinos in the final state.

Invariant mass of the dilepton system $\boxed{m_{\ell\ell}}$

The two leptons are selected under the assumption that they originate from an on-shell Z boson and are thus required to have a value of the combined invariant mass compatible with that of a Z boson, as stated in Sec. 7.1.1. Therefore, signal events clearly reflect the resonance of the Z boson at approximately 91 GeV in the mass range of 80 GeV to 100 GeV the selection of events is restricted to, whereas $t\bar{t}$ background events constitute a flat distribution for mass values in that range.

Separation in η-ϕ-space of the dilepton system and the two b-jets $\boxed{\Delta R(bb, \ell\ell)}$

The two leptons are assumed to originate from the decay of the Z boson, which is either radiated

from one of the two top quarks or recoils against the $t\bar{t}$ system, whereas for dilepton $t\bar{t}$ events the two leptons originate from the decay of the associated W boson. The separation in η-ϕ-space of the dilepton system and the two b-jets are therefore expected to on average be differently distributed in $t\bar{t}Z$ events than in dileptonic $t\bar{t}$ events. The b-jets considered to build this variable were required to have the highest b-tagging output weight of all jets in an event.

Minimum of the sum of the invariant masses of all possible combinations of b-jets and leptons $\boxed{m_{b\ell}^{\text{min}}}$

For dileptonically decaying top quark pairs, two b-jets and two leptons can be combined to partially reconstruct both top quarks. The combination of the b-jet and the lepton which belong to the same top quark feature on average a smaller invariant mass value than the wrongly associated $b\ell$-pairs. This feature is employed for the discrimination against dileptonic $t\bar{t}$ events since in the 2ℓOS channel the leptons cannot be inextricably connected to either of the top quarks. As it is not a priori known which b-jet has to be associated to which lepton, the invariant mass values of all $b\ell$-associations were calculated and then combined aiming to find the minimum of the sums of both pairs for a respective $b\ell$-association:

$$m_{b\ell}^{\text{min}} = \frac{1}{2} \min \left\{ \left(m_{b_1\ell_1} + m_{b_2\ell_2} \right), \left(m_{b_1\ell_2} + m_{b_2\ell_1} \right) \right\} \tag{7.4}$$

The factor of one half was added for convenience in order to keep the outcome of Eq. (7.4) smaller than the top quark mass. The b-jets considered to build this variable were required to have the highest b-tagging output weight of all jets in an event.

Maximum of the minima of the invariant masses of all possible combinations of b-jets and leptons $\boxed{m_{b\ell}^{\text{max-min}}}$

As an extension to the definition of $m_{b\ell}^{\text{min}}$, the differences between the $t\bar{t}Z$ system decaying in the 2ℓOS channel and dileptonic $t\bar{t}$ events were exploited by taking the minimum of the invariant mass of each $b\ell$-pair per association and to take the maximum of both:

$$m_{b\ell}^{\text{max-min}} = \max \left\{ \min \left\{ m_{b_1\ell_1}, m_{b_2\ell_2} \right\}, \min \left\{ m_{b_1\ell_2}, m_{b_2\ell_1} \right\} \right\} \tag{7.5}$$

This variable has already be seen to offer some sensitivity to true dileptonic $t\bar{t}$ events in Ref. [252]. The b-jets considered to build this variable were required to have the highest b-tagging output weight of all jets in an event.

Maximum of the invariant mass of the two systems of a light lepton and the closest b-jet $\boxed{\max m_{b\ell}^{\min \Delta R}}$

For dileptonically decaying top quark pairs, the b-jet and the lepton from the subsequent W boson decay, which belong the same top quark, are expected to on average have an smaller angular distance to each other than to the b-jet and lepton of the other top quark, respectively. As a choice the $b\ell$-pair with the higher value of its invariant mass was taken as a discriminating variable for the BDT training. This variable provides a better discrimination the higher the initial energy of the $t\bar{t}Z$ and $t\bar{t}$ systems, respectively. The b-jets considered to build this variable were required to have the highest b-tagging output weight of all jets in an event.

Two-neutrino scanning method output weight $w_{2\nu\text{SM}}$

The $t\bar{t}Z$ signal process features a lepton pair from the decay of the Z boson, whereas for the $t\bar{t}$ background the two leptons are expected to originate from W boson decay. Consequently, the kinematic reconstruction of the dileptonic $t\bar{t}$ signature provides a significant discrimination to $t\bar{t}Z$. However, the two neutrinos in the final state of dileptonic $t\bar{t}$ events both contribute to a single measured quantity, the missing transverse momentum, but the individual four-momentum vectors of the two neutrinos are a priori unknown. Therefore, a novel technique, referred to as two-neutrino scanning method, which aims to reconstruct the two neutrino four-vectors and the $t\bar{t}$ system, was developed and employed for this analysis.

Numerous different assumptions on the values of the azimuthal angle and pseudorapidity of either neutrino are tested by systematically scanning the η-ϕ-space and, with the set of the respective values at each point in η-ϕ-space of the two neutrinos, the $t\bar{t}$ signature is constructed with the information from the two leptons and b-jets, respectively. The b-jets considered to build this variable were required to have the highest b-tagging output weight of all jets in an event. Of all hypotheses for the neutrino four-momenta, the most probable solution is determined based on the interpolation of the reconstructed top quark mass with an idealised distribution of top quark invariant mass. This distribution was constructed from a lepton and a b-jet matched to the corresponding parton-level objects as well as the parton-level neutrino. In addition, the difference between the sum of the two neutrino momenta in the x-direction, based on parton-level information, and the x-component of $E_{\text{T}}^{\text{miss}}$ is interpolated with an idealised reference distribution which was built in a similar manner than that for the top quark mass. The same procedure is performed for the y-component of the missing transverse energy. A single output weight, $w_{2\nu\text{SM}}$, is defined as the product of the three interpolated values, and the combination of reconstructed objects and neutrino hypotheses yielding the largest output weight is selected as the reconstructed dileptonic $t\bar{t}$ system.

The two-neutrino scanning method has been used for the first time in an analysis and its output weights was seen to be assigned a high rank from the BDT. A more detailed prescription of this method can be found in App. D.

7.2.2.2. Discrimination of the Z+jets background

The BDT input variables which predominantly discriminate against the Z+jets background are explained in the following paragraphs.

Transverse momenta of the second to the sixth jet $p_{\text{T}}^{j_i} \quad (i = 2, 3, 4, 5, 6)$

The $t\bar{t}Z$ signature in the 2ℓOS channel features six jets at tree-level from the decay of the top-antitop quark pair. In contrast, for Z+jets events the jets originate from QCD processes such as gluon radiation or gluon splitting, and therefore have less transverse momentum than the jets induced by the $t\bar{t}$ final state objects, in particular the two b-jets, defined as the two jets in an event with the highest b-tagging output weight. However, high-energy ISR jets are likely present in Z+jets events which is the reason why the transverse momentum of the leading jet does not provide as much discrimination power as the other considered five jets and is thus not employed as an discriminating variable. In case of the 2ℓ-Z-2b5j target region $p_{\text{T}}^{\text{6th jet}}$ cannot be used for the BDT training.

Transverse momentum of the leading b-jet $p_{\text{T}}^{\text{1st }b\text{-jet}}$

For signal events, two b-jet, defined as the two jets in an event with the highest b-tagging output weights, are expected from the decay of either top quark. Since the mass difference between the top quark and its decay products is of the order of the mass of the Z boson, both the b-quark and the W boson receive a considerable boost. In case of the Z+jets background, it is suspected that the main source of heavy-flavour jets is gluon splitting. Jets originating from them, such as the b-jets in Z+jets events, are expected to carry on average less momentum than the b-jets from top quark decays. However, as the subleading b-jet p_{T} turned out to offer much less discrimination power than the leading one, only the latter was used for the BDT training.

Scalar sum of the transverse momenta of all leptons and jets $\boxed{H_{\mathrm{T}}}$

At tree-level, signal events feature six jets which originate from the $t\bar{t}$ final states, whereas for the Z+jets background it is suspected that the main source of jets is QCD radiation. Therefore, the transverse momenta of the individual jets as well as the scalar sum is larger for $t\bar{t}Z$ events than for Z+jets events. Although both processes feature an on-shell Z boson, the leptons on average have an higher transverse momentum in signal events with respect to Z+jets events. As the Z boson is either radiated from one of the two top quarks or recoils against the $t\bar{t}$ system, it receives a considerable boost which is entirely passed to its decay products.

Invariant mass of the two b-jets $\boxed{m_{bb}}$

The two b-jets within $t\bar{t}Z$ events are expected to originate from the decay of the top quarks. For Z+jets events, in contrast, it is suspected that b-jets predominantly originate from the splitting of a gluon into a pair of b-quarks. As the gluon is assumed to be massless, the invariant mass of the two b-jets is considerably smaller for the background than for the signal. However, it is not compatible with zero due to detector effects and the limited jet energy resolution.

Pseudorapidity of the dilepton system $\boxed{|\eta_{\ell\ell}|}$

The value of the pseudorapidity of the dilepton system is on average greater for Z+jets events than for the signal process. As the $t\bar{t}Z$ system comprises three heavy particles, it features a high invariant mass and most of the energy of the initial-state partons is required to produce a $t\bar{t}Z$ system. Therefore, only a small amount of the energy is left to boost the $t\bar{t}Z$ system which is located in the central part of the detector at lower pseudorapidity values. In contrast, much less energy is needed to produce a single Z boson, which is the reason why within Z+jets events the dilepton system receives a greater boost in the direction of the beam axis corresponding to higher values of the pseudorapidity. Since the detector setup is symmetric, the absolute value of the pseudorapidity was used as a discriminating variable.

Average of the separation in η-ϕ-space of all jet pairs $\boxed{\Delta R_{jj}^{\mathrm{avg}}}$

$t\bar{t}Z$ events feature six jets in the $2\ell\mathrm{OS}$ decay channel at tree-level which are correlated in their relative positions. Most jets in Z+jets events, however, originate from QCD radiation and therefore feature a correlation different to that of the jets in signal events. In order to exploit these differences, the discriminating variable was defined as:

$$\Delta R_{jj}^{\mathrm{avg}} = \frac{1}{N_{\mathrm{jets}}} \sum_{n \neq m}^{N_{\mathrm{jets}}} \Delta R(j_n, j_m) \tag{7.6}$$

and was seen to provide some discrimination power. However, a thorough check whether the differences in the shapes of this variable observed for signal and background arise due to the above discussed differences between signal and background has to be performed.

Invariant mass of the jet pair with the highest transverse momentum $\boxed{m_{jj}^{\max p_{\mathrm{T}}}}$

The two b-jets are expected to originate from the decay of the two top quarks for signal events. Thus, due to the big mass difference between the top quark and its decay products, the b-jets feature a high transverse momentum which often exceeds that of the other jets within a given event. For the Z+jets process however, as b-jets are typically produced via gluon splitting, such jets feature a considerably smaller invariant mass than those from a top quark decay.

Number of combinations of one b-jet and two non-tagged jets with an invariant mass compatible with the top quark mass $\boxed{N_{bjj}^{\text{top-mass}}}$

Within the $2\ell\mathrm{OS}$ decay channel of the $t\bar{t}Z$ process, six jets, of which two are b-tagged, are expected to be produced by the top quark decay. As $t\bar{t}Z$ events feature on average a higher jet multiplicity than Z+jets events, a larger number of jet combinations with an invariant mass compatible within $15\,\mathrm{GeV}$ with that of the top quark is expected for signal events. In order for a given combination to be included in the sum, the invariant mass of the two jets in the light-quark positions are required to have an invariant mass within $15\,\mathrm{GeV}$ from that of the W boson, namely $|m_{jj} - m_W| < 15\,\mathrm{GeV}$.

7.2.2.3. Discrimination against both the $t\bar{t}$ and Z+jets backgrounds

In the following, the set of training variables which offer discrimination against both dominant backgrounds are discussed.

Transverse momentum of the dilepton system $\boxed{p_{\mathrm{T}}^{\ell\ell}}$

In signal events, the dilepton system is assumed to originate from the decay of an on-shell Z boson and is suspected to feature on average a higher transverse momentum than for the $t\bar{t}$ background where both leptons originate from W boson decays of dileptonic $t\bar{t}$. Z+jets events do feature an on-shell Z boson, but the initial-state objects are not necessarily high enough in energy as one needs to create a $t\bar{t}Z$ system. In addition, if the Z boson is either radiated from one of the top quarks or recoils against the $t\bar{t}$ system, it receives a considerable boost which increases the transverse momentum of its decay products.

Transverse momentum of the leading b-jet and two closest jets $\boxed{p_{\mathrm{T}}^{b_1 jj}}$

The $t\bar{t}Z$ system decaying in the $2\ell\mathrm{OS}$ channel features six jets, two of which are expected to be b-tagged, in the final state. Due to the large mass difference between the top quark and its decay products, the b-jets in $t\bar{t}Z$ events are expected on average to feature rather high values of transverse momentum which motivates why for the construction of the $p_{\mathrm{T}}^{b_1 jj}$ variable the jet with the highest transverse momentum within an event was assumed to be the b-jet instead of taking the jet with highest b-tagging score. Dileptonically decaying $t\bar{t}$ events feature two b-jets from the top quark decay in the final state as well, but since both W boson decay leptonically, the non-tagged jets are deemed to originate from radiated gluons or pile-up activity and carry on average not as much transverse momentum as the non-tagged jets of the $t\bar{t}Z$ signature. In Z+jets events, most

jets originate from QCD radiation which is the reason why the discrimination is expected to be even larger than for dileptonic $t\bar{t}$.

Number of jet pairs with an invariant mass compatible with a W boson $N_{jj}^{W\text{-mass}}$

In an ideal case, both W bosons decay into a pair of quarks in $t\bar{t}Z$ events which gives rise to two pairs of non-tagged jets with an invariant mass value compatible with that of a W boson. Any additional jet initiated by higher-order interactions increases the combinatorial freedom to find a higher number of jet pairs that are compatible with a W boson. As $t\bar{t}Z$ events feature on average a higher jet multiplicity than events of dileptonic $t\bar{t}$ or Z+jets production, a larger number of jet combinations with an invariant mass compatible with that of W boson is expected for signal events. In order to account for the limited jet energy resolution, a jet pair is considered to be compatible if its invariant mass has a value within the range of $70\,\mathrm{GeV}$ to $100\,\mathrm{GeV}$.

Invariant mass of the system of the leading b-jet and two non-tagged jets closest to the top quark mass $m_{bjj}^{\text{top-mass}}$

Signal events feature a fully-hadronic $t\bar{t}$ decay and, additionally, an on-shell Z boson, which is either radiated from one of the two top quarks or recoils against the $t\bar{t}$ system. In many cases one can expect at least one of the top quarks to receive a considerable boost such that its three final-state objects, one b-jet and two non-tagged jets, are likely to be very close to each other. Since neither background process features this signature and $t\bar{t}Z$ events feature on average a higher jet multiplicity, the probability for signal event having three jets with an invariant mass close to that of the top quark is larger than for background events.

Scalar sum of the transverse momenta of all leptons and b-jets $H_{\mathrm{T}}^{bb\ell\ell}$

Both the leptons as well as the b-jets in the final state of $t\bar{t}Z$ have on average higher transverse momenta than the corresponding objects within Z+jets events. The Z boson is either radiated from one of the two top quarks or recoils against the $t\bar{t}$ system receiving a considerable boost which is entirely transferred to its decay products. Due to the mass difference of the order of the Z boson mass between the top quark and its decay products, the b-quarks are expected to receive a considerable amount of momentum. For Z+jets, however, b-jets are suspected to be induced by gluon radiation, and the Z boson is not related to the heavy $t\bar{t}$ system. The b-jets considered to build this variable were required to have the highest b-tagging output weight of all jets in an event.

Multi-hypothesis reconstruction output weights w_{1W}, w_{1t}, w_{2W}, w_{1t1W}, w_{2t}

At tree-level, six jets from the fully-hadronic decay of the $t\bar{t}$ system are part of the $2\ell\mathrm{OS}$ $t\bar{t}Z$ signature. In contrast, the dileptonic $t\bar{t}$ background features only two b-jets from top quark decay, and the jets within Z+jets events predominantly originate from QCD radiation. Therefore, the reconstruction of the fully-hadronic $t\bar{t}$ signature of the signal process provides significant discrimination power. Due to the finite jet energy resolution and the limited coverage of the detector, for a considerable number of signal events at least one jet from the fully-hadronic $t\bar{t}$ decay can either not be associated to a quark of the $t\bar{t}$ final state or not reconstructed at all. In order to perform a kinematic reconstruction of the $t\bar{t}$ system for such events, several hypotheses of the number of missing jets were taken into account for a given event. Five different scenarios were considered with the reconstruction procedure:

- The full $t\bar{t}$ system can be reconstructed (labelled $2t$)
- One of the two b-jet is missing, thus one top quark and the hadronically decaying W boson of the other top quark can be reconstructed (labelled $1t1W$)
- Both b-jets are missing, thus only the two hadronically decaying W boson can be reconstructed (labelled $2W$)
- Only one of the two top quark and not the second hadronically decaying W boson can be reconstructed (labelled $1t$)
- Only one of the two hadronically decaying W bosons and none of the top quarks can be reconstructed (labelled $1W$)

For each of the five categories multiple assignments of the reconstructed jets to the positions of the final-state quarks are possible. All five scenarios are probed independently from each other such that five distinct outputs, one for each hypotheses, are evaluated. Similarly to the two-neutrino scanning method, for each possibility of jet-quark assignment the resulting mass values of the involved W bosons and top quarks were interpolated with the respective p.d.f. built with parton-level information. For all categories except for $1W$, the individual interpolated values of the involved W bosons and top quarks were combined such that a single output weight is obtained: w_{2t}, w_{1t1W}, w_{2W}, w_{1t} and w_{1W}, with the superscript denoting the category the weight belongs to. From all possible assignments of the reconstructed jets to the positions of the $t\bar{t}$ final-state objects, only the assignment yielding the highest value of the final output weight is considered correct for a given hypothesis. Dedicated vetoes, optimised separately for each hypothesis, were applied on the b-tagging requirements to take advantage of the inherent correlations and overlap between each of the hypotheses.

The multi-hypothesis reconstruction method as presented here has been used for the first time in an analysis and the various output weights were seen to be assigned a high rank from the BDT. Further details on the hypotheses-based top quark reconstruction are given in App. E.

7.2.3. Training and application of Boosted Decision Trees

After the concepts of multivariate analysis techniques in general and of Boosted Decision Trees were discussed in the previous section, this section will be dedicated to the MVA-related analysis parts of the studies in the $t\bar{t}Z$ 2ℓOS channel. As mentioned at the end of Sec. 7.1.1, a two-dimensional multivariate approach was employed in order to separate the signal from the two dominant background processes $t\bar{t}$ and Z+jets. First, the configuration settings of the BDT setup will be given before the training results are discussed in a second step. Afterwards, the trained classifiers are applied to the data and simulated events in order to define signal- and background-enriched phase-space regions, respectively.

7.2.3.1. Training of the BDT classifiers

For each of the targeted phase-space regions defined in Sec. 7.1.1 two separate Boosted Decision Trees were trained each of which focusing on one of the two dominant backgrounds. Either classifier was given a slightly different set of discriminating variables providing the highest separation power with respect to the chosen target region and background process. The lists of variables used for each BDT were given in Sec. 7.2.2 within Tabs. 7.9 and 7.10. However, the configuration settings for the training procedure were the same for all classifiers.

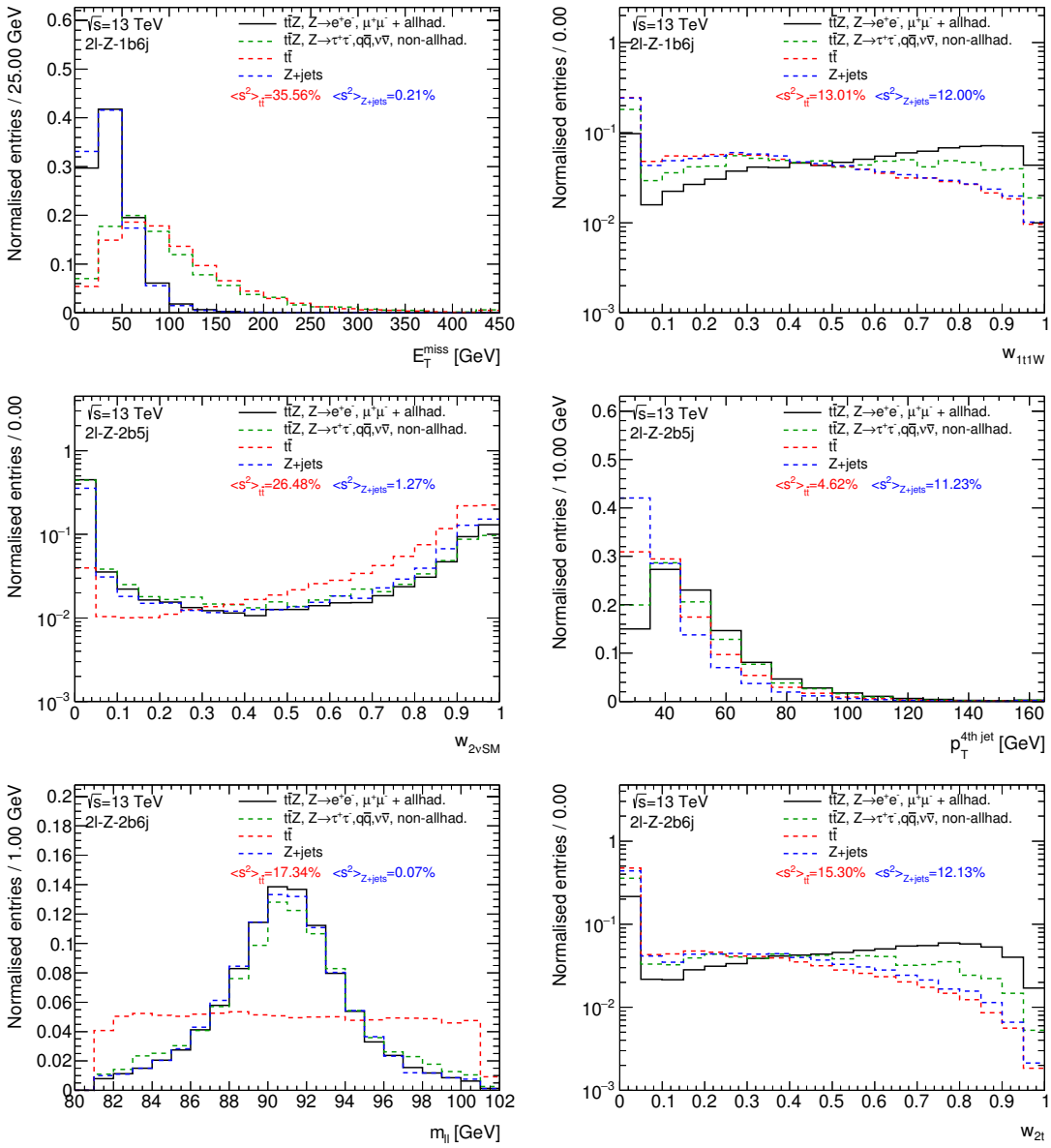


Fig. 7.7.: Distributions of various discriminating variables each normalised to the respective number of entries, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

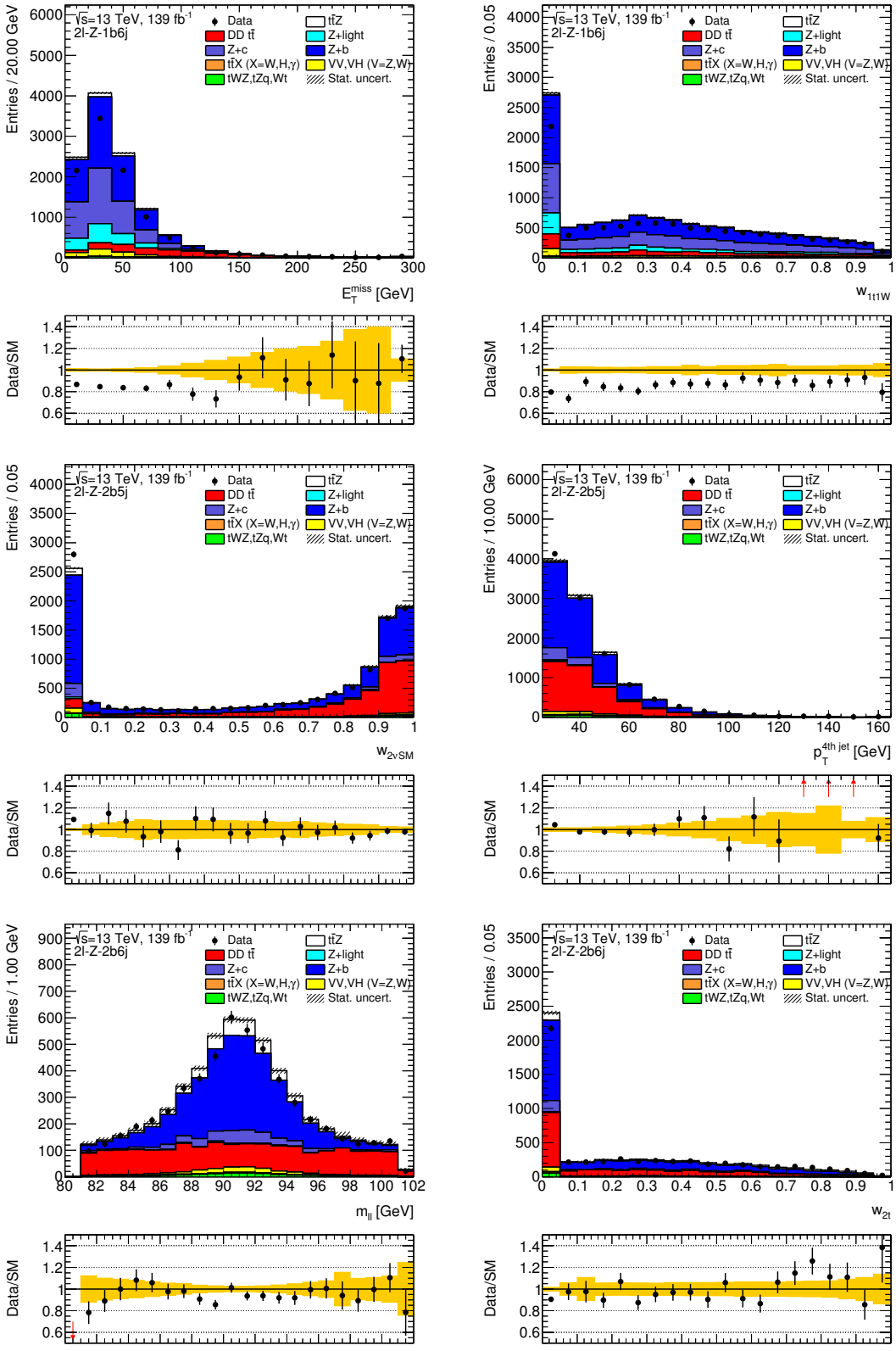


Fig. 7.8.: Distributions of the discriminating variables of which the normalised shapes are shown in Fig. 7.7. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

Rank	2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
1	$E_{\mathrm{T}}^{\mathrm{miss}}$	$w_{2\nu\mathrm{SM}}$	$m_{\ell\ell}$
2	$m_{\ell\ell}$	$m_{\ell\ell}$	$w_{2\nu\mathrm{SM}}$
3	$w_{2\nu\mathrm{SM}}$	$\Delta R(\ell,\ell)$	$E_{\mathrm{T}}^{\mathrm{miss}}$
4	$\Delta R(\ell,\ell)$	$E_{\mathrm{T}}^{\mathrm{miss}}$	w_{1t}
5	w_{1t}	w_{1W}	$\Delta R(\ell,\ell)$
6	w_{1t1W}	w_{1t}	w_{2t}
7	w_{2W}	w_{2W}	w_{1W}
8	w_{1W}	$\Delta R(bb,\ell\ell)$	w_{1t1W}
9	w_{2t}	$m_{b\ell}^{\mathrm{min}}$	$m_{b\ell}^{\mathrm{min}}$
10	$N_{jj}^{W\text{-mass}}$	w_{1t1W}	w_{2W}
11	$\max m_{b\ell}^{\mathrm{min}\,\Delta R}$	$H_{\mathrm{T}}^{bb\ell\ell}$	$\Delta R(bb,\ell\ell)$
12	$H_{\mathrm{T}}^{bb\ell\ell}$	$p_{\mathrm{T}}^{\ell\ell}$	$N_{jj}^{W\text{-mass}}$
13	$p_{\mathrm{T}}^{\text{2nd lepton}}$	$\max m_{b\ell}^{\mathrm{min}\,\Delta R}$	$p_{\mathrm{T}}^{\ell\ell}$
14	$m_{b\ell}^{\mathrm{min}}$	$p_{\mathrm{T}}^{\text{1st lepton}}$	$m_{b\ell}^{\text{max-min}}$
15	$p_{\mathrm{T}}^{\ell\ell}$	$m_{bjj}^{\text{top-mass}}$	$\max m_{b\ell}^{\mathrm{min}\,\Delta R}$
16	$p_{\mathrm{T}}^{\text{1st lepton}}$	$m_{b\ell}^{\text{max-min}}$	$H_{\mathrm{T}}^{bb\ell\ell}$
17	$m_{bjj}^{\text{top-mass}}$	$p_{\mathrm{T}}^{\text{2nd lepton}}$	$p_{\mathrm{T}}^{\text{1st lepton}}$
18	$m_{b\ell}^{\text{max-min}}$	–	$p_{\mathrm{T}}^{\text{2nd lepton}}$
19	–	–	$m_{bjj}^{\text{top-mass}}$

Tab. 7.9.: Ranking of input variables used to train the BDT against the $t\bar{t}$ background as a result of the training. The rankings are based on the how often a specific variable was used for setting a cut when building a decision tree during the training [248].

In order to prevent overtraining, the number of grown trees in a forest was restricted to 500. The depth of a single tree was limited to three consecutive nodes which means that not more than three split decisions were allowed to be made until reaching a final node. Furthermore, the minimum fraction of events within a final nodes with respect to the entire training dataset had to be equal to or greater than 5 %. The best suited variable to cut on in order to make a split decision was determined with the so-called Gini index separation method. An optimal cut value was determined by scanning over the variable range with a granularity of 20 bins. For the boosting the Gradient Boost algorithm was employed featuring a shrinkage of 10 %. In addition, only a subset of half of the training events was given to the boosting algorithm, referred to as stochastic boosting. Finally, a node is considered to be signal-like if it features a signal purity of at least 50 %.

The training dataset consists of simulated events only. As signal only $t\bar{t}Z$ events with the Z boson decaying into an electron-positron or muon-antimuon pair were employed and, additionally, the $t\bar{t}$ system was required to feature a true fully-hadronic decay, employing parton-level information. Depending on the background process that was trained against, either dileptonically decaying $t\bar{t}$ events or Z+jets events with a leptonically decaying Z boson were added. The entire dataset for a specific training was split randomly into two subset of equal size and only one of which was used for the actual training of the classifier. The other subset of events, however, is used for an in-situ test of the training results in order to check if the classifier was trained on too

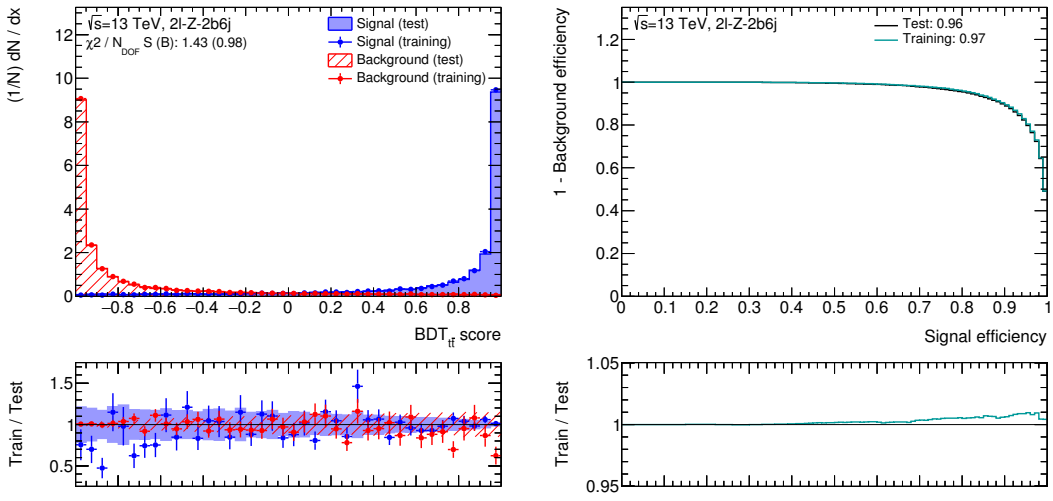
Rank	2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
1	w_{1t}	w_{1t}	$\Delta R_{jj}^{\text{avg}}$
2	w_{2W}	$\Delta R_{jj}^{\text{avg}}$	w_{1t}
3	$\lvert\eta_{\ell\ell}\rvert$	w_{2W}	$\lvert\eta_{\ell\ell}\rvert$
4	w_{1t1W}	w_{1W}	m_{bb}
5	w_{1W}	m_{bb}	$m_{jj}^{\max p_{\text{T}}}$
6	$\Delta R_{jj}^{\text{avg}}$	$p_{\text{T}}^{\ell\ell}$	w_{1t1W}
7	$N_{jj}^{W\text{-mass}}$	$m_{jj}^{\max p_{\text{T}}}$	$m_{bjj}^{\text{top-mass}}$
8	$p_{\text{T}}^{\text{1st } b\text{-jet}}$	$m_{jj}^{\min \Delta R}$	w_{2t}
9	w_{2t}	H_{T}	w_{2W}
10	$p_{\text{T}}^{\text{5th jet}}$	w_{1t1W}	w_{1W}
11	$H_{\text{T}}^{bb\ell\ell}$	$p_{\text{T}}^{\text{1st } b\text{-jet}}$	$H_{\text{T}}^{bb\ell\ell}$
12	$m_{bjj}^{\text{top-mass}}$	$p_{\text{T}}^{\text{3rd jet}}$	$p_{\text{T}}^{\text{1st } b\text{-jet}}$
13	$p_{\text{T}}^{\text{4th jet}}$	$m_{bjj}^{\text{top-mass}}$	$p_{\text{T}}^{\text{5th jet}}$
14	$N_{bjj}^{\text{top-mass}}$	$p_{\text{T}}^{\text{4th jet}}$	$N_{jj}^{W\text{-mass}}$
15	–	$p_{\text{T}}^{\text{5th jet}}$	$p_{\text{T}}^{\text{6th jet}}$
16	–	$p_{\text{T}}^{\text{2nd jet}}$	$p_{\text{T}}^{\text{4th jet}}$
17	–	$p_{\text{T}}^{b_1 jj}$	$N_{bjj}^{\text{top-mass}}$
18	–	$H_{\text{T}}^{bb\ell\ell}$	–

Tab. 7.10.: Ranking of input variables used to train the BDT against the Z+jets background as a result of the training. The rankings are based on the how often a specific variable was used for setting a cut when building a decision tree during the training [248].

many details of the input model.

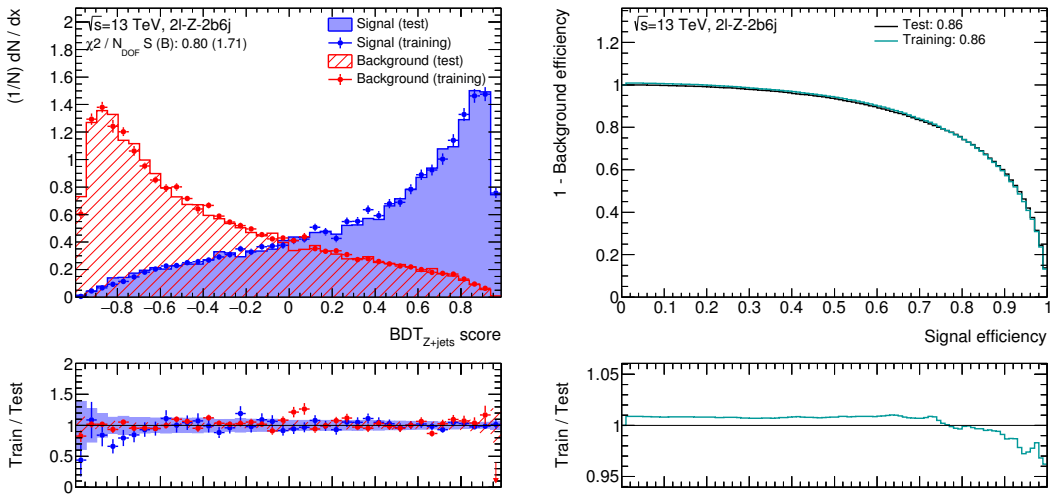
The results of the various training procedures against the $t\bar{t}$ background are depicted in Fig. 7.9 as well as in Figs. F.18 and F.19, and the training results for the discrimination of the Z+jets background are given in Fig. 7.10 as well as in Figs.F.20 and F.21, respectively. All figures show both the normalised output score of the respective classifier on the left side, whereas in the right side the so-called Receiver Operating Characteristic (ROC) is presented. The coloured areas in the left subfigures depict the BDT responses of the test datasets, shown in blue for signal and in red for background events. They are overlayed with the dotted distributions of the corresponding training datasets. Below the plots, the panels shows the ratio of training and test results for the signal and background datasets, respectively. The ROC curves were drawn for the training in dark cyan blue and for the test procedure in black, showing the true positive classification rate on the x-axis versus the false positive selection rate on the y-axis. For a non-discriminating classifier the result would be a straight line from the upper left to the lower right corner. In contrast, a BDT featuring perfect discrimination has a rectangular ROC curve since one would correctly select signal and reject background events all the time. For a given classifier, the higher the so-called ROC value of a classifier given by the integral of the area beneath the ROC curve, the better the separation it provides. The ratio of the training and test curves is shown in the panels below the ROC curves.

For all trained classifiers, the separation between signal and background was found to be very good. The training and test results agree with each other within the uncertainties indicated by



(a) (b)

Fig. 7.9.: Training results of the training against the $t\bar{t}$ background within the 2ℓ-Z-2b6j target region. In Fig. (a) the normalised distributions of the BDT score are shown in blue for the signal and in red for the $t\bar{t}$ background. The test datasets are depicted as colour-shaded areas and are overlayed with the dotted training distributions. On the right, Fig. (b) shows the corresponding ROC curve for both training and test datasets.



(a) (b)

Fig. 7.10.: Training results of the training against the Z+jets background within the 2ℓ-Z-2b6j target region. In Fig. (a) the normalised distributions of the BDT score are shown in blue for the signal and in red for the Z+jets background. The test datasets are depicted as colour-shaded areas and are overlayed with the dotted training distributions. On the right, Fig. (b) shows the corresponding ROC curve for both training and test datasets.

the vertical error bars of the dotted distributions and the coloured bands of the training-test ratios. In addition to the statistical uncertainties, the errors are driven by the boost weights determined during the training. In particular the training routines focusing on the $t\bar{t}$ background performed

very strongly, as can be seen from the high peaks shifted towards the edges of the normalised BDT score distributions as well as from the corresponding ROC curves which feature integral values of above 0.95. The training procedures on the Z+jets background did not perform just as well, but still provide quite clear separation of signal and background events. The corresponding ROC values are in the range of 0.83 to 0.87.

A quantitative measure of the agreement of the BDT score distributions of the test and training subsets was performed employing the so-called χ^2-test, which has been originally developed and is widely used to test the goodness of fits but can be easily adopted to other use cases. The basic idea is to construct the quantity χ^2 such that it reflects the level of agreement between two datasets $\boldsymbol{n} = (n_1, \ldots, n_N)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)$, which correspond to the number of entries of observed and expected entries within the bins i $(i = 1, \ldots, N)$ of a given histogram, respectively:

$$\chi^2 = \sum_{i=1}^{N} \frac{(n_i - \nu_i)^2}{\sigma_i^2} \tag{7.7}$$

The expression σ_i in the denominator of Eq. (7.7) can be identified with either the square root of the expectation value or the error on the measurement[(3)] in bin i, depending on the chosen definition. For a Poisson-distributed variable, however, both definitions are equal to the standard deviation, and χ^2 is therefore the sum of squares of the deviations between observed and expected bin entries, measured in units of the corresponding standard deviations. Consequently, larger values of χ^2 correspond to larger discrepancies between $\boldsymbol{n}$ and $\boldsymbol{\nu}$. The relevant quantity to make statements about the goodness of a fit is the so-called χ^2 probability P, which is defined as

$$P(\chi^2; N_{\text{DOF}}) = \int_{\chi^2}^{\infty} f(z; N_{\text{DOF}})\, \mathrm{d}z \quad , \tag{7.8}$$

and depends on both the χ^2 value and the number of degrees of freedom N_{DOF}, which is identical to the number of bins N of a given histogram. The probability distribution $f(\chi^2; N_{\text{DOF}})$ for χ^2 is given by

$$f(\chi^2; N_{\text{DOF}}) = \frac{2^{-N_{\text{DOF}}/2}}{\Gamma(N_{\text{DOF}}/2)} \chi^{N_{\text{DOF}}-2} e^{\chi^2/2} \quad , \tag{7.9}$$

employing the gamma function

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x}\, \mathrm{d}x \quad . \tag{7.10}$$

Since the expectation value of the χ^2 probability distribution is equal to the number of degrees of freedom, the ratio χ^2/N_{DOF} is commonly quoted as a measure of agreement. If the agreement between $\boldsymbol{n}$ and $\boldsymbol{\nu}$ is very good, a contribution of about one is expected from each term in the sum of Eq. (7.7) and hence χ^2 to be roughly equal to N_{DOF}. However, in the case of comparing two weighted histograms[(4)] the respective weight for each bin entry has to be taken into account for

[(3)] If the errors on the measurement are correlated, Eq. (7.7) becomes

$$\chi^2 = (\boldsymbol{n} - \boldsymbol{\nu})^T (V)^{-1} (\boldsymbol{n} - \boldsymbol{\nu})$$

with V being the covariance matrix for the n_i measurements.

[(4)] The events filled into the bins have weights assigned such that the value an event contributes to the total amount is scaled by a certain factor. Such weights typically have various reasons, for example Monte Carlo generator weights, scale factors from calibration, scaling to the luminosity of the investigated dataset and the theoretical cross section of a certain physics process, or boost weights from multivariate classification.

the expression in the numerator as well as the denominator of Eq. (7.7), which is constructed from the contributions of both histograms for a specific bin [139, 245].

All χ^2/N_{DOF} values in Figs. F.18a–7.10a are in the range of one to two, and thus confirm the excellent agreement of training and test distributions for both signal and background events as it was estimated by eye.

Furthermore, the linear correlation between the input variables was examined for training and test datasets separately. As a measure for the level of correlation of two variables x and y, the linear correlation coefficient ρ was calculated for each possible pairing of the input variables according to

$$\rho_{xy} = \frac{V_{xy}}{\sigma_x \sigma_y} \quad , \tag{7.11}$$

where V_{xy} in the numerator denotes the element of the covariance matrix for the variables x and y. In the denominator, σ_x and σ_y are the standard deviation of x and y, respectively [139]. The corresponding illustrations of the linear correlations for the various BDT training procedures are depicted in Figs. F.22–F.33.

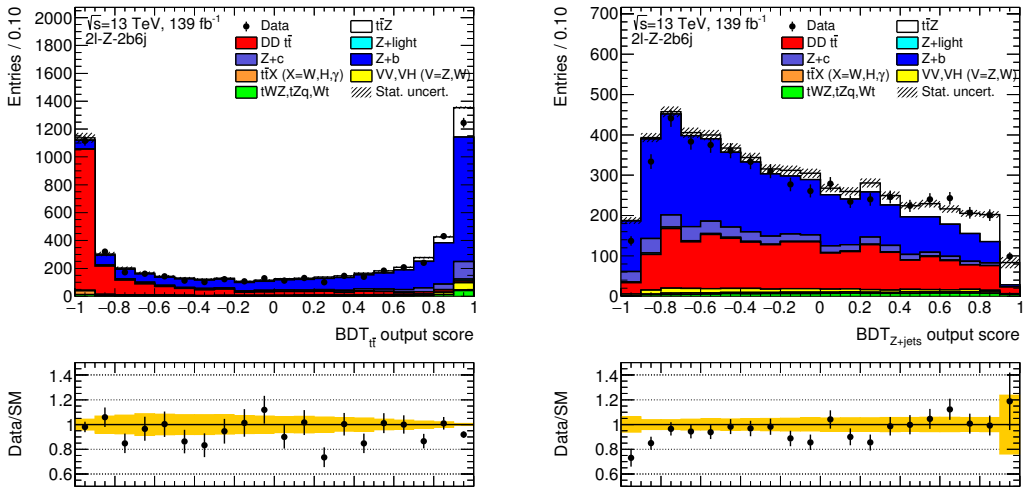
7.2.3.2. Application of the trained classifiers to data and Monte Carlo simulation

The trained Boosted Decision Trees were applied to the entire set of simulated physics processes that was used for the studies presented in this thesis as well as to the data. For those Monte Carlo datasets that have already been used for the BDT training procedures the signal and background composition were treated as unknown which serves as an additional consistency check. The various classifiers were evaluated for each individual event which were thereby classified as either signal- or background-like according to the information from the input variables of a given event. In order to exploit the full discrimination power of the various Boosted Decision Trees, the same set of input variables as those employed for the BDT training procedures within the respective target regions has to be used.

As one BDT was trained separately for each dominant background process and for each targeted phase-space region, a total of six separate BDT configurations were considered. For each event in a given targeted phase-space region, two BDT output scores were evaluated: one for the BDT to reject $t\bar{t}$ and one for the BDT to reject Z+jets, again, specific to the particular targeted phase-space region. The distributions of the BDT scores applied to both the data and simulated events are depicted in Figs. 7.11 within the 2ℓ-Z-2b6j target region. The left panel shows the output scores for the discrimination of the $t\bar{t}$ background, whereas the figure on the right presents the output scores for the BDT trained against Z+jets. The various simulated background processes and the signal were scaled to their respective theoretical cross section and normalised to an integrated luminosity of $139\,\text{fb}^{-1}$. Both MC generator and the various calibration weights were applied. The signal events represented in white are clearly shifted towards high positive values of the BDT score, whereas the events of the background processes against which the respective BDT was trained are accumulated towards the left. The black data points are in general in quite good agreement with the total Standard Model prediction corresponding to the sum of yields based on the various coloured contributions in the stacked histogram. The corresponding distributions for the 2ℓ-Z-1b6j and 2ℓ-Z-2b5j region are shown in Figs. F.64 and F.65, respectively.

7.2.3.3. Definition of 2D-regions

The measurement of a certain physics process and of its cross section is best performed within a phase-space region with a rather high purity of contributions from this process. These regions are typically defined by selection criteria rejecting as many background events as possible while



(a) (b)

Fig. 7.11.: Distributions of the BDT output scores within the 2ℓ-Z-2b6j target region. Fig. (a) on the left shows the score of the classifier trained against $t\bar{t}$, and Fig. (b) on the right presents the BDT for the discrimination against the Z+jets background. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

retaining a high signal efficiency. For the studies in the 2ℓOS channel presented in this thesis, however, an MVA-based method was preferred due to the very limited statistics, since a cut-based analysis would only have allowed for a substantial rejection of background events at the cost of a large loss in signal statistics Two Boosted Decision Trees were trained per target region, each focusing on one of two dominant backgrounds, in order to simultaneously isolate the signal events from the two dominant backgrounds. For each target region the two BDT output scores were combined such that a signal-enriched phase-space region could be defined by setting boundaries in the two-dimensional plane spanned by the two BDT output scores. The three resulting regions are labelled 2ℓ-Z-1b6j-$t\bar{t}Z$, 2ℓ-Z-2b5j-$t\bar{t}Z$ and 2ℓ-Z-2b6j-$t\bar{t}Z$ with the suffix "$t\bar{t}Z$" denoting the high signal purity. The decisions on the boundaries of the respective regions in the plane spanned by the BDT outputs were made based on the purity of the signal as well as of the two dominant background processes in the respective bins. In the same way, regions enriched in $t\bar{t}$ background and another region predominantly populated with Z+jets events were defined, denoted with suffix "$t\bar{t}$" and "Z+jets", respectively. The corresponding illustrations of the purity of the respective processes within the 2ℓ-Z-2b6j target region are presented in Fig. 7.12, where the panels (a), (b) and (c) depict the relative fraction of $t\bar{t}Z$, $t\bar{t}$ and Z+jets, respectively. For a given target region, the signal- and background-enriched regions, also referred to as 2D-regions, were defined such that they do not overlap with each other. The parts of the plane of the BDT outputs which were not seen to be rather pure in either process are labelled as so-called mixed regions.As neither the signal nor any of the dominant backgrounds feature a clear majority in those areas, they are not considered at any further stage of the analysis. The selection criteria for the 2D-regions are indicated by the coloured areas within Fig. 7.12d for the 2ℓ-Z-2b6j target region whereas the phase-space region featuring a mix in signal and dominant background processes is highlighted with the grey cross-

hatched fill style. The definition of the 2D-regions as well the purity plots for the 2ℓ-Z-1b6j and 2ℓ-Z-2b5j target regions are illustrated in Figs. F.66 and F.67, respectively. As the distribution of signal and background events in the BDT plane is different in each of the target regions, the boundaries of the 2D-regions enriched in a particular physics process are not the same for all target regions. The various regions enriched in a specific physics process are located as they were expected to be within the phase space of the BDT output scores. However, a small amount of signal events were classified as background-like by either BDT and vice versa, which is the reason why the areas featuring a high purity of a certain process are not as well separated from each other as indicated in Fig. 7.2.

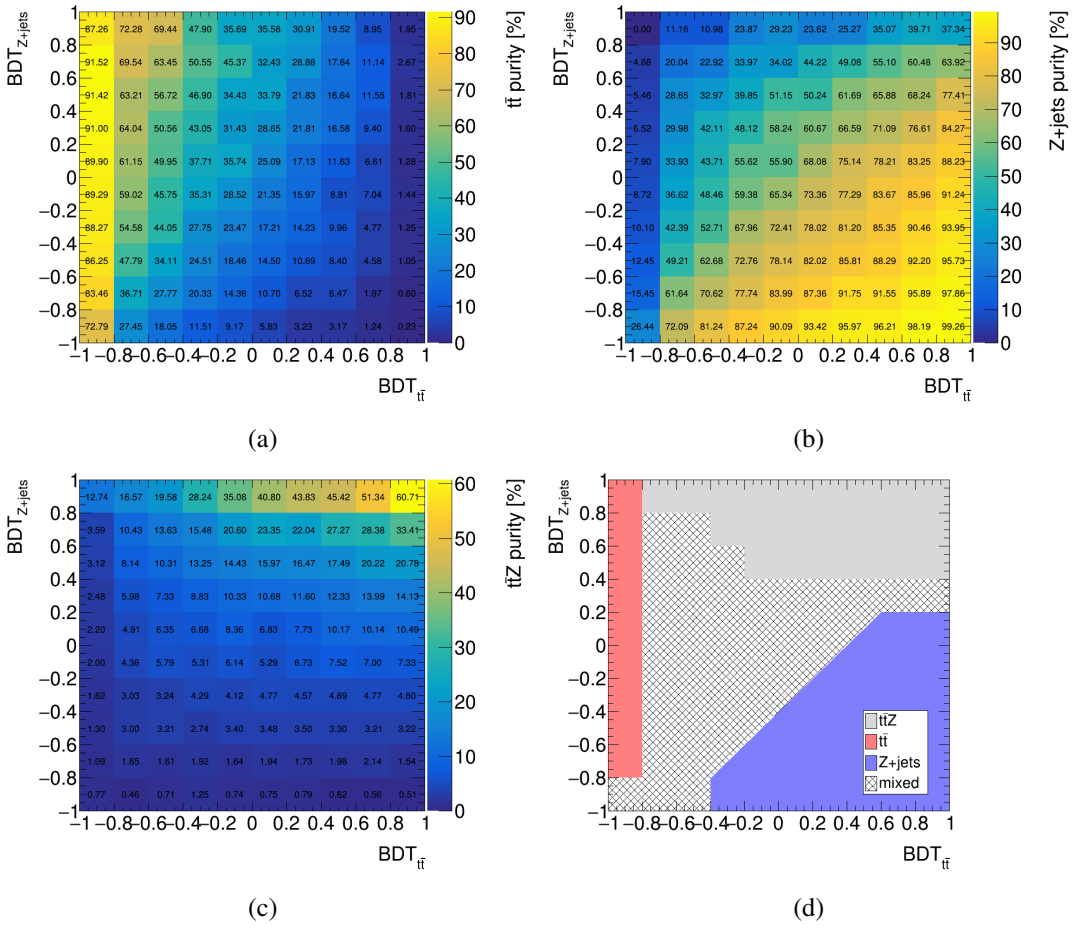


Fig. 7.12.: Depiction of the purities of the various processes and boundary lines for the two-dimensional space constructed from the separate BDT outputs for the 2ℓ-Z-2b6j target region. The Figs. (a), (b) and (c) depict the purity of the $t\bar{t}$, Z+jets and $t\bar{t}Z$ events within the phase space spanned by the two BDT output scores. In the bottom right corner, however, Fig. (d) shows the various phase-space regions enriched in the respective physics processes which were determined based on Figs. (a)–(c).

The expected event yields of the three $t\bar{t}Z$-enriched phase-space regions are listed in Tab. 7.11. The corresponding numbers for the 2D-regions featuring a high purity of either $t\bar{t}$ or Z+jets events can be found in Tabs. G.3 and G.4, respectively. By employing the discrimination power of both Boosted Decision Trees for the definition of the 2D-regions, the signal purity could be significantly increased, reaching a value of about 30 % in case of the 2ℓ-Z-2b6j-$t\bar{t}Z$ region. This is equivalent

to an increase by a factor of five with respect to the most sensitive target region presented in Sec. 7.1.1. Similarly, the background-enriched 2D-regions feature a high purity in $t\bar{t}$ and Z+jets events, respectively. The Z+jets-enriched 2D-regions derived from the target regions featuring at least two b-jets are highly dominated by the Z+b fraction, whereas a considerable amount of Z+c and a certain amount of Z+light flavour events remain within the 2ℓ-Z-1b6j-Z+jets region. The amount of predicted events slightly exceeds the data in the 2ℓ-Z-1b6j-$t\bar{t}Z$ and 2ℓ-Z-2b5j-$t\bar{t}Z$ regions, whereas the numbers of predicted and observed events agree quite well.

Process	2ℓ-Z-1b6j-$t\bar{t}Z$	2ℓ-Z-2b5j-$t\bar{t}Z$	2ℓ-Z-2b6j-$t\bar{t}Z$
$t\bar{t}Z$	100.96 ± 0.85	96.94 ± 0.69	247.7 ± 1.3
DD $t\bar{t}$	42.6 ± 6.7	70 ± 87	114 ± 11
Z+light flavour	48.9 ± 2.4	4.1 ± 1.8	3.54 ± 0.69
Z+c	196.6 ± 4.1	27.9 ± 2.0	48.6 ± 2.1
Z+b	282.4 ± 2.6	262.5 ± 3.0	364.4 ± 3.0
$t\bar{t}X$ $(X = W, H, \gamma)$	2.00 ± 0.17	4.53 ± 0.29	7.48 ± 0.33
VV, VH $(V = Z, W)$	44.9 ± 1.4	18.6 ± 1.2	25.05 ± 0.96
Other (tWZ, tZq, Wt)	17.37 ± 0.64	19.11 ± 0.65	37.00 ± 0.93
Total SM	735.6 ± 8.8	503.8 ± 9.7	848 ± 12
Data	663	457	861
Data/SM	0.90	0.91	1.02
$t\bar{t}Z$ purity	14 %	19 %	29 %

Tab. 7.11.: Estimated and measured yields from simulation as well as those from measured data of the $t\bar{t}Z$-enriched 2D-regions of the 2ℓOS analysis. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. Only the Monte Carlo statistical errors are shown.

In order to check the modelling within the 2D-regions, several kinematic variables were investigated. For example, Fig. 7.13 shows the distributions of the invariant mass of the lepton pair, $m_{\ell\ell}$, and the transverse momentum of the leading b-jet, $p_{\mathrm{T}}^{\text{1st } b\text{-jet}}$, within the $t\bar{t}Z$-enriched 2D-regions. All simulated processes were normalised to their respective cross section and scaled to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. Both the generator and calibration weights were applied, and the overflow bin was added to the various processes. The black data points agree reasonably well with the distributions of simulated events within the statistical uncertainties indicated by the striped and yellow error bands. Due to the reduced statistics in the 2D-regions with respect to the target regions, the fluctuation of the data points around the expected number of simulated events per bin is more pronounced, as expected. The disagreement of the data with the Monte Carlo prediction in the 2ℓ-Z-1b6j-$t\bar{t}Z$ and 2ℓ-Z-2b5j-$t\bar{t}Z$ regions is recognisable from the corresponding distributions and it appears not to have a significant slope. Further checks of the modelling for signal-enriched as well as background-enriched phase-space regions can be found in Figs. F.68–F.72.

7.3. Statistical data analysis

At the LHC, the quantitative interpretation of results of searches and measurements is performed with the so-called frequentist approach which interprets probability as the relative frequency of a

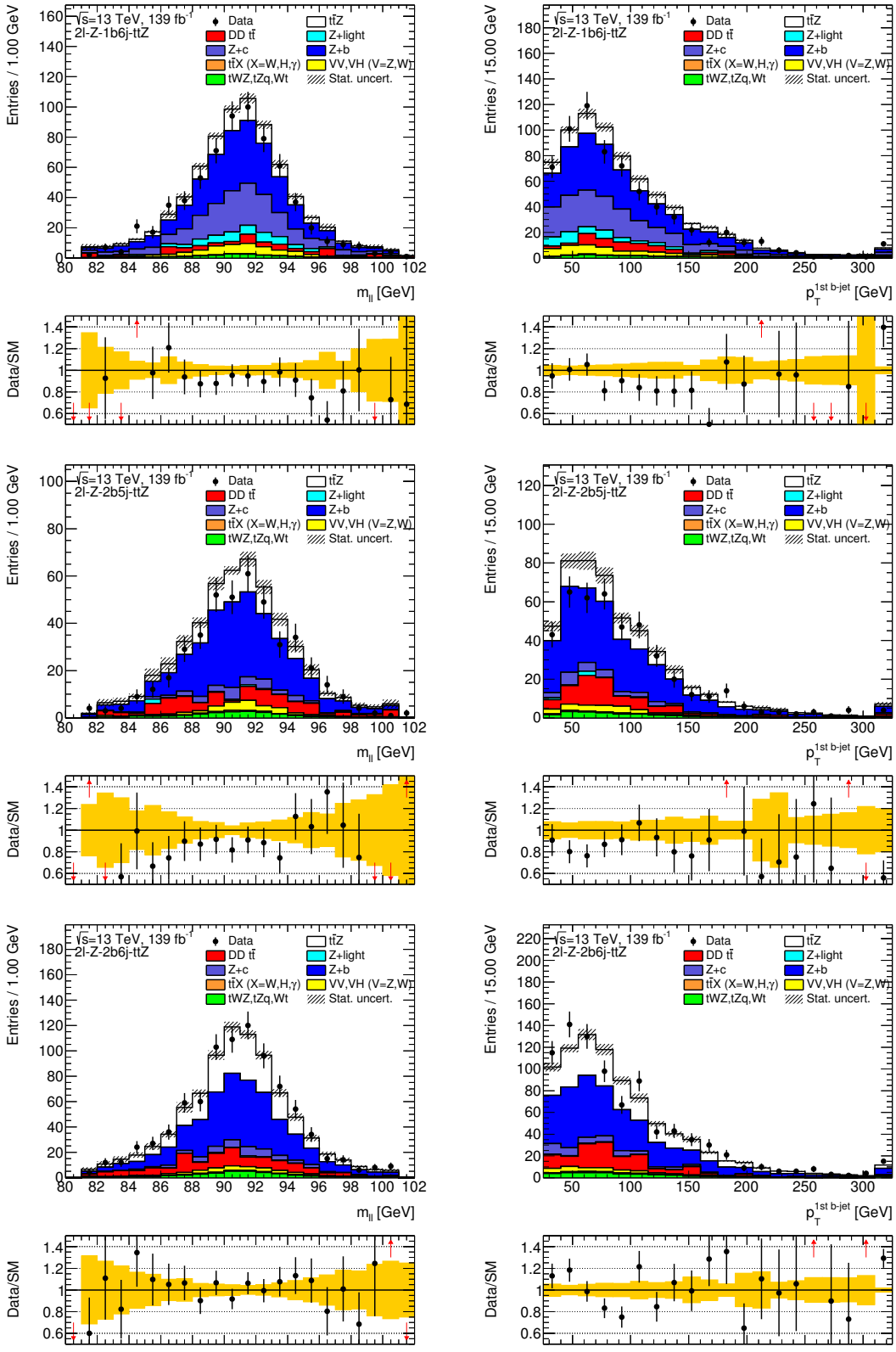


Fig. 7.13.: Distributions of the invariant mass of the lepton pair in the left column and of the first b-jet p_{T} in the right column, shown for the three $t\bar{t}Z$-enriched 2D-regions. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

possible outcome of an experiment. The claim as to whether a physics process was observed in experiment or, otherwise, its observation could be rejected relies on various statistical concepts such as hypothesis tests, significance and the estimation of parameters of the probability density function describing the observed data.

This section is dedicated to the statistical interpretation of the observed data and the relevant methods that were used for the $t\bar{t}Z$ cross-section measurement in the 2ℓOS channel.

7.3.1. Statistical tools

In the following, the statistical concepts and tools that were used for the measurement in this thesis are presented. The technical implementation was accomplished using the TRExFitter framework [253, 254]. It employs the HistFactory package [255] to create statistical models which are then passed to the statistical tools provided by the RooStats package incorporated in the ROOT software.

7.3.1.1. Parameter estimation and the maximum likelihood method

One of the main goals of statistics is to infer the properties of a probability distribution function $f(\boldsymbol{x})$ describing a random variable x, based on a set of $n \in \mathbb{N}$ independent observations $\boldsymbol{x} = x_1, \ldots, x_n$ which is commonly referred to as a sample. Usually the measurements are assumed to be independent from each other, thus the sample p.d.f. $f(\boldsymbol{x})$ can be factorised into the probability density functions of the individual measurements:

$$f(\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i) \tag{7.12}$$

In many cases, however, the corresponding probability density function is not known. As briefly discussed in Sec. 6.3.2, a set of functions $f(x;\theta)$ can be constructed which depend on an unknown parameter θ, or in the more general cases a set $m \in \mathbb{N}$ of unknown parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$, in order to estimate the functional behaviour of the observations x_i and thereby the underlying true p.d.f. $f(x_i)$. Such a function, if all parameters θ_j are known, is called a statistic and, if it is particularly used to estimate some properties of the underlying probability density function, it is denoted an estimator. In order to distinguish between the true value θ and the estimator for this quantity, the latter is typically written with a hat, $\hat{\theta}$. The approach of estimating the mean value of an estimator, referred to as estimate, is commonly denoted as parameter fitting, given a set of measurements $\boldsymbol{x}$ [139].

A common technique to estimate the values of the parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ of the p.d.f. $f(x;\theta)$ is the method of maximum likelihood (ML). The probability for the first of the n measurements to be within the interval $[x_1, x_1 + \mathrm{d}x]$ is given by $f(x_1;\theta)\,\mathrm{d}x$ under the hypothesis of $f(x;\theta)$. Since all individual measurements are typically assumed to be independent, the probability that x_i can be found within $[x_i, x_i + \mathrm{d}x]$ is given by

$$\prod_{i=1}^{n} f(x_i;\theta)\,\mathrm{d}x \tag{7.13}$$

for all i. A high probability for the measured data is expected if the hypothesised p.d.f. and parameter values are correct, and vice versa. The probability for the entire datasets $\boldsymbol{x}$ is then given

by the joint p.d.f. of the individual measurements,

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta), \tag{7.14}$$

referred to as likelihood L which is a function of the parameters and not of the data. The intervals $\mathrm{d}x_i$ typically do not depend on the parameters which is the reason why the likelihood function does not depend on them either. The ML estimators for the parameters $\boldsymbol{\theta}$ are defined to maximise the likelihood function. The requirement for an individual estimator is therefore

$$\left[\frac{\partial L}{\partial \theta_j}\right]_{\theta=\hat{\theta}} = 0, \quad (j = 1, \ldots, m) \tag{7.15}$$

as long as the likelihood function is differentiable in all θ_j and its maximum with respect to any parameter is not at the boundary of the parameter range. Additionally, given a function $a(\theta)$ of some parameter, Eq. (7.15) becomes

$$\left[\frac{\partial L}{\partial \theta_j}\right]_{\theta=\hat{\theta}} \left[\frac{\partial L}{\partial a}\frac{\partial a}{\partial \theta_j}\right]_{\theta=\hat{\theta}} = 0, \quad (j = 1, \ldots, m) \tag{7.16}$$

unless $\partial a/\partial\theta = 0$. The ML estimator is thus obtained by evaluating the function a with the original estimator:

$$\hat{a} = a(\hat{\theta}) \tag{7.17}$$

Often, rather than using the likelihood function as defined in Eq. (7.14) it is more convenient to use its logarithm,

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i; \theta) \quad , \tag{7.18}$$

which has the advantage that the product of within L is converted into a sum, and the parameter values maximising L will also maximise $\ln L$ since the logarithm is a monotonically increasing function [139, 245].

In high energy physics, the measured values of a random variable x, for example the transverse momentum of some physics object, are typically represented as a histogram yielding a certain number of entries $\boldsymbol{n} = (n_1, \ldots, n_N)$ within N bins. This becomes particularly beneficial when dealing with large data samples since the number of summands in Eq. (7.18) reduces from the number of measured events to the number of bins. The expectation values $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)$ of the number of entries can be written as

$$E[n_i] = \nu_i = \mu s_i + b_i \quad , \tag{7.19}$$

where the number of entries of signal and background, or, more precisely, the probability for an event to be found in the ith bin are given by

$$s_i = s_{\text{total}} \int_{\text{bin } i} f_s(x; \boldsymbol{\theta}_s)\, \mathrm{d}x \quad , \tag{7.20a}$$

$$b_i = b_{\text{total}} \int_{\text{bin } i} f_b(x; \boldsymbol{\theta}_b)\, \mathrm{d}x \quad . \tag{7.20b}$$

The parameter μ determines the strength of the signal process such that $\mu = 0$ corresponds to absent signal, whereas $\mu = 1$ clearly indicates the presence of signal events. The probability

density functions of the variable x for signal and background events are given by $f_s(x;\boldsymbol{\theta}_s)$ and $f_b(x;\boldsymbol{\theta}_b)$ which contain the set of parameters $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_b$ characterising the probability density function for signal and background events, respectively[(5)]. The total number of expected signal and background events are represented by s_{total} and b_{total} and are typically determined by Monte Carlo simulation. Each bin of the histogram is described by a Poisson distribution such that the likelihood function for the N bins is defined as the product of the Poisson probabilities for each bin:

$$L_{\text{Poisson}}(\mu,\boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \tag{7.21}$$

In addition to the parameters of interest such as the strength of the signal process μ, both signal and background models typically contain several so-called nuisance parameters which have to be determined by the fit as well. Systematic effects of either experimental or theoretical nature which have to be taken into account for the measurement are each described by a single nuisance parameter that follows a unit Gaussian distribution of the form

$$G(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\theta-\theta_0)}{2\sigma^2}\right] \quad . \tag{7.22}$$

The mean of the Gaussian distribution, θ_0, is defined by the nominal prediction, and its standard deviation σ is determined by the size of the systematic uncertainty which is typically assigned by the user based on prior knowledge on the result of an auxiliary measurement. With the substitution of $a = (\theta - \theta_0)/\sigma$ the mean of the newly defined nuisance parameter function $a(\theta)$ is set to zero and the corresponding standard deviation is assigned a value of unity [139, 256].

In addition, the nuisance parameters can be further constrained with the help of supplementary information statistically disconnected from the targeted signal region of interest but with some commonality which can be exploited. For example, another histogram $\boldsymbol{m} = (m_1, \ldots, m_M)$ of M bins and an expectation value of

$$E[m_i] = u_i(\boldsymbol{\theta}) \tag{7.23}$$

can be constructed by selecting a control samples which consists mainly of background events. Such a measurement designed to exploit additional information on the background normalisation parameter in the likelihood function in Eq. (7.21) will be extended by an additional term:

$$L(\boldsymbol{\theta}) = \prod_{j=1}^{M} \frac{u_j^{m_j}}{m_j!} e^{-u_j} \tag{7.24}$$

The statistical uncertainties within a bin i, however, are taken into account as additional nuisance parameters γ_i parametrising the standard deviation of the number of expected events ν_i [255].

Finally, with the ingredients of Eqs. (7.21) and (7.22) assuming K sources of systematic uncertainties, the overall likelihood functions reads:

$$L(\mu,\boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \prod_{j=1}^{M} \frac{u_j^{m_j}}{m_j!} e^{-u_j} \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{a^2(\theta_k)}{2}\right] \tag{7.25}$$

[(5)] The total set of parameters is comprised of the parameters which are specific to signal and background, $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_b$ and background normalisation factor(s):

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_s, \boldsymbol{\theta}_b, b_{\text{total}})$$

Since multiple nuisance parameters may affect the expected number of events of a single bin and, conversely, a single nuisance parameter may affect multiple bins, the nuisance parameters can be expected to become correlated to some degree even if the individual systematic uncertainties themselves are not a priori.

7.3.1.2. Statistical tests and p-values

The goal of statistical tests is to make a quantitative statement about how well observed data stand in agreement with a given prediction or expectation, referred to as a hypothesis. Typically, the hypothesis under consideration is referred to as null-hypothesis H_0 which defines a probability density function $f(x)$ of a random variable x, for example. The comparison with alternative hypotheses $H_1, H_2, \ldots$ allows one to make a quantitative statement about the validity of H_0. Given a dataset consisting of n measurements $\boldsymbol{x} = (x_1, \ldots, x_n)$ and a set of hypotheses H_0 and H_1, each of which specifying a given p.d.f. $f(\boldsymbol{x}|H_0)$ and $f(\boldsymbol{x}|H_1)$, for the investigation of the measure of agreement between the observed data and a given hypothesis, a function of the measured variables $t(\boldsymbol{x})$, referred to as a test statistic, is constructed. Each of the hypotheses implies a given p.d.f., $g(t|H_0)$ and $g(t|H_1)$, for the test statistic. The statement about the compatibility between the data and H_1 is usually equal to making a decision whether to accept or reject the null-hypothesis. For this purpose, the parameter range of the test statistic t is split into a so-called critical region and an acceptance region at some value t_{cut} which is denoted as a decision boundary. If the observed value of t falls into the critical region, the null-hypothesis is rejected, and otherwise H_0 is accepted. The decision boundary is chosen such that the probability to observe t within the critical region equals some value α which is referred to as the p-value of the test:

$$\alpha = \int_{t_{\text{cut}}}^{\infty} g(t|H_0)\,\mathrm{d}t \tag{7.26}$$

The null-hypothesis is then accepted, or, equivalently, not rejected, if the observed value of t is smaller than t_{cut}. By construction, there is a non-zero probability to reject H_0 although it is true, which is called error of the first kind or type-I error. Similarly, the error of the second kind or type-II error describes the situation that the null-hypothesis is accepted although the true hypothesis was rather an alternative hypothesis H_1. The probability for such a case is given by

$$\beta = \int_{-\infty}^{t_{\text{cut}}} g(t|H_1)\,\mathrm{d}t \tag{7.27}$$

with $1 - \beta$ called the power of the test to discriminate against the alternative hypothesis H_1 [139].

If the aim of an analysis is to find new phenomena or to measure a certain physics process, the null-hypothesis H_0 is usually defined as describing only the background which is to be tested against the alternative hypothesis H_1 including both background and the signal process under investigation. In contrast, if the purpose of an analysis is to set limits on the production rate of some BSM model, H_0 and H_1 swap places such that the null-hypothesis describing both signal and background is tested against the background-only hypothesis H_1. In order to quantify the outcome of such an experiment, the level of agreement of the observed data with some hypothesis is given by the p-value which is the probability that, under the assumption of H, data featuring an equal or greater incompatibility with the prediction of H can be found. The hypothesis under investigation is regarded as excluded if the p-value is found to be smaller than a given threshold. This measure of incompatibility can be either based on the number of observed events within designated phase-space regions or on the corresponding likelihood ratio for signal and background as it will be further elaborated later in this section. For a more intuitive interpretation, the p-value is typically

converted into a significance Z, which is defined in units of standard deviations σ of a Gaussian distribution[6]:

$$Z = \Phi^{-1}(1-p) \tag{7.28}$$

The function Φ^{-1} is the inverse of the cumulative distribution of a Gaussian distribution with a mean value of zero and a standard deviation equal to one. By convention, a significance of at least $Z = 5$ was taken as the appropriate level to regard the background-only hypothesis as rejected and hence to constitute an observation of the signal process which corresponds to a p-value of $2.87 \cdot 10^{-7}$[7]. At a significance of at least $3\,\sigma$ the signal process is not claimed to be observed, although evidence for the signal process is deemed to be found. For the purpose of excluding the signal hypothesis, however, a threshold p-value of 0.05 is commonly used which corresponds to $Z = 1.64$[8] in the one-sided definition. In addition to making statements about the observation or exclusion of a physics process, the significance is an important instrument for the characterisation of the sensitivity of an experiment to the signal, by reporting the expected significance that one would obtain for a variety of signal hypotheses. Specifically, if the aim of an analysis is to observe a signal process, the expected significance is defined as the median value for which the background-only hypothesis would be rejected, under the assumption of the presence of the signal. The corresponding p-value is calculated for the probability density function of the background-only hypothesis at the place of the median of the null-hypothesis describing both signal and background [256].

In order to construct a statistical test of good discrimination power, one has to choose the critical and acceptance regions and hence the decision boundary wisely in advance. According to the Neyman-Pearson lemma, the best test statistic in the sense of maximum power for a given p-value is given by the so-called likelihood ratio [257]:

$$f(\boldsymbol{x}) = \frac{f(\boldsymbol{x}|H_0)}{f(\boldsymbol{x}|H_1)} \tag{7.29}$$

The test statistic t is a function of the dataset $\boldsymbol{x} = (x_1, \ldots, x_n)$ and is defined as the ratio of the two probability density function $f(\boldsymbol{x})$ of the data given the null- and the alternative hypothesis, respectively. In many use cases, the p.d.f. under a certain hypothesis is determined from Monte Carlo simulation and is represented by a histogram [139].

In the frequentist approach used by the experiments at the LHC, statistical models are assumed to be known as they provide statements about the probability of data. When defining a hypothesis, all unknown parameters of the statistical model are to be determined from the data by a likelihood fit as previously described in Sec. 7.3.1.1. The test statistic for tests based on such models is defined as the so-called profile likelihood ratio,

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{7.30}$$

which is a function of the parameter of interest, the signal strength parameter μ that scales the total signal rate. The double-hatted set of nuisance parameters $\hat{\hat{\boldsymbol{\theta}}}$ denotes the set of estimators that maximise the likelihood function for a hypothesised value of μ, denoted as the conditional ML estima-

[6] The significance Z is defined such that the observed value of a Gaussian distributed variable found Z standard deviations above its expectation value has an upper-tail probability equal to the p-value [256].

[7] This number is valid for the one-sided case which is defined such that $Z = 0$ for $p = 0.5$. If one takes the two-sided definition which allows for fluctuations of a Gaussian variable on both sides of the p.d.f., the p-value corresponding to a significance of $5\,\sigma$ equals $5.7 \cdot 10^{-7}$.

[8] For the two-sided case $p = 0.05$ corresponds to two Gaussian standard deviations.

tor, and is thus a function of μ. The denominator, however, represents the unconditional likelihood function in which $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the estimators that maximise the likelihood function [256].

The profile likelihood ratio in Eq. (7.30) is constructed such that its value will always be in the range $0 \leq \lambda \leq 1$ with values near one implying good agreement between the data and the hypothesised value of μ. According to Wilk's theorem, for large datasets the distribution of the test statistic approaches a χ^2-distribution under the hypothesis $\mu = 0$ [258]. Since the χ^2 p.d.f. has the form of an asymptotic distribution for a sufficiently low number of degrees of freedom, the regime of large datasets is also referred to as the asymptotic regime. In most cases there is only one degree of freedom, the signal strength μ. Therefore, the higher the obtained values of an asymptotically distributed test statistic, the smaller the p-value for the background-only hypothesis and the compatibility of the data with the background-only hypothesis. A test statistic which fulfils this requirement is given by

$$t_\mu = -2 \ln \lambda(\mu) \quad , \tag{7.31}$$

where $\lambda(\mu)$ is the profile likelihood ratio defined in Eq. (7.30). The factor two in Eq. (7.31) is motivated by the fact that in the asymptotic regime μ follows a unit Gaussian distribution of the form

$$L(\mu, \hat{\mu}) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp\left[-\frac{1}{2}\left(\frac{\mu - \hat{\mu}}{\sigma_\mu}\right)^2\right] \tag{7.32}$$

according to the central limit theorem [245]. Inserting this Gaussian into Eq. (7.31) yields the χ^2-structure predicted by Wilk's theorem:

$$-2 \ln \lambda(\mu) = -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{7.33a}$$

with

$$L(\mu, \hat{\hat{\boldsymbol{\theta}}}) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp\left[-\frac{1}{2}\left(\frac{\mu - \hat{\hat{\boldsymbol{\theta}}}}{\sigma_\mu}\right)^2\right] \tag{7.33b}$$

and

$$L(\hat{\mu}, \hat{\boldsymbol{\theta}}) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \tag{7.33c}$$

which is the maximised likelihood function of Eq. (7.32). The likelihood ratio then clearly shows a χ^2-structure:

$$-2 \ln \lambda(\mu) = \left(\frac{\mu - \hat{\hat{\boldsymbol{\theta}}}}{\sigma_\mu}\right)^2 \tag{7.33d}$$

A test of the hypothesised value of μ using the statistic t_μ consequently yields high values for observations in favour of the alternative hypothesis. The level of disagreement of data and a given hypothesis for μ using the statistic t_μ is thus given by the p-value p_μ which is calculated according to

$$p_\mu = \int_{t_{\mu,\mathrm{obs}}}^{\infty} f(t_\mu|\mu)\, \mathrm{d}t_\mu \quad , \tag{7.34}$$

where $t_{\mu,\text{obs}}$ denotes the observed value of the statistic t_μ and $f(t_\mu|\mu)$ represents the probability density function of t_μ under the assumption of some hypothesis for μ [256].

In many particle physics analyses such as the counting experiments performed at the LHC, the contribution of the signal process to the mean number of events is assumed to be non-negative which implies that the estimator for the signal strength must be non-negative as well. For the observation or, equivalently, discovery of a non-negative signal contribution, the statistical test is defined as one-sided, such that unphysical downward fluctuations yielding fewer events than the expected number for the background-only hypothesis do not constitute evidence against the null-hypothesis. In order to take into account this additional constraint, an alternative test statistic q_0 is defined as

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{7.35}$$

where $\lambda(0)$ corresponds to the profile likelihood ratio as given by Eq. (7.30) for $\mu = 0$. Any lack of agreement with the background-only hypothesis is only considered for $\hat{\mu} \geq 0$, whereas values of $\hat{\mu} < 0$ are interpreted as originating from systematic uncertainties. Consequently, only observed values for $\hat{\mu} \geq 0$ can lead to increasingly large values of q_0 which correspond to an increasing level of incompatibility between the data and the background-only hypothesis. The level of disagreement is quantified, using the observed value of $q_{0,\text{obs}}$, by the p-value p_0, which is defined as

$$p_0 = \int_{q_{0,\text{obs}}}^{\infty} f(q_0|0)\,\mathrm{d}q_0 \quad , \tag{7.36}$$

where $f(q_0|0)$ denotes the p.d.f. of the statistic q_0 under the assumption of the background-only hypothesis $\mu = 0$. In the asymptotic limit of large datasets, p_0 depends only on the value of the test statistic such that the corresponding significance is, according to Eq. (7.28), given by

$$Z_0 = \sqrt{q_0} \quad . \tag{7.37}$$

The assumption that negative estimators for μ are unphysical does not apply in general. For example, experiments investigating neutrino oscillations, in which the presence of new phenomena could lead to an increase as well as a decrease in the observed numbers of events, have to consider both a greater and a lower rate than the expected number of events as signal hypothesis and therefore employ an appropriate form of the test statistic.

7.3.2. Cross-section measurement

The production cross section of the $t\bar{t}Z$ process was determined by a fit to the data within the three signal-enriched 2D-regions or, more simply, the signal regions as they will be referred to for the remainder of the thesis, each of them representing a single bin within the likelihood function. Within the fit, the ratio of the measured value of the $t\bar{t}Z$ production rate to the corresponding theoretical prediction of the Standard Model, referred to as signal strength $\mu_{t\bar{t}Z}$, was treated as the parameter of interest which was allowed to float freely. The value obtained for $\mu_{t\bar{t}Z}$ is able to then be converted into an actual cross section by

$$\sigma_{t\bar{t}Z} = \mu_{t\bar{t}Z} \cdot \sigma_{\text{MC}} \tag{7.38}$$

which relates the measured cross section to the production cross section employed during the generation of the simulated signal sample, as motivated in Sec. 6.3.2. As was seen in Sec. 7.2.3.3, Z+b constitutes the most dominant background contribution that remains within the signal regions

which motivates to take its normalisation from data. Therefore, in addition to $\mu_{t\bar{t}Z}$, the normalisation of the Z+b background was extracted simultaneously in the fitting procedure from the two Z+jets-enriched 2D-regions defined for the 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target regions, respectively. Since it was observed that for Z+jets-dominated regions the ratio of the observed yields from data to those from the total Monte Carlo prediction was larger than one for 5-jet events, but less than one for 6-jet events, it was decided to make use of two separate normalisation parameters for each of these two respective processes. They are referred to as N_{Z+b}^{2b5j} for the 2b5j Z+jets-enriched region and N_{Z+b}^{2b6j} for the 2b6j Z+jets-enriched region. Both normalisation parameters were connected to the Z+b contributions in the corresponding $t\bar{t}Z$-enriched regions, but were otherwise allowed to float with no further constraints on their values. A similar effect in the modelling was already observed in the 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target regions before the Z+jets-enriched 2D-regions had been defined based on the distribution of the respective BDT output scores. The modelling of the events with exactly five and at least six selected jets in association with at least two b-jets are shown in Fig. F.73a in the left and right bin, respectively. It appears that the ratio of selected data to simulated events is greater than one in the left bin, whereas the right bin it yields a value of smaller than one. In addition, the data-MC ratio with respect to the jet multiplicity, as depicted in Fig. F.73b, was seen to decrease towards higher number of selected jets. This is assumed to be driven by the most dominant Z+b background contribution, as the $t\bar{t}$ contribution shows the opposite behaviour in terms of the modelling which can be seen in Figs. F.10 and F.69. The various Boosted Decision Trees were independently trained and optimised within the respective target regions and were also seen to demonstrate good behaviour in terms of both overtraining and post-application modelling. Consequently, the different mismodelling with respect to the selected number of jets in the target regions translates into the Z+jets-enriched 2D-regions which is depicted in Fig. F.70. In contrast, although the three signal regions feature different kinematics as well, a single signal strength parameter was employed for the $t\bar{t}Z$ process targeted by this measurement. The measured value of $\mu_{t\bar{t}Z}$ is an overall scale factor for the expected signal yield and thus can be interpreted as an effective cross section which has to be independent from the acceptance of signal events within the phase-space regions featuring different kinematics. With a value of below $50\,\%$, the purity of Z+b is too small in the 2ℓ-Z-1b6j-Z+jets region which is the reason why this region was not employed as a control region for Z+b contribution in the 2ℓ-Z-1b6j-$t\bar{t}Z$ signal region, and a fixed uncertainty in the form of a nuisance parameter on the normalisation of Z+b was applied instead. The Z+light flavour background fraction, however, which was seen to be problematic for the target regions, could be sufficiently suppressed due to discrimination provided by the BDT such that the 2ℓ-Z-0b5j and 2ℓ-Z-0b6j regions were ultimately not used in the fit as dedicated control regions.

Several so-called validation regions were used in order to check the impact of the obtained values for the nuisance parameters and particularly of the free floating parameters on the modelling of the data. The extrapolation of the various nuisance parameters, in addition to the N_{Z+b} fitted parameter values, from the control regions into the signal regions was validated within the third Z+jets-enriched phase region, 2ℓ-Z-1b6j, which was not employed as a control region for Z+b due to the large amount of Z+c events in that region. Additionally, the three $t\bar{t}$-enriched regions were re-classified as validation regions as well since the $t\bar{t}$ process has already been estimated from data using the method described in Sec. 7.1.2. As previously mentioned in Sec. 7.2.3.3, the mixed phase-space regions were not considered for the measurement at all.

The total observed and expected event yields of all signal and control regions prior to having performed the fit are presented in Fig. 7.14 with each bin representing one region. The corresponding illustration of the pre-fit event yields for the validation regions can be found in Fig. F.74. In contrast to the way they were originally introduced, the labels of the regions were shortened in

order to properly fit into the bin widths. The signal regions, depicted on the left, are labelled with the suffix "$t\bar{t}Z$", followed by the two Z+jets control regions, 2b5j-Z+jets and 2b6j-Z+jets each labelled with the suffix "Z+jets". The subsequent bin corresponds to the Z+jets validation region, and the three rightmost bins represent the three $t\bar{t}$-enriched regions which were employed as validation regions as well. The yields of the simulated processes were normalised to their respective cross section and scaled to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The striped bands represent the combined statistical and systematic uncertainties on the expectation. In the ratio pad below the ratio of the events from data and Monte Carlo prediction are depicted and two were seen to be in reasonable agreement within the given uncertainties. The total pre-fit observed and expected yields of all 2D-regions that were employed for the fit are quoted in Tab. G.6–G.8.

The statistical uncertainties are treated as described in the Sec. 7.3.1.1, and a Gaussian function was added to the overall likelihood function for each of the uncertainties to be discussed in the subsequent section.

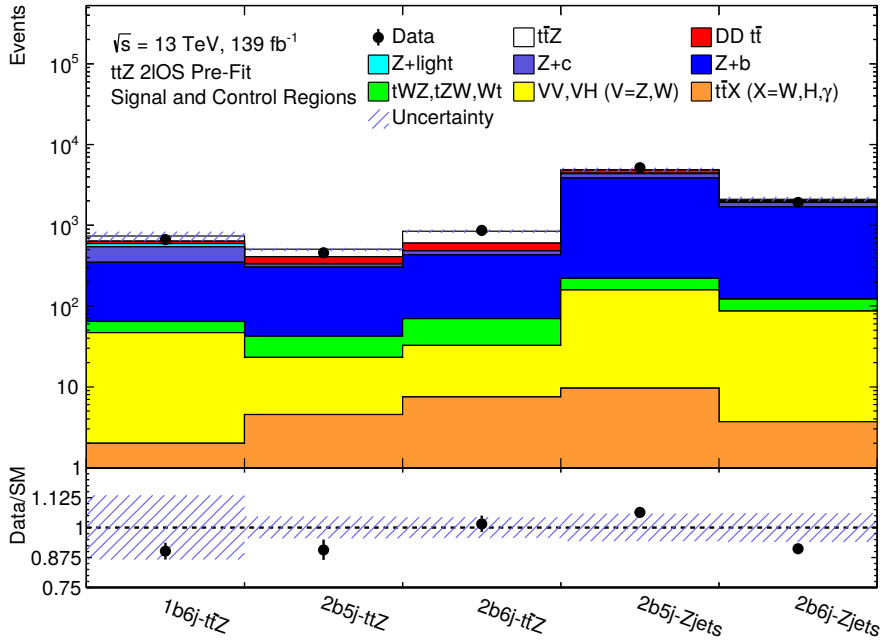


Fig. 7.14.: Observed and expected event yields in the 2ℓOS signal and control regions prior to the fit. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped bands represent the combined statistical and systematic uncertainties of the Monte Carlo prediction.

7.3.2.1. Systematic uncertainties

The simulated processes are affected by several sources of uncertainties, referred to as systematic uncertainties. The systematic uncertainties sources considered can be separated into experimental uncertainties, which are related to the reconstruction and calibration of physics objects in the detector, and theoretical uncertainties originating from the modelling of the various processes. In the following, the systematics which were considered for the measurement of the $t\bar{t}Z$ production cross section will be described.

There are several systematic uncertainties typically incorporated to a standard ATLAS analysis which will not be mentioned in this section. Since the evaluation of those systematics, which are expected to be rather small in this analysis, is very intensive in computing power and requires significant time and additional checks to be performed, they were not included into the approach presented in this thesis due to time constraint.

Experimental uncertainties

The value of the integrated luminosity of $139\,\mathrm{fb}^{-1}$ has an uncertainty $1.7\,\%$ which was derived by the methodology described within Ref. [259]. The uncertainty was applied to all simulated processes.

An uncertainty to account for residual differences between the pile-up distributions of Monte Carlo and data after the pile-up distribution of the Monte Carlo was reweighted to match those of the data was considered for the simulated processes.

Various uncertainties associated with the efficiencies of the triggers and with the efficiency of the reconstruction, identification and isolation of electrons and muons give rise to further uncertainties which were also included [195, 200–204].

In addition, uncertainties accounting for the choice of the Monte Carlo generator and the statistical uncertainty for the estimation of the residual contamination from pile-up jets which were not rejected by the JVT requirement are included in the fit. The efficiencies for the flavour-tagging procedure were measured through the use of dedicated control samples of both data and simulated events in order to derive factors which were applied to the Monte Carlo simulations.

The correction factors for b-tagged jets as well as the corresponding uncertainties are estimated using dileptonic $t\bar{t}$ events [260], whereas those for c-jets are derived from the jets of a hadronically decaying W boson from $t\bar{t}$ events [261]. The correction factors for the rejection of light-flavoured jets the correction factors are taken from dijet events [262]. All correction factors were derived as a function of the transverse momentum and the pseudorapidity of the respective jet and separate uncertainties are assigned to the extrapolation from the phase-space region where the various correction factors were determined from other phase-space regions. The uncertainties on these derived scale factors are propagated to the yields in simulation and the resulting differences in yields are taken as the corresponding source of uncertainty due to the flavour-tagging procedure.

From the data-driven estimation of the $t\bar{t}$ background described in Sec. 7.1.2, the uncertainty on the correction factor $C_{t\bar{t}}$ accounting for the different detector acceptance and selection efficiency between the SF and DF selection for dileptonic $t\bar{t}$ events is added as a systematic uncertainty for the $t\bar{t}$ events. The values of $C_{t\bar{t}}$ for the different selected jet and b-jet multiplicities are quoted in Tab. 7.5 with the corresponding uncertainties which are built from the statistical uncertainty of the $t\bar{t}$ events selected with either an DF or SF criterion, and an overall scale difference between data and prediction within the respective validation region.

Theoretical uncertainties

The normalisations of the Z+b background fraction was obtained from the fit to data within the two Z+b control regions. The prediction for Z+b in the 1b6j regions as well as the Z+light flavour and Z+c fractions are constrained within the fit to their respective prediction from Monte Carlo simulation. A normalisation uncertainty of $25\,\%$ and $30\,\%$ was assigned to the Z+light flavour and Z+c background fraction, respectively, in order to account for differences in the flavour composition of the Z+jets background between the control and the signal regions. The value for the uncertainty associated with Z+light flavour was motivated from Refs. [263] and [264], whereas

the uncertainty assumed for Z+c was inferred from the measurement of the ratio of the cross sections of Z+light flavour and Z+c, presented within Ref. [265]. For the Z+b background a flat uncertainty of 25 % was assumed based on the results of the measurement of the Z+b production cross section quoted in Ref. [266]. This value was chosen significantly higher than the uncertainty stated in the reference. As can be seen in Ref. [263], the relative uncertainty on the Z+light flavour cross section increases with a growing number of jets. However, the results are quoted inclusive in both jet and b-jet multiplicity such that uncertainties include the much larger contributions of the uncertainties from the events with a high jet-multiplicity. Therefore, a value of 25 % seemed to be a reasonable trade-off as the relative normalisation uncertainty of Z+b for the 1b6j regions. The tZq background was assigned a 30 % normalisation uncertainty which was motivated by the various measurements of this processes performed by the ATLAS and CMS collaborations, respectively [267–269]. An uncertainty on the cross section of single top quark events produced in association with a W boson of 15 % was assumed according to Ref. [270]. For the various processes labelled $t\bar{t}X$ as well as for tWZ, VV and VH backgrounds a conservative overall normalisation uncertainty of 50 % was considered for each process.

7.3.2.2. Results

The amount of signal and Z+b events were scaled according to the best values of $\mu_{t\bar{t}Z}$ and N_{Z+b} extracted from the fit, respectively. The yields of the other simulated processes were adjusted during the fitting procedure to their best-fit values within the provided uncertainties. The agreement between the data and the fitted prediction are shown for all signal and control regions in Fig. 7.15 and for the validation regions in Fig. F.75. The total uncertainty represented by the striped bands was noticeably reduced in each bin with respect to the pre-fit situation shown in Fig. 7.14. The total post-fit observed and expected yields of all 2D-regions that were employed for the fit are quoted in Tab. G.9–G.11.

The statistical and systematic uncertainties were parametrised, as described in the Sec. 7.3.1.1, by a Gaussian function each. The deviation of the estimate for a given nuisance parameter value as determined by the fit ($\hat{\theta}$) with respect to its pre-fit value (θ), divided by its assigned uncertainty ($\Delta\theta$), is referred to as pull P which is defined as:

$$P = \frac{\hat{\theta} - \theta}{\Delta\theta} \tag{7.39}$$

All nuisance parameters, including the γ factors accounting the for statistical uncertainties, were found to be constrained within one standard deviation of their Gaussian parametrisation, which implies that the model given to the fit is statistically stable as well as physically meaningful. The largest pulls were observed for the nuisance parameters associated to the uncertainties on the Z+c and VV/VH backgrounds as well as for the Z+b fraction within the 2ℓ-Z-1b6j-$t\bar{t}Z$ region. The best-fit values for those three nuisance parameters are situated almost half a Gaussian standard deviation below the pre-fit value, thus scaling down the expected number of events of the corresponding background processes. The largest pulls in terms of experimental uncertainties were observed for the nuisance parameters associated with the luminosity and one of the various flavour-tagging uncertainties. Both parameters are constrained within one standard deviation and are closer to their respectively pre-fit values than the three parameters accounting for the normalisation of certain background processes described above. A graphical overview of all nuisance parameters related to experimental or theoretical uncertainties can be found in Fig. F.76.

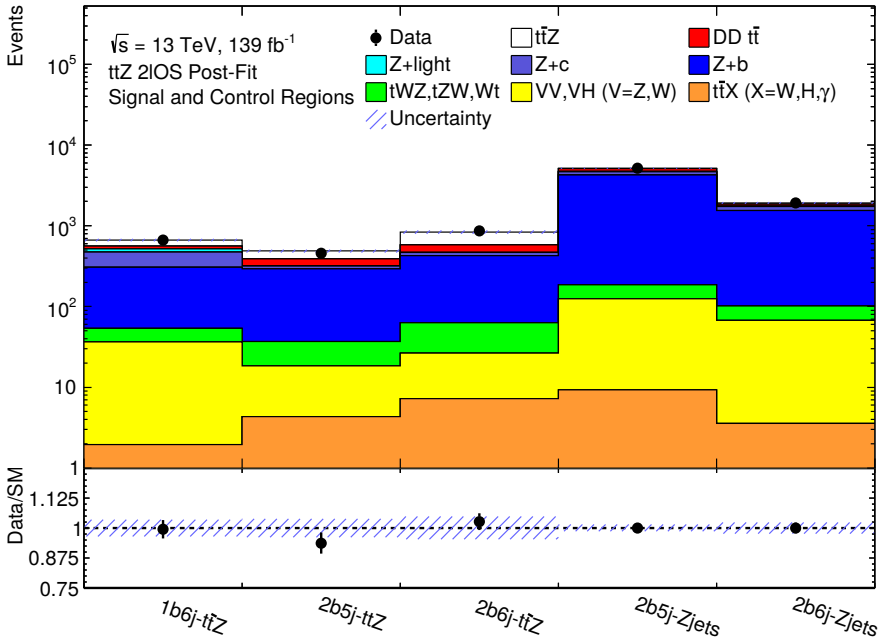


Fig. 7.15.: Observed and expected event yields in the 2ℓOS signal, control and validation regions after the fit. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped bands represent the combined statistical and systematic uncertainties of the Monte Carlo prediction.

The parameter of interest, the signal strength $\mu_{t\bar{t}Z}$, is found to have a best-fit value of

$$\mu_{t\bar{t}Z} = 1.04 \pm 0.10\,(\text{stat.}) \pm 0.16\,(\text{syst.}) = 1.04 \pm 0.18 \tag{7.40}$$

which is compatible with $\hat{\mu} = 1$. The values for the Z+b normalisation factors obtained from data are

$$N_{Z+b}^{\text{2b5j}} = 1.13 \pm 0.07 \tag{7.41a}$$
$$N_{Z+b}^{\text{2b6j}} = 0.93 \pm 0.07 \tag{7.41b}$$

thus scaling the total amount of predicted Z+b events up by approximately $10\,\%$ in the 2ℓ-Z-2b5j-$t\bar{t}Z$ signal region, whereas in the 2ℓ-Z-2b6j-$t\bar{t}Z$ the Z+b fraction was lowered by roughly the same percentage. The corresponding observed significance has a value of $5.9\,\sigma$, given an expected significance of $5.6\,\sigma$. The background-only hypothesis of $\hat{\mu} = 0$ is therefore able to be rejected in favour of the signal-hypothesis of $\hat{\mu} = 1$ and the observation of the $t\bar{t}Z$ process in 2ℓOS channel was established.

The measured value of $\mu_{t\bar{t}Z}$ and its uncertainty were then converted into an inclusive measurement of the $t\bar{t}Z$ cross section using the theoretically predicted value of $0.88\,\text{pb}$ [172, 173] from the nominal MC simulated signal sample. According to Eq.(7.38) the production cross section of the $t\bar{t}Z$ process was determined to be:

$$\sigma_{t\bar{t}Z} = 0.91 \pm 0.08\,(\text{stat.}) \pm 0.14\,(\text{syst.})\,\text{pb} = 0.91 \pm 0.16\,\text{pb} \tag{7.42}$$

The impact of the systematic uncertainties on the total uncertainty of the measured cross section was found to be greater than that associated with the limited data statistics. Since several systematic sources of uncertainties were not considered for this measurement, the quoted systematic uncertainty quoted here should therefore be interpreted as an underestimate. The contributions from the relevant systematic uncertainties of the measured $t\bar{t}Z$ cross section are summarised in Tab. 7.12 stating the relative uncertainty $\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z}$, expressed as a percent, on the measured cross section based on the particular systematic source. The quoted values are shown for the categories described in Sec. 7.3.2.1 and, since none of the nuisance parameters were seen to exhibit any significant asymmetry, were symmetrised. The dominant uncertainty was found to be attributed to b-tagging, followed by the uncertainties on the normalisation of the VV/VH background and the luminosity measurement. The value related to the data-driven $t\bar{t}$ estimation is compatible with zero at the quoted level of precision. As the uncertainty on the luminosity affects both the signal and the background normalisation, the related value after the fit is seen to be larger than the pre-fit value. The total relative uncertainty amounts to approximately $18\,\%$ and is clearly dominated by the various sources of systematic errors.

Uncertainty	$\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z}$ [%]
Flavour tagging	7.9
VV/VH normalisation	4.9
Luminosity	4.8
Z+c normalisation	4.4
Z+b normalisation (1b6j)	3.9
Lepton efficiencies	3.9
tWZ normalisation	3.6
JVT	2.2
Pile-up modelling	1.7
$t\bar{t}X$ normalisation	1.6
tZq normalisation	1.4
Z+light flavour normalisation	0.3
Wt normalisation	0.2
Data-driven $t\bar{t}$ estimation	0.0
Total systematic uncertainty	15.0
Statistical uncertainty	9.2
Total uncertainty	17.6

Tab. 7.12.: List of relative uncertainties on the measured $t\bar{t}Z$ production cross section. The quoted values were symmetrised and grouped into the categories described in Sec. 7.3.2.1. Due to correlations among various systematics sum in quadrature of the individual contributions in not equal to the total uncertainty.

The measured cross section was seen to be in good agreement with the NLO+NNLL theoretical prediction of the $t\bar{t}Z$ production cross section for a centre-of-mass energy of $13\,\mathrm{TeV}$,

$$\sigma_{t\bar{t}Z}^{\mathrm{NLO+NNLL}} = 0.863^{+8.5\,\%}_{-9.9\,\%}\,(\mathrm{scale})\ \pm 3.2\,\%\,(\mathrm{PDF}+\alpha_s)\,\mathrm{pb} \tag{7.43}$$

within the uncertainties [87].

As mentioned at the beginning of this chapter, the $t\bar{t}Z$ production cross section has already been measured in 2ℓOS channel at ATLAS [39]. At the time, the observed signal strength was

compatible with the signal hypothesis as well, although the best-fit value had a significantly higher relative uncertainty of more than $30\,\%$. However, a $3\,\sigma$ evidence could be constituted in favour of the signal hypothesis. The main difference of the new attempt presented in this thesis with respect to the previous measurement is, aside from the increased amount of available data, the method used to isolate the signal events from the vast background contribution of dileptonically decaying $t\bar{t}$ events and the Z+jets production. The approach of training a separate BDT against each dominant background process allows for the selection of $t\bar{t}Z$-enriched phase-space regions featuring values of the signal purity of up to three times higher than for the previous measurement which employed a single BDT to discriminate against the inclusive set of background processes. As a consequence, the sensitivity and thus the power of rejecting the background-only hypothesis increased significantly. However, the new approach does certainly benefit from the tremendous amount of data recorded by the ATLAS detector of the full Run 2 dataset which is almost four times larger than the dataset of $36.1\,\mathrm{fb}^{-1}$ that was employed in the previous measurement. With the 2D-selection a large amount of the available events is excluded from the measurement which is the reason why the $t\bar{t}Z$-enriched regions feature only about two or three times more signal events than the amount of signal selected by the previous measurement. If one were to apply the same technique as presented in this thesis for a smaller dataset of $36.1\,\mathrm{fb}^{-1}$, the statistical uncertainty would increase considerably, thereby reducing the sensitivity of the measurement.

8. Conclusion

For the first time, the full LHC Run 2 dataset recorded by the ATLAS detector was employed for a measurement of the $t\bar{t}Z$ production cross section by an analysis targeting in the 3ℓ and 4ℓ channels [46]. This thesis describes several dedicated studies which contributed to the global analysis effort. The two $t\bar{t}Z$ decay channels with either exactly three or four isolated leptons in the final state feature the highest signal purity within the hadronic environment of the LHC collisions. As final state of the Z boson pairs of electron and muons were considered, whereas taus contributed indirectly if they decayed leptonically. The additional leptons originate from the decay of one or both W bosons from the $t\bar{t}$ system. The inclusive cross section of $t\bar{t}Z$ production was measured by a simultaneous profile-likelihood fit of the Monte Carlo prediction to the observed data within the 3ℓ and 4ℓ signal regions. Normalisation factors for the light-flavour components of the WZ+jets and ZZ+jets background processes were estimated in dedicated control regions. Background contributions arising from non-prompt leptons or jets faking a lepton were estimated with the data-driven matrix method. Furthermore, differential cross sections of the $t\bar{t}Z$ process were measured as functions of several variables which are sensitive to the $t\bar{t}Z$ production.

In the first part of the presented studies for this measurement the simulated samples used to estimate the systematic uncertainties related to choice of a certain Monte Carlo generator were investigated. For the modelling of hard-scatter interactions of the signal samples, the nominal MADGRAPH5_AMC@NLO generator was compared to the alternative choice PYTHIA 8 and both were found to be reasonably consistent with each other. An important background for the 3ℓ channel is the associated electroweak production of a top quark with both a Z and a W boson. In order to prevent interference with the signal process during the Monte Carlo generation, the so-called diagram removal technique was applied to prevent such an overlap. The impact of systematic uncertainties from the application of this method was estimated with simulated samples generated with a slightly different implementation than the nominal samples. The validation of those samples revealed that a small subset of the events was assigned an incorrect value of the MC generator weight. As those events were identified to be the reason for the extremely high deviations in the predicted yields of the nominal and alternative samples, they were removed from the analysis. These studies consequently led to an improvement in the overall precision of the cross-section measurement.

The second part of the studies related to the 3ℓ and 4ℓ channels analysis were dedicated to an attempt to reduce the uncertainties on the measured $t\bar{t}Z$ cross section by exploiting the kinematic reconstruction of the top or antitop quark which decayed into a b-quark, a charged lepton and a neutrino. In the 3ℓ decay channel, two leptons originate from the Z boson decay, whereas the third lepton is assumed to be the decay product of a W boson from top quark decay. All possible combinations of associations of reconstructed objects to parton-level final states of the top quark decay were probed for a given event. The decision which of those combinations correspond to the true top quark decay products was made by interpolating its invariant mass with a reference distribution of the top quark mass. An optimal value to set a cut on the resulting distribution of the reconstruction output weight was looked for using an approximation to the mathematical description of the relative uncertainty on the measured cross section. Scanning a range from $70\,\%$ to $100\,\%$ of signal selection efficiency a minimum value could be found corresponding to a signal efficiency of $84\,\%$. All values of relative cross section uncertainty scatter within a range of $9.4\,\%$ to $9.9\,\%$ such that the found cut value does not appear to be statistically robust. This method was

seen to provide strong discrimination power, therefore the current implementation can serve as a starting point for a refinement of the method which, for example, involves the reconstruction of the $t\bar{t}Z$ system. A natural extension to the studies presented would be to employ the full fit machinery when performing the sweep through the cut values, including an evaluation of both statistical and the full set of systematic uncertainties. Such a study was beyond the available timescale, but would allow one to apply such a cut to the reconstruction output which would result in a clear improvement in precision considering all aspects.

The second part of this thesis focused on a novel approach for a measurement of the $t\bar{t}Z$ production cross section in the 2ℓOS channel. The general strategy is based on the latest ATLAS measurement of the $t\bar{t}Z$ production cross section in the 2ℓOS channel which employed a dataset of $36.1\,\text{fb}^{-1}$ [39]. The 2ℓOS channel was not targeted by the previously mentioned ATLAS cross-section measurement in the 3ℓ and 4ℓ channels as the sensitivity of this channel was seen to be significantly lower. In this thesis, prospects for a measurement of the $t\bar{t}Z$ production cross section solely targeting this decay channel were introduced. The studies were performed with the full LHC Run 2 dataset recorded by the ATLAS detector in the years from 2015 to 2018. Three phase space regions, denoted as target regions, with different numbers of selected jets and b-jets were defined and optimised for a high signal purity. The amount of expected $t\bar{t}$ events was estimated from data by subtracting the total Monte Carlo non-$t\bar{t}$ background from the data within phase space regions made orthogonal to the target regions by the requirement on the two leptons to feature different flavour.

As the background contribution in the target regions was found to be too large in order to perform a measurement of the $t\bar{t}Z$ process, a multivariate analysis technique was employed for the separation of signal and background processes. Two distinct Boosted Decision Trees per targeted phase space region were independently trained, each focusing on one of the two dominant background processes, dileptonic $t\bar{t}$ and the associated production of a Z boson with jets. For the training of the various classifiers, different sets of variables featuring the best discriminating power between the signal and the respective background process within the various target regions, including the output of two reconstruction techniques. The two-neutrino scanning method was employed in order to kinematically reconstruct dileptonically decaying $t\bar{t}$ events which, in contrast to $t\bar{t}Z$, feature two neutrinos in the final state. Furthermore, a kinematic reconstruction of the fully-hadronically decaying $t\bar{t}$ signature of the signal events was performed. In order to increase the sensitivity to the signal process, several hypotheses of the number of missing jets for the targeted signature in an event were taken into account to cover those events for which not all jet objects belonging to the $t\bar{t}$ final state were reconstructed. Both reconstruction techniques employ the interpolation of the reconstructed quantities such as the top quark mass with the corresponding reference distributions in order to find the probability of how consistent a reconstructed top quark candidate is with a top quark of dileptonic or fully-hadronically decaying $t\bar{t}$.

Various phase space regions, enriched in either the signal or in one of the two background processes, were defined by cutting in the two-dimensional plane spanned by the two BDT output scores. The $t\bar{t}Z$ production cross section was then measured within the signal-enriched phase space regions with a simultaneous profile-likelihood fit to the data in each of those regions. In addition to the signal strength parameter, the normalisation of the most dominant background fraction in those regions, Z+b, was separately determined within two of the Z+jets-enriched regions. In addition to various detector-related systematic uncertainties, the uncertainties of the important background processes were taken from the results of several measurements performed by the ATLAS and CMS collaborations.

The background-only hypothesis was excluded with a statistical significance of $5.9,\sigma$ and the signal strength of the $t\bar{t}Z$ process obtained from the fit was seen to be compatible with the Standard

Model prediction. The measured signal strength was then converted into a measured cross section of

$$\sigma_{t\bar{t}Z} = 0.91 \pm 0.16\,\text{pb} \tag{8.1}$$

for the observed $t\bar{t}Z$ signal which is in agreement with the most precise theoretical prediction. This measurement is dominated by the systematic uncertainties which amount to approximately $17\,\%$, whereas the statistical uncertainties were found to be $11\,\%$. In total the relative uncertainty on the measured cross section amounts to $20\,\%$.

It should be highlighted that only a partial set of the systematic uncertainties which are typically considered for such a cross section measurement was included in the measurement presented in this thesis. Aside from several detector-related systematics which would have required a comprehensive validation, the systematic uncertainties related to the choice of a specific Monte Carlo generator, which are estimated by the comparison of simulated samples produced with different Monte Carlo generators, were not taken into account. In contrast, most of the systematics related to the normalisation of a certain background processes were estimated rather conservatively such that the overall systematic uncertainty, despite featuring only a partial set, may nonetheless be sufficiently conservative. However, a quantitative statement cannot be provided until the missing systematics are included in the measurement.

Using the full LHC Run 2 dataset recorded by the ATLAS detector, the $t\bar{t}Z$ process could be observed in the $2\ell\text{OS}$ channel with a statistical significance of more than five standard deviations. A large increase of sensitivity with respect to Ref. [39] was seen to come from the multivariate framework. The two Boosted Decision Trees were each trained entirely focusing on a single background process, and were then combined to mark phase-space regions which are rather pure in either the signal events or in one of the two dominant background processes. The BDT classifiers themselves benefit from the outputs of the two kinematic reconstruction techniques targeting either the dileptonic or fully-hadronic $t\bar{t}$ signature, which were used as training variables. Due to the large amount of data collected in the years from 2015 to 2018, tight selection requirements for the signal-enriched phase space regions could be applied while maintaining a sufficiently high level of statistics. As the amount of collected data during LHC Run 3 is predicted to correspond to an integrated luminosity of about $300\,\text{fb}^{-1}$ and to even grow bigger up to $3000\,\text{fb}^{-1}$ in the context of the high-luminosity LHC up to 2035, a tremendous amount of data awaits to be analysed and to perform high-precision measurements. The novel approaches and methods introduced in this thesis still feature some aspects which could be refined in future measurements, but would certainly bring improvements to the overall precision, particularly when combined with 3ℓ and 4ℓ channels. In general, the technique presented in this thesis might serve as a model for the measurement of various properties of physical processes which are similarly or even more contaminated by a vast background contribution such as $t\bar{t}Z$ in the $2\ell\text{OS}$ channel.

A. Top quark pair production at the LHC

In Sec. 2.3.1 two mechanisms of top quark production have been discussed, the electroweak production of single top quarks and the top quark pair production via the strong force. At hadron colliders, top quark pair production is by far the main source of top quarks. It is illustrative example to compare the two mechanisms of top-antitop quark production, quark-antiquark annihilation and gluon-gluon fusion, and to identify the dominant one at the LHC operating at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. The centre-of-mass energy of a system, $\sqrt{s}$ is a Lorentz-invariant quantity. For a system of two colliding particles with four-momenta p_a and p_b, the quantities are related via

$$s = (p_a + p_b)^2 \quad . \tag{A.1}$$

The momenta of the partons a and b can be expressed in terms of the fractions x_a and x_b of the respective parent proton momenta. As the momentum of a proton is equal to the beam energy E_beam in the relativistic limit, the parton momenta can be rewritten[(1)] as

$$p_a = E_\text{beam} \begin{pmatrix} x_a \\ 0 \\ 0 \\ x_a \end{pmatrix} \qquad p_b = E_\text{beam} \begin{pmatrix} x_b \\ 0 \\ 0 \\ -x_b \end{pmatrix} \quad . \tag{A.2}$$

In order to produce a top quark pair, the centre-of-mass energy has to be equal to or greater than the squared rest mass of the top quark pair system,

$$s = (p_a + p_b)^2 \geq \left(2m_\text{top}\right)^2 \quad . \tag{A.3}$$

If it is assumed that both partons carry a similar momentum fraction ($x_a \approx x_b$), the required momentum fraction to create a top quark pair is given by

$$x \geq \frac{2m_\text{top}}{\sqrt{s}} = \frac{m_\text{top}}{E_\text{beam}} \tag{A.4}$$

where the centre-of-mass energy is half the beam energy $\sqrt{s} = 2E_\text{beam}$ and $x = x_a = x_b$. For a centre-of-mass energy of $13\,\text{TeV}$ and an assumed top quark mass of $172.5\,\text{GeV}$, the required momentum fraction would be approximately 0.027. At such a low momentum fraction the probability for gluons is by far higher than for (sea) quarks as was discussed in Sec 2.2, hence the main top quark pair production mechanism at the LHC is the gluon-gluon fusion which accounts for roughly $90\,\%$. For the associated production of top quark pairs with a Z boson the mass of the target system is higher than for top quark pair production alone, leading to a slightly higher momentum fraction of about 0.034.

The total number of top quark pairs expected to be produced at the LHC can be calculated according to

$$N = \sigma \cdot \mathcal{L} \tag{A.5}$$

[(1)] The general four-momentum vector has the form $\begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix}$

by multiplying the production cross section with the time-integrated luminosity. With a luminosity of $139\,\mathrm{fb}^{-1}$ and the theoretical cross section for top quark pair production given in Sec. 2.3.1.2 roughly 116 million top quark pair events were produced over the course of the Run-2 data-taking period by the ATLAS detector. Similarly, about $120\,000$ top quark pairs in association with a Z boson were produced, employing the theoretical cross section stated in Sec. 2.4.

B. Neutrino kinematics in the leptonic-side top quark reconstruction

The starting point for the leptonic-side top quark reconstruction is an assumption that the combined invariant mass of the non-Z lepton and neutrino four-vectors be equal to that of the W boson:

$$(p_\ell + p_\nu)^2 = m_W^2 \tag{B.1}$$

Expanding the left-hand side gives:

$$p_\ell^2 + p_\nu^2 + 2p_\ell p_\nu = m_W^2 \tag{B.2a}$$

$$E_\ell E_\nu - \vec{p}_\ell \cdot \vec{p}_\nu = \frac{m_W^2 - m_\ell^2}{2} \tag{B.2b}$$

$$E_\ell E_\nu - p_{\ell,x}p_{\nu,x} - p_{\ell,y}p_{\nu,y} - p_{\ell,z}p_{\nu,z} = \frac{m_W^2 - m_\ell^2}{2} \tag{B.2c}$$

In the above expression, the kinematic quantities of the lepton, E_ℓ and $\vec{p}_\ell = (p_{\ell,x}, p_{\ell,y}, p_{\ell,z})$, as well as the W boson mass are known. The transverse momentum of the neutrino, $p_{\mathrm{T},\nu}$, and its angle in the transverse plane, ϕ_ν, are in principle unknown as well, but assuming that the neutrino is the dominant source of missing energy in an event, the magnitude $E_\mathrm{T}^\mathrm{miss}$ and the corresponding azimuthal angle $\phi_\mathrm{T}^\mathrm{miss}$ of $E_\mathrm{T}^\mathrm{miss}$ can be taken as an appropriate substitute, respectively. Consequently, the z-component of the neutrino momentum, $p_{\nu,z}$, remains the only unknown. According to the relativistic energy-momentum relation [48],

$$p^2 = E^2 - \vec{p}^2 = m^2 \tag{B.3}$$

the squared invariant mass of a particle is equal to the square of its four-momentum. As the neutrinos can be treated as massless, the relation that

$$\begin{aligned} E^2 &= \vec{p}^2 \\ &= p_\mathrm{T}^2 + p_z^2 \quad , \end{aligned} \tag{B.4}$$

derived from Eq. (B.3), can be used to rearrange Eq. (B.2c) in order to isolate $p_{\nu,z}$:

$$E_\ell E_\nu = \frac{m_W^2 - m_\ell^2}{2} + p_{\ell,x}p_{\nu,x} + p_{\ell,y}p_{\nu,y} + p_{\ell,z}p_{\nu,z} \tag{B.5a}$$

$$E_\ell \sqrt{p_{\mathrm{T},\nu}^2 + p_{\nu,z}^2} = \frac{m_W^2 - m_\ell^2}{2} + p_{\ell,x}p_{\nu,x} + p_{\ell,y}p_{\nu,y} + p_{\ell,z}p_{\nu,z} \tag{B.5b}$$

$$E_\ell \sqrt{p_{\mathrm{T},\nu}^2 + p_{\nu,z}^2} = \frac{m_W^2 - m_\ell^2}{2} + p_{\mathrm{T},\nu}\left(p_{\ell,x}\cos\phi_\nu + p_{\ell,y}\sin\phi_\nu\right) + p_{\ell,z}p_{\nu,z} \tag{B.5c}$$

In the last step the Cartesian coordinates were substituted by polar coordinates using the following relations:

$$p_x = p_\mathrm{T}\cos\phi \tag{B.6a}$$

$$p_y = p_\mathrm{T}\sin\phi \tag{B.6b}$$

with the transverse momentum defined as

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2} \tag{B.6c}$$

Defining

$$\alpha = \frac{m_W^2 - m_\ell^2}{2} + p_{\mathrm{T},\nu}\left(p_{\ell,x}\cos\phi_\nu + p_{\ell,y}\sin\phi_\nu\right) \tag{B.7}$$

and squaring both sides of Eq. (B.5c) yields

$$E_\ell^2\left(p_{\mathrm{T},\nu}^2 + p_{\nu,z}^2\right) = \alpha^2 + 2\alpha p_{\ell,z}p_{\nu,z} + p_{\ell,z}^2 p_{\nu,z}^2 \quad . \tag{B.8}$$

If the terms of Eq.(B.8) are collected in powers of $p_{\nu,z}$, a quadratic equation of the form

$$Ap_{\nu,z}^2 + Bp_{\nu,z} + C = 0 \quad , \tag{B.9}$$

can be found with the following definitions:

$$A = E_\ell^2 - p_{\ell,z}^2 \tag{B.10a}$$
$$B = -2\alpha p_{\ell,z} \tag{B.10b}$$
$$C = p_{\mathrm{T},\nu}^2 E_\ell^2 - \alpha^2 \tag{B.10c}$$

The values of A, B, C and α are constant for a given event and are only governed by the non-Z lepton kinematics and the missing transverse energy, which are measured quantities, as well as from the W boson mass which is assigned a value of $80.385\,\mathrm{GeV}$ [51].

If the discriminant of Eq. (B.9) satisfies $B^2 - 4AC > 0$, two real solutions can be directly evaluated:

$$p_{\nu,z} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \tag{B.11}$$

Since for a given event it is not known which of the two solutions is the more appropriate one, both solutions are considered for the reconstruction of the top quark candidates. For a negative discriminant $B^2 - 4AC < 0$, no real solutions can be found. In such cases $E_{\mathrm{T}}^{\mathrm{miss}}$, or rather $p_{\mathrm{T},\nu}$, is solved for analytically by requiring $B^2 = 4AC$. This leads to a second quadratic equation in powers of $p_{\mathrm{T},\nu}$:

$$C_2 p_{\mathrm{T},\nu}^2 + C_1 p_{\mathrm{T},\nu} + C_0 = 0 \tag{B.12}$$

where the coefficients are defined as

$$C_2 = E_\ell^2 - p_{\ell,z}^2 - \left(p_{\ell,x}\cos\phi_\nu + p_{\ell,y}\sin\phi_\nu\right)^2 \tag{B.13a}$$
$$C_1 = -\left(m_W^2 - m_\ell^2\right)\left(p_{\ell,x}\cos\phi_\nu + p_{\ell,y}\sin\phi_\nu\right) \tag{B.13b}$$
$$C_0 = -\left(\frac{m_W^2 - m_\ell^2}{2}\right) \tag{B.13c}$$

The quadratic equation (B.12) yields in all cases two solutions for $p_{\mathrm{T},\nu}$ of which one is positive and the other one is negative. Only the positive solution is physically meaningful and is then taken to be then neutrino transverse momentum. The values of A, B and C of Eq. (B.8) have to re-evaluated given the new $p_{\mathrm{T},\nu}$ in order to solve for the single value of $p_{\nu,z}$.

C. Boosted Decision Trees

In this appendix two important features of Boosted Decision Trees, boosting algorithms and data preprocessing, are going to be explained in detail.

C.1. Boosting algorithms

The performance of a so-called weak classifier [250] can be enhanced by sequentially reweighting misclassified events after each training cycle such that they are preferentially treated by the subsequent training cycle hoping to refine the wrong decisions. In this section, the two boosting algorithms provided by TMVA, AdaBoost and Gradient Boost, are going to be presented. Additionally, a resampling technique referred to as bagging will be introduced although it is not necessarily a genuine boosting algorithm.

C.1.1. AdaBoost

One of the most widespread boosting algorithms is the so-called adaptive boosting method, shortly referred to as AdaBoost [271, 272], which will be explained in the following using the example of a BDT. Before the training of a classifier, each event out of the set of N training events, $\boldsymbol{x}$, is assigned an initial classification value $y_i = \pm 1$ with $+1$ and -1 corresponding to signal-like and background-like, respectively, as well as an initial weight $w_{(1),i} = 1/N$ $(i = 1, \ldots, N)$ such that the sum of all weights is normalised to unity:

$$\sum_{i=1}^{N} w_{(1),i} = 1 \tag{C.1}$$

If an event was misclassified by the BDT in a certain training cycle, it reweighted with a relatively high value such that it has a higher priority in the training of the subsequent decision tree. The first decision tree, $f(\boldsymbol{x}; w_{(1)})$, is evolved using the set of initial weights $w_{(1)}$. All other trees use the same input events which have been reweighted with a boost weight α_k, which is derived from the misclassification rate ε from the preceding training of the kth cycle:

$$\alpha_k = \frac{1}{2}\left(\frac{1-\varepsilon}{\varepsilon}\right) \tag{C.2}$$

The event weights for the subsequent training are then updated with the boost weights:

$$w_{(k),i} \to w_{(k+1),i} = w_{(k),i} \cdot \exp\left(\frac{\alpha_k f(\boldsymbol{x}; w_{(k)}) y_i}{Z_k}\right) \tag{C.3}$$

The denominator Z_k in Eq. (C.3) is a normalisation factor which is chosen such that the sum of weights equals unity:

$$\sum_{i=1}^{N} w_{(k+1),i} = 1 \tag{C.4}$$

The final classifier, $F(\boldsymbol{x})$, is built from the weighted average of all K decision trees trained:

$$F(\boldsymbol{x}) = \sum_{k=1}^{K} \alpha_k f(\boldsymbol{x}; w_{(k)}) \tag{C.5}$$

With the result of an individual decision tree for a given event defined as $h_k(\boldsymbol{x}) = \pm 1$ with $+1$ and -1 corresponding to signal-like background-like classification, respectively, the final classification of that event, $H(\boldsymbol{x})$, is given by the sum of all individual classifications:

$$H(\boldsymbol{x}) = \sum_{k=1}^{K} \alpha_k h_k(\boldsymbol{x}) \tag{C.6}$$

Values of $H(\boldsymbol{x})$ between 0 and $+1$ indicate a signal-like classification of an event, and values between -1 and 0 identify a given event as background-like. The sketched procedure employed by the AdaBoost algorithm performs best on weak classifiers which would feature a very poor classification performance without the application of a boosting algorithm. A further enhancement of the performance can be achieved by regularising the learning rate with a control parameter, β $(0 < \beta < 1)$, in the exponent of the boost weights:

$$\alpha_k \to \alpha_k^{\beta} \quad . \tag{C.7}$$

C.1.2. Gradient Boost

Alternatively to the AdaBoost algorithm, the concept of gradient descent was adapted which is the reason why this technique is referred to as Gradient Boost [273, 274]. The basic idea is estimate a function $F(\boldsymbol{x}$, which parametrises the BDT, with an additive expansion of the form

$$F(\boldsymbol{x}; P) = \sum_{k=1}^{K} a_k f(\boldsymbol{x}; w_k) \quad , \tag{C.8}$$

where each base function $f(\boldsymbol{x}; w_k)$ corresponds to a weak classifier and a_k are real parameters. The set of parameters P is comprised of the a_k parameters and the event weight w_k of the kth iteration:

$$P \in \{\alpha_k; w_k\}_0^K \tag{C.9}$$

By applying the boosting procedure P is adjusted such that the deviation between the the model responsive $F(\boldsymbol{x})$ and the true value y obtained from the training sample for an event is minimised. This deviation is measured by the so-called loss function $L(F, y)$ which is often chosen to be of the form of squared loss:

$$L(F, x) = (F(\boldsymbol{x}) - y)^2 \tag{C.10}$$

For example, AdaBoost is typically based on a exponential loss,

$$L(F, y) = \exp\left(-F(\boldsymbol{x})y\right) \tag{C.11}$$

which is the reason why Eq. (C.3) has a exponential form. However, the exponential loss is not as robust as with respect to large fluctuations in the input dataset such that it is expected to lose a bit of performance in noisy settings.

In contrast, the Gradient Boost algorithm, as implemented in TMVA, employs the more robust

so-called binomial log-likelihood loss function,

$$L(F, y) = \ln\left(1 + \exp\left(-2F(\boldsymbol{x})y\right)\right) \tag{C.12}$$

for which one has to resort to the method of steepest or gradient decent in order to perform the minimisation step. After calculating the current gradient of the loss function, a regression tree is grown such that its outcome matches the calculated gradient. The initial model in the loss function, $F(\boldsymbol{x})$ is then updated with the functional relation defined by the regression tree and the procedure is repeated until the desired set of decision trees that minimises the loss function is found.

The goal of the Gradient Boost algorithm is to find an approximation $\hat{F}(\boldsymbol{x})$ to $F(\boldsymbol{x})$ that minimises the value of a differentiable loss function for a given set of training events and which has the form of a weighted sum of individual weak classifiers $h_k(\boldsymbol{x})$:

$$F(\boldsymbol{x}) = \sum_{k=1}^{K} \gamma_k h_k(\boldsymbol{x}) + C \tag{C.13}$$

where C is real-valued constant. For the initialisation of the model a constant value, γ, is taken for the set of N input events:

$$F_0(\boldsymbol{x}) = \arg\min_{\gamma} \sum_{i=1}^{N} L(y_i, \gamma) \tag{C.14}$$

The model is then incrementally expanded by calculating so-called pseudo-residuals r_{ik} for each event i $(i = 1, \ldots, N)$ and each step k $(k = 1, \ldots, K)$:

$$r_{ik} = -\left[\frac{\partial L(y_i, F_{k-1}(x_i))}{\partial F_{k-1}(x_i)}\right] \tag{C.15}$$

Afterwards, a weak classifier $h_k(\boldsymbol{x})$ is fitted to the set of the kth residuals by means of training a regression tree using the r_{ik} as true outcomes instead of the y_i. The multiplicative factor γ_k is computed by solving the one-dimensional optimisation problem:

$$\gamma_k = \arg\min_{\gamma} \sum_{i=1}^{N} L(y_i, F_{k-1}(x_i) + \gamma h_k(x_i)) \tag{C.16}$$

As a last step, the model is updated:

$$F_k(\boldsymbol{x}) = F_{k-1}(\boldsymbol{x}) + \gamma_k h_k(\boldsymbol{x}) \tag{C.17}$$

This procedure is then repeated until the last step K is reached.

A big advantage of Gradient Boost is that its concept can in principle be applied to any arbitrary loss function, given its differentiability. As for AdaBoost, the performance can be further enhanced by limiting the learning rate within each iteration via a so-called shrinkage parameter which controls the amount of information gained in each iteration.

C.1.3. Bagging

Bagging refers to a resampling technique which is often applied instead of or in addition to boosting. A classifier is repeatedly trained using a random subset of all training events each time such that the final classifier represents an average of all individual classifiers. The main difference to a

boosting algorithm is its different approach to enhance the performance of a weak classifier. Instead of focusing on straightening the impact of statistical fluctuations of the input dataset on the training, it effectively smears over them. This results into an increases of both the stability and the performance of a classifier.

In addition, boosting and bagging techniques can be combined such as it was done for the method called stochastic gradient boosting [275], for which not the entire input dataset but only a subset of the training events is employed by he boosting algorithm.

C.2. Data preprocessing

Before discriminating variables are put into the training of a classifier, it is possible and sometimes beneficial to reduce correlations among some variables or to bring the shapes of the variables in an appropriate form.

C.2.1. Variable normalisation

Normalising a input variable means to apply a linear scaling such that its range is confined between -1 and $+1$. This transformation allows for a direct comparison of the classification weights that are assigned to the variables [248].

C.2.2. Variable decorrelation

Variables are often seen to be correlated causing problems for several MVA techniques which either underperform in presence of such correlations or ignore them at all. Linear correlations can be entangled using the square root of the covariance matrix for a couple of variables. This is equivalent to finding the matrix C' that yields the covariance matrix C when being multiplied with itself [276]. The covariance matrix can be determined by diagonalising the covariance matrix such that

$$D = S^T C S \qquad \overset{C=(C')^2}{\Longleftrightarrow} \qquad C' = S\sqrt{D}S^T \tag{C.18}$$

where D is diagonal matrix and S as well as its transposed matrix S^T are orthogonal. The term $\sqrt{D}$ corresponds to taking the square root of all entries d_i $(i = 1, \ldots, n)$ of the matrix D:

$$\sqrt{D} = \text{diag}\left(\sqrt{d_1}, \ldots, \sqrt{d_n}\right) \tag{C.19}$$

The linear decorrelation of the selected variables, $\boldsymbol{x}'$, is obtained from the multiplication of the original tuple $\boldsymbol{x}$ with the inverse of the square root of the covariance matrix:

$$\boldsymbol{x} \rightarrow \boldsymbol{x}' = (C')^{-1}\boldsymbol{x} \tag{C.20}$$

This procedure targets only linear correlations of the input variables. Non-linear correlations cannot be solved and may even be reinforced by this method [248].

A quite similar approach is the so-called principal component decomposition or principal component analysis (PCA) [277, 278]. It represents a linear transformation which rotates the training dataset points in the phase space of the input variables such that the maximum variability is visible. This procedure is equal to identifying the most important gradients within the dataset phase space. Within in the PCA-transformed space, the first coordinate is identified with the largest variance irrespective of a particular projection, referred to as first principal component. Accordingly, the second coordinate is identified with the second largest variance and so forth until the

dimensionality of the PCA-space is equal to that of the original phase space. As only the first few PCA-components with the largest variances are of importance, the given problem can be reduced from N dimensions to a few dimensions, where N equals the number of input variables. Therefore, the tuples of principal components,

$$\boldsymbol{x}_U^{\mathrm{PC}}(i) = \left(x_{U,1}^{\mathrm{PC}}(i), \ldots, x_{U,N}^{\mathrm{PC}}(i)\right) \quad , \tag{C.21}$$

of a tuple of input variables,

$$\boldsymbol{x}(i) = (x_1(i), \ldots, x_N(i)) \tag{C.22}$$

, measured for the ith event to belong to either signal ($U = S$) or background ($U = B$) are obtained by the following transformation rule:

$$x_{U,k}^{\mathrm{PC}}(i) = \sum_{l=1}^{N} \left(x_{U,l}(i) - \bar{x}_{U,l}(i)\right) v_{U,l}^{(k)} \quad , \qquad (k = 1, \ldots, N) \tag{C.23}$$

The $bold\bar{s}ymbolx_U$ and $v_U^{(k)}$ represent the tuples of the sample mean values and the eigenvectors of the variable phase space, respectively. The matrix of eigenvectors

$$V_U = \left(\boldsymbol{v}_U^{(1)}, \ldots, \boldsymbol{v}_U^{(N)}\right) \tag{C.24}$$

is then used to formulate the eigenvalue equation

$$C_U V_U = D_U C_U \tag{C.25}$$

for the covariance matrix C_U of the sample U, where D_U represents the diagonal matrix of the eigenvalues. With this transformation one is able to eliminate linear correlations of the input variables, but non-linear correlations cannot be entangled at all [248].

C.2.3. Variable transformation

The two previously described methods require linear correlations in order to yields reasonable results. As for many applications the input variables are correlated at higher orders, one has to transform the input variables such that their distributions become Gaussian before applying a decorrelation technique.

In a first step, a variable x has to be transformed into a uniform distribution with the help of the corresponding cumulative distribution function $F(x)$, which is defined as

$$F(x) = \int_{\infty}^{x} f(x')\,\mathrm{d}x' \tag{C.26}$$

with $f(x)$ being the probability density function of the variable x. Then, the uniform distribution is further transformed into a Gaussian shape with zero mean and unity width using its inverse error function, such that a input variable x with probability density function $f(x)$ is transformed as follows:

$$x \to \sqrt{2}\,\mathrm{erf}^{-1}\left(2\int_{-\infty}^{x} f(x')\,\mathrm{d}x' - 1\right) \tag{C.27}$$

The transformed input variable can then be passed to decorrelation method but may still feature non-linear correlations [248].

D. Two-neutrino scanning method

The aim of the two-neutrino scanning method is to perform a kinematic reconstruction of $t\bar{t}$ events which decay via the dileptonic channel and thus feature two neutrinos in the final state from intermediate W boson decays. The top quarks and the associated W bosons to which they decay can be reconstructed from their respective decay products such that the combined mass of the decay products is equal to the resonance mass in each case:

$$(p_\ell + p_\nu)^2 = m_W^2 \tag{D.1a}$$

$$(p_\ell + p_\nu + p_b)^2 = m_t^2 \tag{D.1b}$$

In order to perform a kinematic reconstruction of the dileptonic $t\bar{t}$ system, at least two jets as well as at least two charged leptons are needed. Since the top quark decays almost exclusively in a W boson and a b-quark, for the jets the two with the highest output score from the b-tagging algorithm are chosen for a given event. For the leptons only electrons and muons are considered. The leptons are required to be identified as an OSSF pair as the goal of this reconstruction is to improve the discrimination of signal events from dileptonic $t\bar{t}$ production. In such events need the amount of missing transverse energy is typically higher than for the $t\bar{t}Z$ signal.

As mentioned in Sec. 5.7, the individual four-momenta of the two neutrinos cannot be directly measured in the detector as only the combined information of the sum of the neutrino transverse momenta in the transverse plane is available. Therefore, the system in Eq. (D.1) is under-constrained and can thus not be solved analytically. Additional information about the angular separation of the two neutrinos and therefore about the portion either neutrino contributes to the observed value of $E_{\mathrm{T}}^{\mathrm{miss}}$ is required in order to reconstruct the $t\bar{t}$ system unambiguously. However, since the angular information of the two neutrino candidates is a priori unknown, many different assumptions of the values of both the pseudorapidity and the azimuthal angle of either neutrino candidate have to be tested.

The method adopted begins by scanning the possible values for η and ϕ are scanned within a range of $[-5, +5]$ and $[-\pi, +\pi]$, respectively, using a step size of 0.4 in either case. This results into a total of 160 thousand possible solutions for the values of the pseudorapidity (η_1, η_2) and azimuthal angle (ϕ_1, ϕ_2) of the two neutrino candidates.

In order to reduce this vast amount of possible combinations, regions in η-ϕ-space can be excluded in which the two neutrino candidates are very unlikely to be found given the kinematics of a $t\bar{t}$ event. Based on the values of some measured quantities, limits can be put on the kinematics of either of the reconstructed neutrino itself or its relation to other more directly measured objects. For this purpose, two-dimensional veto maps had been constructed for η and ϕ of the individual neutrinos separately, where η and the angular difference between the parton-level neutrino and its corresponding charged lepton from W boson decay, $\Delta R(\ell\nu)$, are treated as a function of a constraining variable. Those veto maps act as a probability density function f for the phase space spanned by the angular coordinates of the neutrino candidates and the corresponding constraint variable and therefore provide an indication of how likely a given hypothesis of η and ϕ is to be correct. For the construction of the veto maps parton-level information of simulated $t\bar{t}$ events was used such that the bare neutrino information was accessible.

The first constraint is set on the values of the pseudorapidity values of the two neutrino candidates, (η_1, η_2), based on the pseudorapidity of the joint system of the two leptons and the two

b-jets, $\eta^{bb\ell\ell}$, evaluated with detector-level information only. A given (η_1, η_2)-hypothesis is rejected if 95 % of all neutrinos from true simulated $t\bar{t}$ events at that given value of $\eta^{bb\ell\ell}$ hypotheses are above or below the chosen value η_0 of either neutrino candidate:

$$\int_{\eta_0}^{\infty} f(\eta^{bb\ell\ell}, \eta)\, \mathrm{d}\eta \geq 0.95 \tag{D.2a}$$

or

$$\int_{-\infty}^{\eta_0} f(\eta^{bb\ell\ell}, \eta)\, \mathrm{d}\eta \geq 0.95 \tag{D.2b}$$

A second constraint is applied on the set of all possible combinations $(\eta_{\nu 1}, \phi_{\nu 1}, \eta_{\nu 2}, \phi_{\nu 2})$ which pass the first requirement. The veto map was built for the angular difference between both parton-level neutrinos and either lepton, $\Delta R(\ell\nu)$, as a function of the sum of transverse momenta of the two leptons, $H_{\mathrm{T}}^{\ell\ell}$, which is the constraint variable. The same double-sided veto defined by Eqs. (D.2a) and (D.2b) is applied for a given (ϕ_1, ϕ_2)-hypothesis using $H_{\mathrm{T}}^{\ell\ell}$ as constraining variable. As the difference of the variables at detector and parton level are small particularly for the lepton-related $H_{\mathrm{T}}^{\ell\ell}$, and $\eta^{bb\ell\ell}$ is less impacted by jet response and resolution overall than some variable strictly related to the energy or p_{T} of the b-jets themselves, evaluating the veto maps on parton level by applying them to the detector-level values of the neutrino angular coordinates is a approach. The kinematic constraint for η is illustrated in Fig. D.1 and the constraint for $\Delta R(\ell, \nu)$ in Fig. D.2. Both figures depict the respective distribution of the events in the plane spanned by the two variables using parton-level information from the Monte Carlo generator in the left panel. The right panel, however, shows the corresponding phase-space regions that are vetoed in white and those which are not in red when applying the constraint to the objects on detector-level. Once the

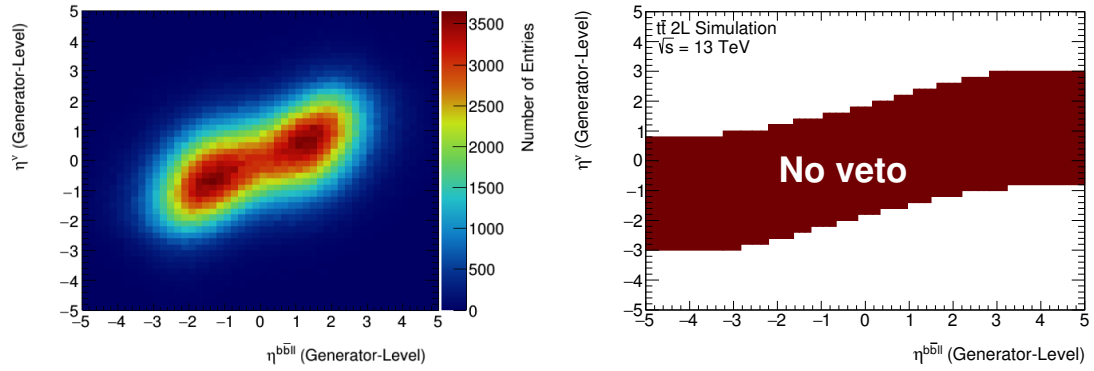


Fig. D.1.: illustration of the kinematic constraint on the reconstructed neutrino η. In the left panel, the occupancy of the events in the spanned phase space is shown. This translates into a veto for events outside the 95 %-area depicted in the right panel.

values of η and ϕ are selected, the transverse momentum of each of the two neutrino candidates can be calculated according to

$$p_{\mathrm{T},\nu} = \frac{\frac{1}{2}\left(m_W^2 - m_\ell^2\right)^2}{E_\ell \cosh\eta_\nu - p_{\ell,z}\sinh\eta_\nu - p_{\ell,x}\cos\phi_x - p_{\ell,y}\cos\phi_\nu} \quad , \tag{D.3}$$

derived by requiring the squared sum of of the four-momenta of a neutrino candidate and the corresponding lepton to be equal to the squared mass of the W boson. This approach is similar to what

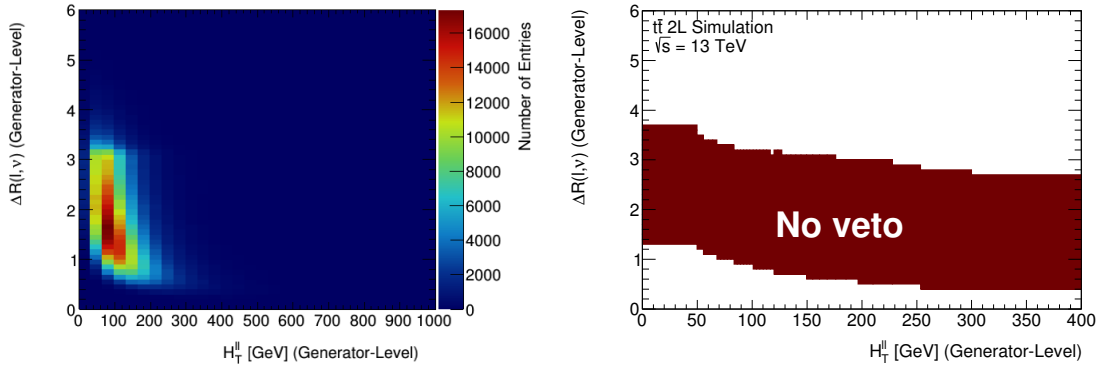


Fig. D.2.: illustration of the kinematic constraint on the $\Delta R(\ell, \nu)$ of the reconstructed neutrino and the corresponding charged lepton. In the left panel, the occupancy of the events in the spanned phase space is shown. This translates into a veto for events outside the $95\,\%$-area depicted in the right panel.

was done in Eq. (B.1) for the leptonic-side top quark reconstruction described in Sec. 6.3.1, with auxiliary information given in App. B. For the mass of the W boson a value of $80.385\,\text{GeV}$ [51] was used.

The final step of the reconstruction of two top quark candidates is the selection of the detector-level jets to be associated to the b-quark and $\bar{b}$-quark from the top and antitop quark decay, respectively. According to Eq. D.1, the two top quark candidates are to be built from the neutrino candidates which pass the constraining requirements as well as from the two reconstructed leptons and the selected b-jet candidates. All possible combinations to build a top quark from the reconstructed objects and the two neutrinos are probed and assigned some weight. The compatibility of the given solution of the $t\bar{t}$ reconstruction with the hypothesis that a given event features a dileptonic $t\bar{t}$ system is tested with an output weight:

$$w_{2\nu\text{SM}} = P_{m_t1} \cdot P_{m_t2} \cdot P_{\Delta E_x} \cdot P_{\Delta E_y} \tag{D.4}$$

The weight is defined based on the invariant mass distributions of the two top quarks and the distributions of the differences in kinematics between the parton-level neutrino and E_T^miss. The first two terms of Eq. (D.4), P_{m_t1} and P_{m_t2}, represent the probability of the invariant mass values of the reconstructed top quark candidates are with leptonic-side top quarks which are obtained from interpolation with a reference top quark mass distribution of $m_{b\ell\nu}$ built from parton-level neutrinos and reconstructed leptons and jets which have been matched to the corresponding parton-level objects. This idealised distribution of the top quark mass was scaled such that it has its maximum has a value of one. The second part of Eq. (D.4), $P_{\Delta E_x} \cdot P_{\Delta E_y}$, accounts for agreement of the kinematics of the reconstructed neutrinos given a set of η and ϕ coordinates with the true neutrinos at parton level. To measure the amount of agreement of reconstructed and generated neutrinos, a weight is derived from the interpolation of the observed value of the difference between the parton-level neutrino p_T and the measured E_T^miss in the x- and y-direction with a corresponding reference distribution:

$$\Delta E_x = E_{x,\nu1} + E_{x,\nu2} - E_x^\text{miss} \tag{D.5a}$$

$$\Delta E_y = E_{y,\nu1} + E_{y,\nu2} - E_y^\text{miss} \tag{D.5b}$$

The combination of reconstructed objects and neutrino hypotheses yielding the largest product of

all weights is selected as the reconstructed dileptonic $t\bar{t}$ system. The idealised reference distributions of the top quark mass and the neutrino energy difference can be found in Fig. D.3. Both top-related weights are derived from the distribution shown in the left panel. As the differences between the neutrino energy difference in x-direction to those in y-direction are very small, a single reference distribution was employed which is depicted in the right panel.

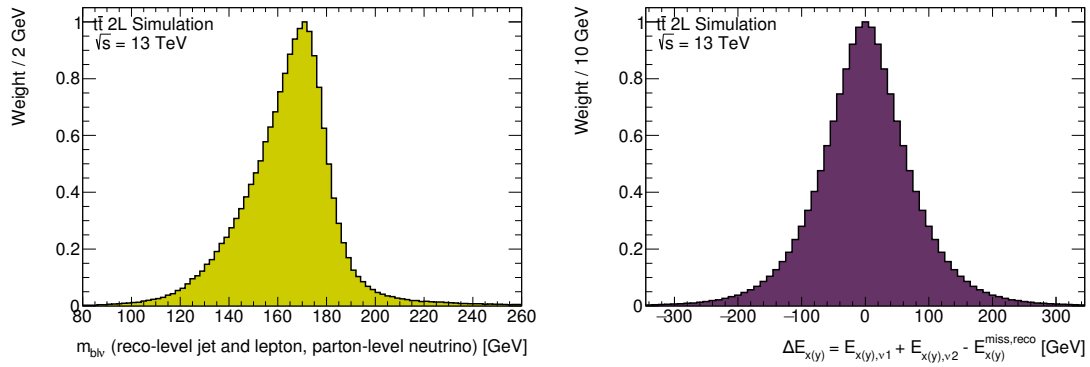


Fig. D.3.: Distribution of correctly reconstructed top quarks on the left ($m_{b\ell\nu}$) and of the difference between the energy of a reconstructed and a parton-level neutrino in simulated $t\bar{t}$ events. The reconstructed objects used to fill the distribution are matched to the corresponding parton level objects.

As much larger uncertainties are associated with the jet kinematics with respect to leptons, the reconstructed top quark mass has a reduced resolution even if one could access the exact kinematics of the two neutrinos. Simulated events can be used to determine from detector-level jets matched which have been matched to b-quarks in simulated events the ratio of the transverse momenta of the b-quark and the b-jet:

$$R_b = \frac{p_{\mathrm{T}}^{b\text{-quark}}}{p_{\mathrm{T}}^{b\text{-jet}}} \tag{D.6}$$

The spectrum of the b-jet p_{T} was divided into ten bins for each of which a distribution of R_b was filled.A b-jet scale factor (bJSF) is defined for each those distributions as either the point on the x-axis that corresponds to the most probable value, or as unity. The final output weight is calculated for each possible combination of reconstructed objects and neutrino hypotheses with both the unscaled and scaled b-jet four-momentum, thereby doubling the number of combinations. The choice that yields a higher value of $w_{2\nu\mathrm{SM}}$ is then taken as bJSF. Fig. D.4 presents an example of the R_b distribution and also shows how it is binned in terms of the unscaled b-jet p_{T}. However, there are cases where one specific ℓb-association appears to be highly preferred over the other such that the evaluation of the final output weight can be restricted to a single ℓb-association. Defining the minimum average invariant mass of the ℓb-system,

$$\min\langle m_{\ell b}\rangle = \min\left\{m_{\ell_1 b_1} + m_{\ell_2 b_2}, m_{\ell_1 b_2} + m_{\ell_2 b_1}\right\} \quad , \tag{D.7}$$

the minimum value was found to correspond to the correct set of ℓb-associations in $90\,\%$ of cases based on simulated $t\bar{t}$ events for sufficiently high values of $H_{\mathrm{T}}^{\ell\ell}$. These particular values of $HT_{\ell\ell}$ were chosen differently for cases in which the two b-jets were the same ($\eta^b \times \eta^{\bar{b}} \geq 0$ or opposite ($\eta^b \times \eta^{\bar{b}} < 0$) hemispheres with respect to the transverse plane in the detector. In the first case the threshold value for $H_{\mathrm{T}}^{\ell\ell}$ was chosen to be at $150\,\mathrm{GeV}$, whereas in the latter case it has much

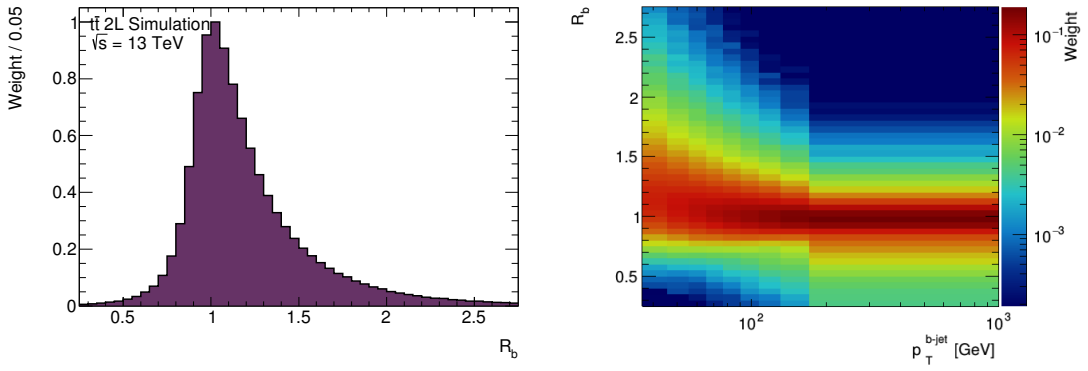


Fig. D.4.: Distribution of the the p_{T} ratio of the b-quark and the reconstructed b-jet in simulated $t\bar{t}$ events. The distribution in the left panel shows the R_b distribution that is used to decide on the value of bJSF and is given inclusively for all p_{T} bins. In the right panel the distribution of R_b as a function of the unscaled b-jet p_{T} is presented which shows the chosen binning. The reconstructed b-jets used to fill the distribution are matched to the corresponding parton-level b-quarks.

higher value of $300\,\mathrm{GeV}$. If in either case the $H_{\mathrm{T}}^{\ell\ell}$ variable has a value smaller than the respective threshold, both combinations are taken into account for the evaluation of the final output weight. This additional constraint limits the number of iterations and therefore the combinatorial freedom for events that do not feature a dileptonic $t\bar{t}$ system. However, there are still $10\,\%$ of $t\bar{t}$ events incorrectly reconstructed since the other ℓb-association was in fact correct.

As the neutrino scanning method was specifically designed to reconstruct dileptonic $t\bar{t}$ events, a sufficiently high value for the final output weight is expected for such events. For $t\bar{t}Z$ events decaying in the $2\ell\mathrm{OS}$ channel, however, a rather low value of $w_{2\nu\mathrm{SM}}$ is expected as its final-state signature does not feature any neutrino at tree-level and, in addition, the two leptons originate from the leptonically decaying Z boson rather than from a leptonically decaying W boson. The presented reconstruction method does feature many parallels to the so-called neutrino weighting method [279] which has been applied by numerous ATLAS analyses. The key difference between the two methods is the treatment of the top quark mass which is constrained from the beginning within the neutrino weighting method according to Eq. (D.1). However, the neutrino scanning method leaves the compatibility test for the reconstructed top quark to the final step due to response and resolution effects.

E. Multi-hypothesis hadronic top quark and W boson reconstruction

The $t\bar{t}Z$ events in the 2ℓOS channel feature a fully-hadronic $t\bar{t}$ decay, whereas its two dominant backgrounds, the production of dileptonic $t\bar{t}$ and the production of a Z boson in association with jets, do not feature hadronically decaying top quarks nor W bosons. For a full kinematic reconstruction of fully-hadronic $t\bar{t}$ events at least six jets are required. For a given number of jets in an event, N_{jets}, the number of unique permutations of the jets is defined as follows:

$$\binom{N_{\text{jets}}}{2}\binom{N_{\text{jets}}-2}{2}\binom{N_{\text{jets}}-4}{2} \tag{E.1}$$

The first in the expression corresponds to the choice of two jets to be in the positions of the two b-quarks. From the remaining $N_{\text{jets}}-2$ jets, two are selected to form the W boson from one of the two top quarks. Two more jets of the $N_{\text{jets}}-4$ jets which are left are needed to form the W boson of the other top quark.

However, due to detector acceptance as well as response and resolution effects, for a substantial number of events at least one of the six primary quarks from the $t\bar{t}$ decay will not lead to the reconstruction of a hadronic jet In such cases, a full kinematic reconstruction of the $t\bar{t}$ cannot be performed. In order to recover these events for the analysis, the standard reconstruction procedure was extended to a set of independent algorithms which perform either a full or partial reconstruction of the hadronically decaying $t\bar{t}$ system. This offers additional discrimination between signal and background than a sole reconstruction of the full hadronic $t\bar{t}$ decay. For this purpose, a reconstruction technique based on several hypothesis for the number of missing jets in an event was employed, referred to as multi-hypothesis top quark and W boson reconstruction. A total of five different scenarios was considered, depending on a given hypothesis about how many quarks of the fully-hadronic $t\bar{t}$ signature can be matched to a reconstructed jet. Aside from the full $t\bar{t}$ system, labelled with $2t$, cases with only one or two missing b-jets are considered, referred to as $1t1W$ and $2W$, respectively. Furthermore, events where only one of the top quarks ($1t$) or even only the associated W boson ($1W$) are accessible were targeted. Depending on the targeted signature, the minimum requirement on the jet multiplicity changes accordingly. All targeted hypotheses with the corresponding number of unique combinations are summarised in Tab. E.1. In addition, the condition on the jet multiplicity for a given hypothesis is shown in order to be considered for a given event. For a given hypothesis, all possible associations of the set of available reconstructed jet to the top quarks and W bosons have to be probed. An output weight was evaluated for each jet pair and jet triplet from the interpolation with reference distributions of the invariant masses of fully-hadronic top quarks (m_{bjj}) and hadronically decaying W boson (m_{jj}), respectively. The reconstructed top quarks and W boson used to fill the reference distribution are formed from jets which have been matched to the corresponding parton-level objects and therefore represent idealised distributions that can be achieved with a perfect assignment of reconstructed objects to their corresponding parton-level objects aside from residual corrections to the object kinematics. The reference distributions of m_{bjj} and m_{jj} were scaled such that their maxima has a value of one such that the interpolation values are representative of the probability of how consistent a reconstructed top quark or Z boson candidate is with a hadronically decaying top quark or W boson. The respective reference distributions used for the reconstruction are depicted in Fig. E.1. The weights

Targeted object(s)	Number of unique combinations	Condition
$2t$ $[b,q,q]\,[b,q,q]$	$\binom{N_{\text{jets}}}{2}\binom{N_{\text{jets}}-2}{2}\binom{N_{\text{jets}}-4}{2}$	$N_{\text{jets}} \geq 6$
$1t + 1W$ $[b,q,q]\,[q,q]$	$\binom{N_{\text{jets}}}{1}\binom{N_{\text{jets}}-1}{2}\binom{N_{\text{jets}}-3}{2}$	$N_{\text{jets}} \geq 5$
$2W$ $[q,q]\,[q,q]$	$0.5 \cdot \binom{N_{\text{jets}}}{2}\binom{N_{\text{jets}}-2}{2}$	$N_{\text{jets}} \geq 4$
$1t$ only $[b,q,q]$	$\binom{N_{\text{jets}}}{1}\binom{N_{\text{jets}}-1}{2}$	$N_{\text{jets}} \geq 3$
$1W$ only $[q,q]$	$\binom{N_{\text{jets}}}{2}$	$N_{\text{jets}} \geq 2$

Tab. E.1.: List of the various categories targeted by the multi-hypothesis top quark reconstruction. In the left column, below the label of the respective scenarios the corresponding jet pairs and triplets are shown. The number of unique conditions based on the jet multiplicity of a given event can be derived from the expression in the central column. According to the number of final-state objects that are assumed to be accessible, a minimum jet multiplicity is required for an event to order to be considered for the respective hypotheses.

of the individual jet pairs and triplets, w_t and w_W, were combined separately for each category in order to build final output weights indicating the compatibility of a given permutation with the true signature, given the respective hypothesis. The final outputs are defined as the geometric weight of all individual weights in each category:

$$w_{2t} = \sqrt[4]{w_t(m_{b1j1j2}) \cdot w_W(m_{j1j2}) \cdot w_W(m_{b2j3j4}) \cdot w_W(m_{j3j4})} \tag{E.2a}$$

$$w_{1t1W} = \sqrt[3]{w_W(m_{b1j1j2}) \cdot w_W(m_{j1j2}) \cdot w_W(m_{j3j4})} \tag{E.2b}$$

$$w_{2W} = \sqrt{w_W(m_{j1j2}) \cdot w_W(m_{j3j4})} \tag{E.2c}$$

$$w_{1t} = \sqrt{w_W(m_{b1j1j2}) \cdot w_W(m_{j1j2})} \tag{E.2d}$$

$$w_{1W} = \sqrt{w_W(m_{j1j2})} \tag{E.2e}$$

From all possible permutations of the reconstructed jets only the one yielding the highest value of the final output weight is considered correct for a given hypothesis. Irrespective of the actual number of jets in an event, the reconstruction assuming each hypothesis is performed since it is a priori unknown what requisite jets are present from the $t\bar{t}$ decay. In the case an events has fewer jets than required for a certain scenario, a compatibility of zero will be assumed for this hypothesis. The distributions of the final output weights of the various scenarios can then be used to discriminate between signal events featuring fully-hadronic $t\bar{t}$ final states and background events of processes with different final states.

A fully reconstructable event which has all correct jets are present such that the $t\bar{t}$ can be correctly reconstructed would typically achieve a high score not just for the $2t$ category but for all other categories. This introduces additional correlations between the individual categories. In order to reduce the combinatorial freedom for both signal and background events and to increase the

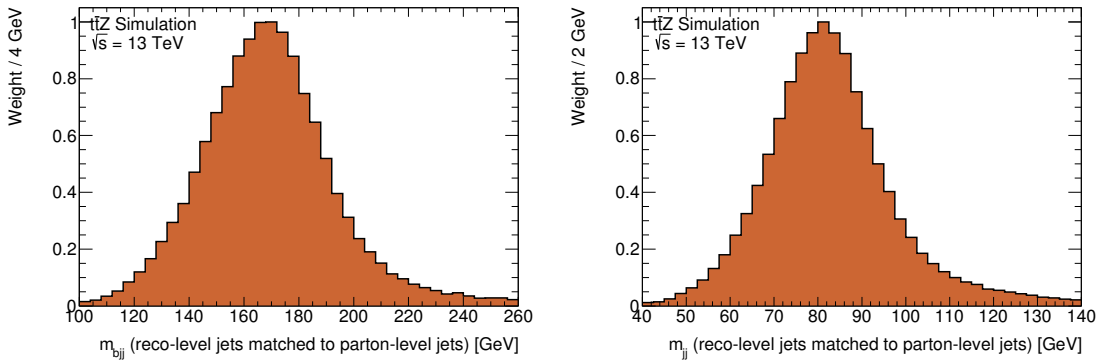


Fig. E.1.: Invariant mass distributions from correctly reconstructed top quarks on the left (m_{bjj}) and of correctly reconstructed hadronically decaying W bosons on the right (m_{jj}) in simulated $t\bar{t}Z$ events. The reconstructed top quarks and W bosons used to fill the distribution are formed from the reconstruction-level jets that are matched to the corresponding parton level objects.

Targeted object(s)	Constraint(s)
$2t$	The two jets with the highest output score from the b-tagging algorithm must be used, and the highest-valued jet must strictly be in the b-quark position of one of the two top quark candidates.
$1t$ + $1W$	The jet with the highest output score from the b-tagging algorithm must be in the b-quark position.
$1t$ only	One of the two jets with the highest output score from the b-tagging algorithm must be in the b-quark position for the cases with five to seven jets. If $N_{\text{jets}} \geq 8$, the highest-valued jet must be in the b-quark position. For fewer than five jets no constraint is applied.

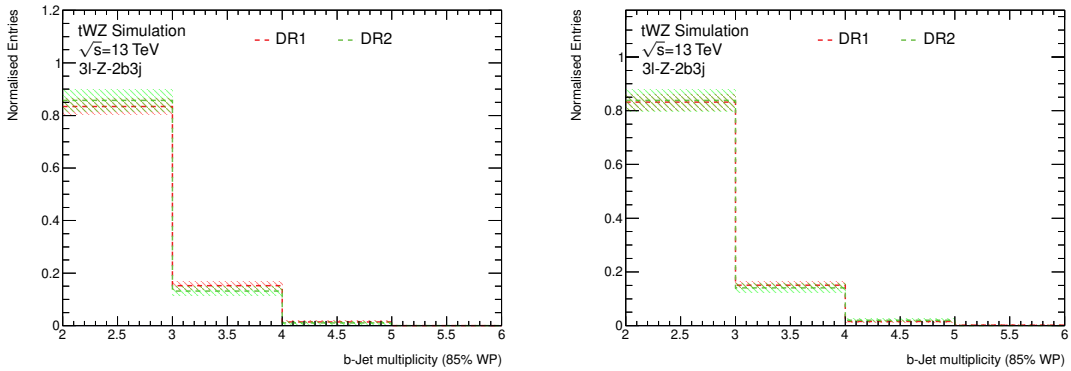
Tab. E.2.: List of constraints from b-tagging which are applied to categories shown in the left column. A detailed description of the course of action in a given scenario can be found in the right column.

fraction of correctly reconstructed events at the same time, a set of constraints from b-tagging was applied to those scenarios with a at least one top quark assumed reconstructable. The constraints are detailed in Tab E.2 for the respective scenarios and were optimised based on simulated $t\bar{t}Z$ events.

F. Additional figures

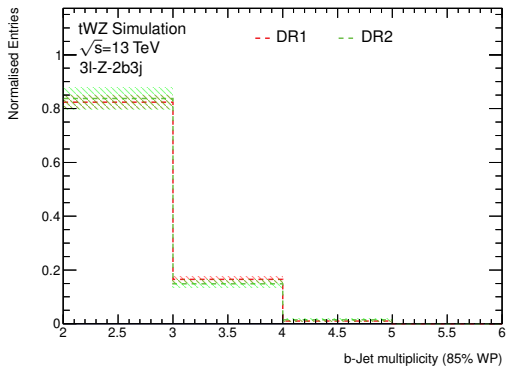
F.1. Comparison of the two different diagram removal schemes

In addition to Fig. 6.3 in Sec. 6.2.2, normalised distributions of the two diagram removal schemes for the tWZ background are shown in Figs. F.1 and F.2, respectively.



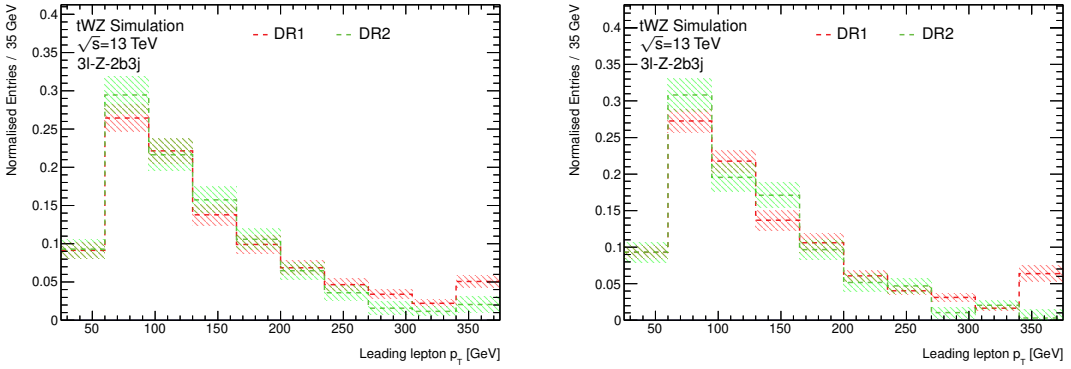
(a)

(b)



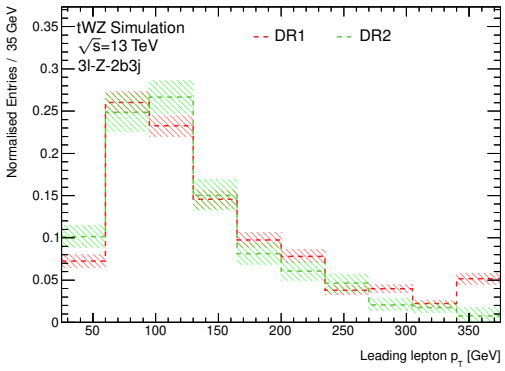
(c)

Fig. F.1.: Normalised distributions of the b-jet multiplicity for mc16a, mc16d and mc16e in Figs. (a), (b) and (c), respectively, shown for the 3ℓ differential signal region. The red line corresponds to the DR1 sample, and the green line to the DR2 tWZ sample. The error bars in all panels include only MC statistical uncertainties.



(a)

(b)

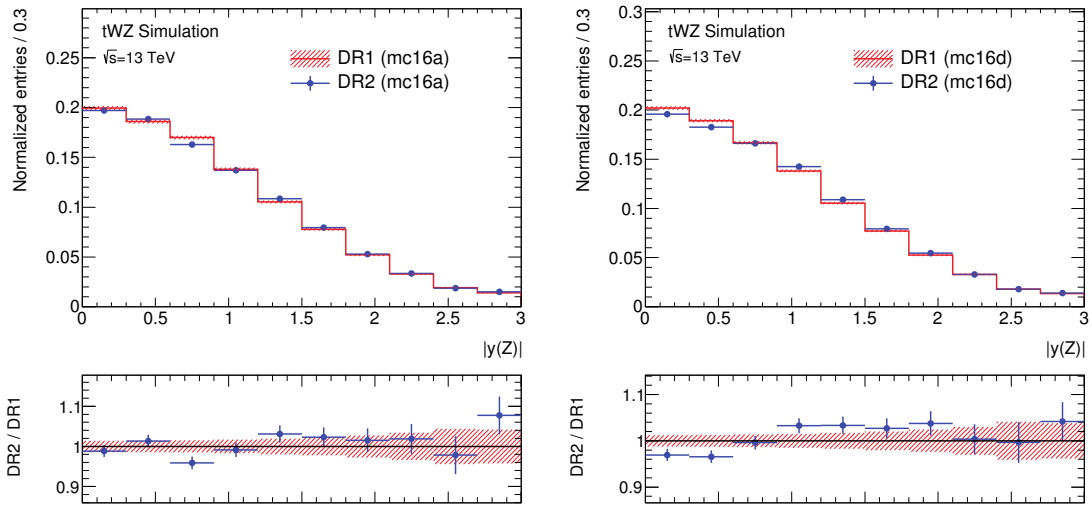


(c)

Fig. F.2.: Normalised distributions of the leading lepton p_{T} for mc16a, mc16d and mc16e in Figs. (a), (b) and (c), respectively, shown for the 3ℓ differential signal region. The red line corresponds to the DR1 sample, and the green line to the DR2 tWZ sample. The error bars in all panels include only MC statistical uncertainties.

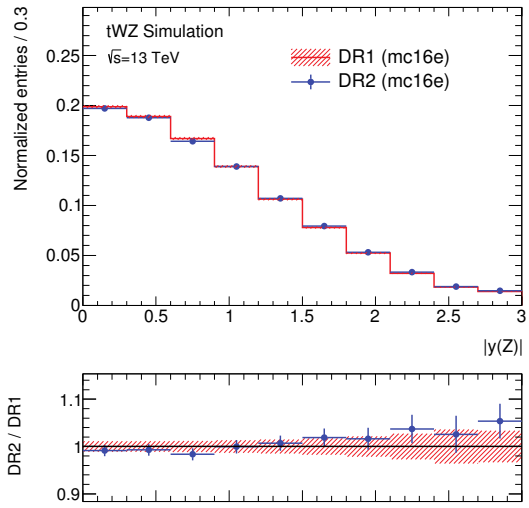
F.2. Differential variables at parton-level

In addition to Fig. 6.3 in Sec. 6.2.2, more normalised distributions of the differential variables used for the differential cross section measurement in the $t\bar{t}Z$ 3ℓ channel are shown in Figs. F.3 and F.4, respectively. All variables were built from parton-level information of the Monte Carlo generators and are presented for both diagram removal schemes.



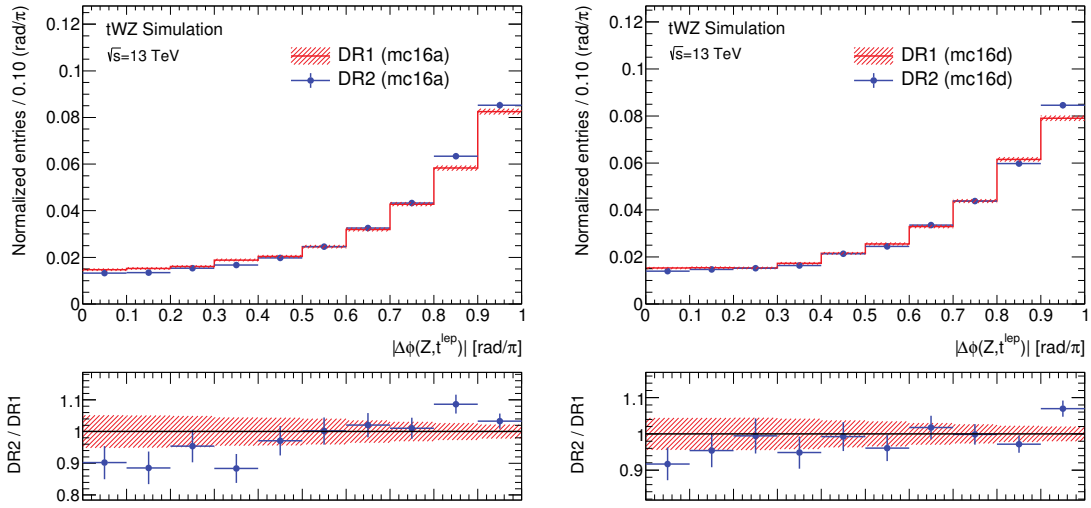
(a)

(b)



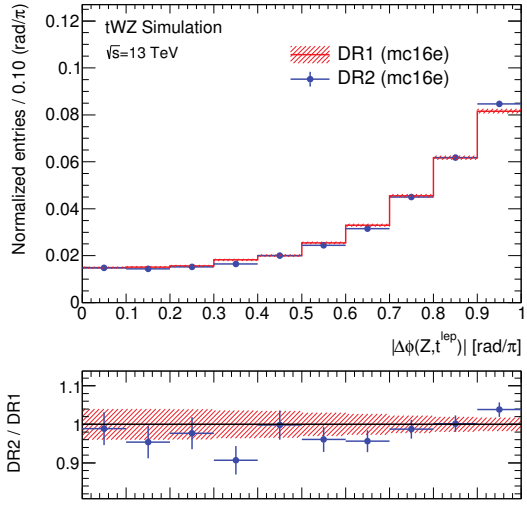
(c)

Fig. F.3.: Normalised distributions of the Z boson rapidity at parton level. The various panels (a), (b) and (c) corresponds to the mc16a, mc16d and mc16e campaigns, respectively. All distributions for DR1 in red and DR2 in blue are normalised to their respective number of entries. The ratio pads below each plot show the ratio of DR2 to DR1. The error bars only statistical MC uncertainties.



(a)

(b)



(c)

Fig. F.4.: Normalised distributions of the azimuthal difference between the Z boson and the top quark with a subsequent leptonic W boson decay at parton level. The various panels (a), (b) and (c) corresponds to the mc16a, mc16d and mc16e campaigns, respectively. All distributions for DR1 in red and DR2 in blue are normalised to their respective number of entries. The ratio pads below each plot show the ratio of DR2 to DR1. The error bars only statistical MC uncertainties.

F.3. Data-MC modelling in the targeted analysis regions

In addition to those shown in Fig. 7.1, several kinematic variables were investigated in order to check the modelling within the 2ℓOS target regions.

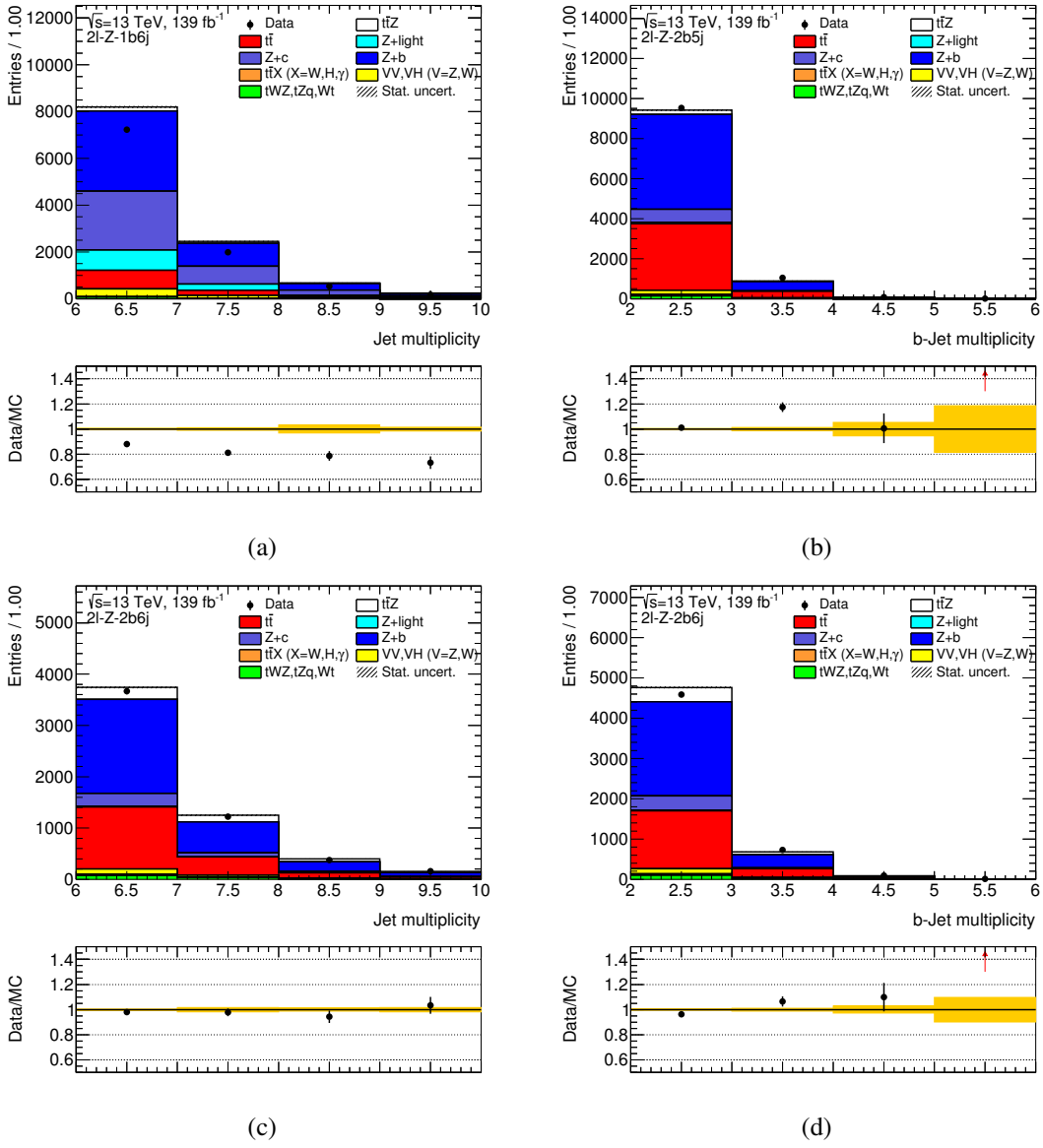


Fig. F.5.: Distributions of the jet multiplicity shown in Fig. (a) for the 2ℓ-Z-1b6j region and in Fig. (c) for the 2ℓ-Z-2b6j region. The b-jet multiplicity is presented in Figs. (b) and (d), depicting the 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

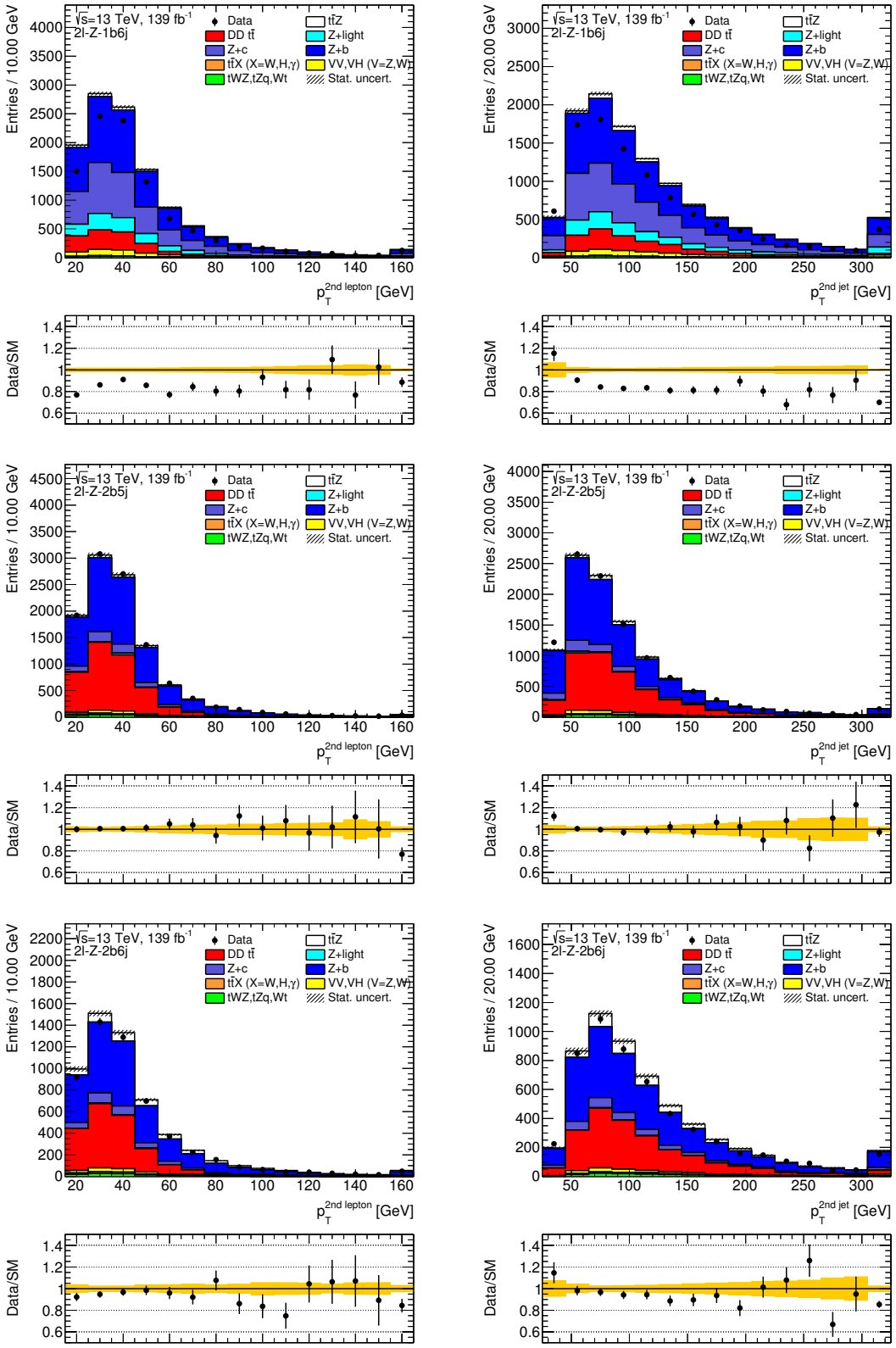


Fig. F.6.: Distributions of the subleading lepton p_{T} in the left column and the subleading jet p_{T} in the right column, shown for the three target regions. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of 139 fb^{-1}. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

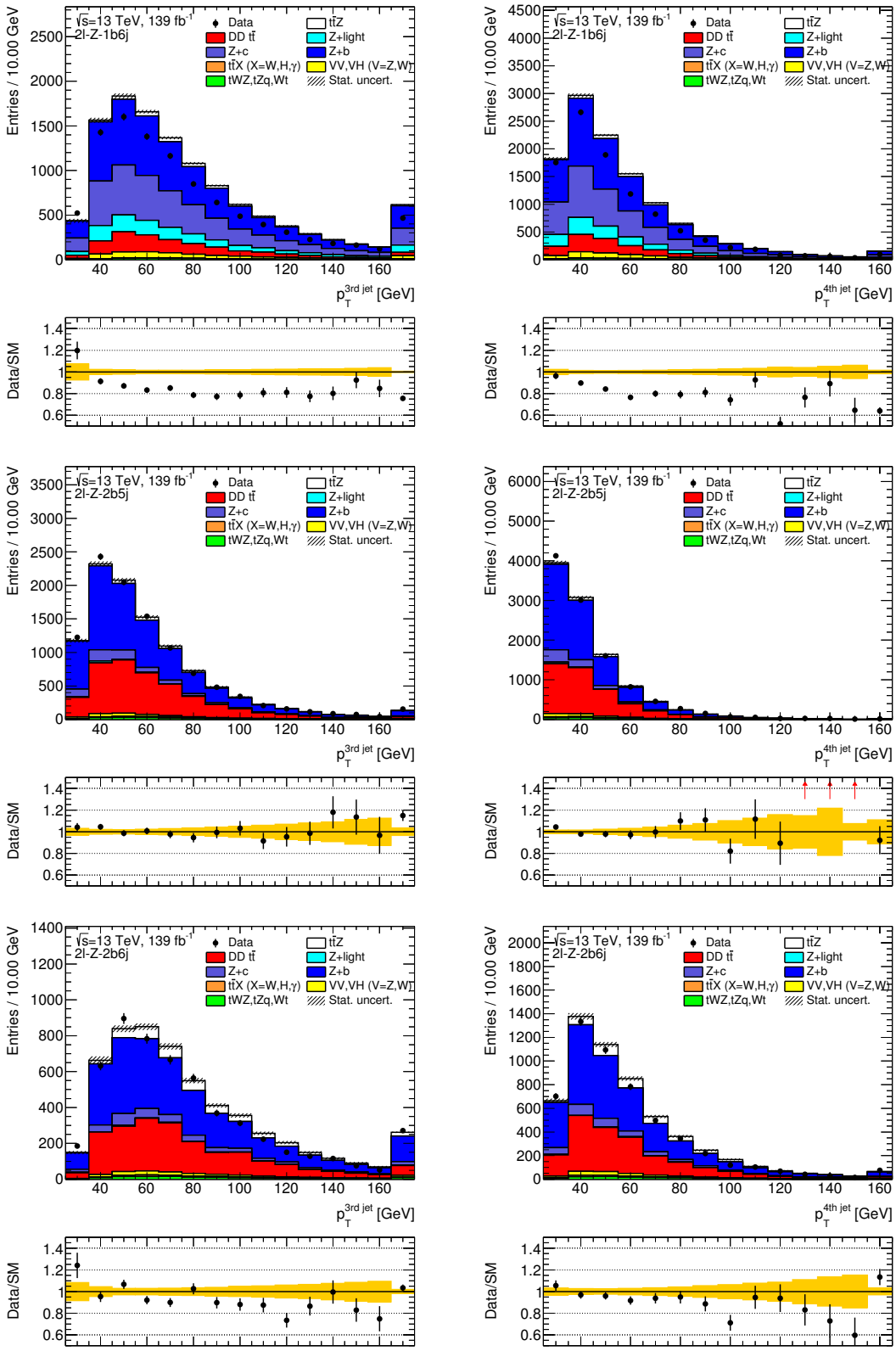


Fig. F.7.: Distributions of the p_{T} of the third jet in the left column and the p_{T} of the fourth jet in the right column, shown for the three target regions. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

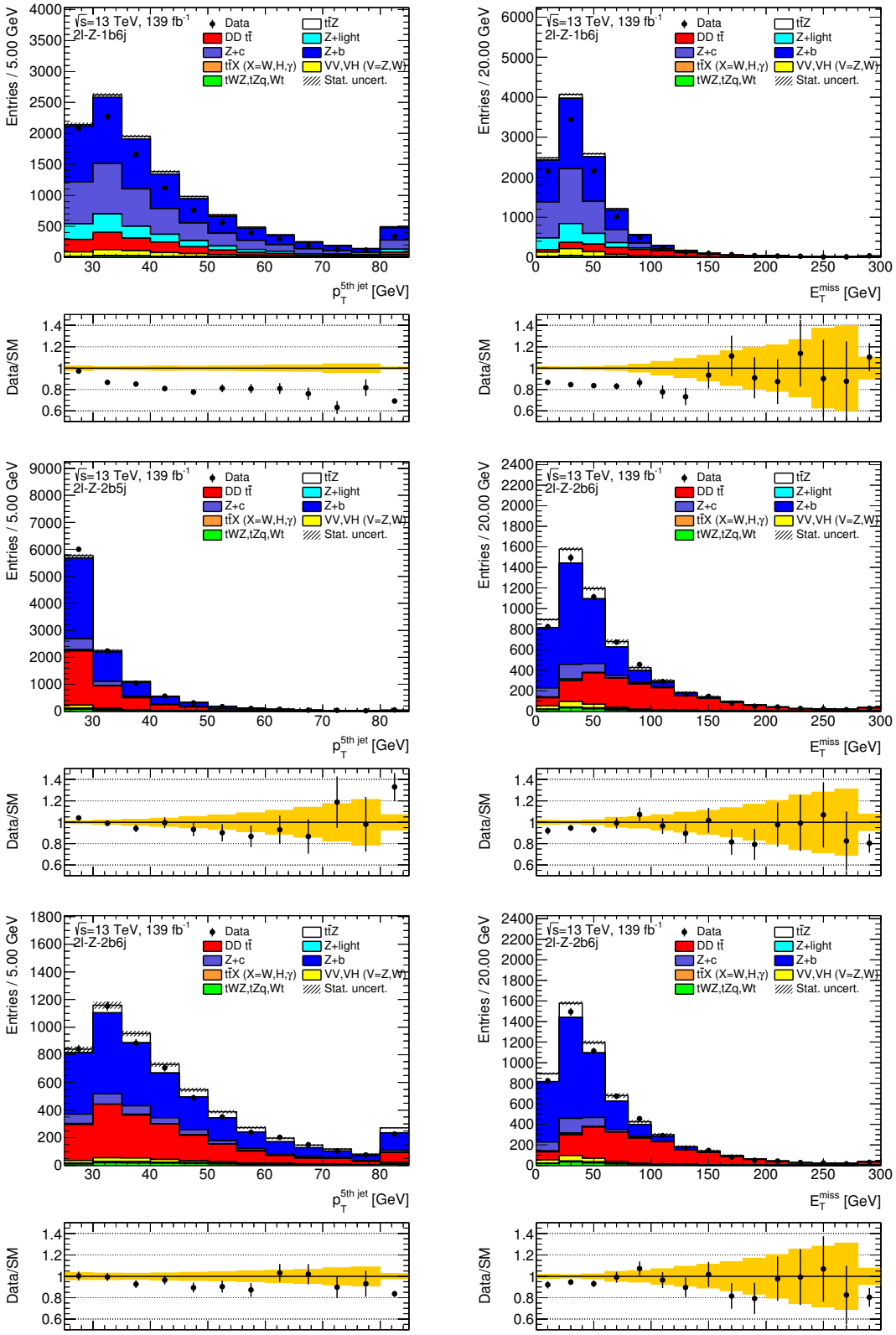


Fig. F.8.: Distributions of the jet p_{T} of the fifth jet in the left column and of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the right column, shown for the three target regions. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

F.4. Data-MC modelling in the Z+light control regions

In addition to those shown in Fig. 7.5, a few more kinematic variables were investigated in order to check the modelling within the Z+light flavour control regions.

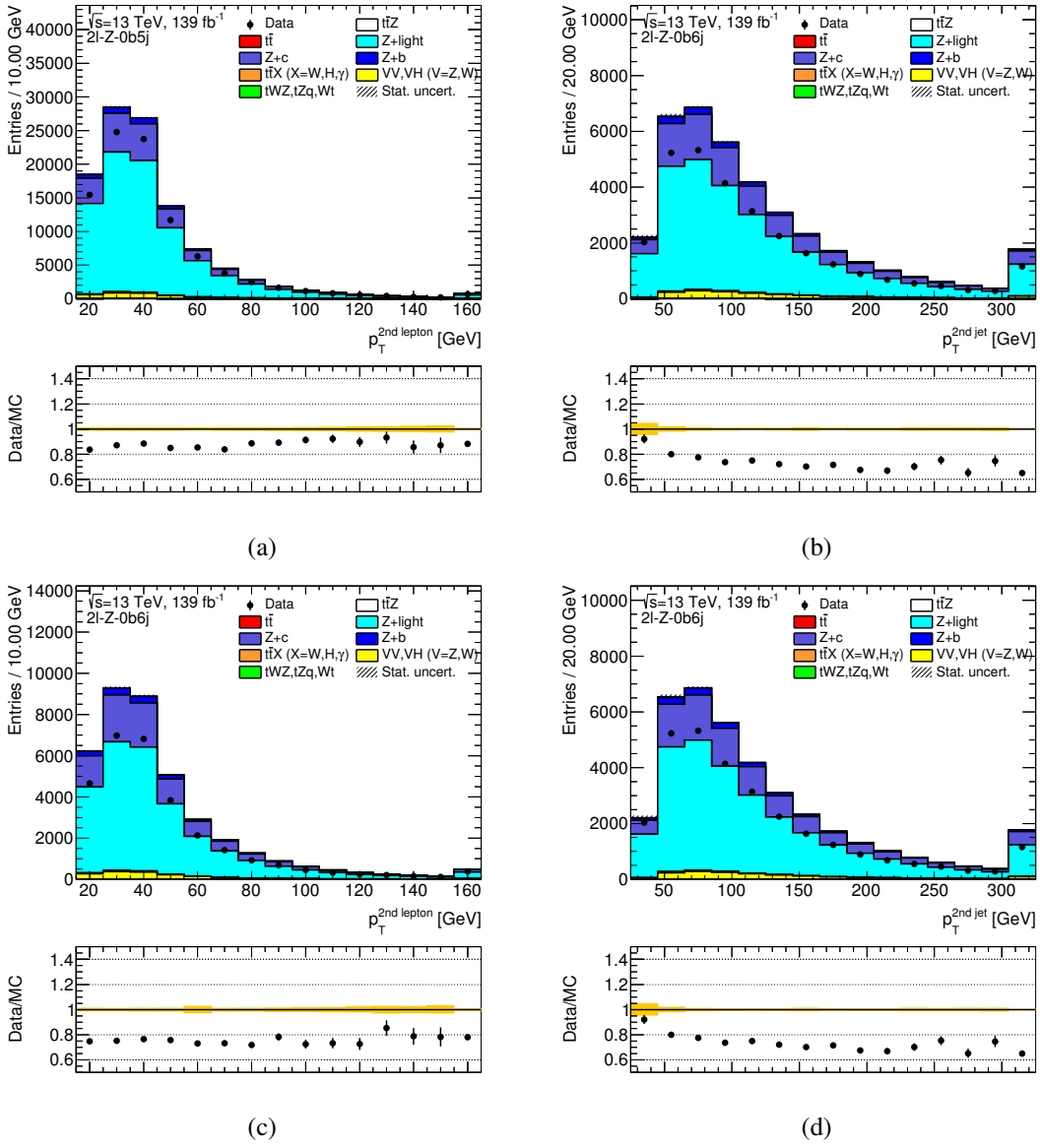
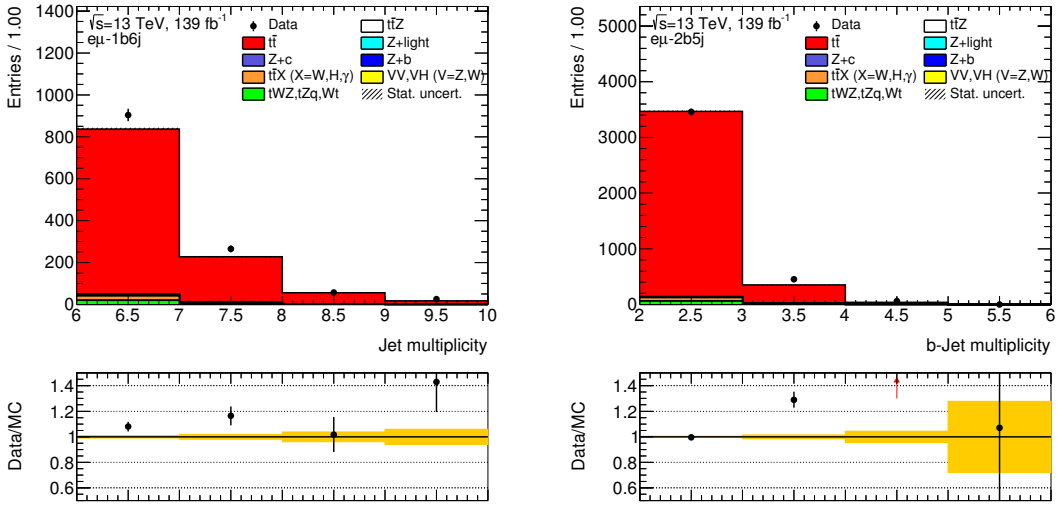


Fig. F.9.: Distributions of the subleading lepton p_{T} shown in Fig. (a) for the 2ℓ-Z-0b5j region and in Fig. (c) for the 2ℓ-Z-0b6j region. The subleading jet p_{T} is presented in Figs. (b) and (d), depicting the 2ℓ-Z-0b5j and 2ℓ-Z-0b6j target region, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

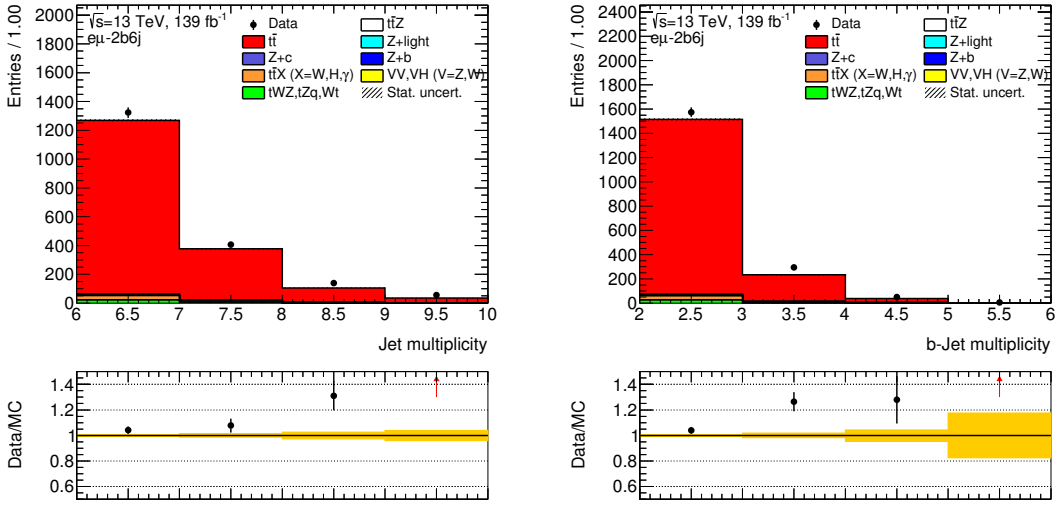
F.5. Data-MC modelling in the $t\bar{t}$ validation regions

In addition to those shown in Fig. 7.3, several kinematic variables were investigated in order to check the modelling within the $t\bar{t}$ validation regions.



(a)

(b)



(c)

(d)

Fig. F.10.: Distributions of the jet multiplicity shown in Fig. (a) for the $e\mu$-1b6j region and in Fig. (c) for the $e\mu$-2b6j region. The b-jet multiplicity is presented in Figs. (b) and (d), depicting the $e\mu$-2b5j and $e\mu$-2b6j validation region, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

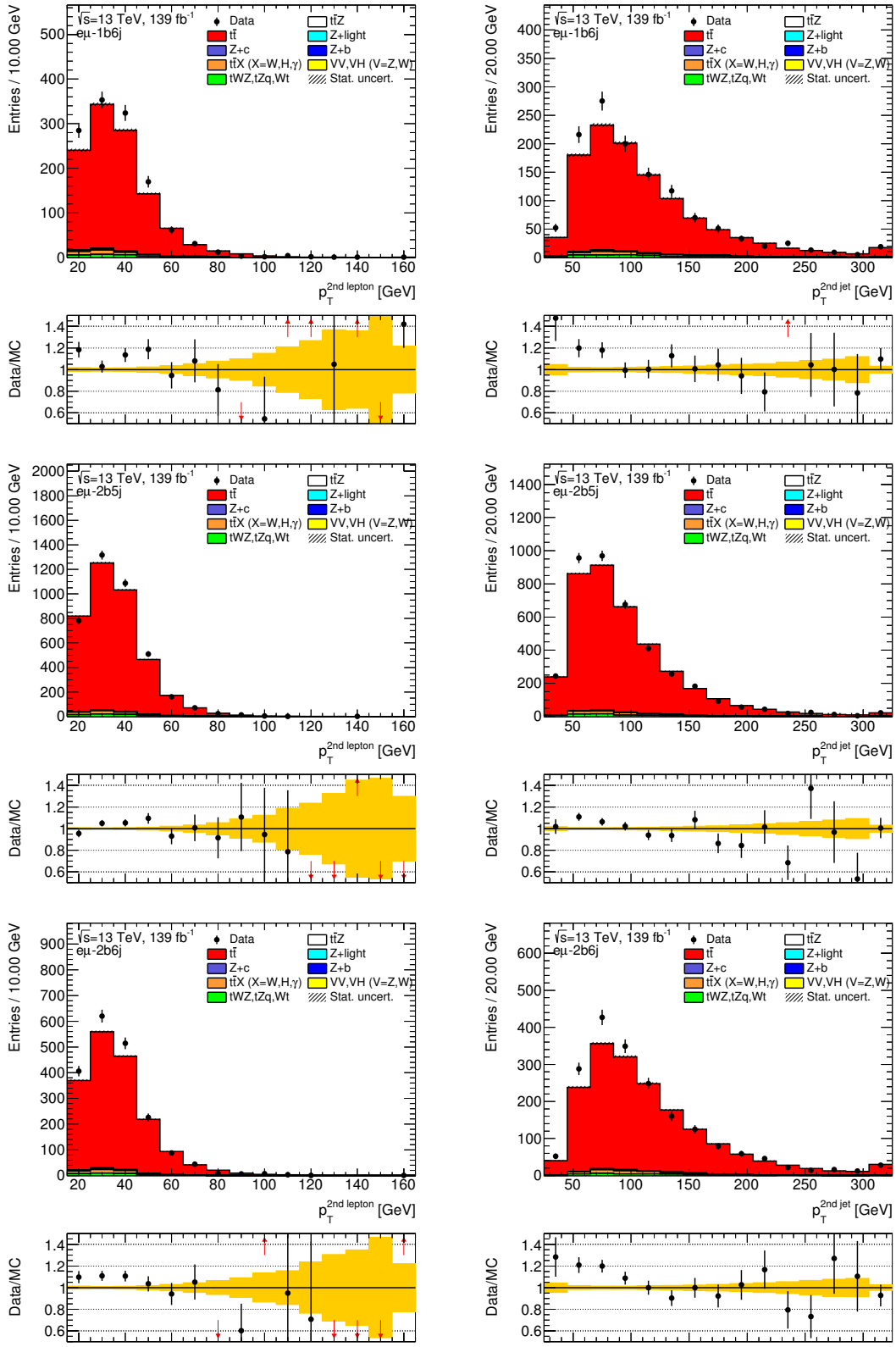


Fig. F.11.: Distributions of the subleading lepton p_{T} in the left column and the subleading jet p_{T} in the right column, shown for the three $t\bar{t}$ validation regions. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of 139 fb^{-1}. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

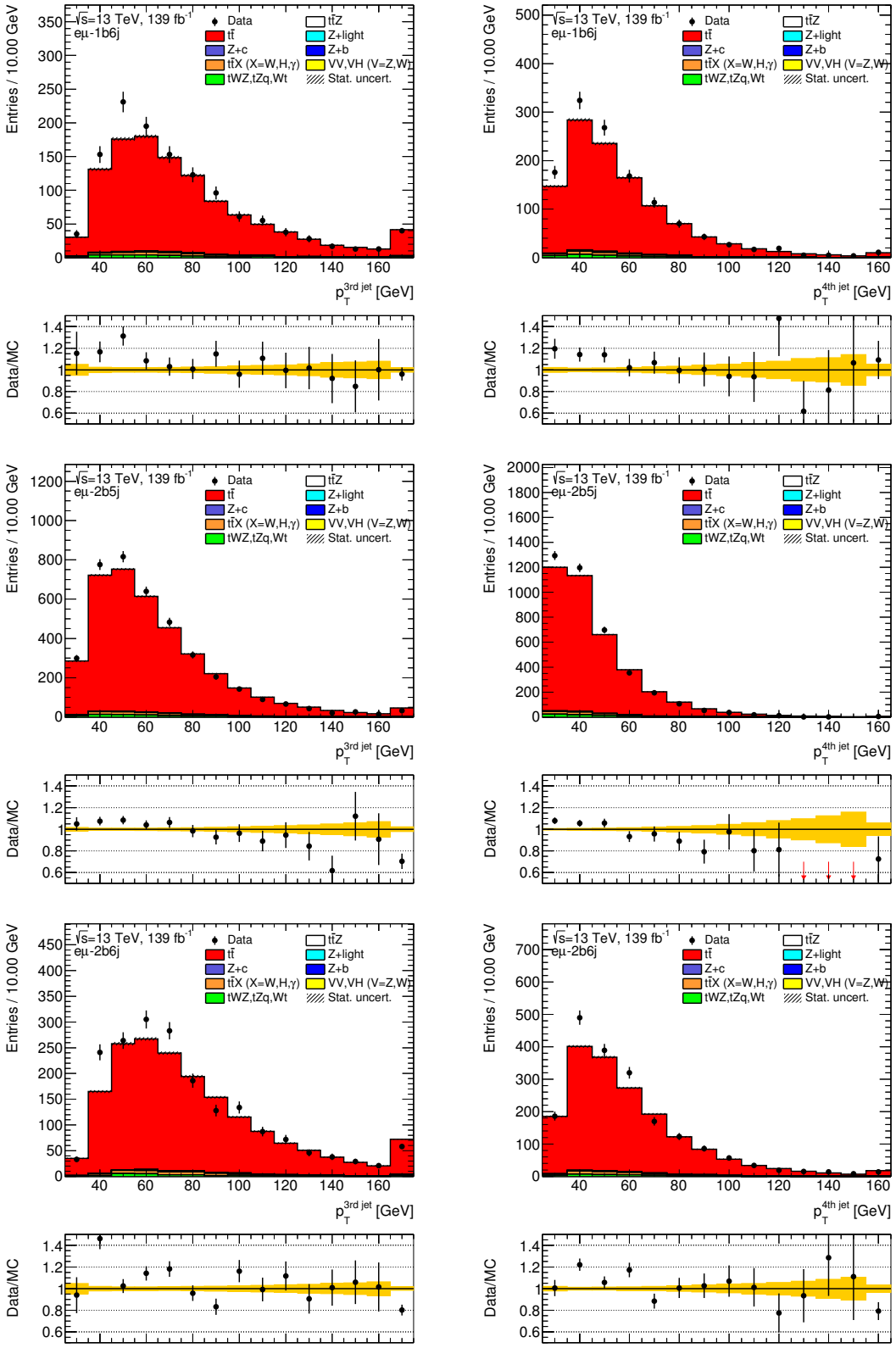


Fig. F.12.: Distributions of the p_{T} of the third jet in the left column and the p_{T} of the fourth jet in the right column, shown for the three $t\bar{t}$ validation regions. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

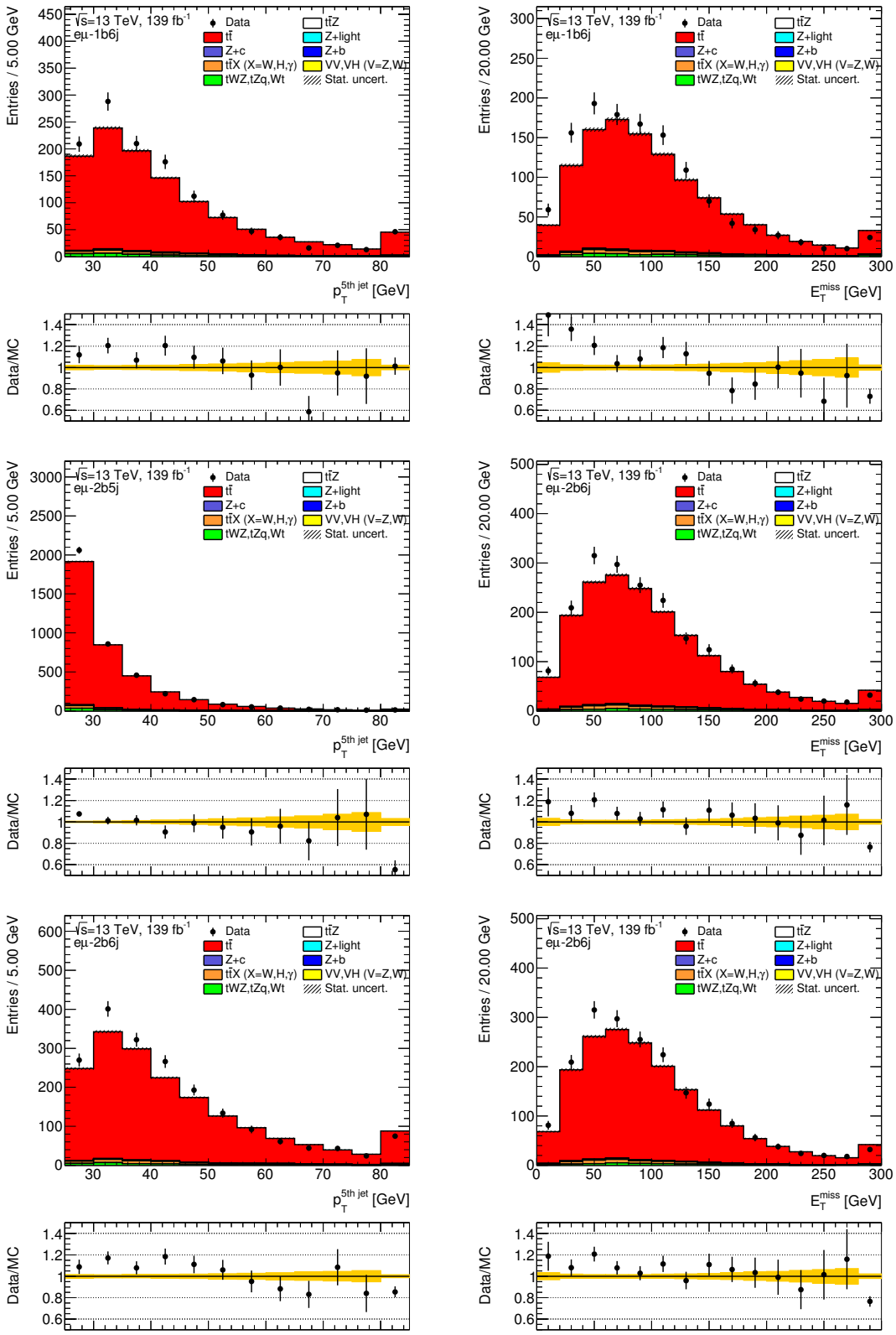


Fig. F.13.: Distributions of the jet p_{T} of the fifth jet in the left column and of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the right column, shown for the three $t\bar{t}$ validation regions. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

F.6. Comparison of $t\bar{t}$ selected with a OF and SF lepton criterion

In addition to those shown in Fig. 7.4, several kinematic variables were investigated in order to check the modelling within the $t\bar{t}$ validation regions.

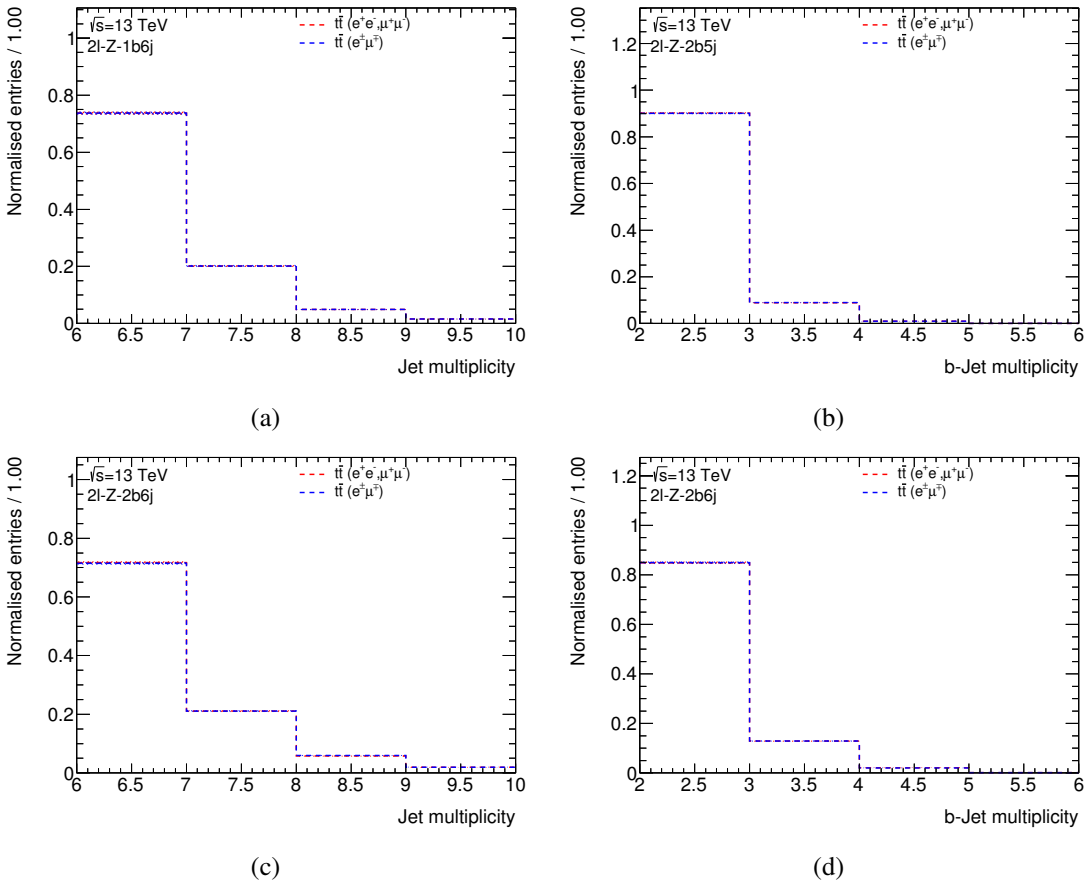


Fig. F.14.: Distributions of the jet multiplicity shown in Fig. (a) for the $e\mu$-1b6j region and in Fig. (c) for the $e\mu$-2b6j region. The b-jet multiplicity is presented in Figs. (b) and (d), depicting the $e\mu$-2b5j and $e\mu$-2b6j validation region, respectively. The $t\bar{t}$ events selected with a SF-criterion are drawn red whereas the $t\bar{t}$ events selected with an OF-criterion are depicted by a blue line. All distributions were normalised to their respective number of entries. The errors indicated by the striped bands represent the statistical uncertainties from Monte Carlo simulation.

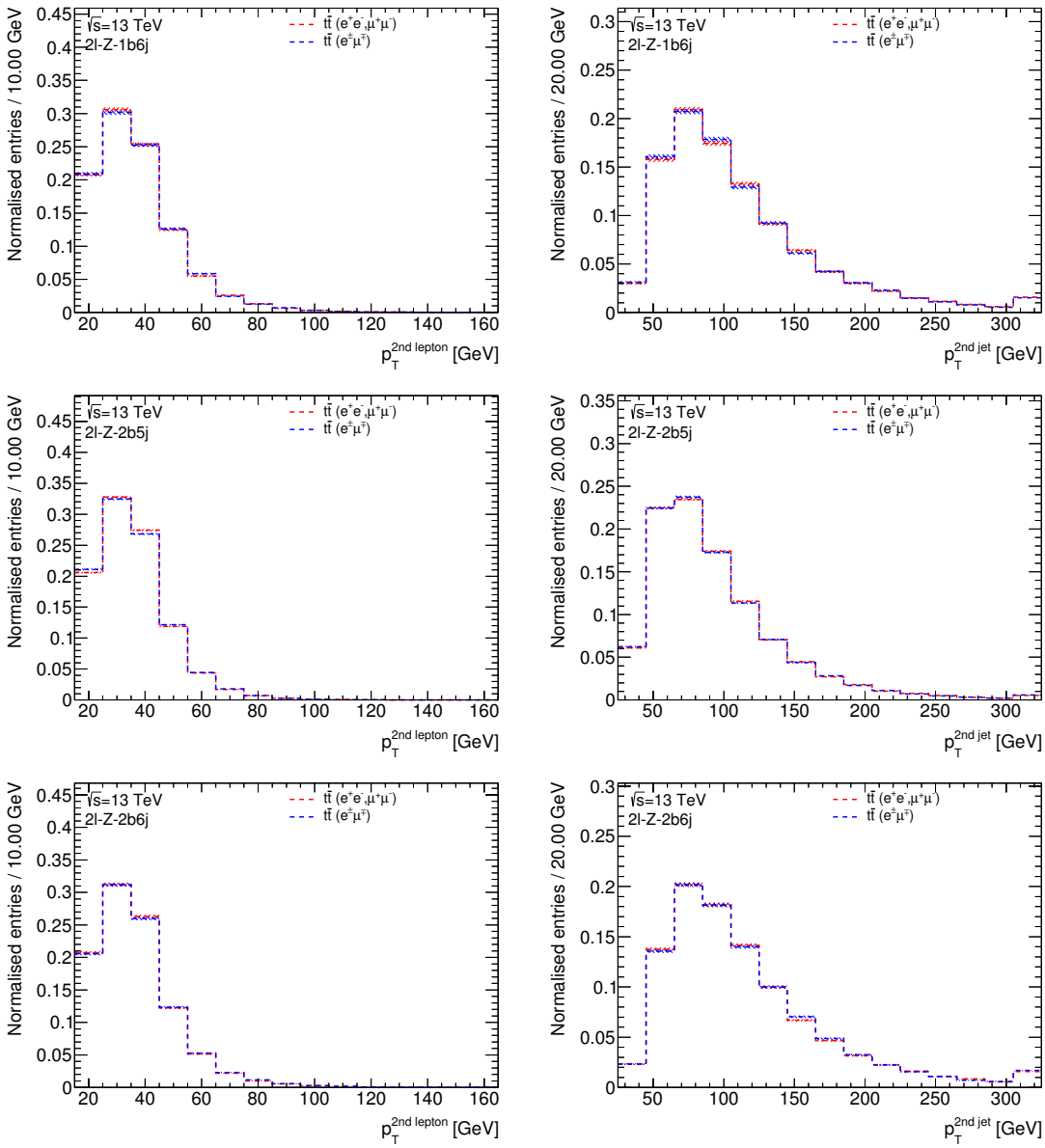


Fig. F.15.: Distributions of the subleading lepton p_{T} in the left column and the subleading jet p_{T} in the right column, shown for $t\bar{t}$ events selected with a SF-criterion in red and for $t\bar{t}$ events selected with an OF-criterion in blue. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. All distributions were normalised to their respective number of entries. The errors indicated by the striped bands represent the statistical uncertainties from Monte Carlo simulation.

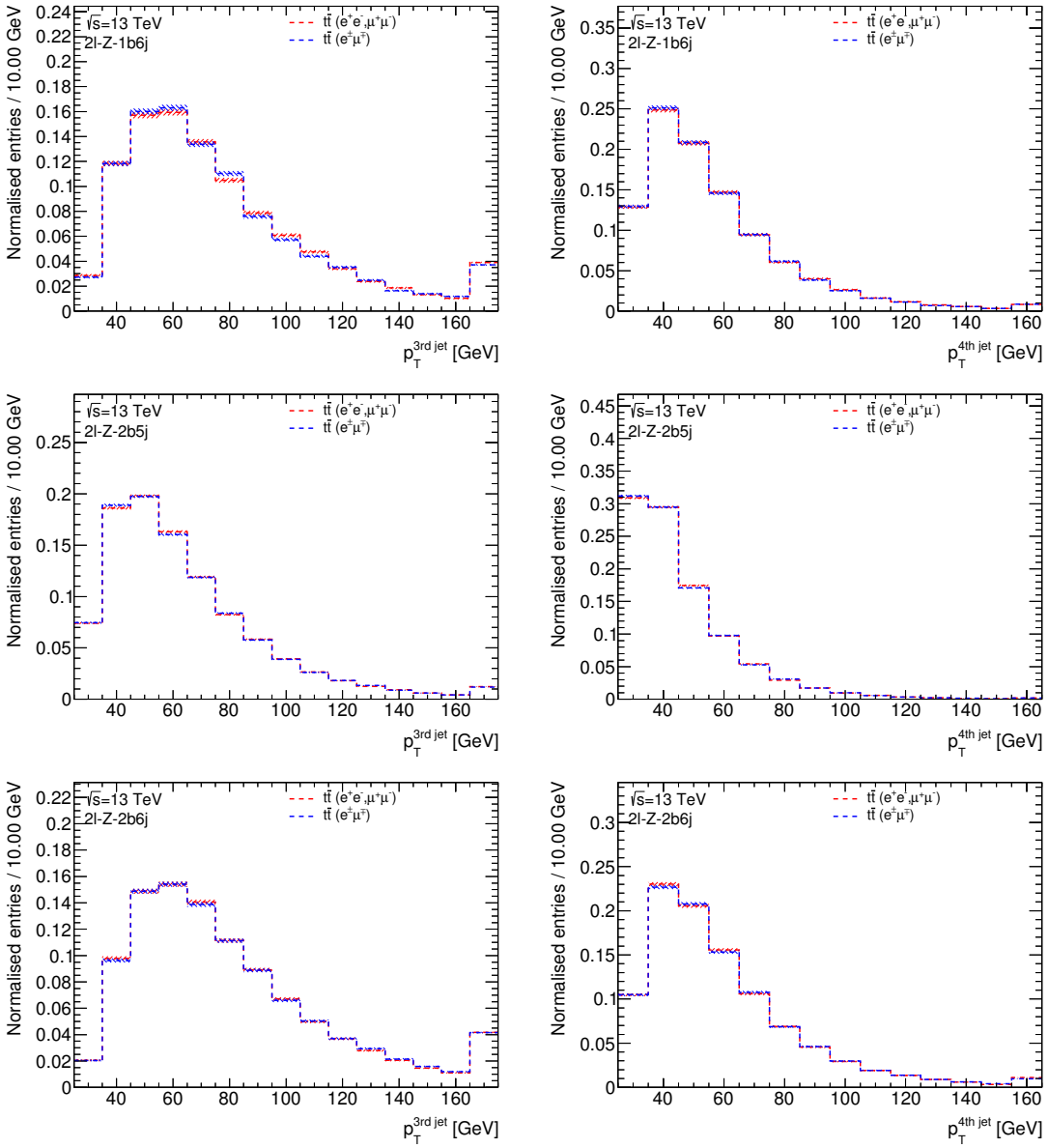


Fig. F.16.: Distributions of the p_{T} of the third jet in the left column and the p_{T} of the fourth jet in the right column,shown for $t\bar{t}$ events selected with a SF-criterion in red and for $t\bar{t}$ events selected with an OF-criterion in blue. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. All distributions were normalised to their respective number of entries. The errors indicated by the striped bands represent the statistical uncertainties from Monte Carlo simulation.

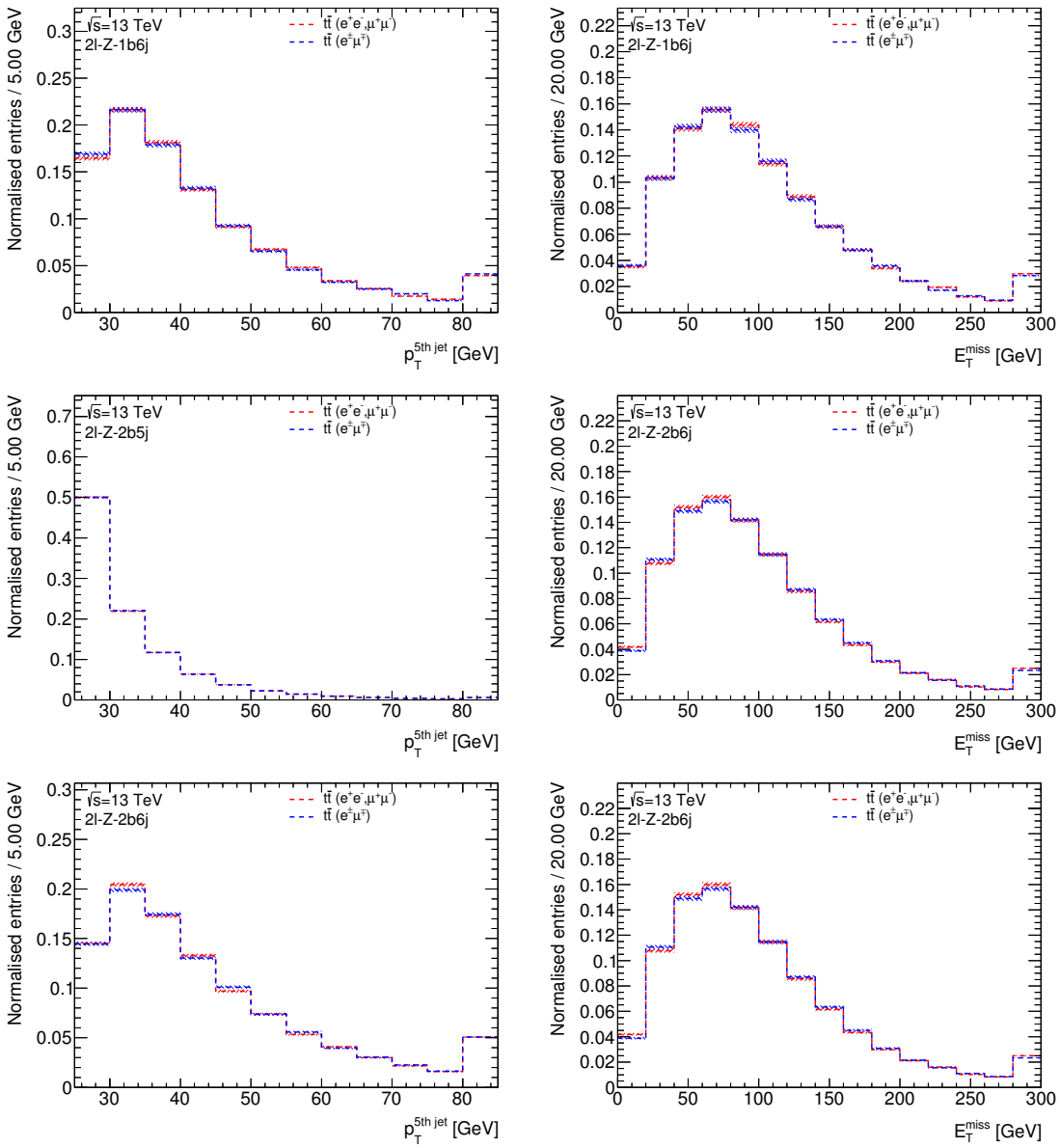
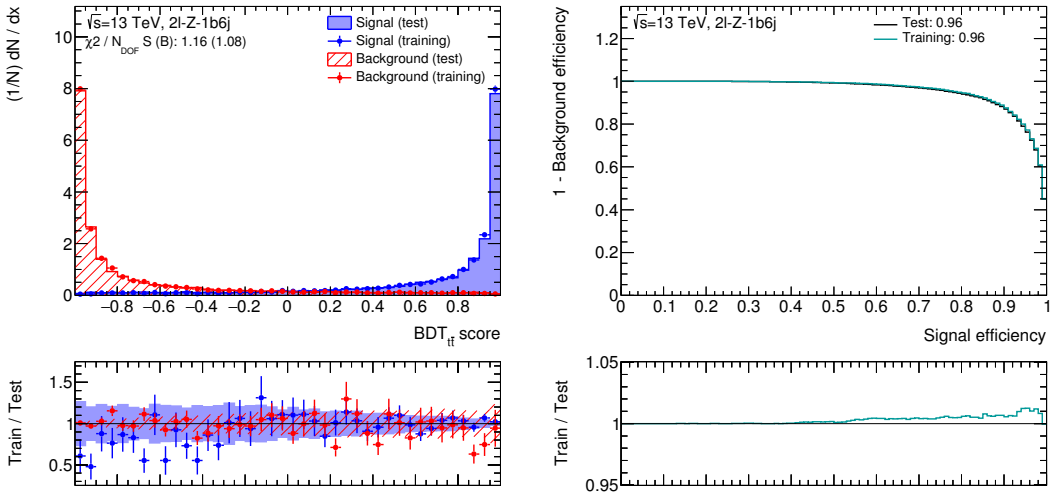


Fig. F.17.: Distributions of the jet p_{T} of the fifth jet in the left column and of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the right column, shown for $t\bar{t}$ events selected with a SF-criterion in red and for $t\bar{t}$ events selected with an OF-criterion in blue. The $e\mu$-1b6j, $e\mu$-2b5j and $e\mu$-2b6j regions are depicted in the top, middle and bottom row, respectively. All distributions were normalised to their respective number of entries. The errors indicated by the striped bands represent the statistical uncertainties from Monte Carlo simulation.

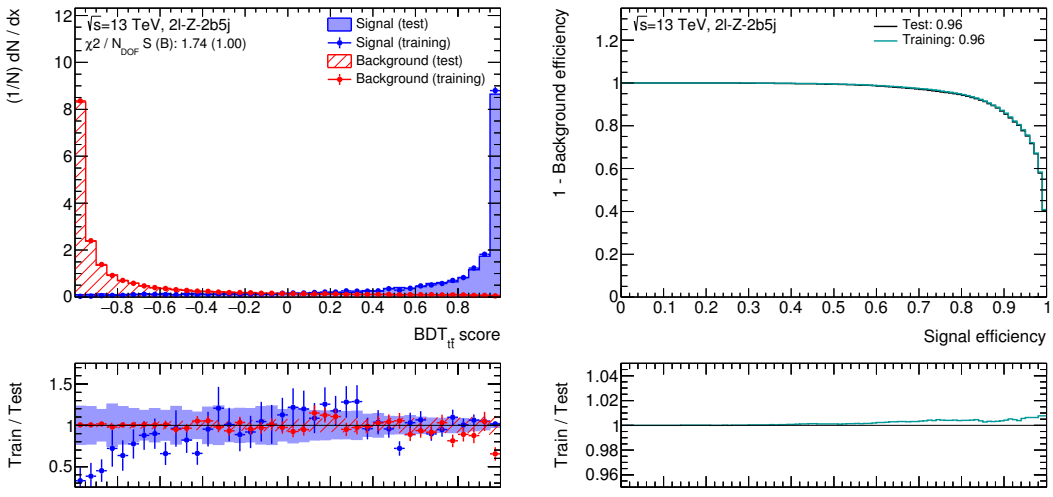
F.7. BDT training results

The BDT training results within the 2ℓ-Z-1b6j and the 2ℓ-Z-2b5j target regions are shown in this section, whereas the corresponding figures for the 2ℓ-Z-2b6j region were shown in Sec. 7.2.3.1.



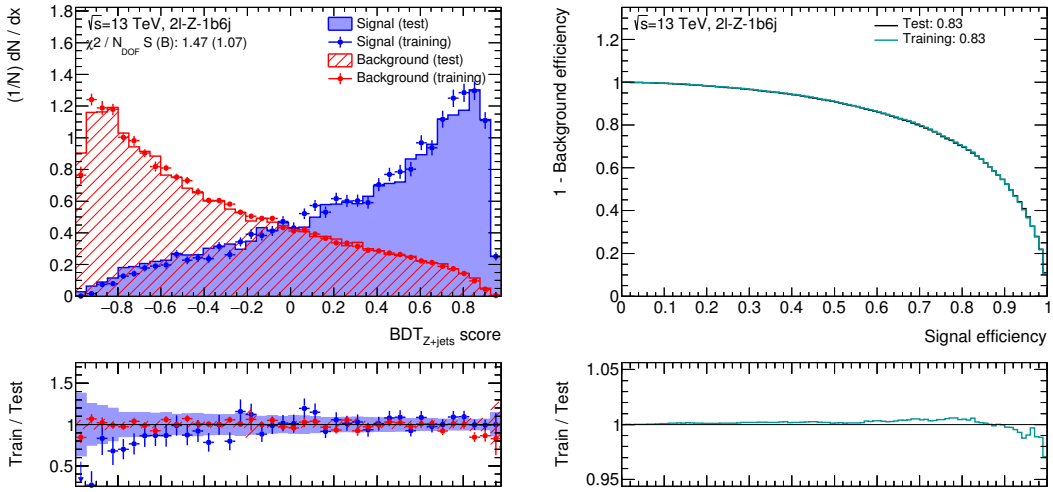
(a) (b)

Fig. F.18.: Training results of the training against the $t\bar{t}$ background within the 2ℓ-Z-1b6j target region. In Fig. (a) the normalised distributions of the BDT score are shown in blue for the signal and in red for the $t\bar{t}$ background. On the right, Fig. (b) shows the corresponding ROC curve for both training and test datasets.



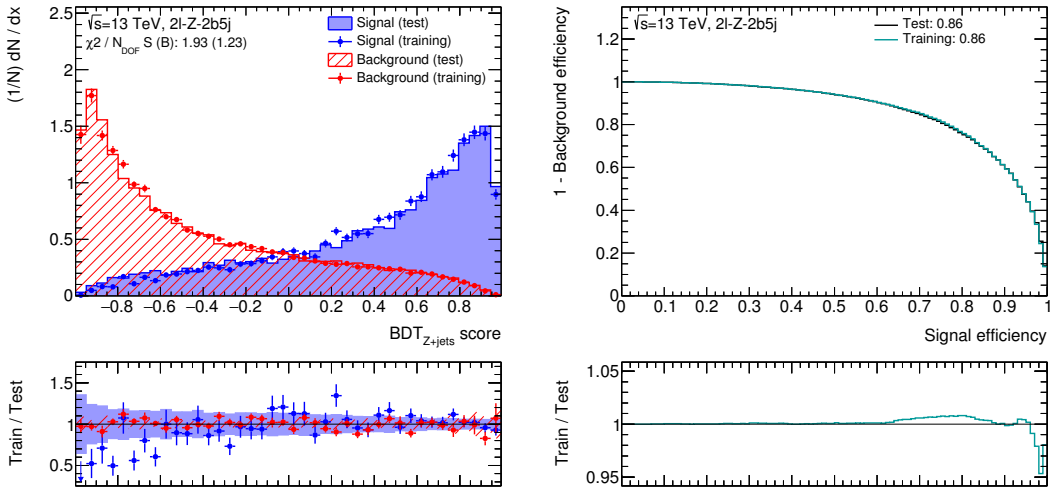
(a) (b)

Fig. F.19.: Training results of the training against the $t\bar{t}$ background within the 2ℓ-Z-2b5j target region. In Fig. (a) the normalised distributions of the BDT score are shown in blue for the signal and in red for the $t\bar{t}$ background. On the right, Fig. (b) shows the corresponding ROC curve for both training and test datasets.



(a) (b)

Fig. F.20.: Training results of the training against the Z+jets background within the 2ℓ-Z-1b6j target region. In Fig. (a) the normalised distributions of the BDT score are shown in blue for the signal and in red for the Z+jets background. On the right, Fig. (b) shows the corresponding ROC curve for both training and test datasets.



(a) (b)

Fig. F.21.: Training results of the training against the Z+jets background within the 2ℓ-Z-2b5j target region. In Fig. (a) the normalised distributions of the BDT score are shown in blue for the signal and in red for the Z+jets background. On the right, Fig. (b) shows the corresponding ROC curve for both training and test datasets.

F.8. Linear correlations of the discriminating variables

In the following the linear correlations for the various BDT training procedures for the different background processes as well as target regions are shown. A comprehensive listing of the variables stated on the axes of the correlation figures can be found within Tab. G.5 given in Sec. G.3.

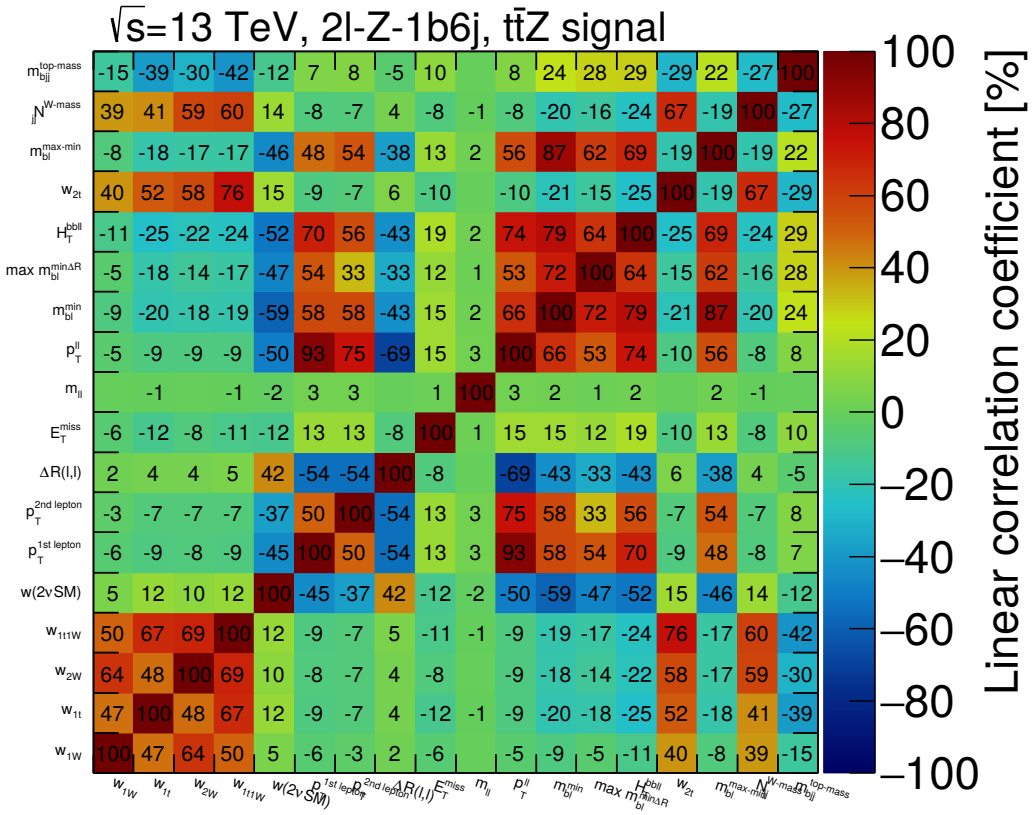


Fig. F.22.: Linear correlations of the input variables for the signal dataset employed for the BDT training against the $t\bar{t}$ background in the 2ℓ-Z-1b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

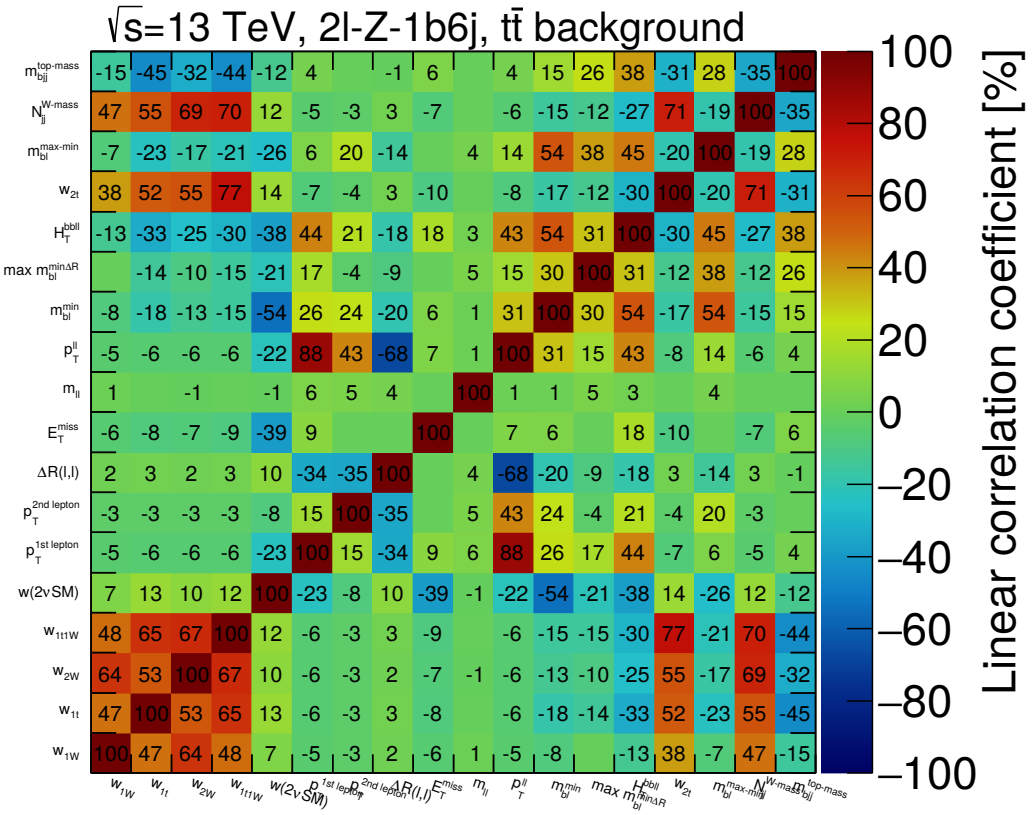


Fig. F.23.: Linear correlations of the input variables for the background dataset employed for the BDT training against the $t\bar{t}$ background in the 2ℓ-Z-1b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

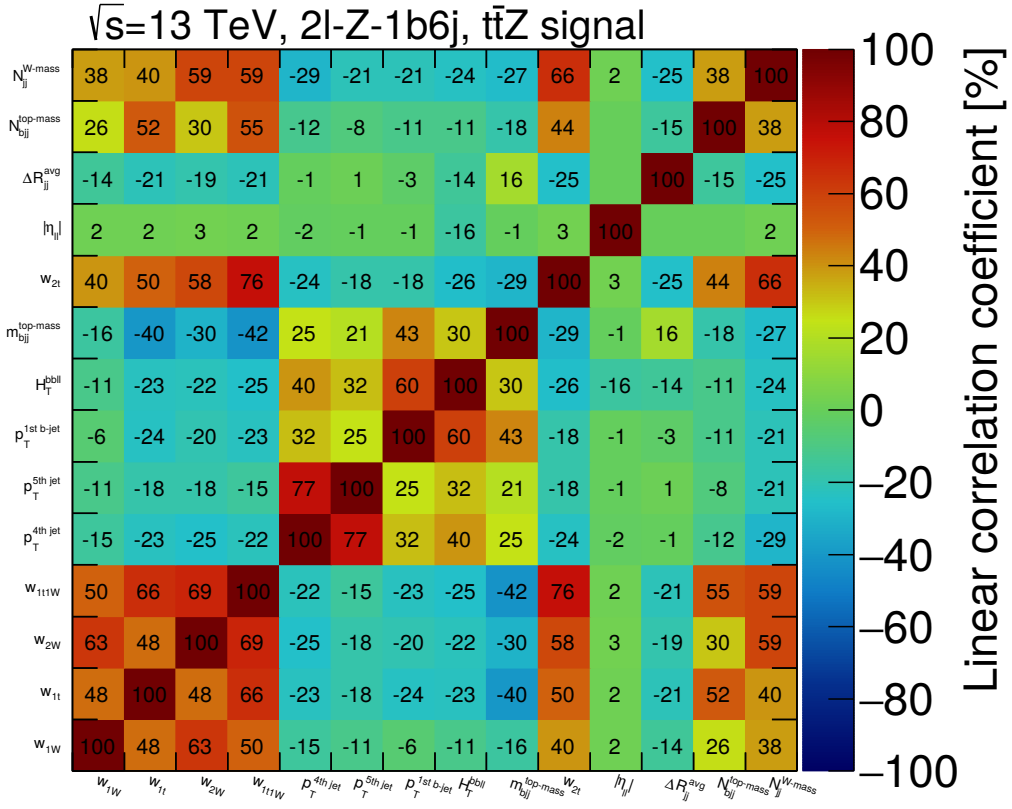


Fig. F.24.: Linear correlations of the input variables for the signal dataset employed for the BDT training against the Z+jets background in the 2ℓ-Z-1b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

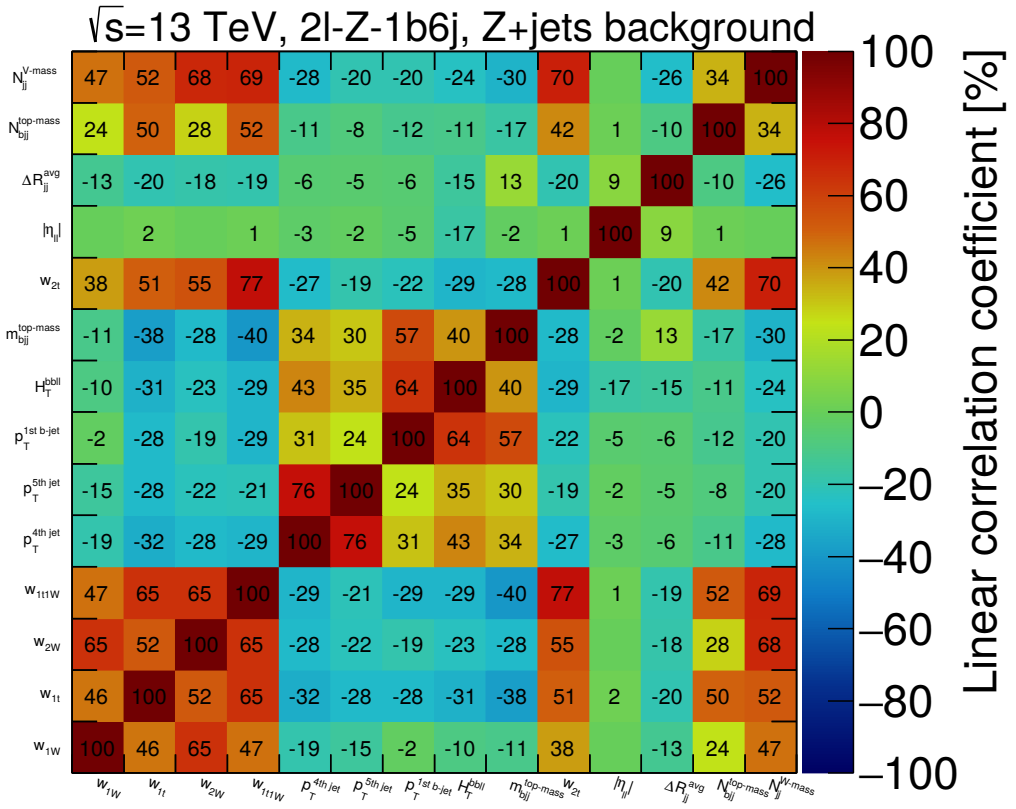


Fig. F.25.: Linear correlations of the input variables for the background dataset employed for the BDT training against the Z+jets background in the 2ℓ-Z-1b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

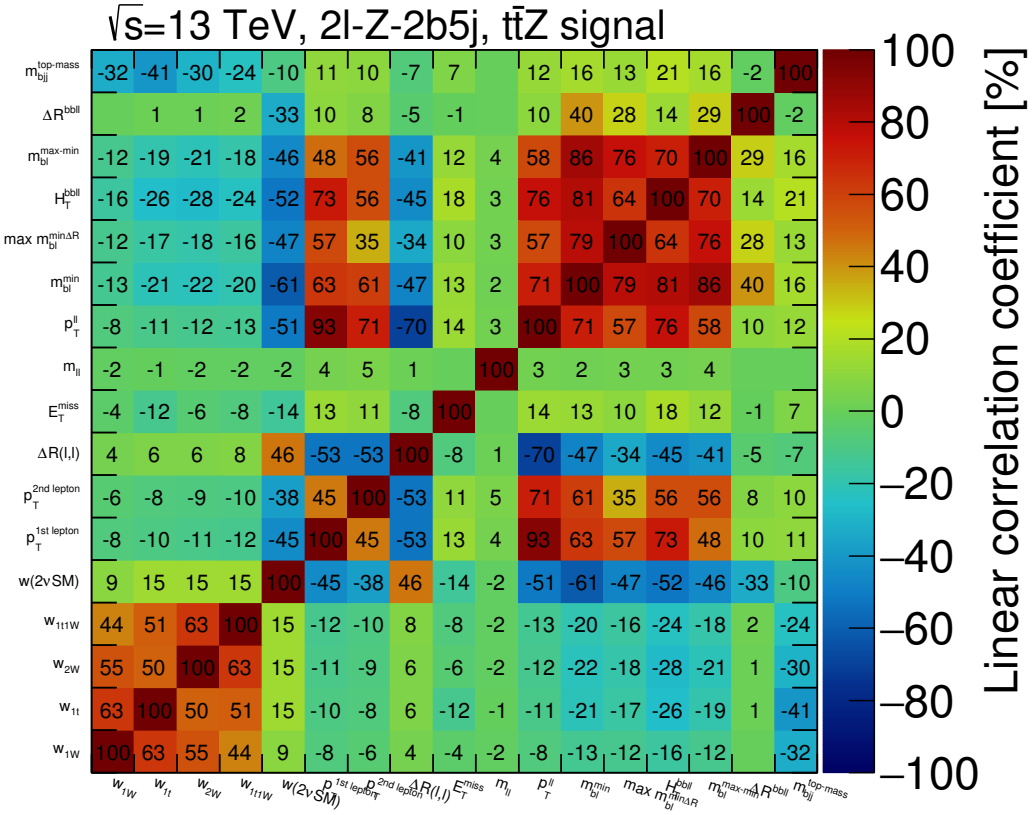


Fig. F.26.: Linear correlations of the input variables for the signal dataset employed for the BDT training against the $t\bar{t}$ background in the 2ℓ-Z-2b5j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

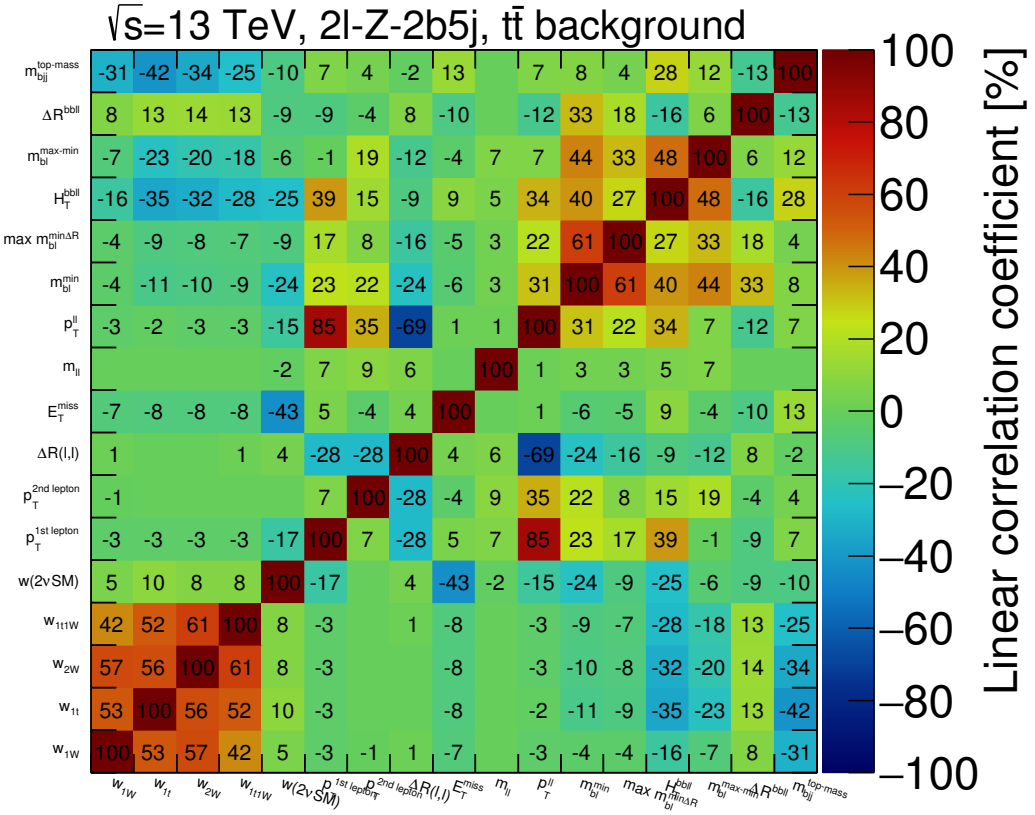


Fig. F.27.: Linear correlations of the input variables for the background dataset employed for the BDT training against the $t\bar{t}$ background in the 2ℓ-Z-2b5j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

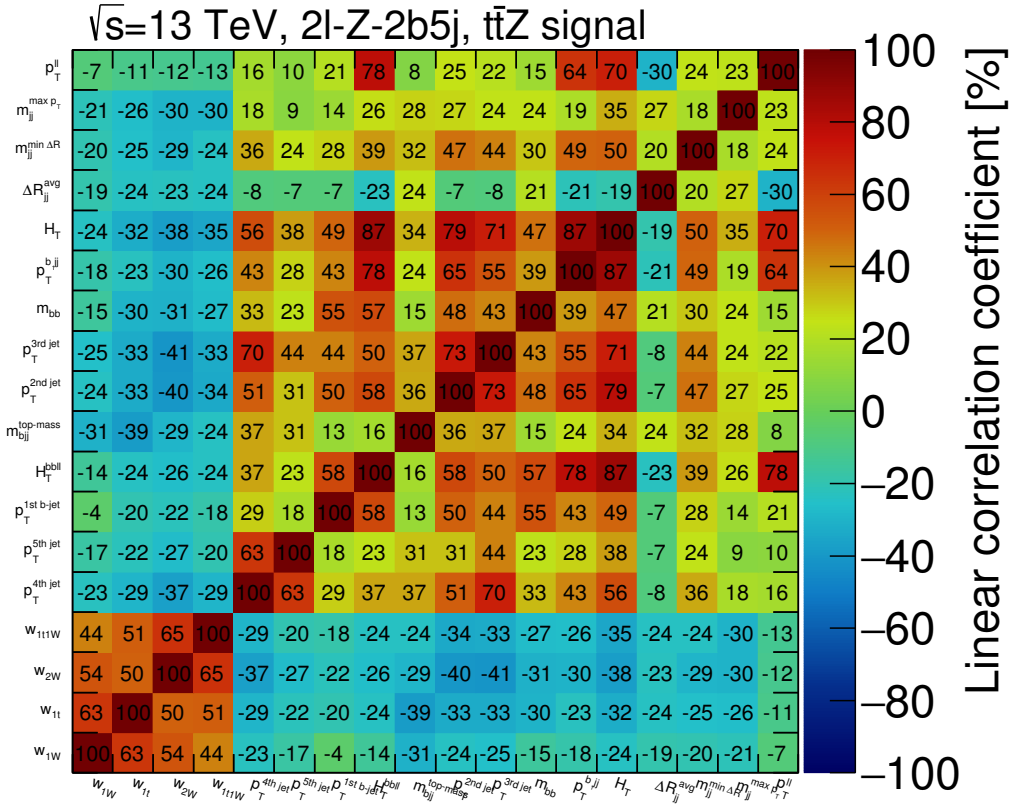


Fig. F.28.: Linear correlations of the input variables for the signal dataset employed for the BDT training against the Z+jets background in the 2ℓ-Z-2b5j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

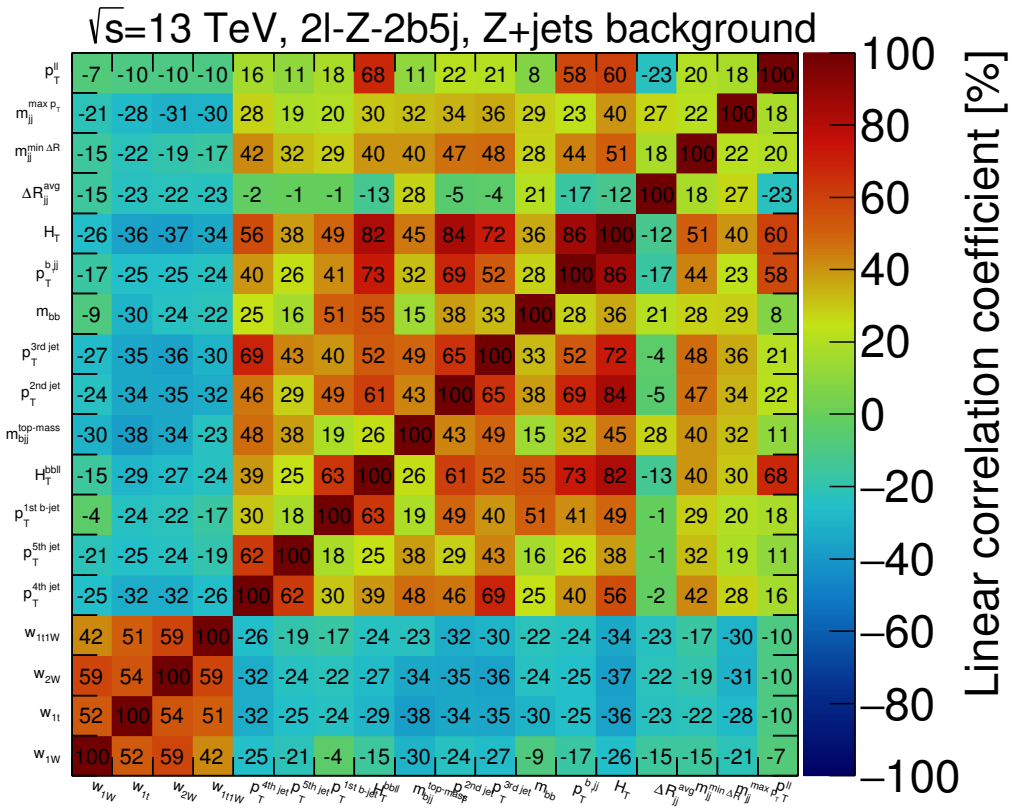


Fig. F.29.: Linear correlations of the input variables for the background dataset employed for the BDT training against the Z+jets background in the 2ℓ-Z-2b5j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

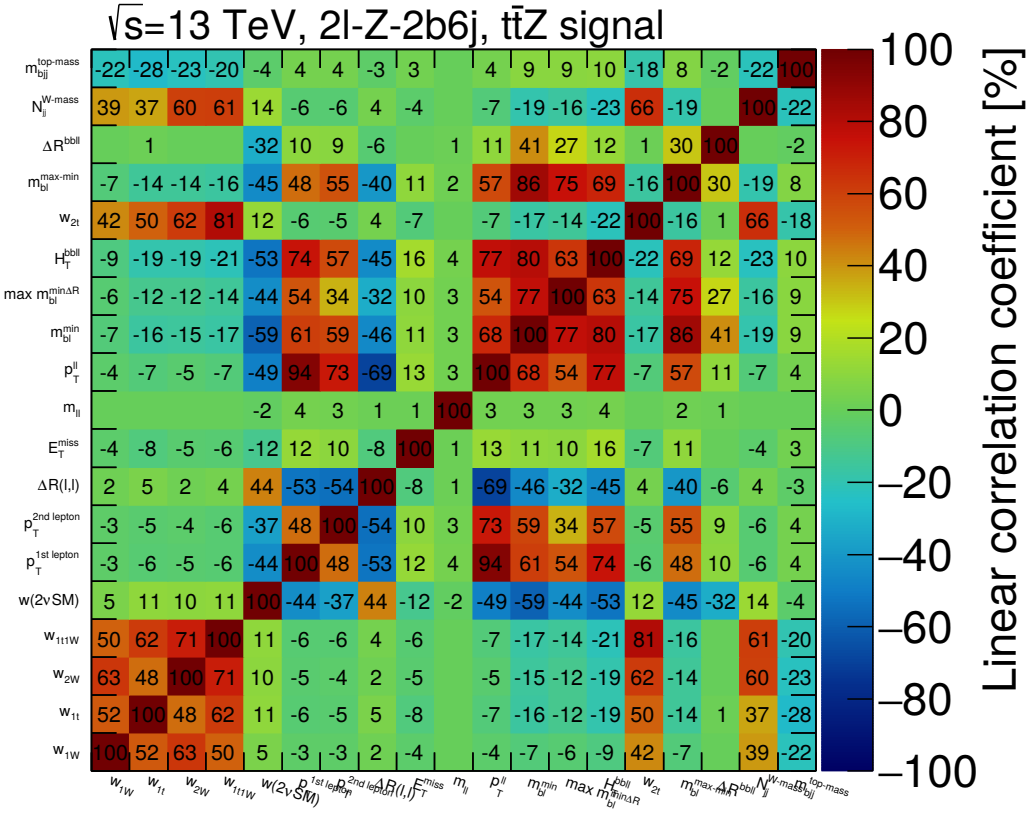


Fig. F.30.: Linear correlations of the input variables for the signal dataset employed for the BDT training against the $t\bar{t}$ background in the 2ℓ-Z-2b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

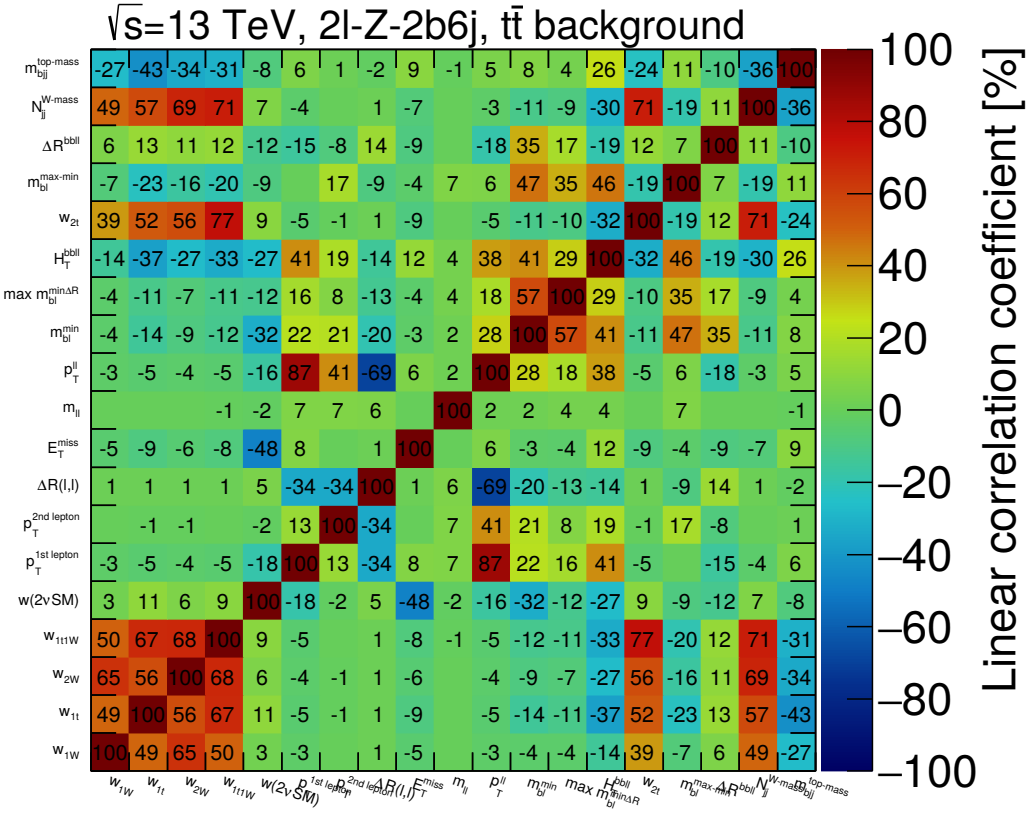


Fig. F.31.: Linear correlations of the input variables for the background dataset employed for the BDT training against the $t\bar{t}$ background in the 2ℓ-Z-2b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

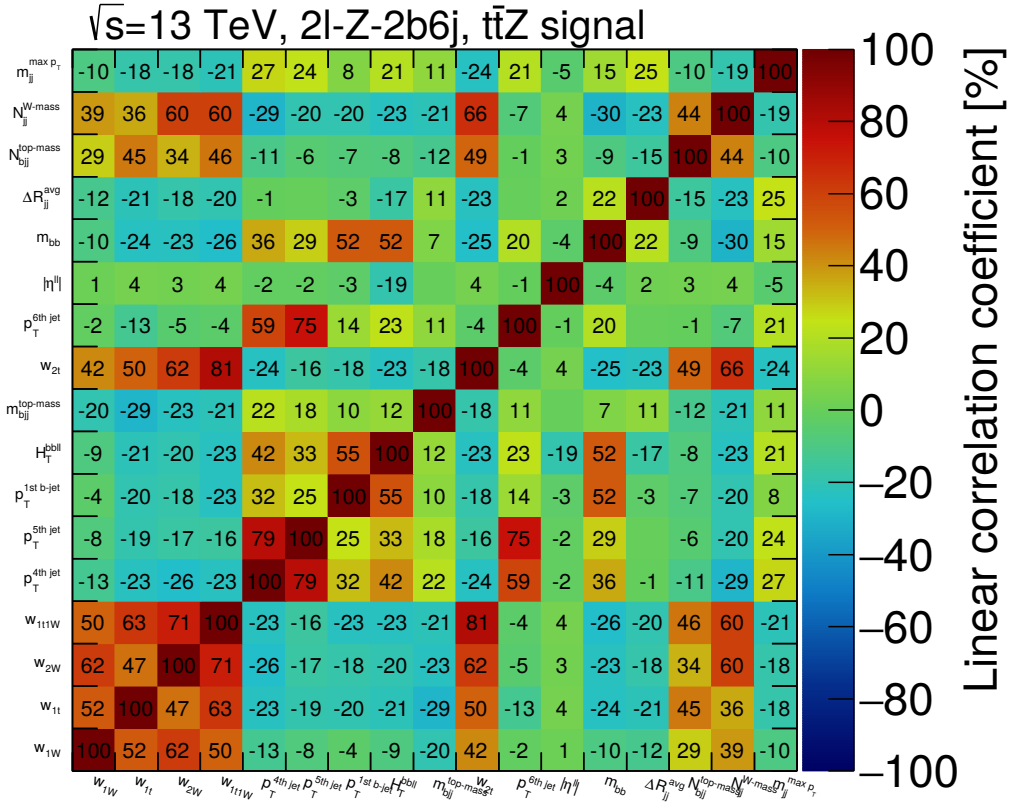


Fig. F.32.: Linear correlations of the input variables for the signal dataset employed for the BDT training against the Z+jets background in the 2ℓ-Z-2b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

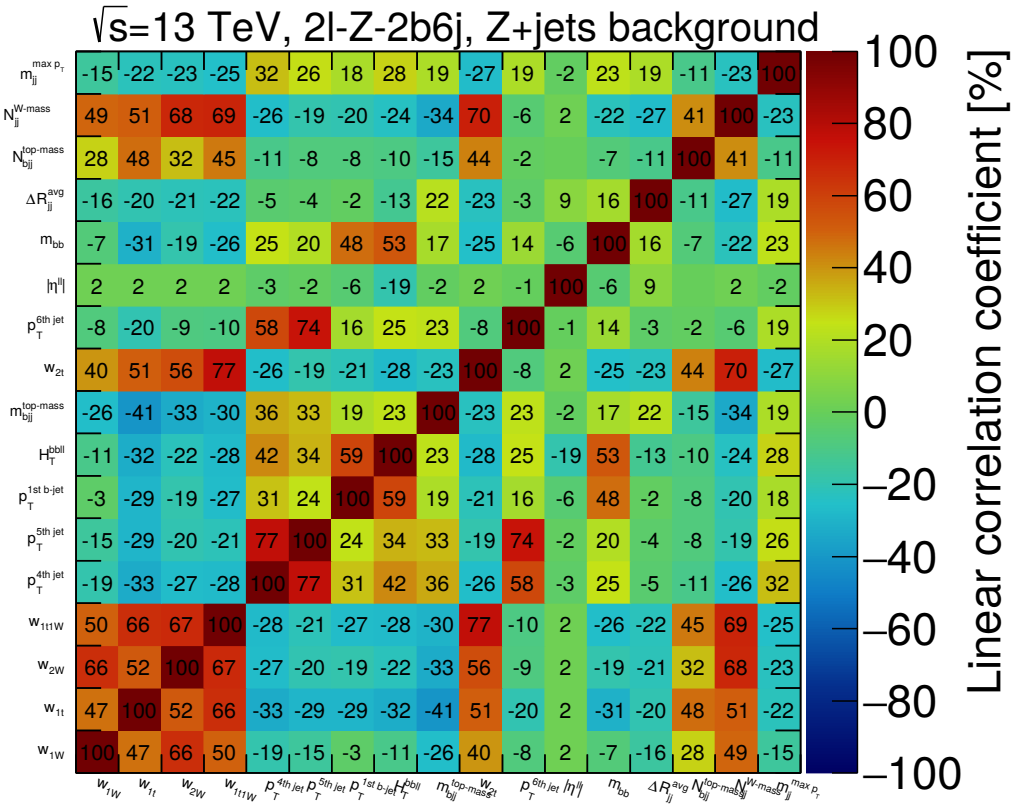


Fig. F.33.: Linear correlations of the input variables for the background dataset employed for the BDT training against the Z+jets background in the 2ℓ-Z-2b6j target region. The top figure shows the linear correlations for the events of the signal model, whereas the linear correlation coefficients for the background events are depicted in the bottom panel. The values are expressed in percent as the correlation coefficient itself is defined within the range of $[-1, 1]$.

F.9. Normalised distributions of the discriminating variables

In this section, the discriminating variables are shown for all regions in which they were employed for the BDT training.

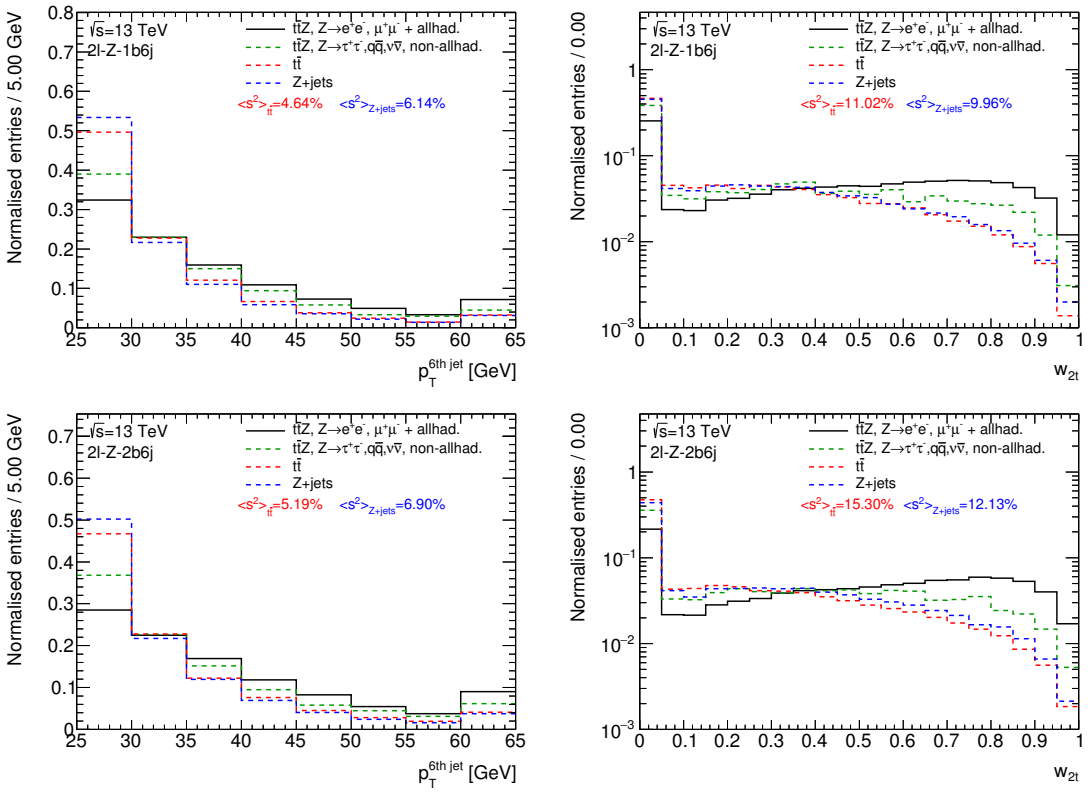


Fig. F.34.: Distributions of $p_{\mathrm{T}}^{\text{6th jet}}$ on the left and of w_{2t} on the right, shown for the 2ℓ-Z-1b6j and 2ℓ-Z-2b6j target region in the top and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

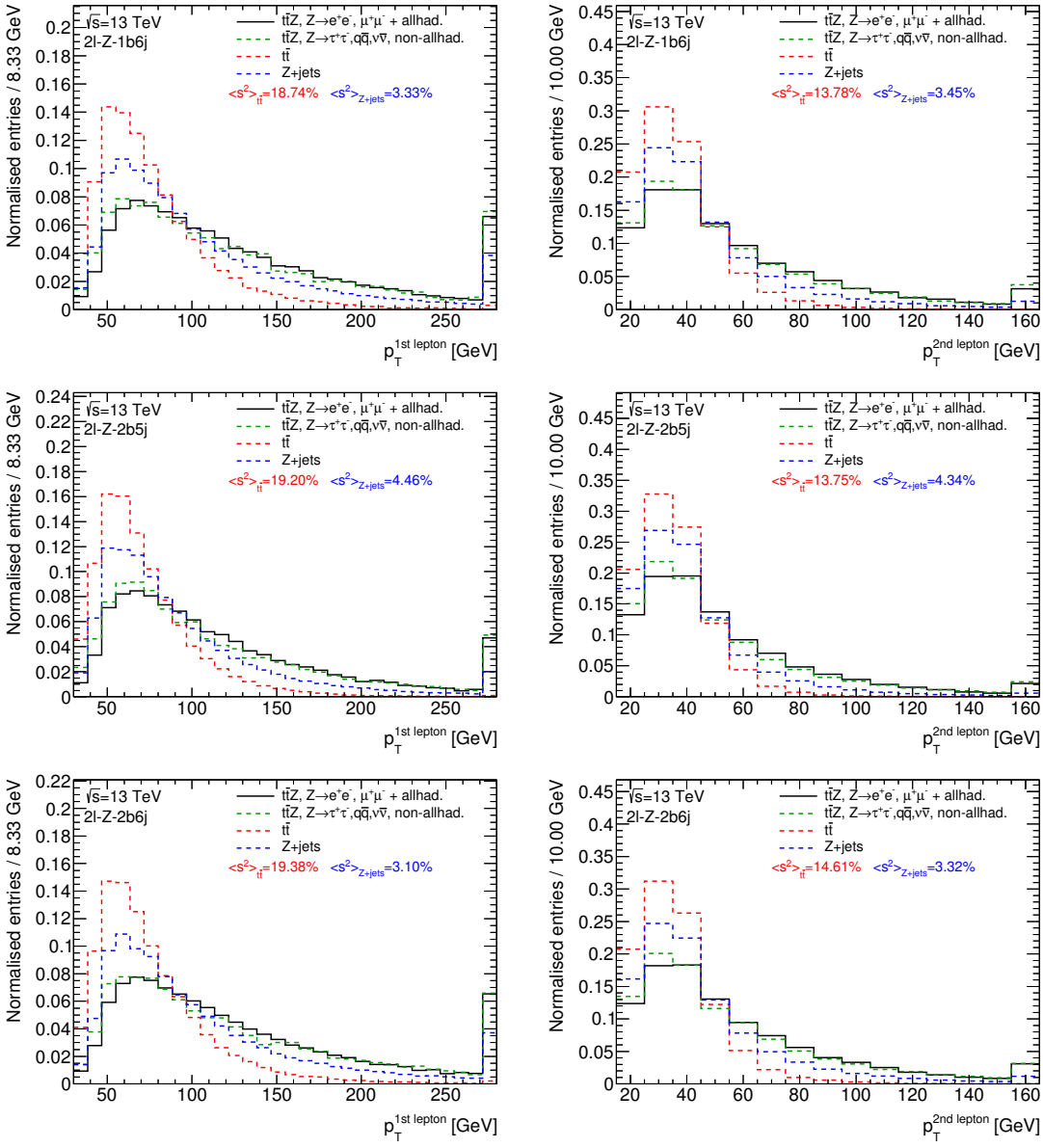


Fig. F.35.: Distributions of $p_{\mathrm{T}}^{\text{1st lepton}}$ on the left and of of $p_{\mathrm{T}}^{\text{2nd lepton}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

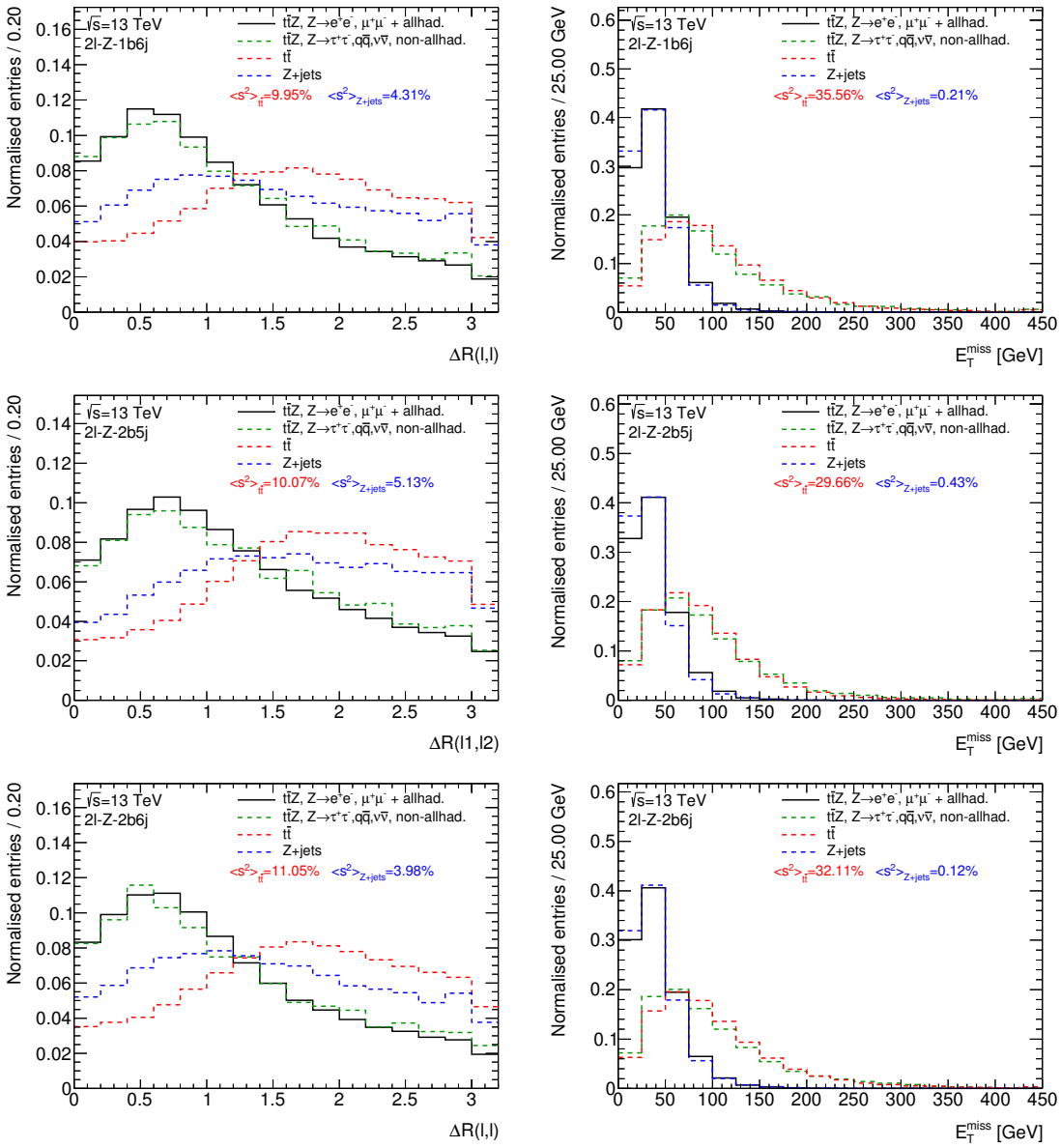


Fig. F.36.: Distributions of $\Delta R(\ell, \ell)$ on the left and of $E_{\mathrm{T}}^{\mathrm{miss}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

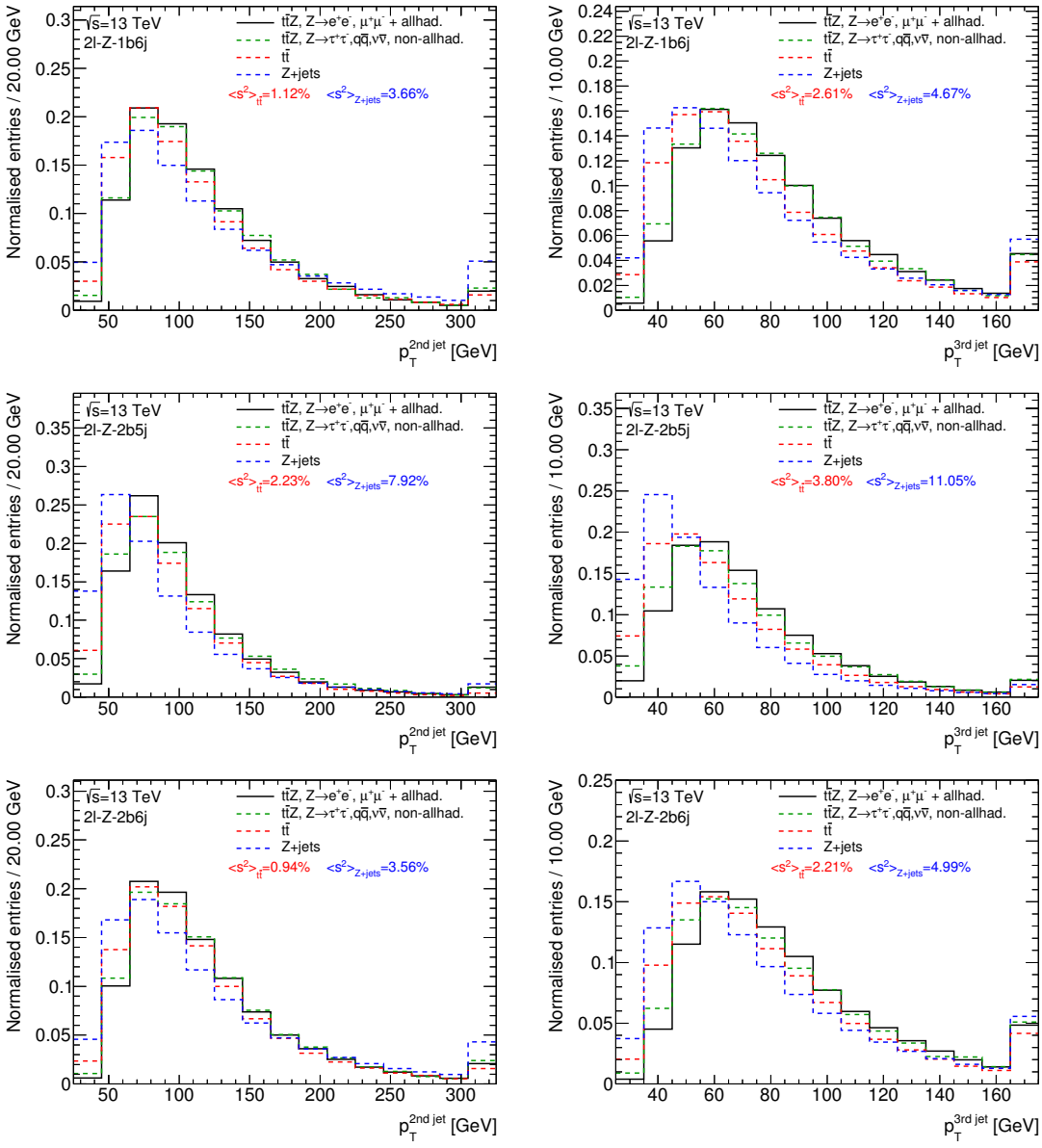


Fig. F.37.: Distributions of $p_{\mathrm{T}}^{\text{2nd jet}}$ on the left and of $p_{\mathrm{T}}^{\text{3rd jet}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

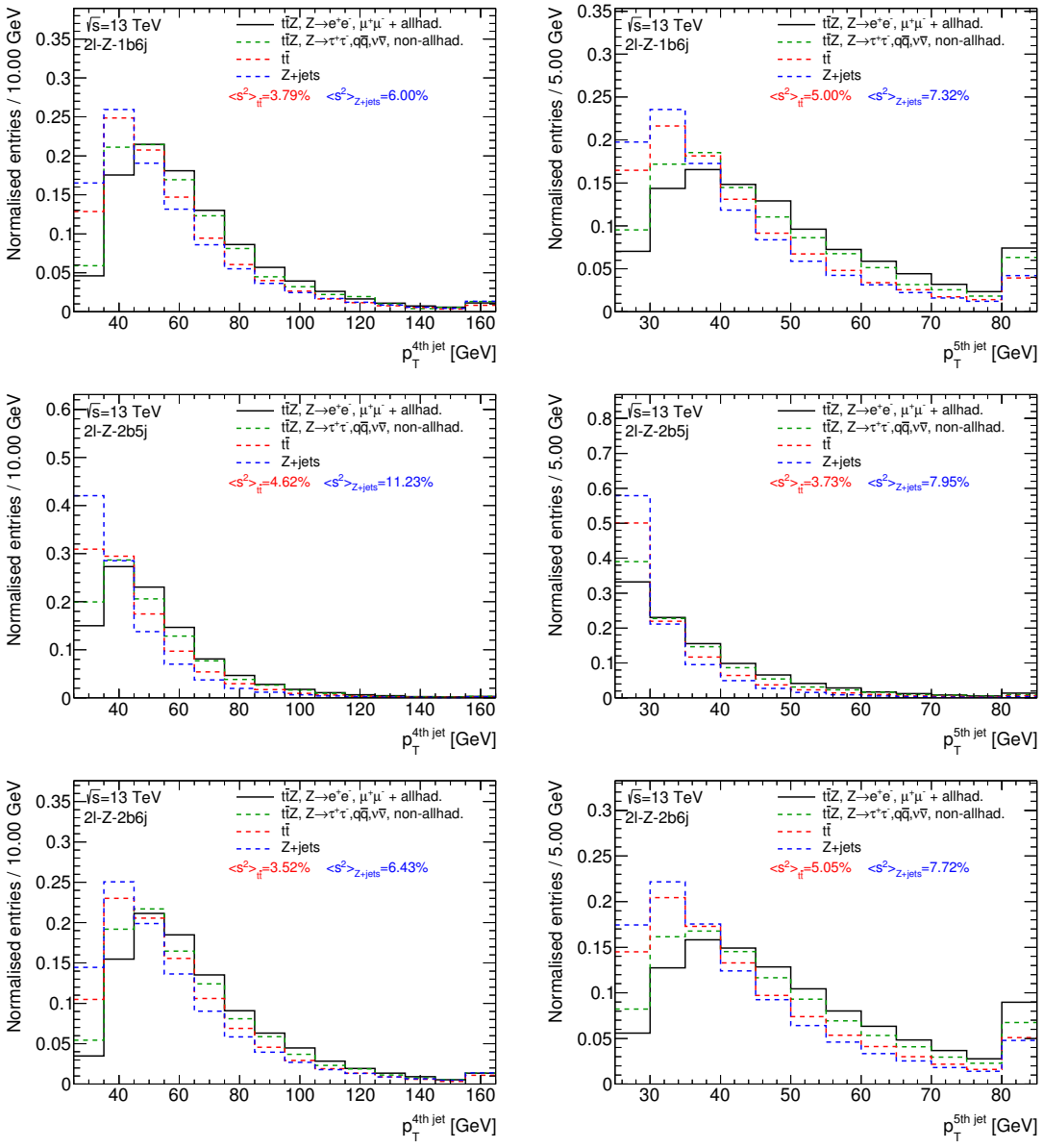


Fig. F.38.: Distributions of $p_{\mathrm{T}}^{\text{4th jet}}$ on the left and of $p_{\mathrm{T}}^{\text{5th jet}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

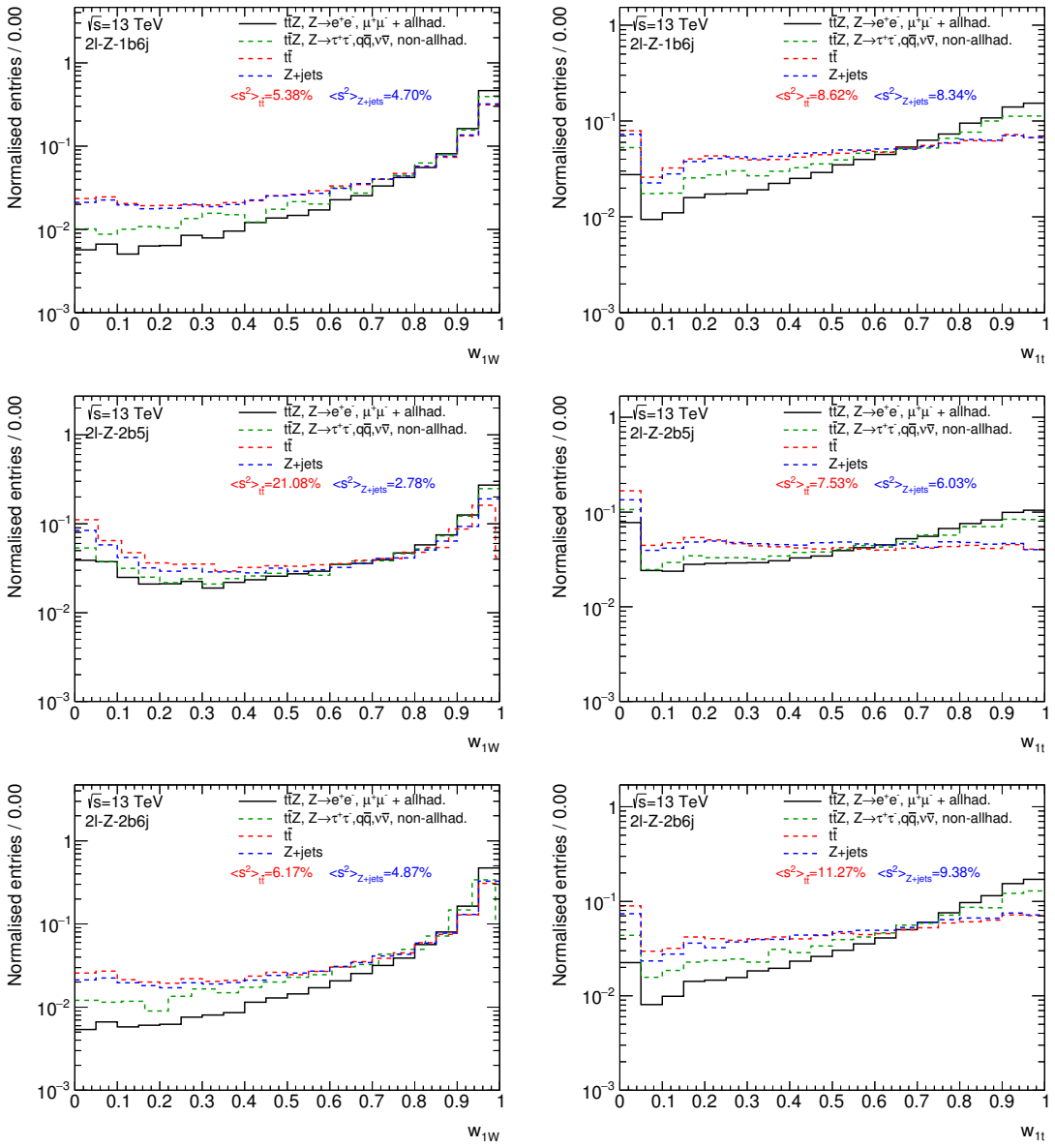


Fig. F.39.: Distributions of w_{1W} on the left and of w_{1t} on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

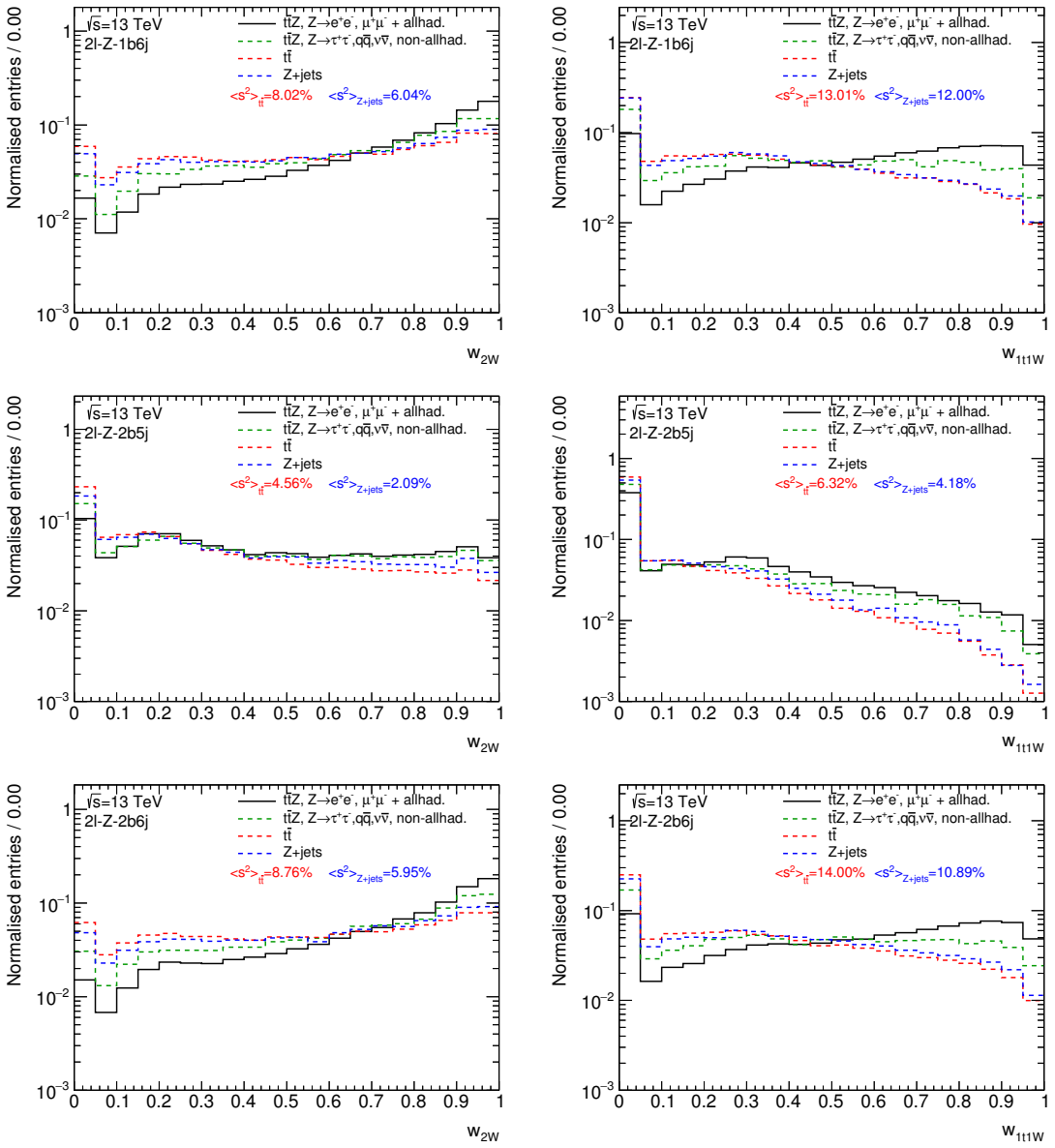


Fig. F.40.: Distributions of w_{2W} on the left and of w_{1t1W} on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

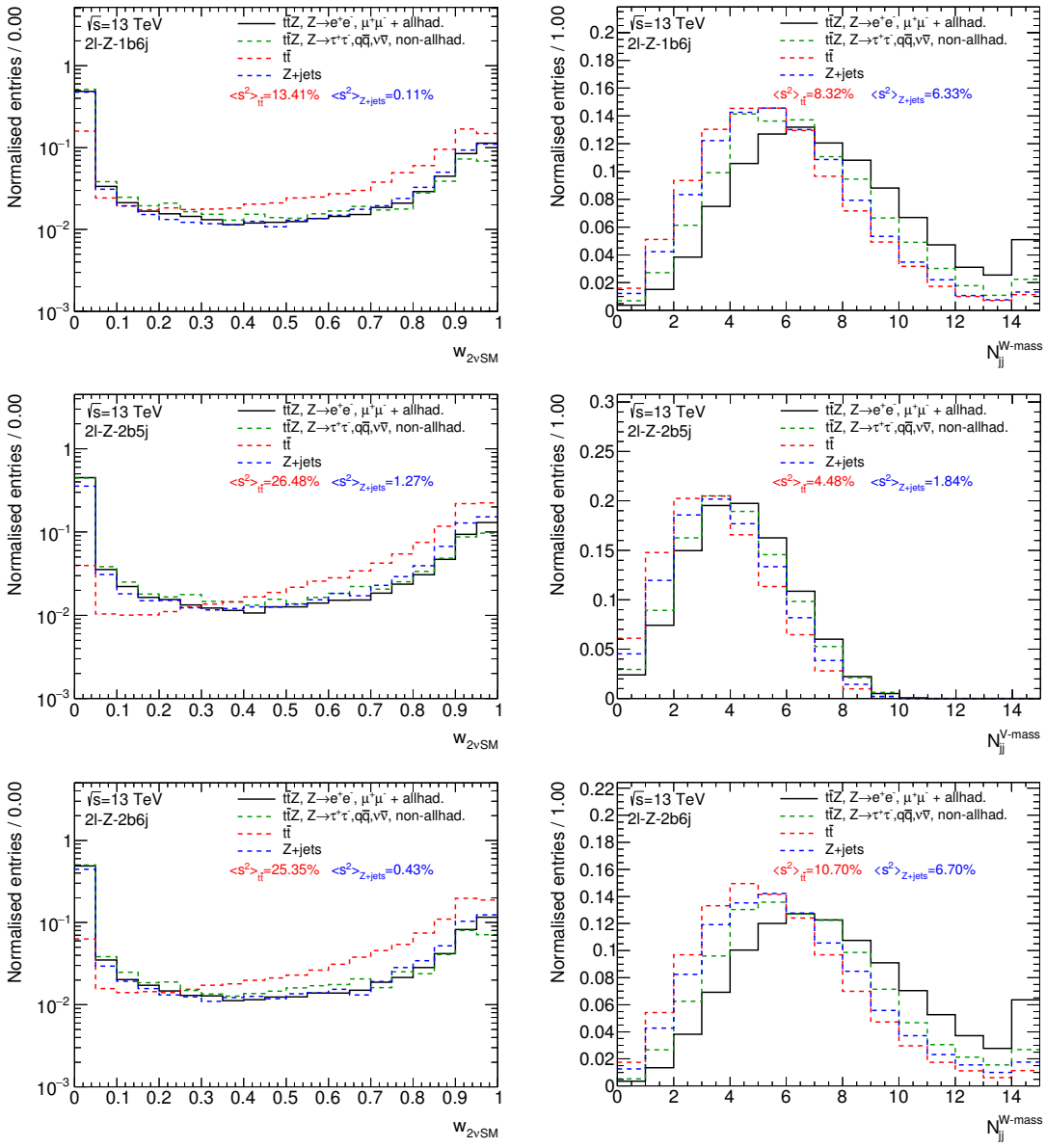


Fig. F.41.: Distributions of $w_{2\nu\text{SM}}$ on the left and of $N_{jj}^{W\text{-mass}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

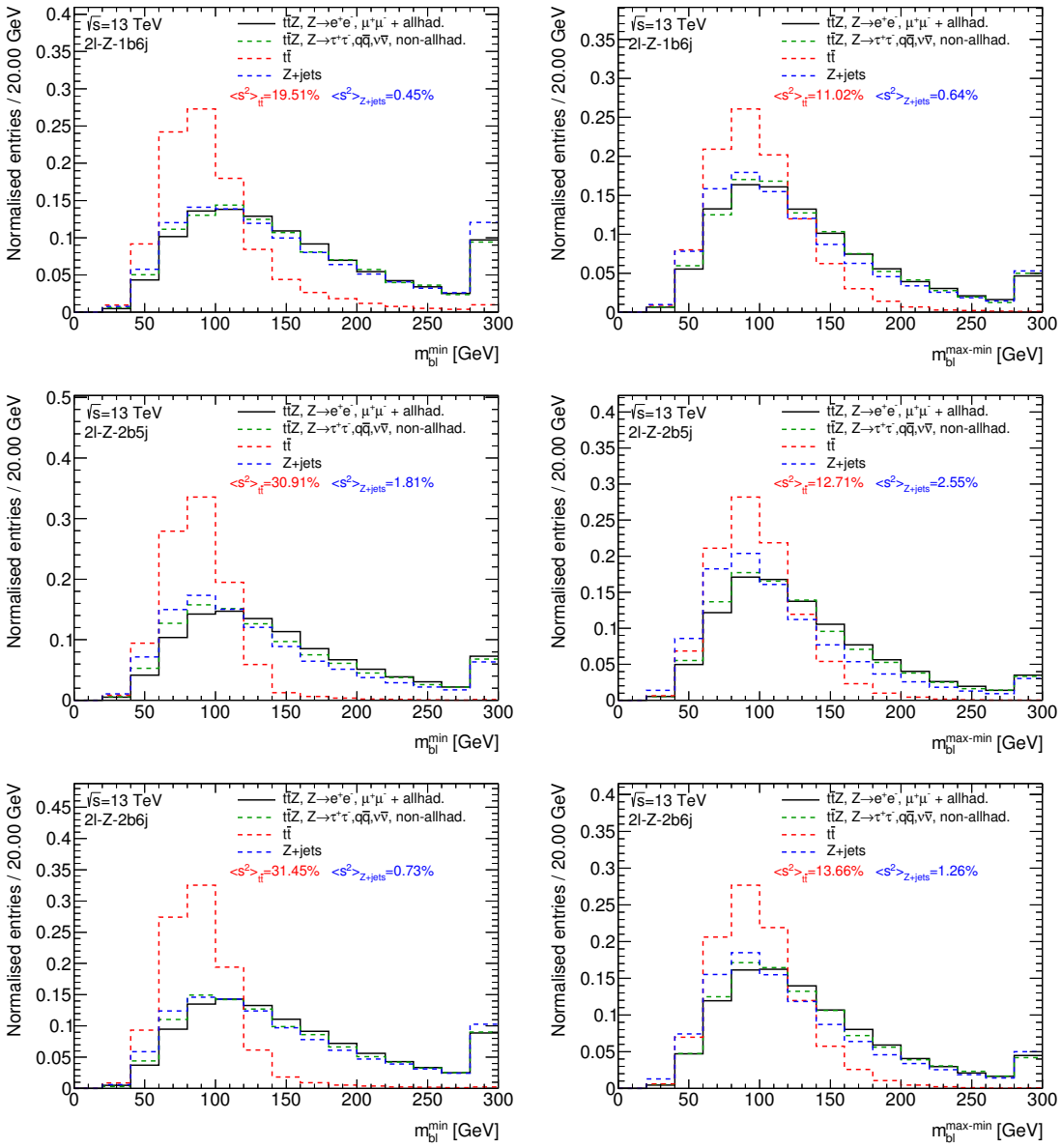


Fig. F.42.: Distributions of $m_{b\ell}^{\text{min}}$ on the left and of $m_{b\ell}^{\text{max-min}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

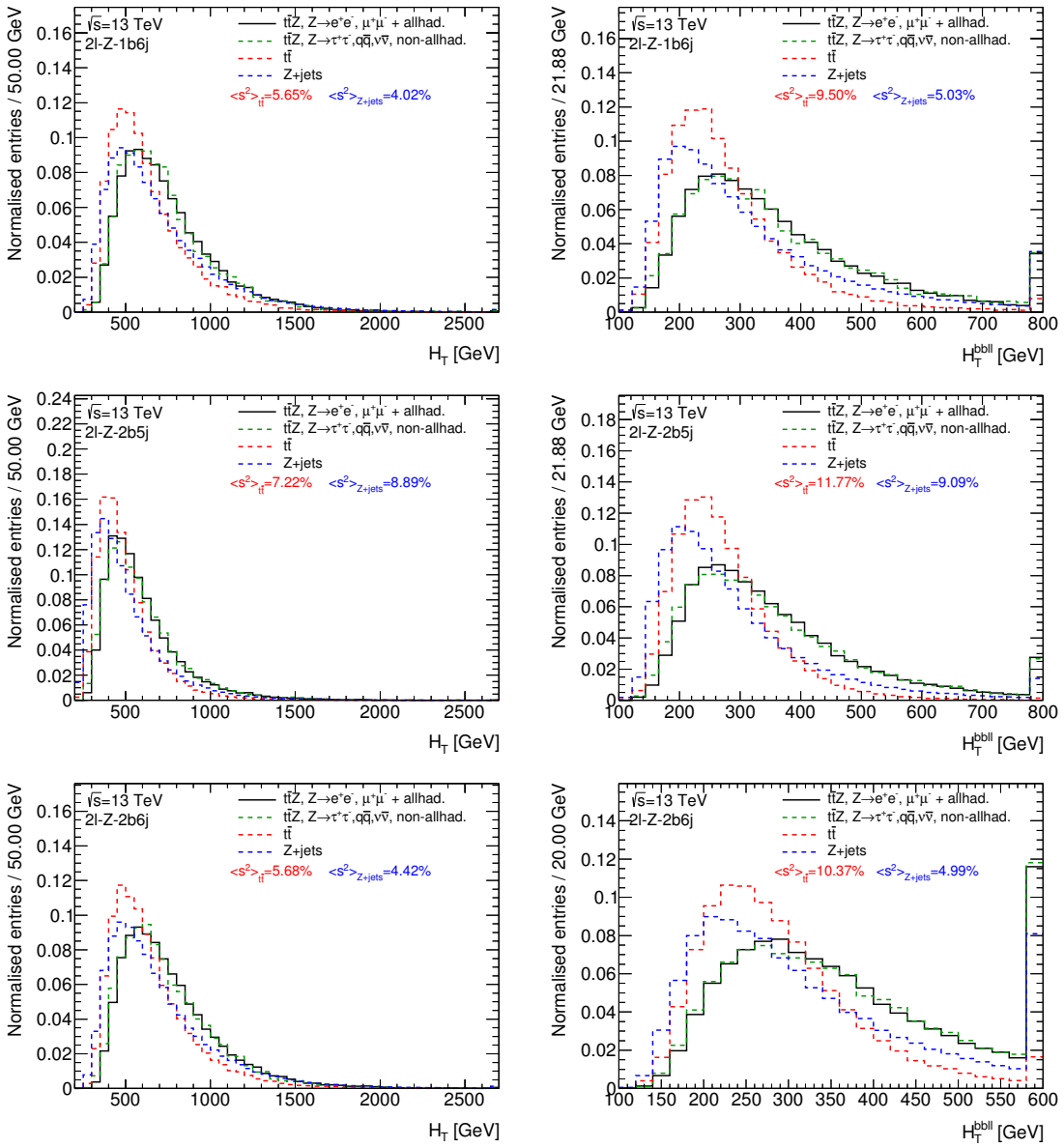


Fig. F.43.: Distributions of H_{T} on the left and of $H_{\mathrm{T}}^{bb\ell\ell}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

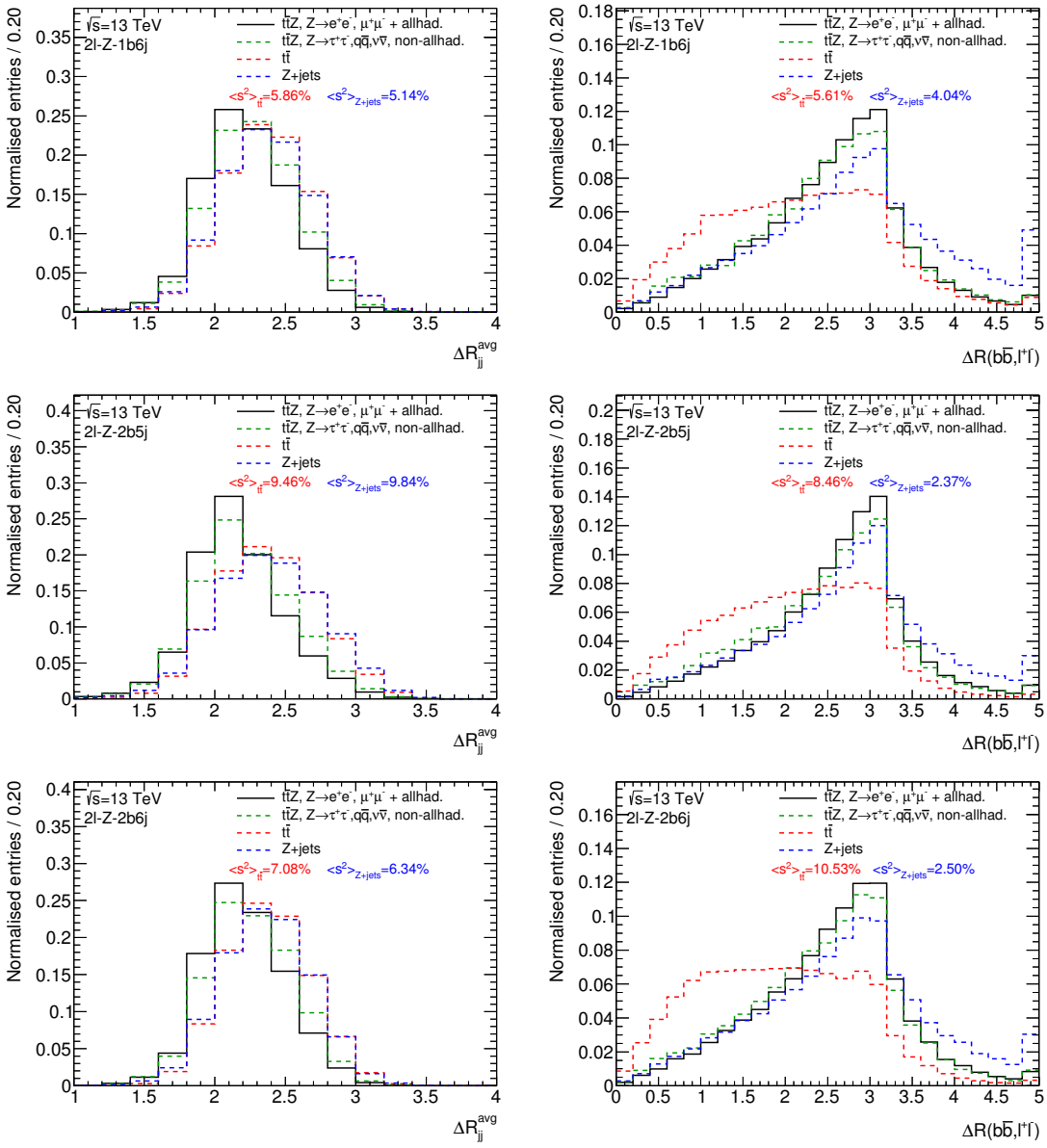


Fig. F.44.: Distributions of $\Delta R_{jj}^{\text{avg}}$ on the left and of $\Delta R(bb, \ell\ell)$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

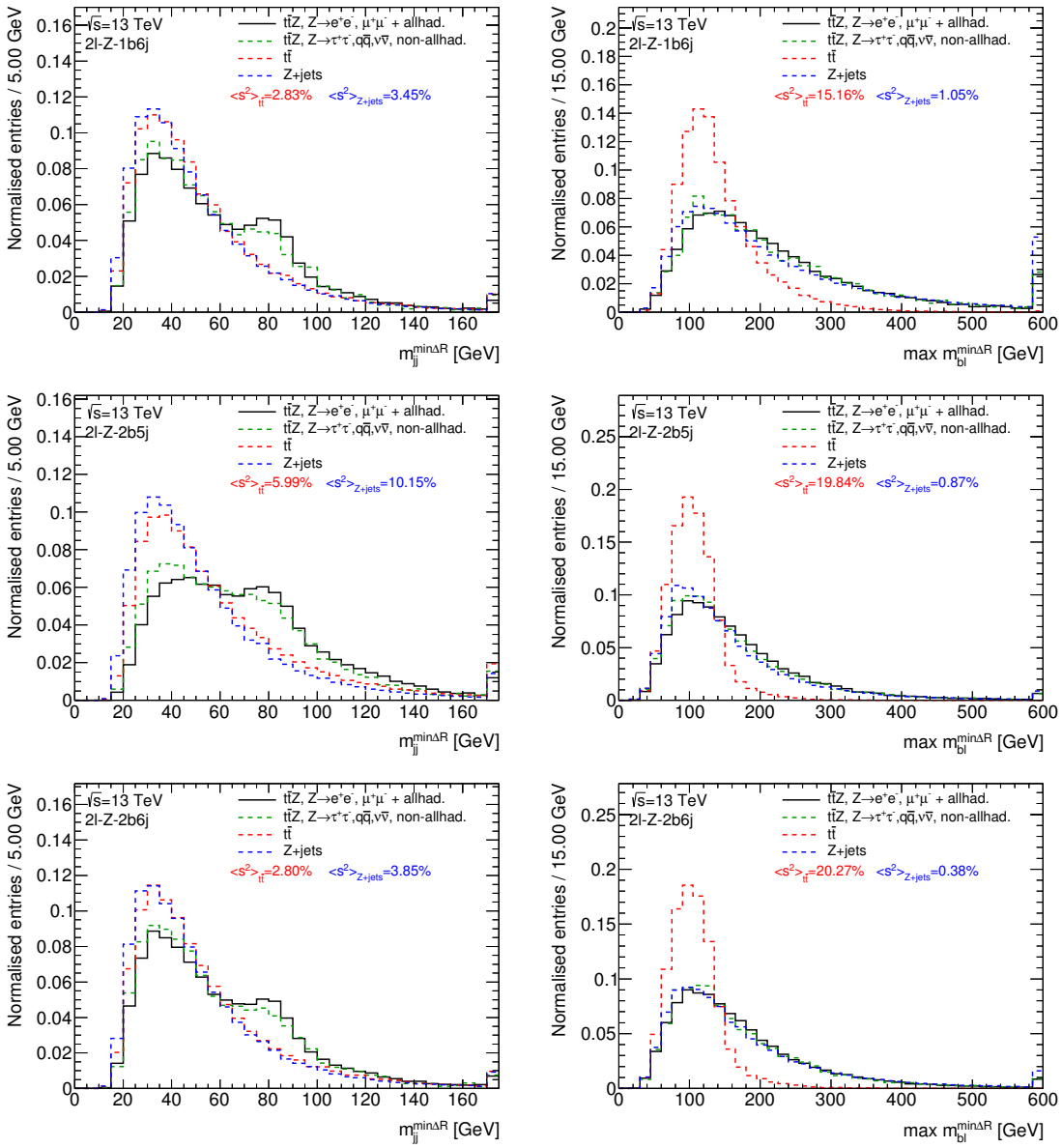


Fig. F.45.: Distributions of $m_{jj}^{\min \Delta R}$ on the left and of $\max m_{b\ell}^{\min \Delta R}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

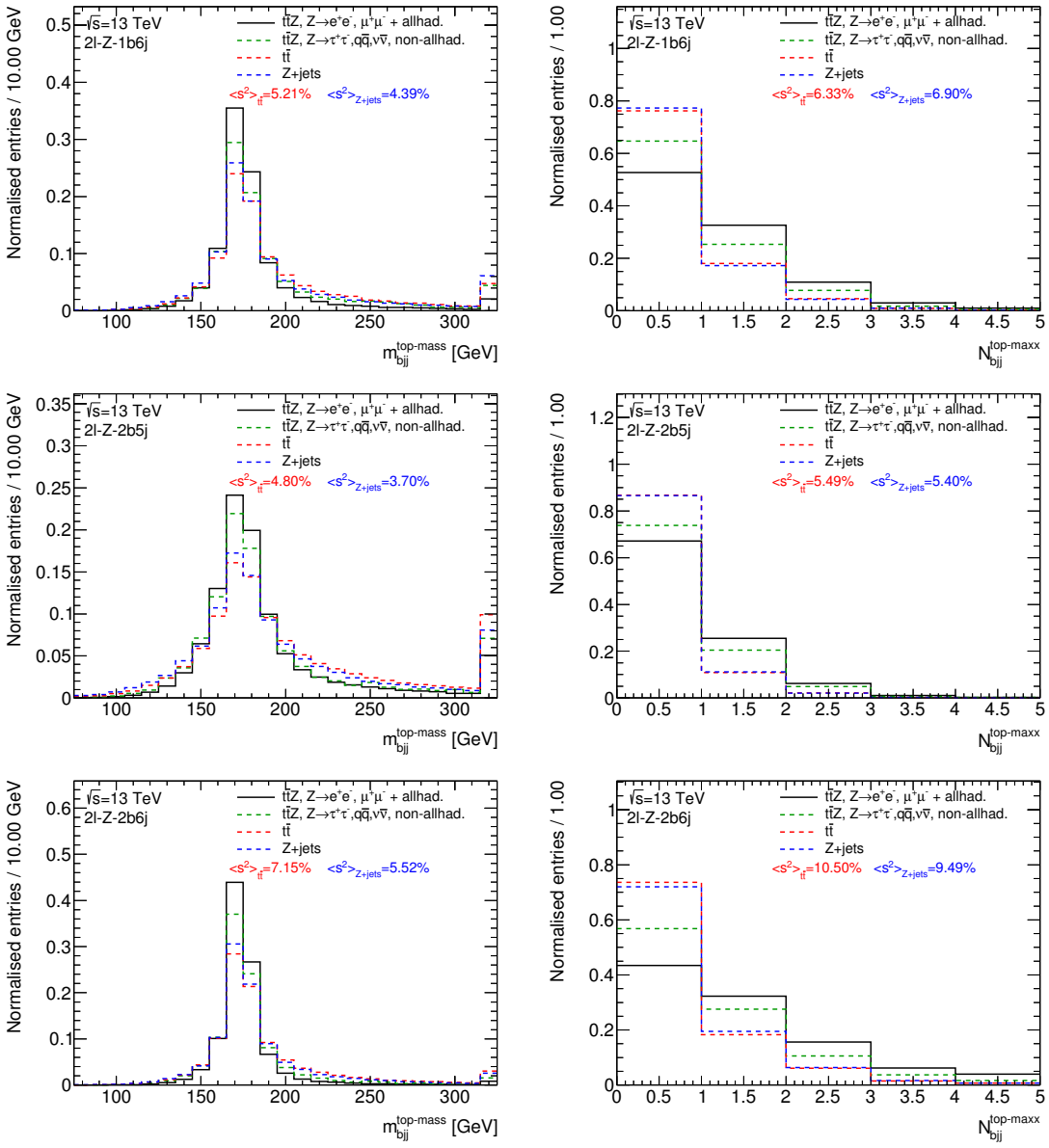


Fig. F.46.: Distributions of $m_{bjj}^{\text{top-mass}}$ on the left and of $N_{bjj}^{\text{top-mass}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

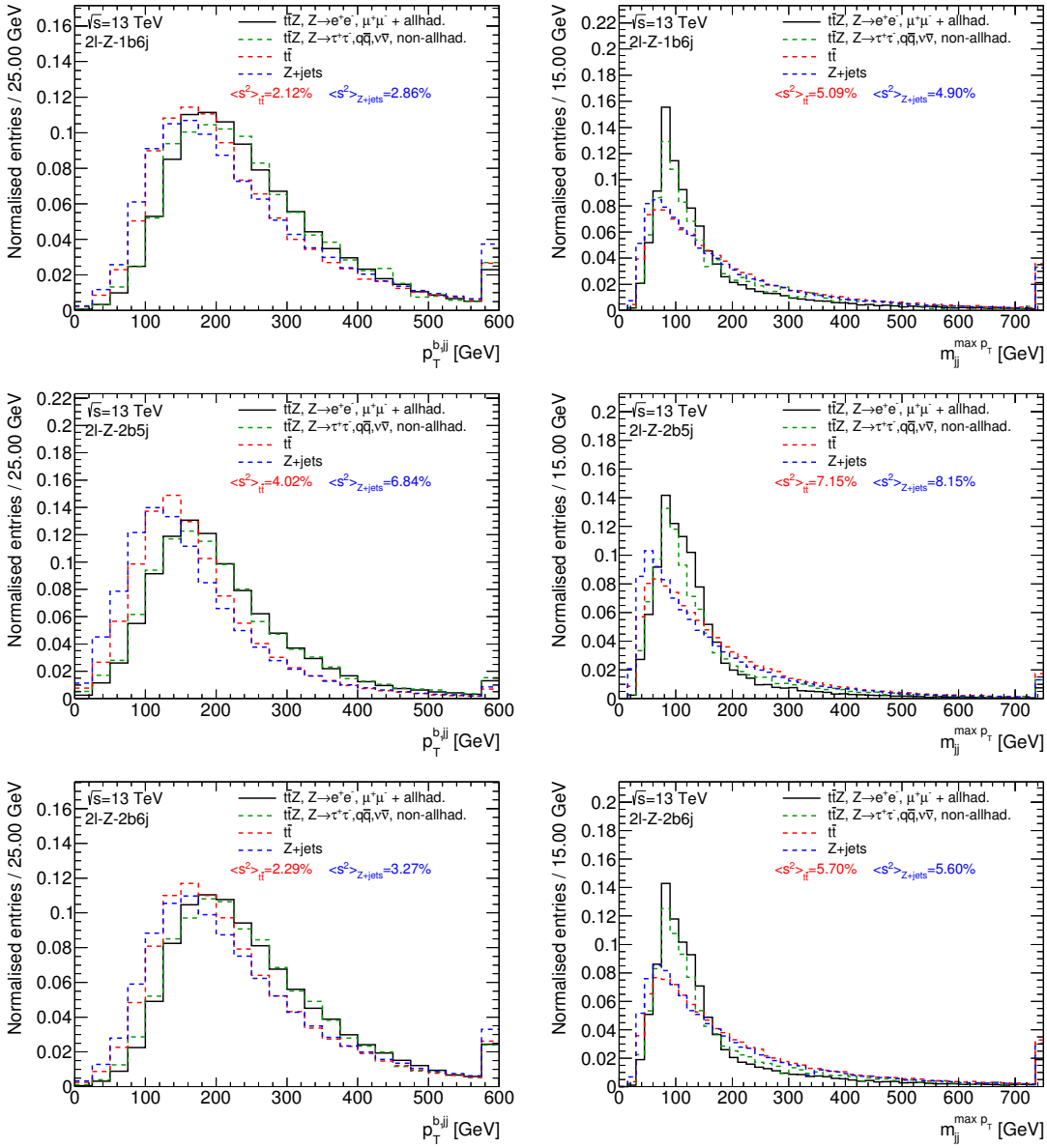


Fig. F.47.: Distributions of $p_{\mathrm{T}}^{b_1 jj}$ on the left and of $m_{jj}^{\max p_{\mathrm{T}}}$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

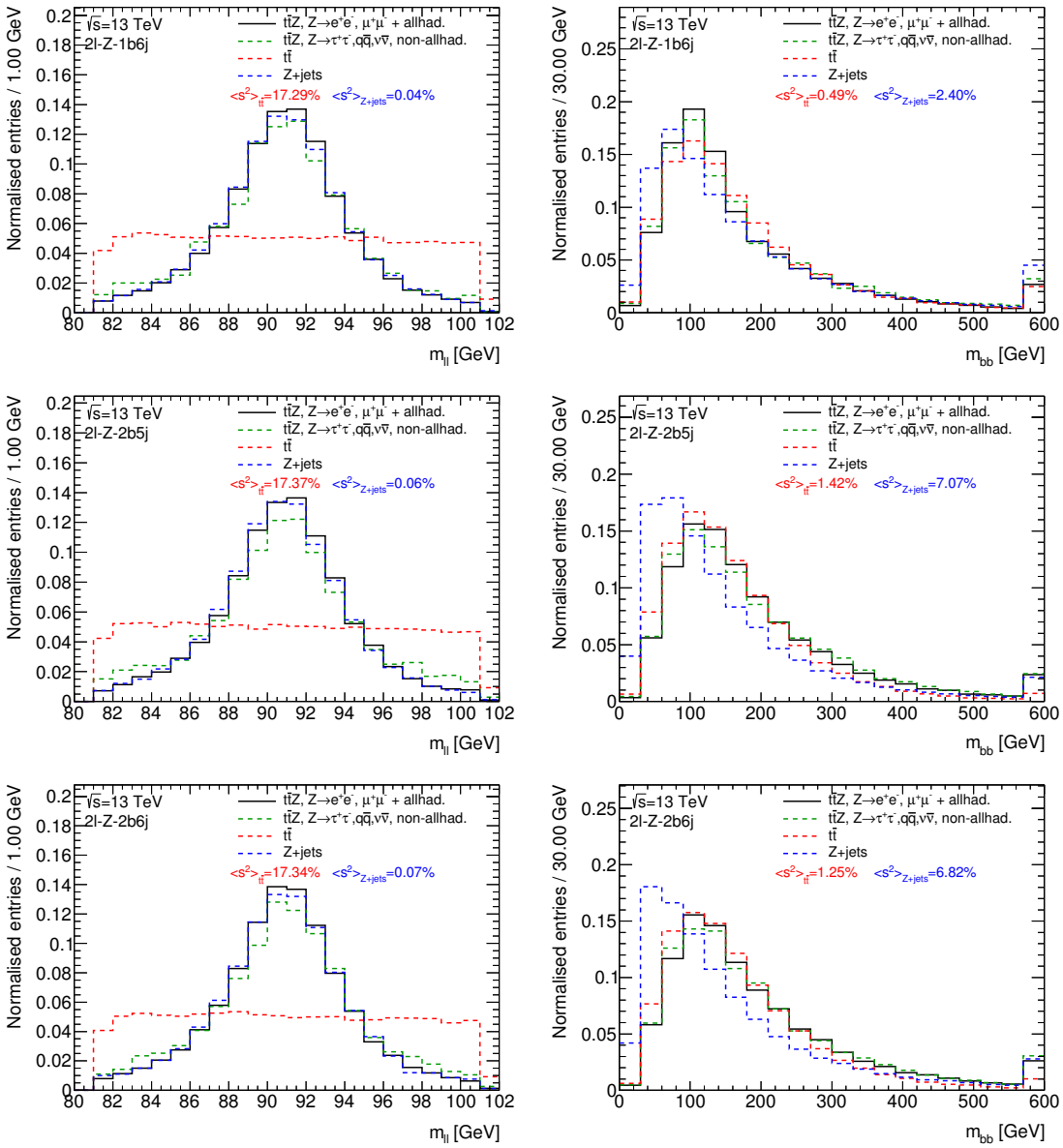


Fig. F.48.: Distributions of $m_{\ell\ell}$ on the left and of m_{bb} on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

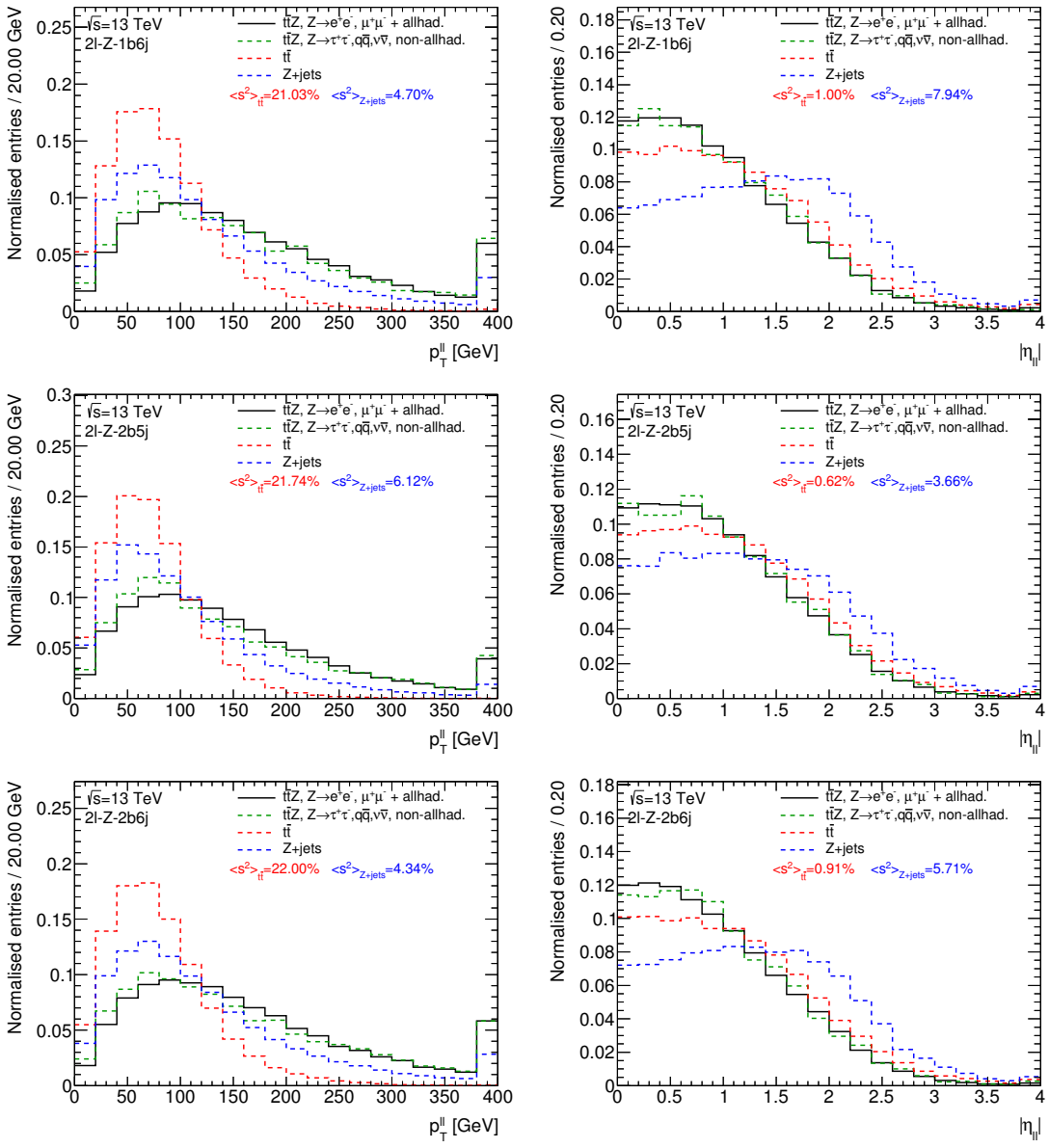


Fig. F.49.: Distributions of $p_{\mathrm{T}}^{\ell\ell}$ on the left and of $|\eta_{\ell\ell}|$ on the right, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

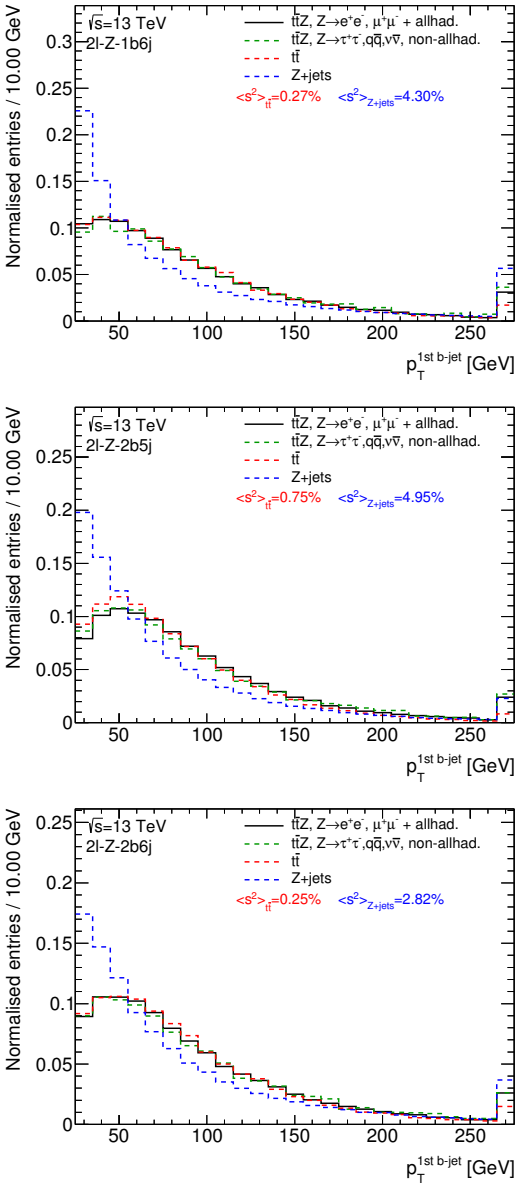


Fig. F.50.: Distributions of $p_{\mathrm{T}}^{1\text{st } b\text{-jet}}$, shown for the 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j target region in the top, middle and bottom row, respectively. The $t\bar{t}Z$ signal is depicted in black for events featuring a fully-hadronic $t\bar{t}$ decay and a Z boson decaying either to an electron-positron or a muon-antimuon pair. The other $t\bar{t}Z$ events are depicted in green. The $t\bar{t}$ and Z+jets backgrounds are shown as red and blue dashed lines, respectively. All distributions were normalised to the respective number of entries. The separation power between the fully-hadronic signal and the respective background processes are calculated according to Eq. (7.3). For all distributions the overflow bin was added.

F.10. Data-MC modelling of the discriminating variables

The figures presented in this section show the modelling of the discriminating variables presented in the previous section, App. F.9.

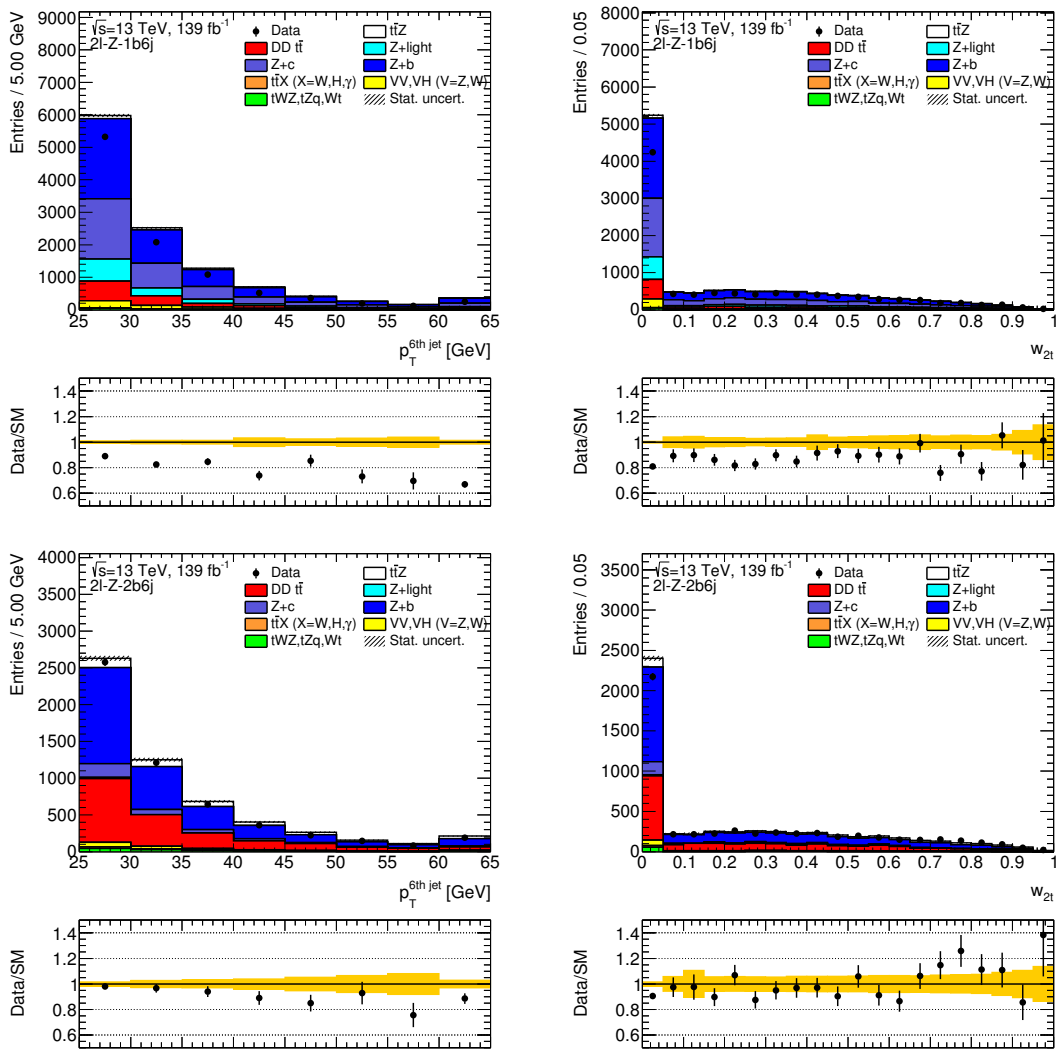


Fig. F.51.: Distributions of $p_{\mathrm{T}}^{\text{6th jet}}$ on the left and of w_{2t} on the right. The 2ℓ-Z-1b6j and 2ℓ-Z-2b6j regions are depicted in the top and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

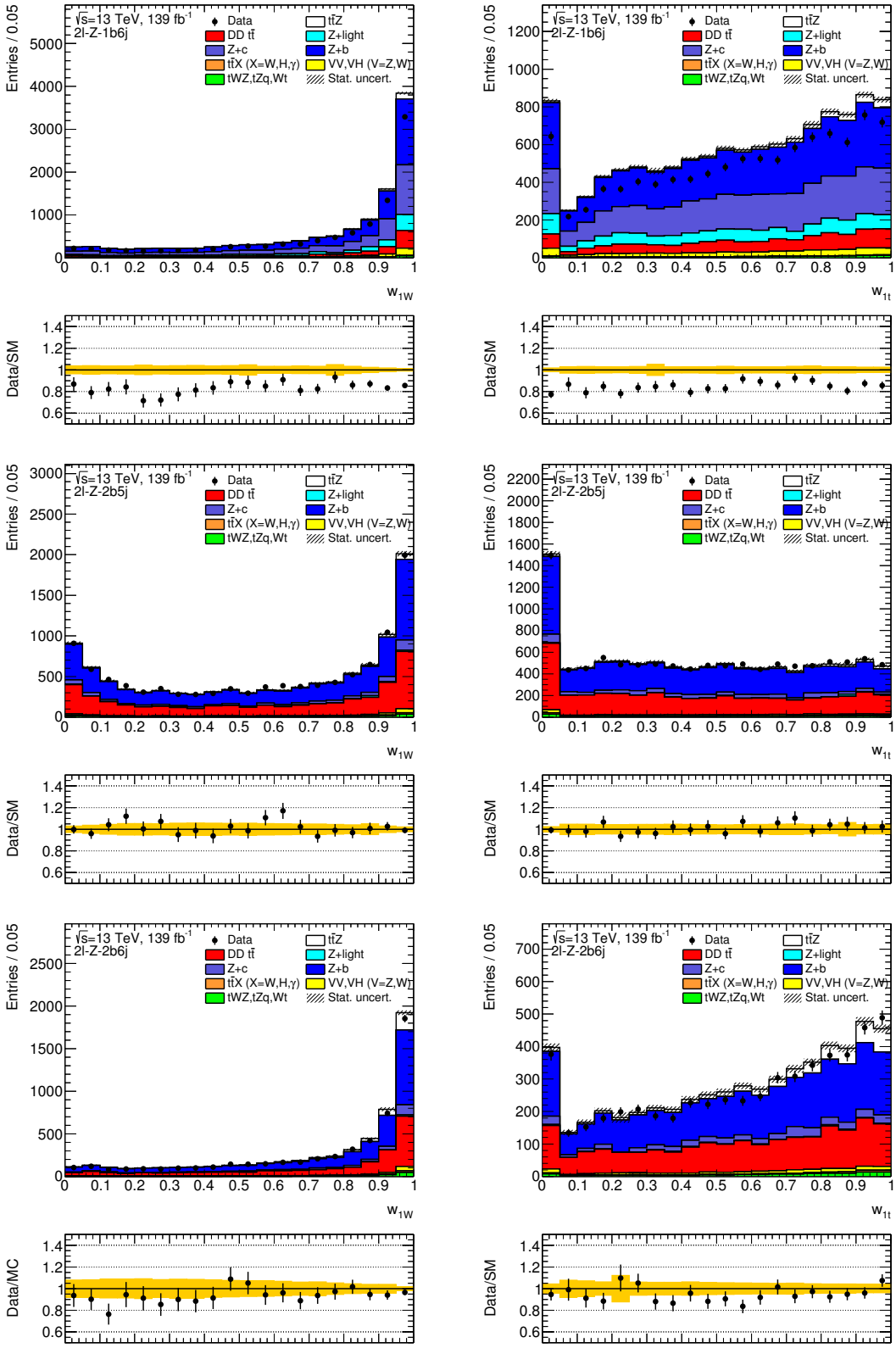


Fig. F.52.: Distributions of w_{1W} on the left and of w_{1t} on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

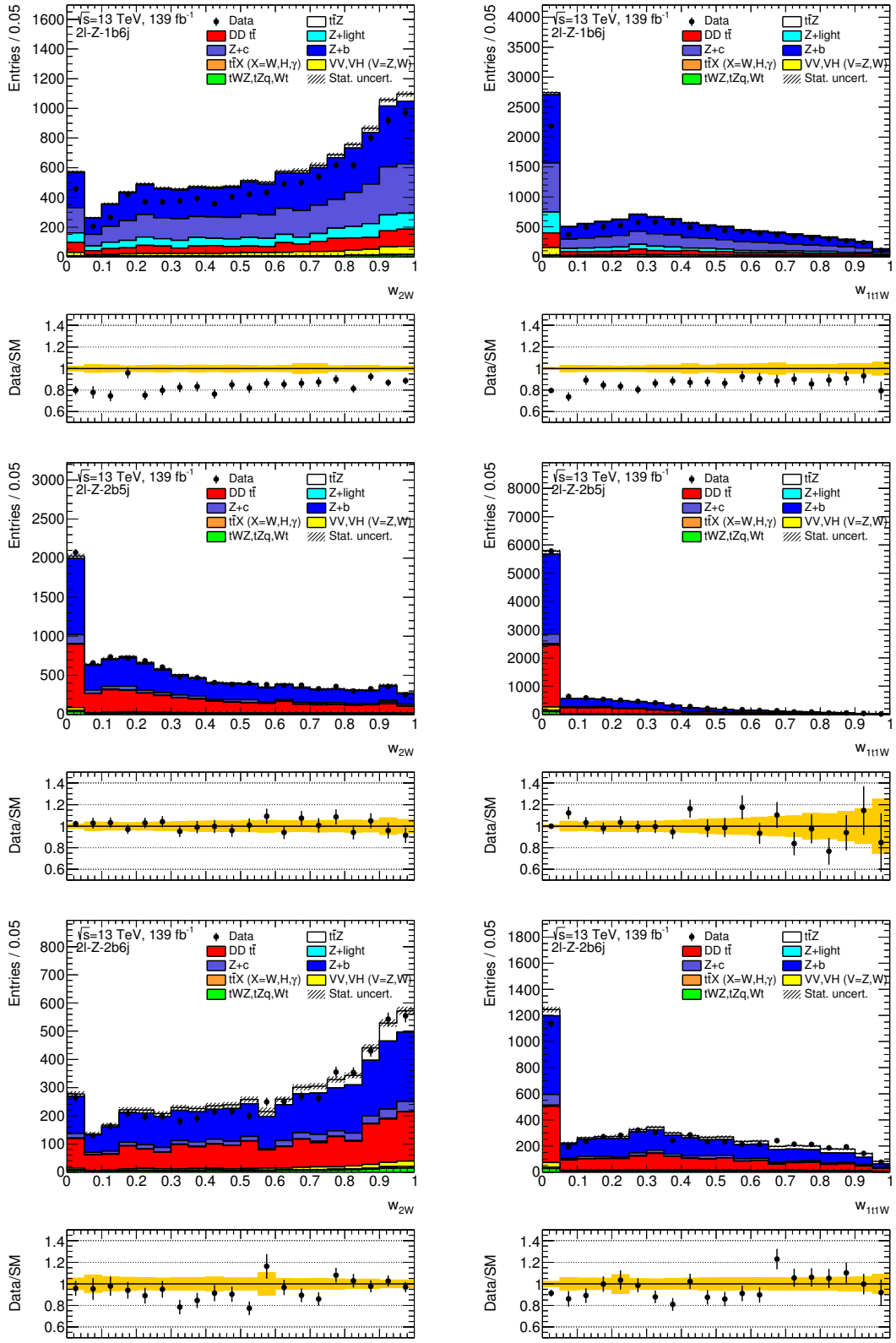


Fig. F.53.: Distributions of w_{2W} on the left and of w_{1t1W} on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

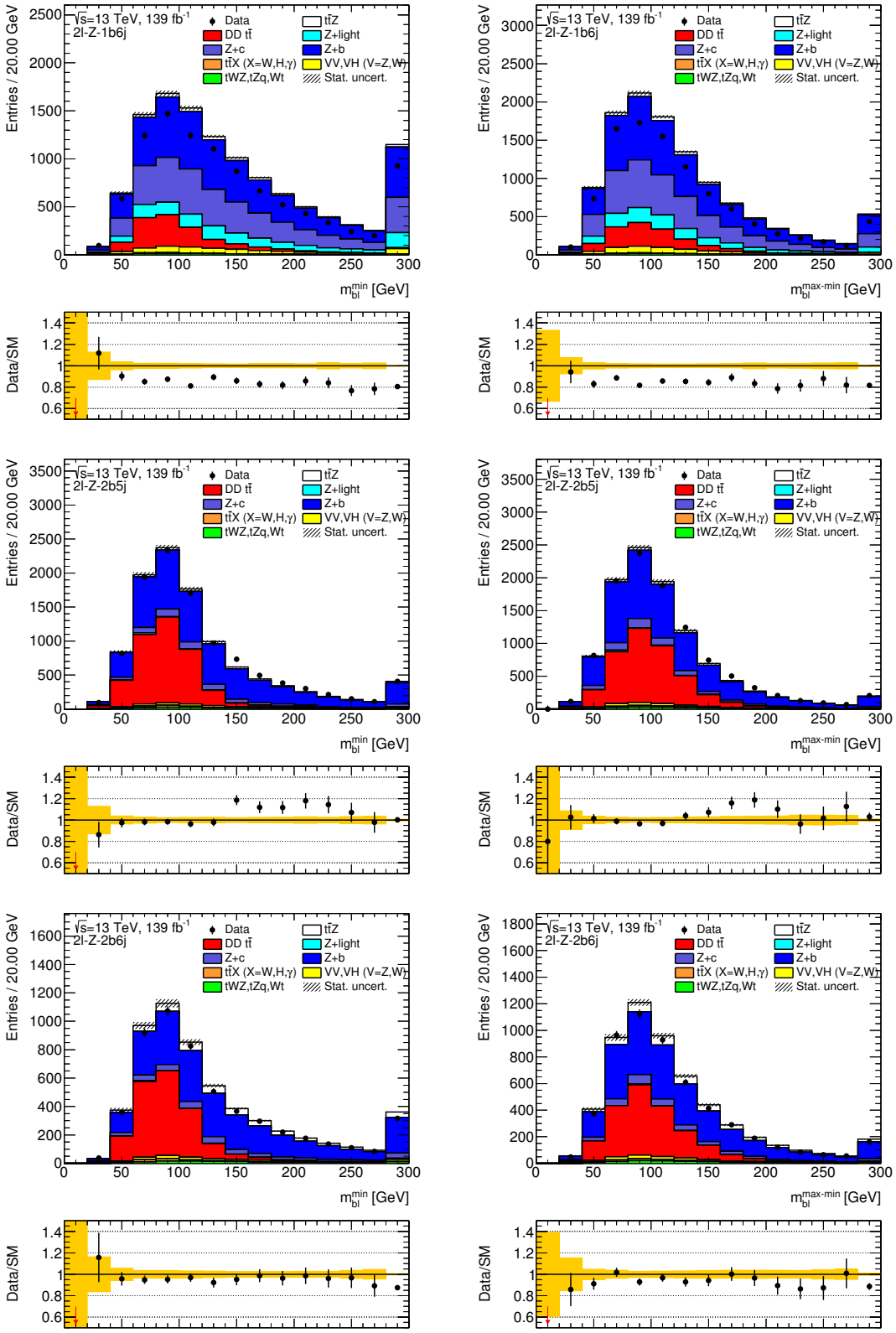


Fig. F.54.: Distributions of $m_{b\ell}^{\text{min}}$ on the left and of $m_{b\ell}^{\text{max-min}}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\text{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

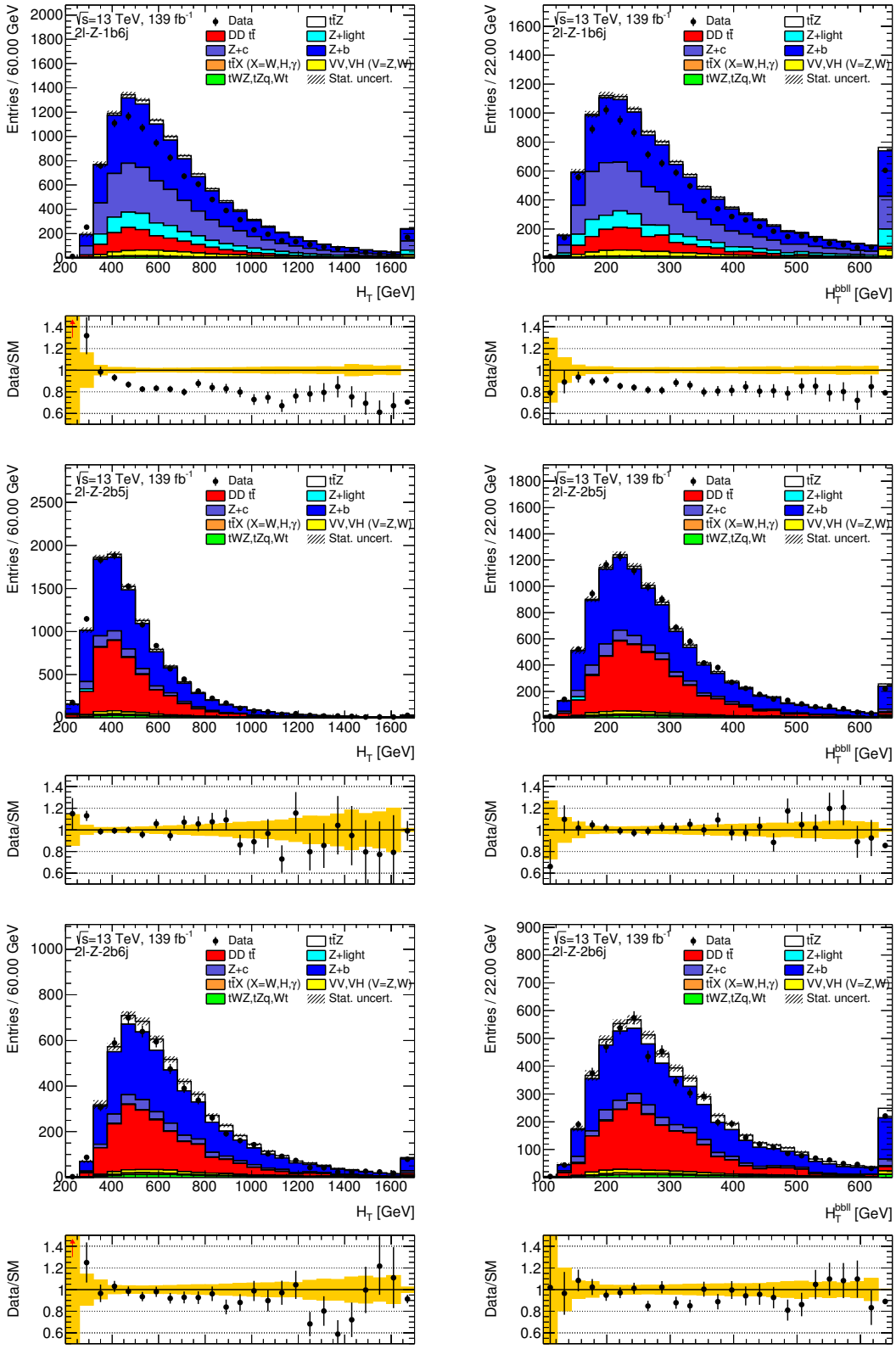


Fig. F.55.: Distributions of H_{T} on the left and of $H_{\mathrm{T}}^{bb\ell\ell}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

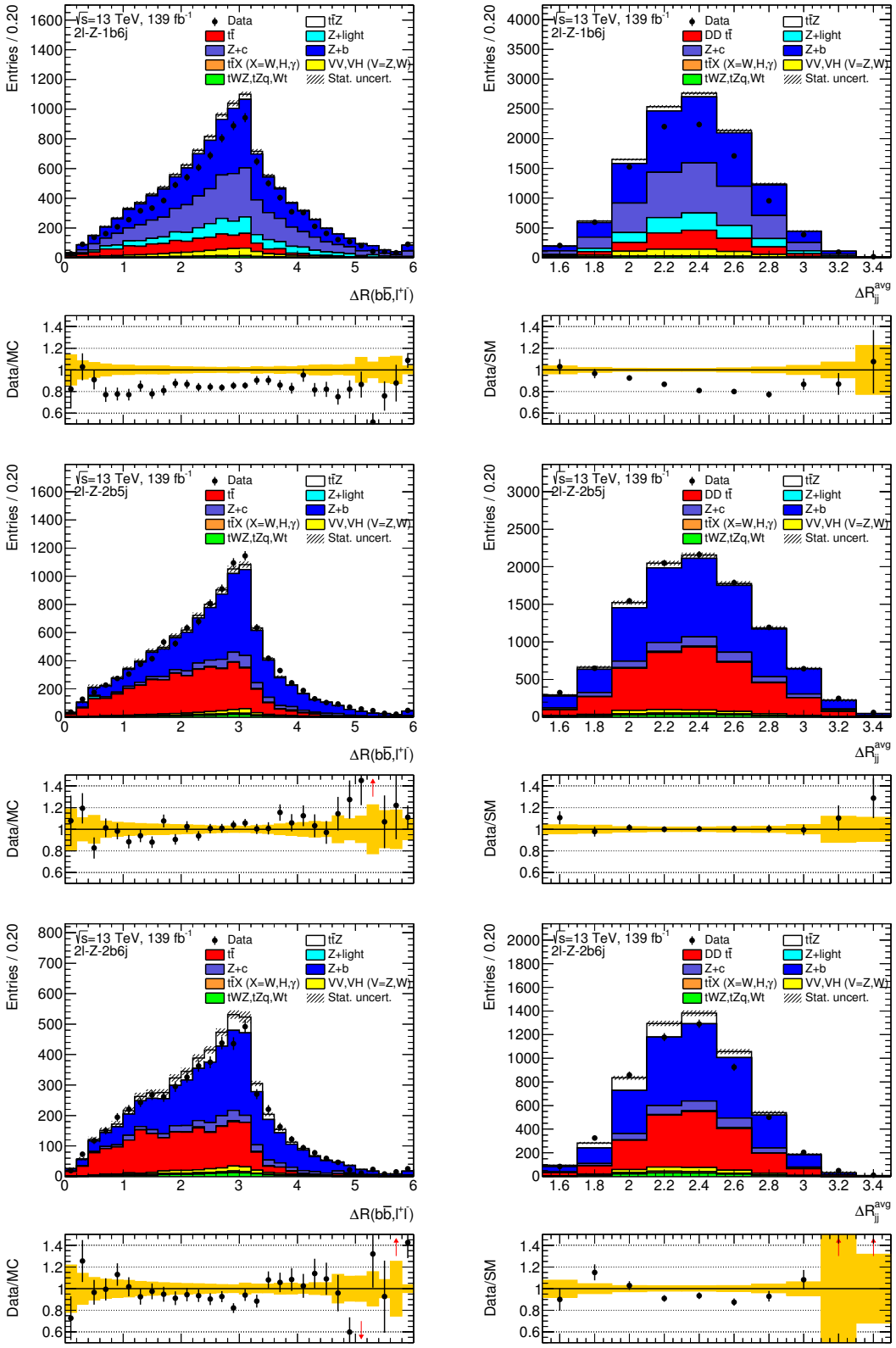


Fig. F.56.: Distributions of $\Delta R(bb, \ell\ell)$ on the left and of $\Delta R_{jj}^{\text{avg}}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\text{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

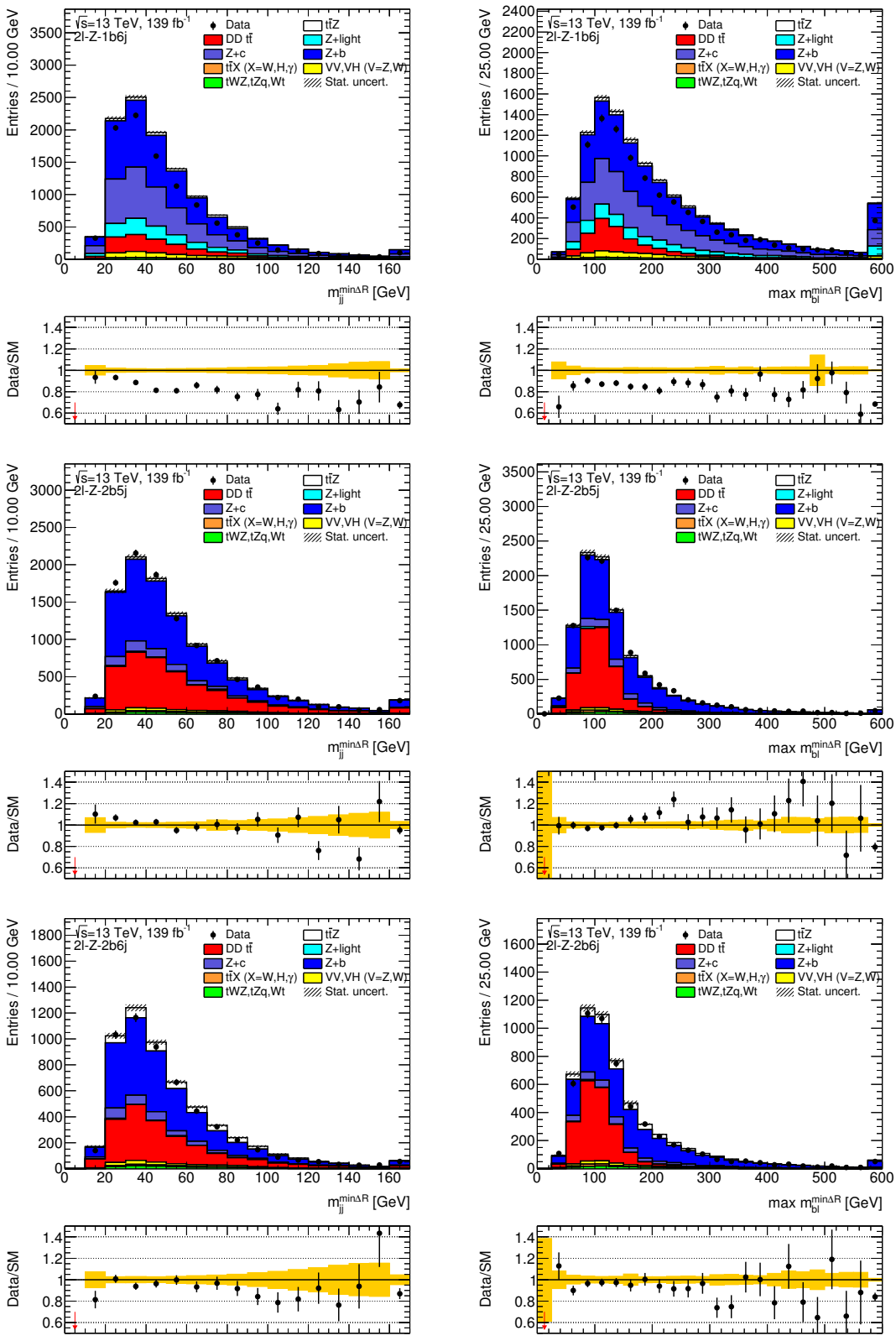


Fig. F.57.: Distributions of $m_{jj}^{\min \Delta R}$ on the left and of $\max m_{b\ell}^{\min \Delta R}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

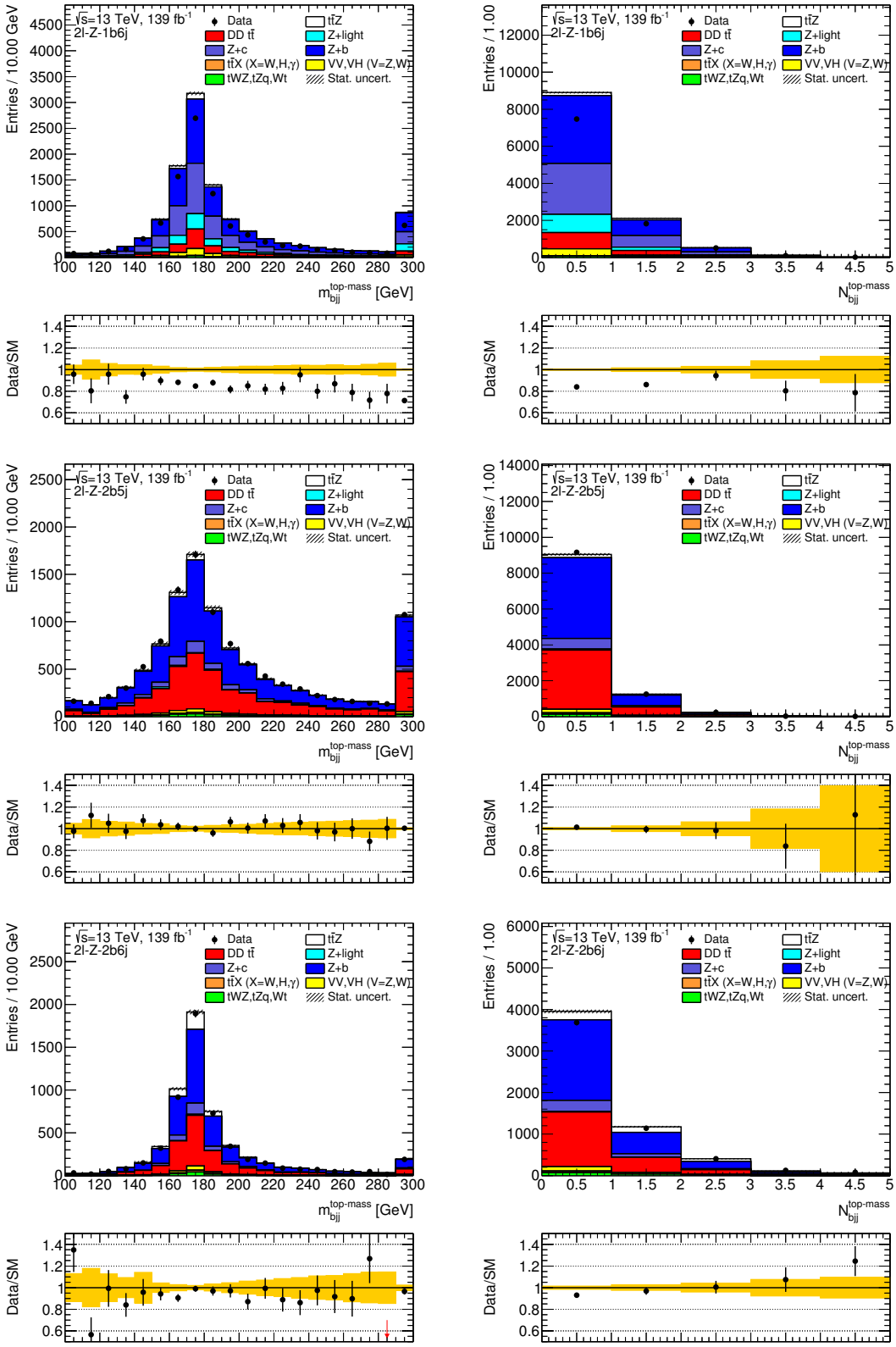


Fig. F.58.: Distributions of $m_{bjj}^{\text{top-mass}}$ on the left and of $N_{bjj}^{\text{top-mass}}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\text{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

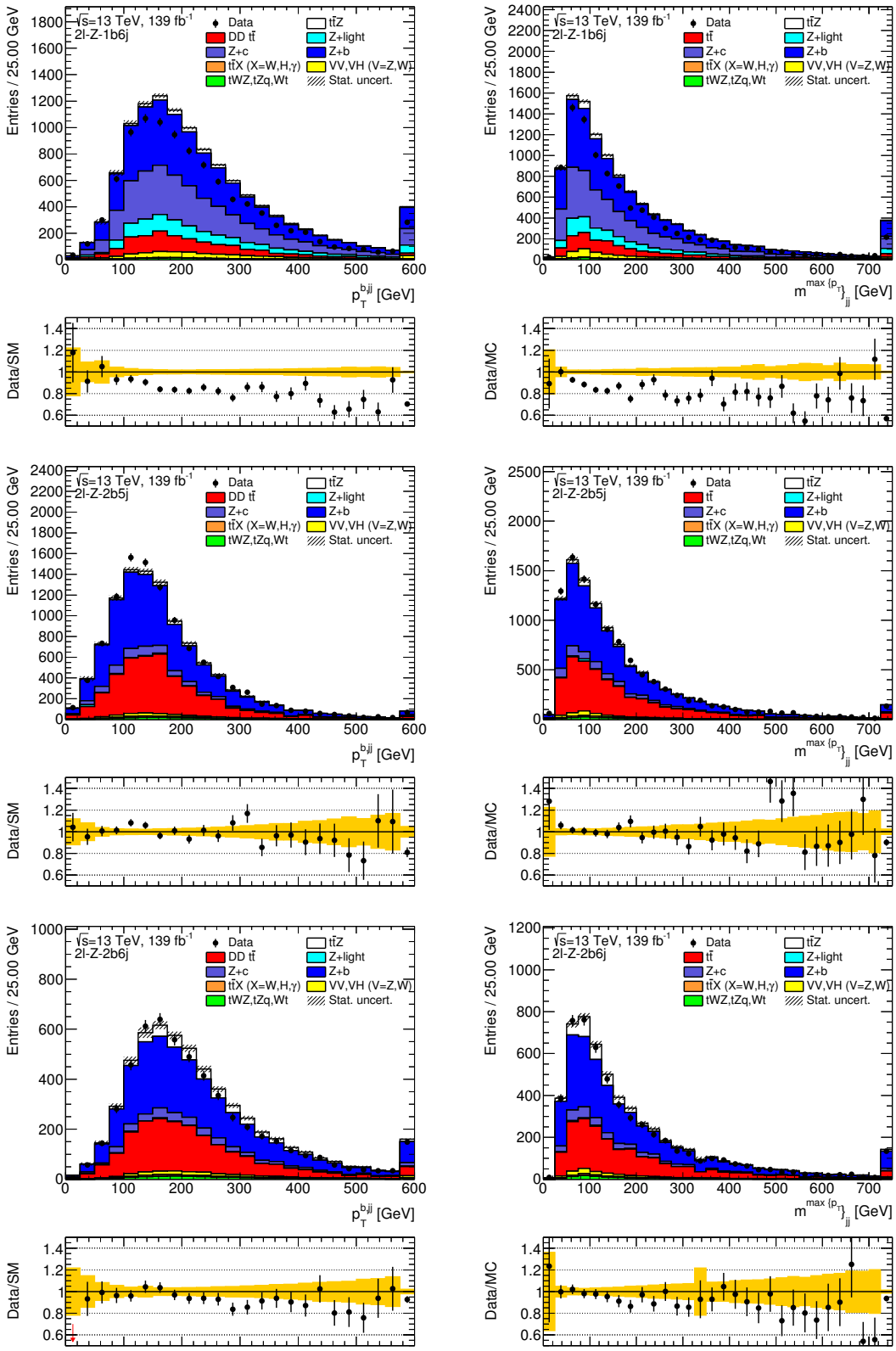


Fig. F.59.: Distributions of $p_{\mathrm{T}}^{b_1 jj}$ on the left and of $m_{jj}^{\max p_{\mathrm{T}}}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of 139 fb^{-1}. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

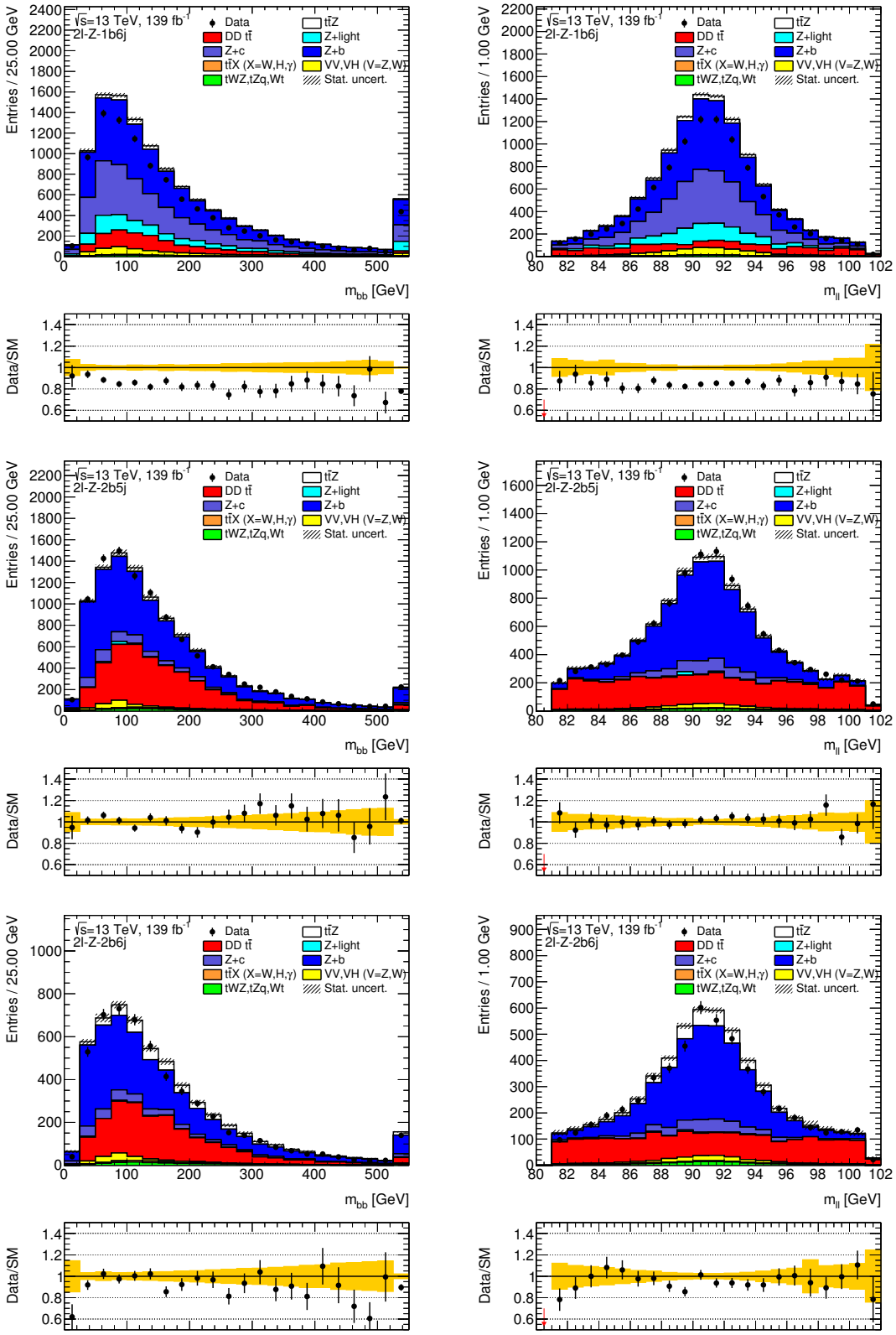


Fig. F.60.: Distributions of m_{bb} on the left and of $m_{\ell\ell}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

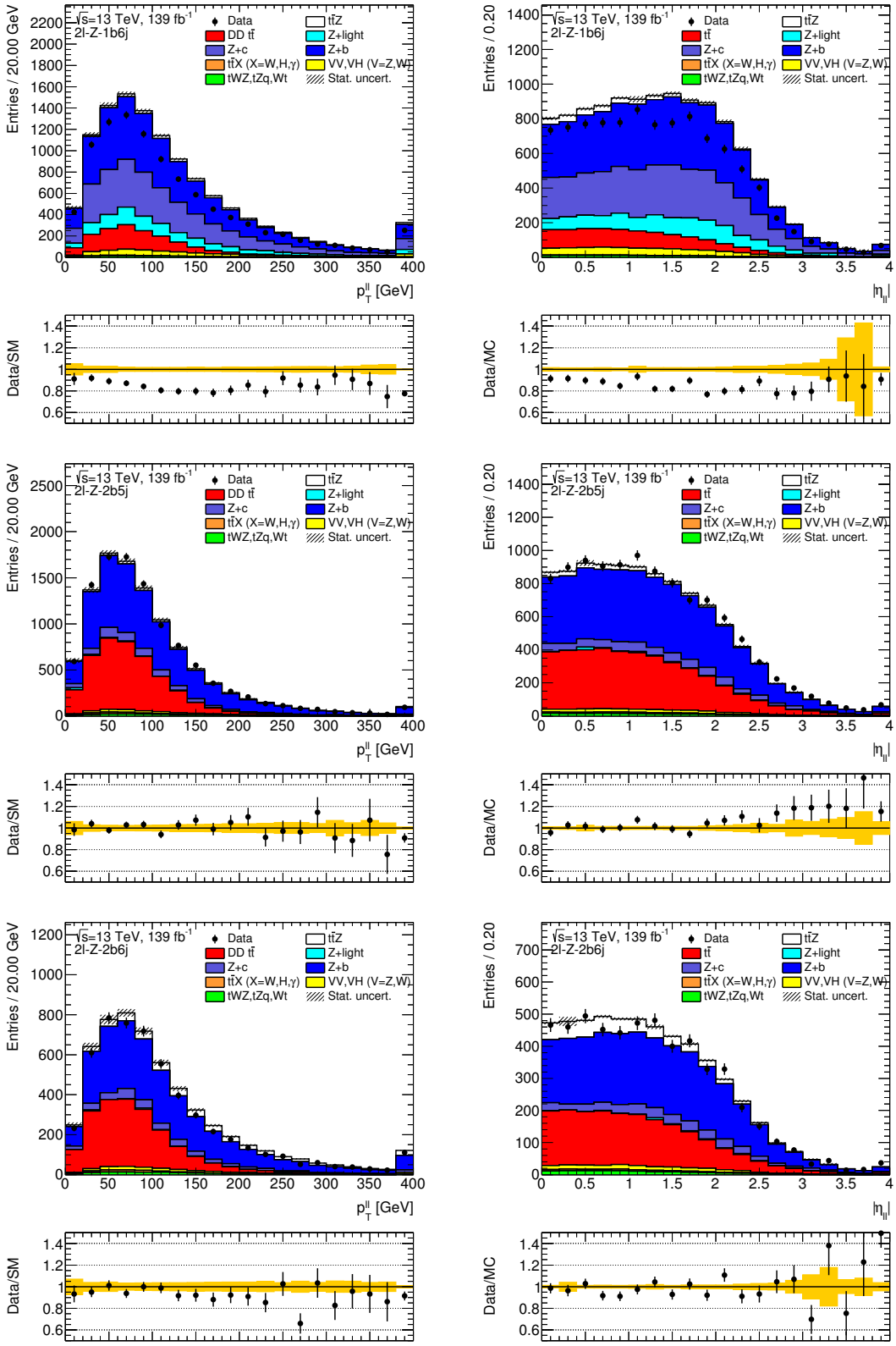


Fig. F.61.: Distributions of $p_{\mathrm{T}}^{\ell\ell}$ on the left and of $|\eta_{\ell\ell}|$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

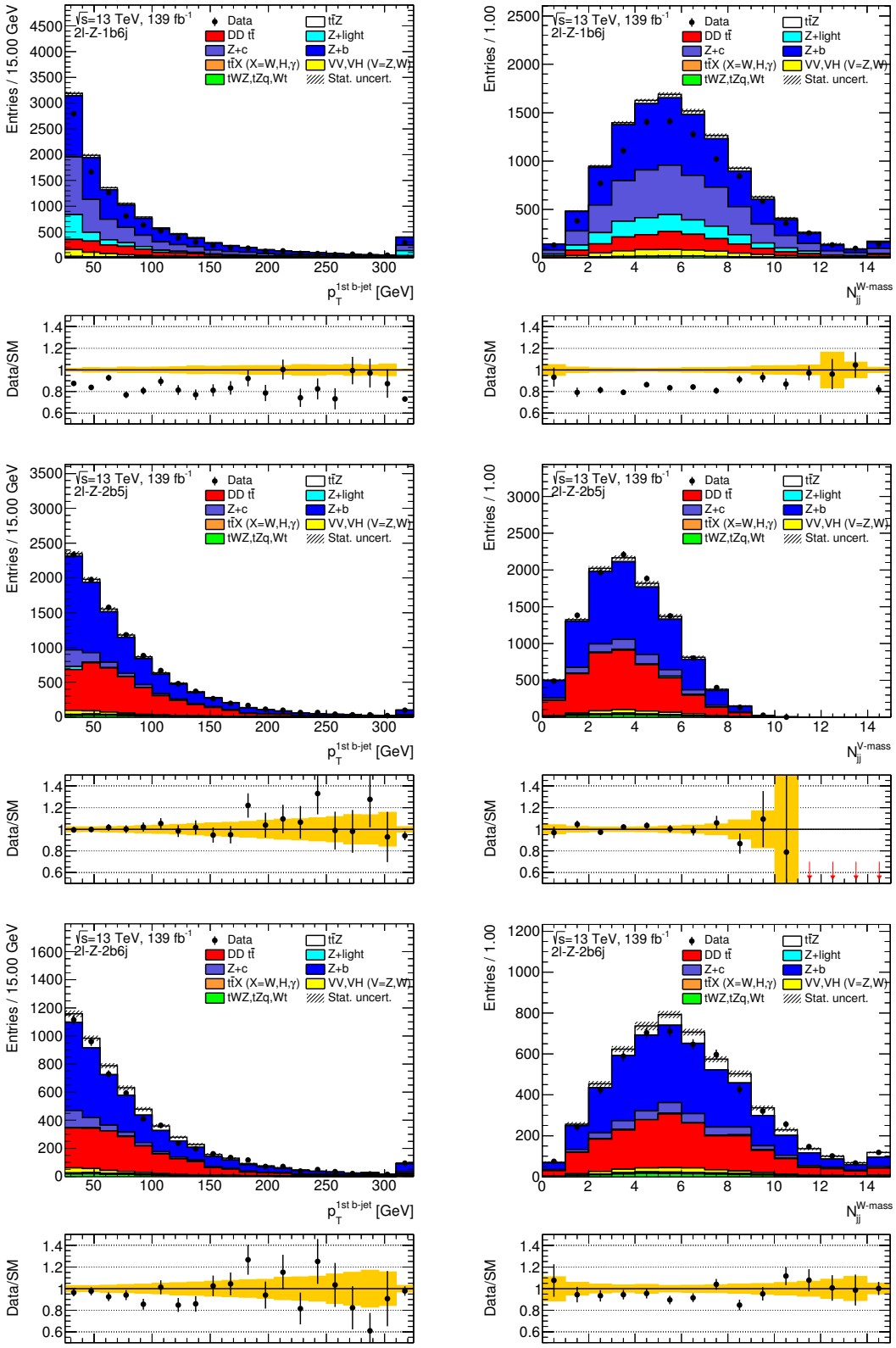


Fig. F.62.: Distributions of $p_{\mathrm{T}}^{1\text{st }b\text{-jet}}$ on the left and of $N_{jj}^{W\text{-mass}}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

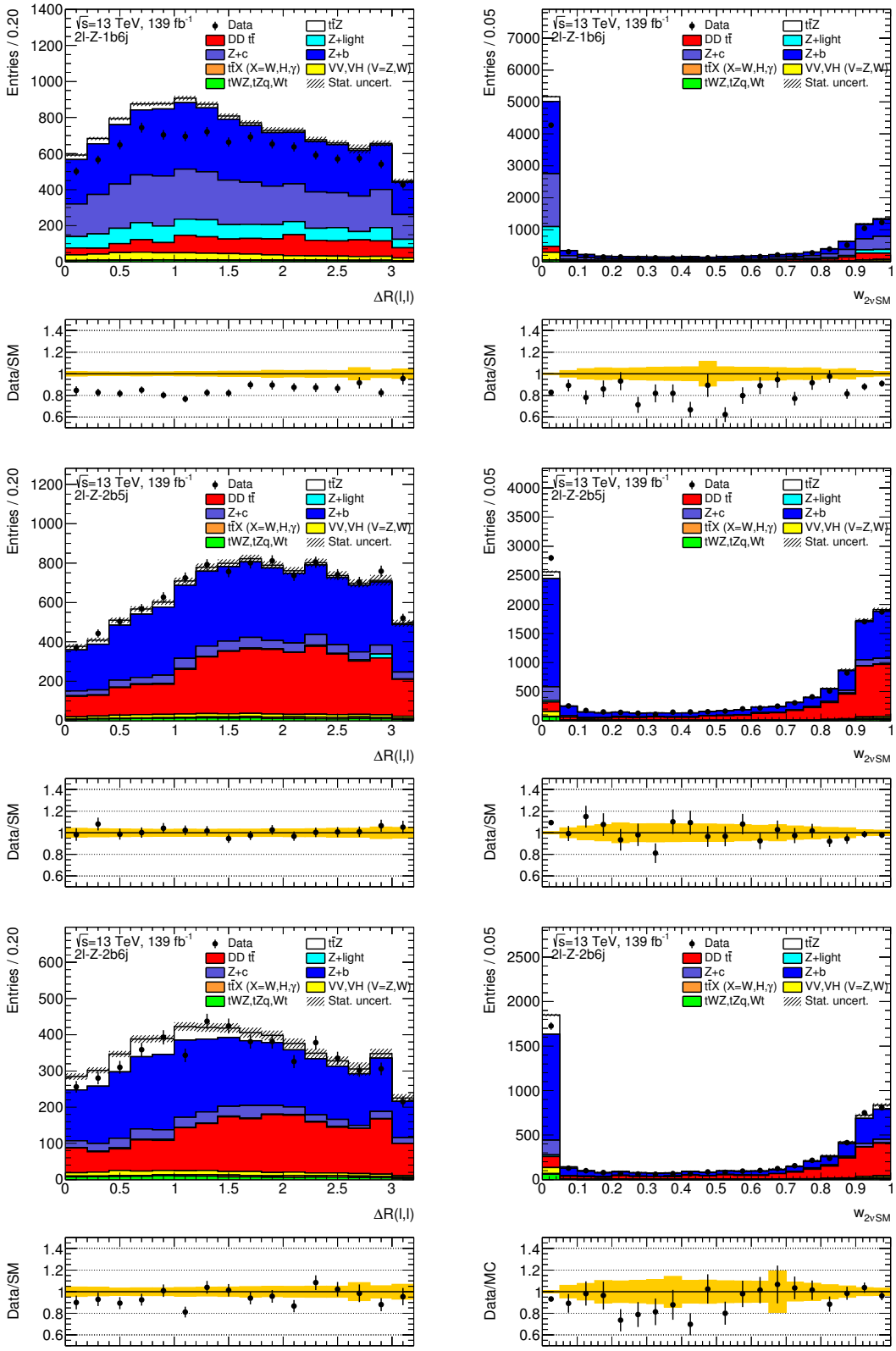
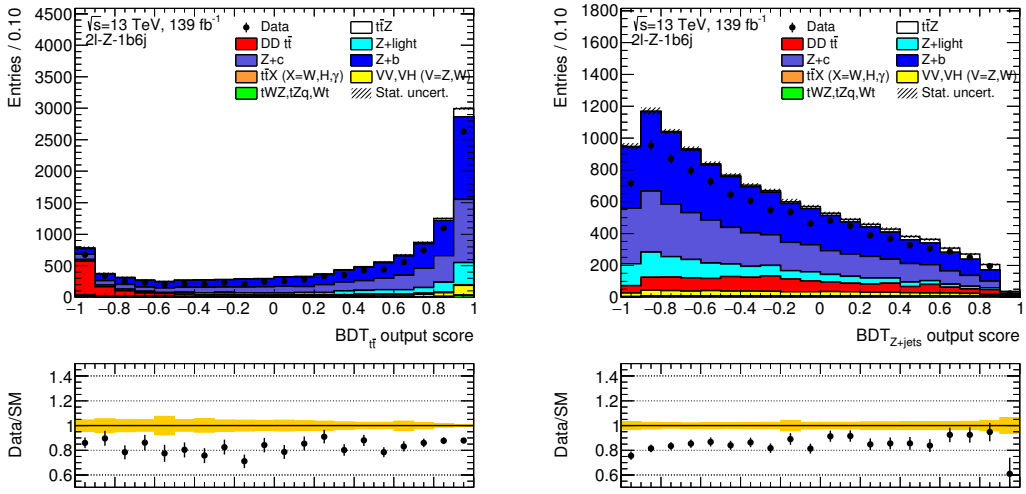


Fig. F.63.: Distributions of $\Delta R(\ell,\ell)$ on the left and of $w_{2\nu\mathrm{SM}}$ on the right. The 2ℓ-Z-1b6j, 2ℓ-Z-2b5j and 2ℓ-Z-2b6j regions are depicted in the top, middle and bottom row, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of 139 fb^{-1}. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

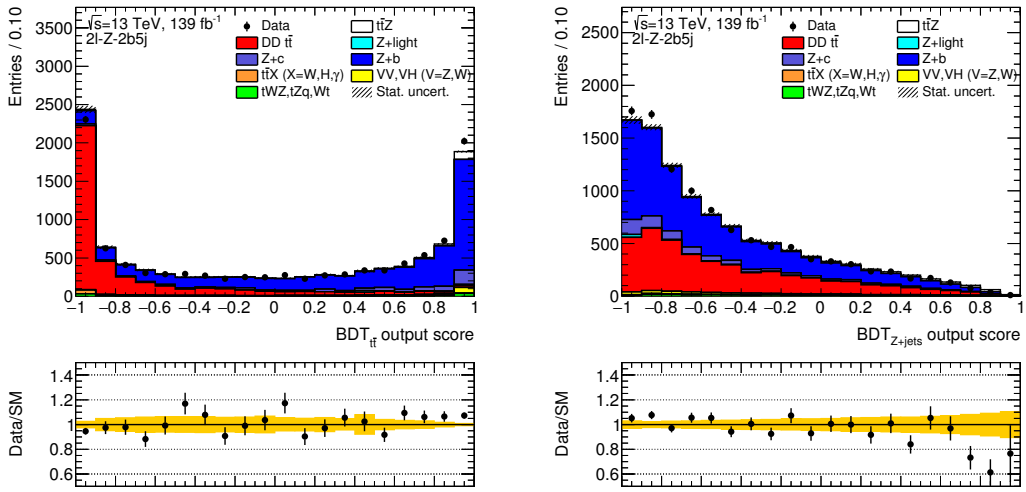
F.11. Data-MC modelling of the BDT output scores

The distributions of the BDT output score after the trained classifiers were applied to data and simulated events were shown in Fig. 7.11 for the 2ℓ-Z-2b6j target region. In this section, the corresponding figures for the 2ℓ-Z-1b6j and 2ℓ-Z-2b5j regions are presented.



(a) (b)

Fig. F.64.: Distributions of the BDT output scores within the 2ℓ-Z-1b6j target region. Fig. (a) on the left shows the score of the classifier trained against $t\bar{t}$, and Fig. (b) on the right presents the BDT for the discrimination against the Z+jets background. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.



(a) (b)

Fig. F.65.: Distributions of the BDT output scores within the 2ℓ-Z-2b5j target region. Fig. (a) on the left shows the score of the classifier trained against $t\bar{t}$, and Fig. (b) on the right presents the BDT for the discrimination against the Z+jets background. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

F.12. Definition of 2D-regions

In this section, the definition of the 2D-regions for the 2ℓ-Z-1b6j and 2ℓ-Z-2b5j target regions are illustrated. In addition, the purity of both the signal events and of either of the two dominant background process within the two-dimensional plane spanned by the BDT output scores are shown.

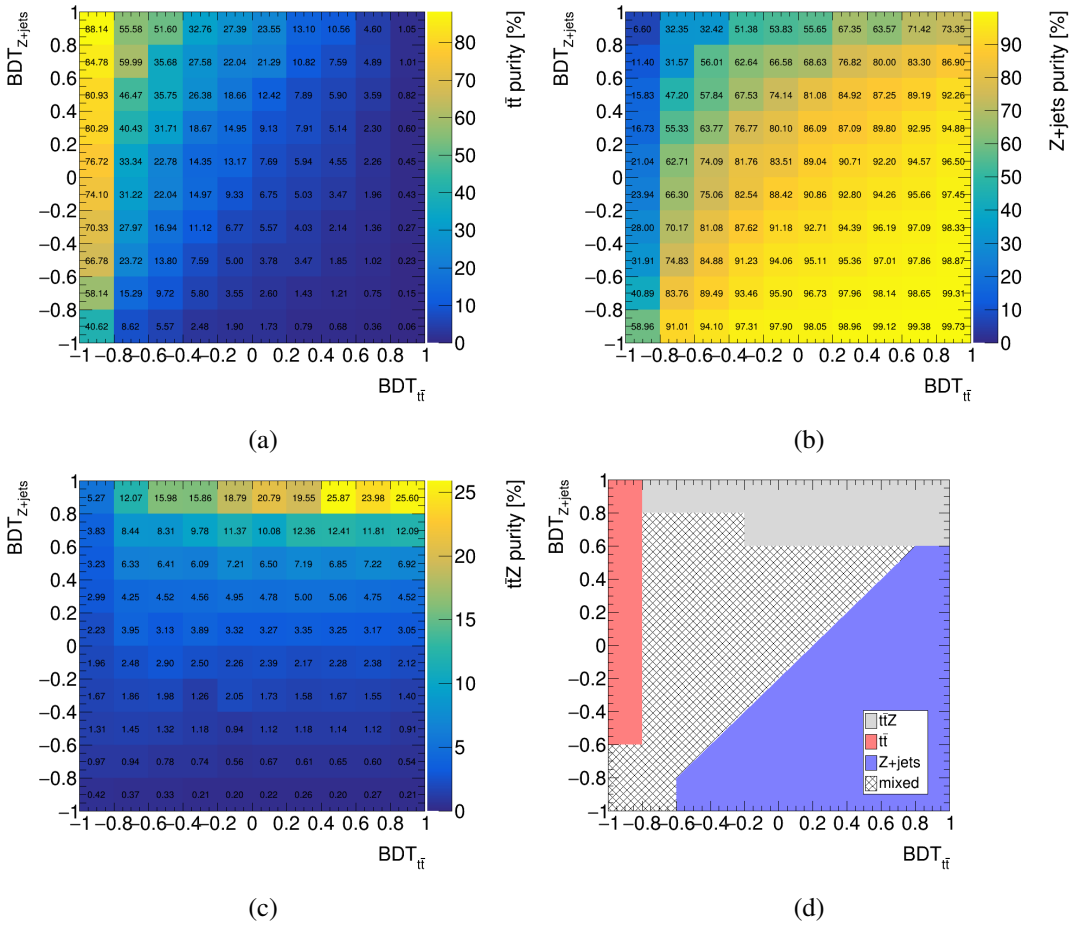


Fig. F.66.: Depiction of the purities of the various processes and boundary lines for the two-dimensional space constructed from the separate BDT outputs for the 2ℓ-Z-1b6j target region. The Figs. (a), (b) and (c) depict the purity of the $t\bar{t}$, Z+jets and $t\bar{t}Z$ events within the phase space spanned by the two BDT output scores. In the bottom right corner, however, Fig. (d) shows the various phase space regions enriched in the respective physics processes which were determined based on Figs. (a)–(c).

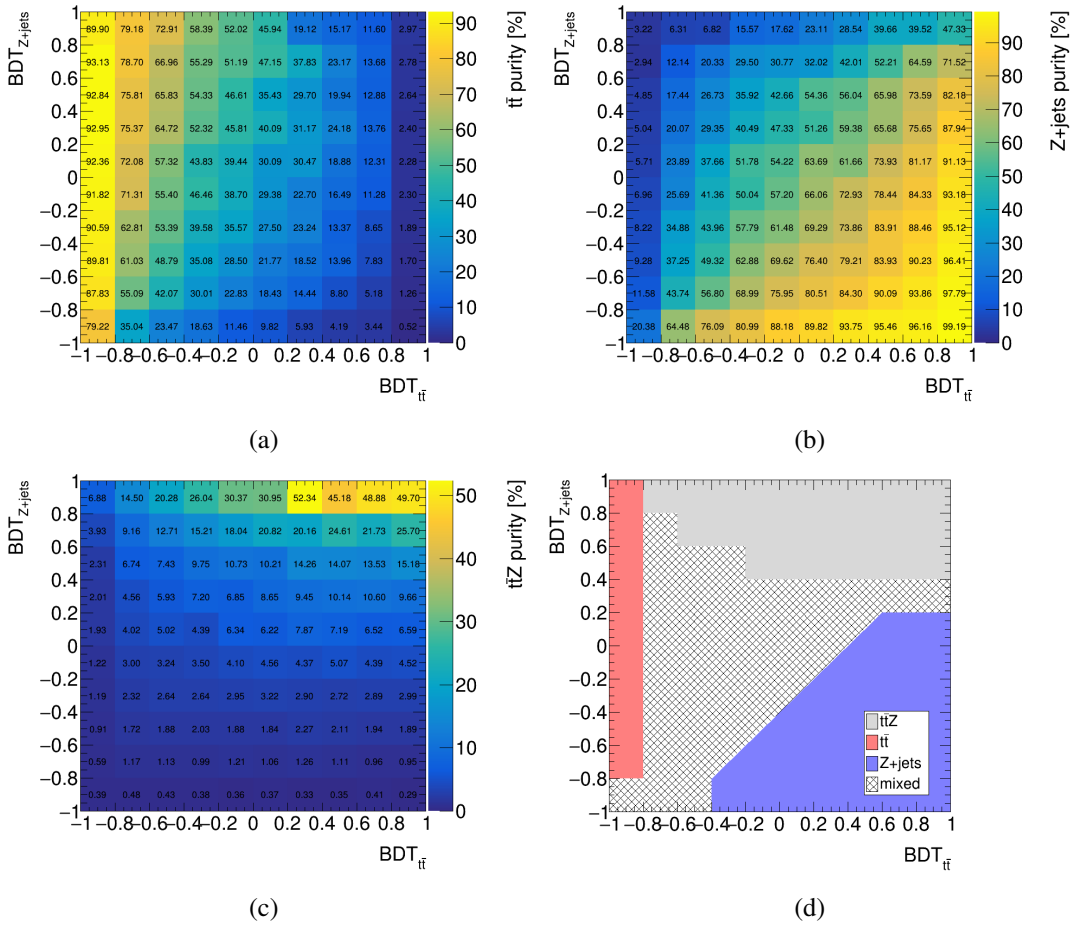


Fig. F.67.: Depiction of the purities of the various processes and boundary lines for the two-dimensional space constructed from the separate BDT outputs for the 2ℓ-Z-2b5j target region. The Figs. (a), (b) and (c) depict the purity of the $t\bar{t}$, Z+jets and $t\bar{t}Z$ events within the phase space spanned by the two BDT output scores. In the bottom right corner, however, Fig. (d) shows the various phase space regions enriched in the respective physics processes which were determined based on Figs. (a)–(c).

F.13. Data-MC modelling in the 2D-regions

In addition to those shown in Fig. 7.13, several kinematic variables were investigated in order to check the modelling within the 2D-regions defined in the two-dimensional of the BDT output scores.

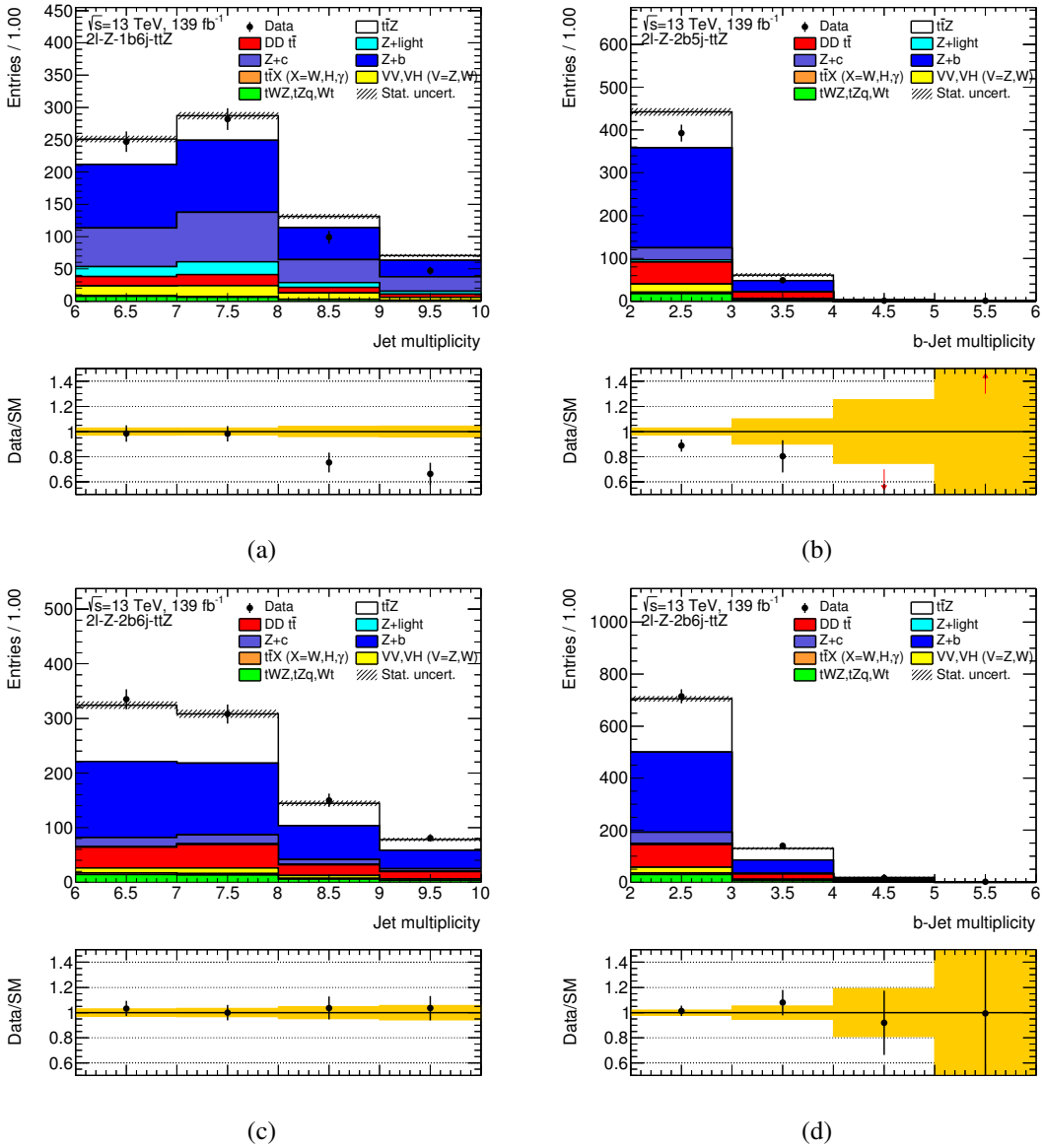


Fig. F.68.: Distributions of the jet multiplicity shown in Fig. (a) for the 2ℓ-Z-1b6j-$t\bar{t}Z$ region and in Fig. (c) for the 2ℓ-Z-2b6j-$t\bar{t}Z$ region. The b-jet multiplicity is presented in Figs. (b) and (d), depicting the 2ℓ-Z-2b5j-$t\bar{t}Z$ and 2ℓ-Z-2b6j-$t\bar{t}Z$ 2D-region, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

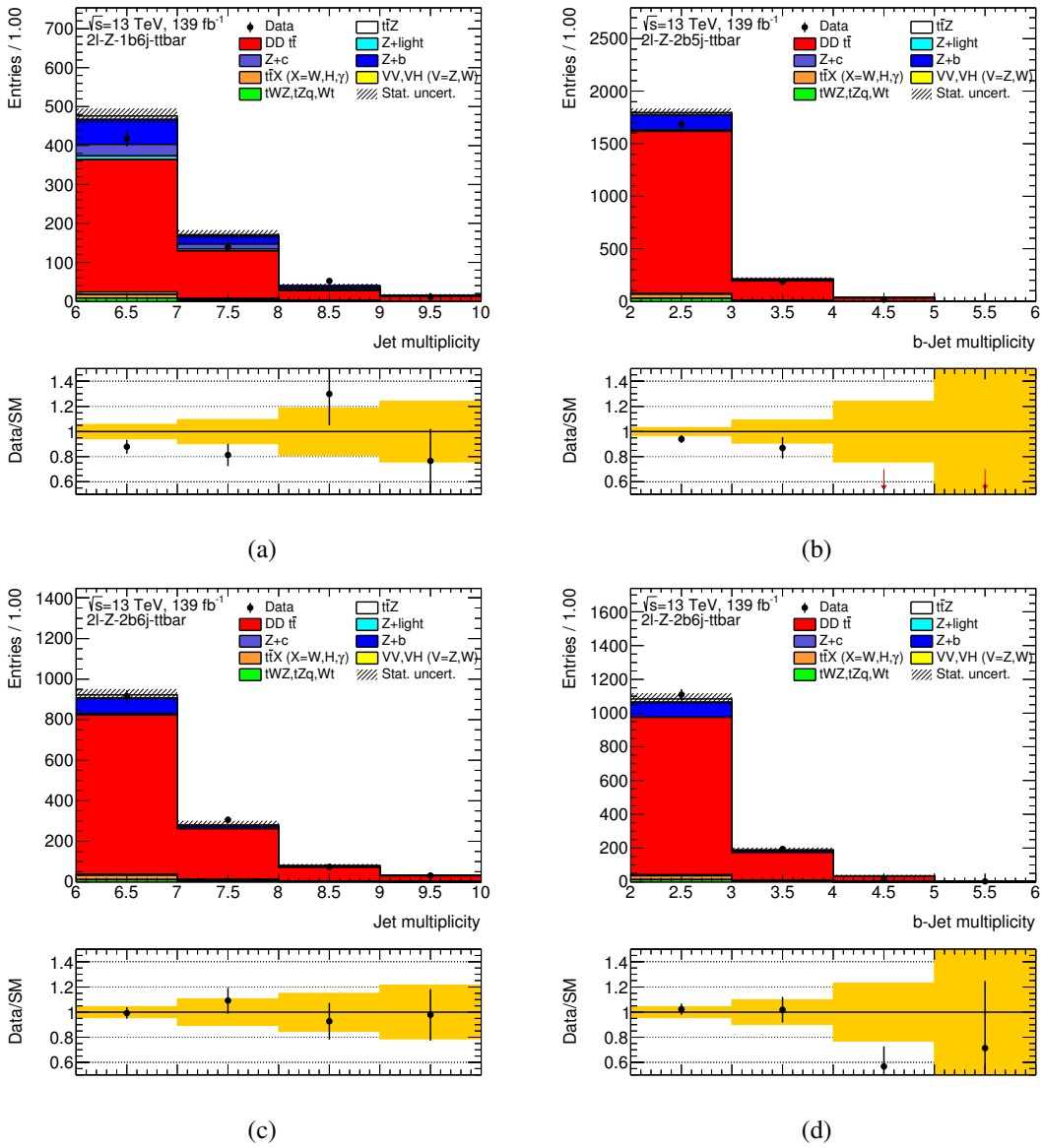


Fig. F.69.: Distributions of the jet multiplicity shown in Fig. (a) for the 2ℓ-Z-1b6j-$t\bar{t}$ region and in Fig. (c) for the 2ℓ-Z-2b6j-$t\bar{t}$ region. The b-jet multiplicity is presented in Figs. (b) and (d), depicting the 2ℓ-Z-2b5j-$t\bar{t}$ and 2ℓ-Z-2b6j-$t\bar{t}$ 2D-region, respectively. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

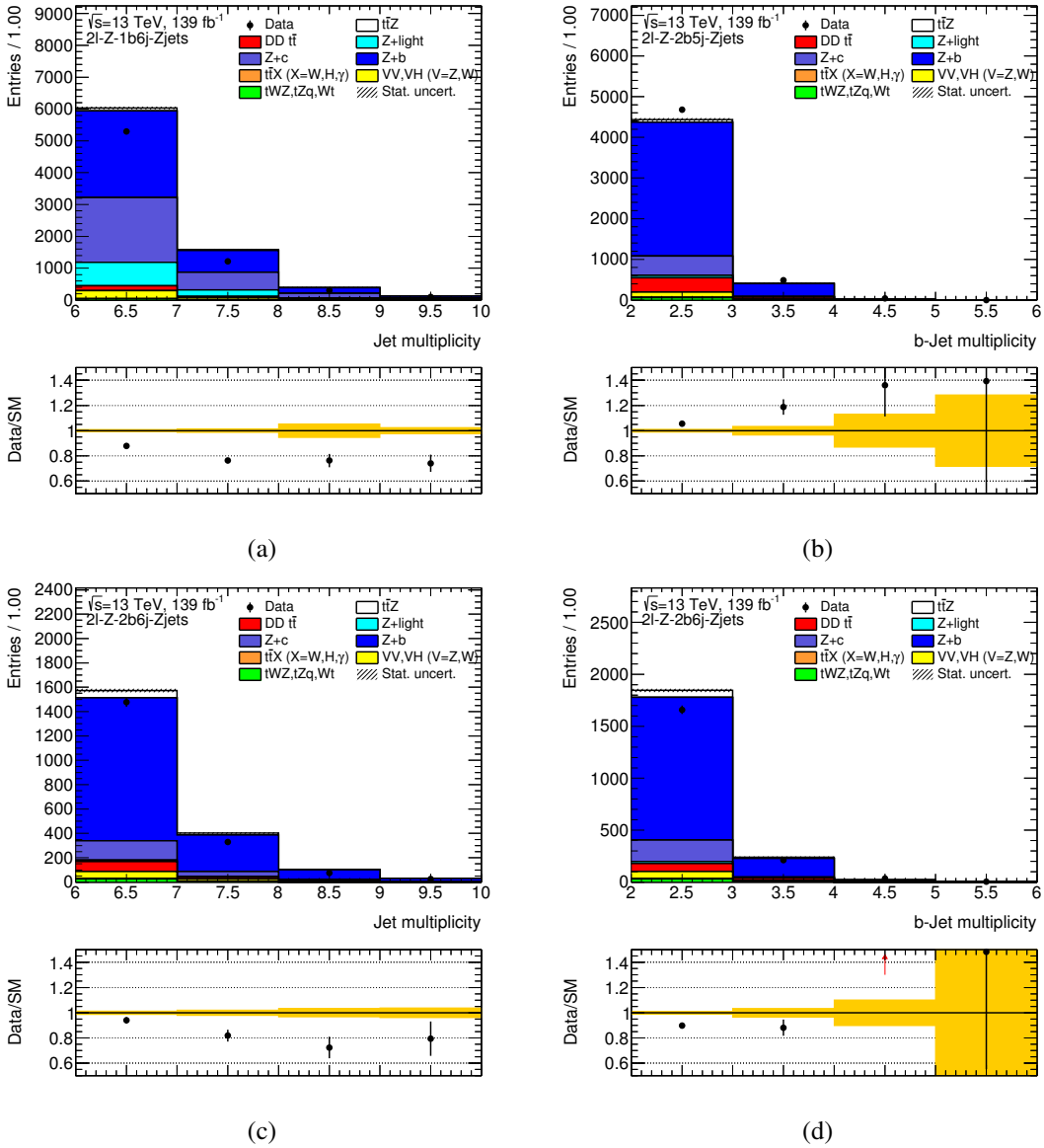


Fig. F.70.: Distributions of the jet multiplicity shown in Fig. (a) for the 2ℓ-Z-1b6j-Z+jets region and in Fig. (c) for the 2ℓ-Z-2b6j-Z+jets region. The b-jet multiplicity is presented in Figs. (b) and (d), depicting the 2ℓ-Z-2b5j-Z+jets and 2ℓ-Z-2b6j-Z+jets 2D-region, respectively.

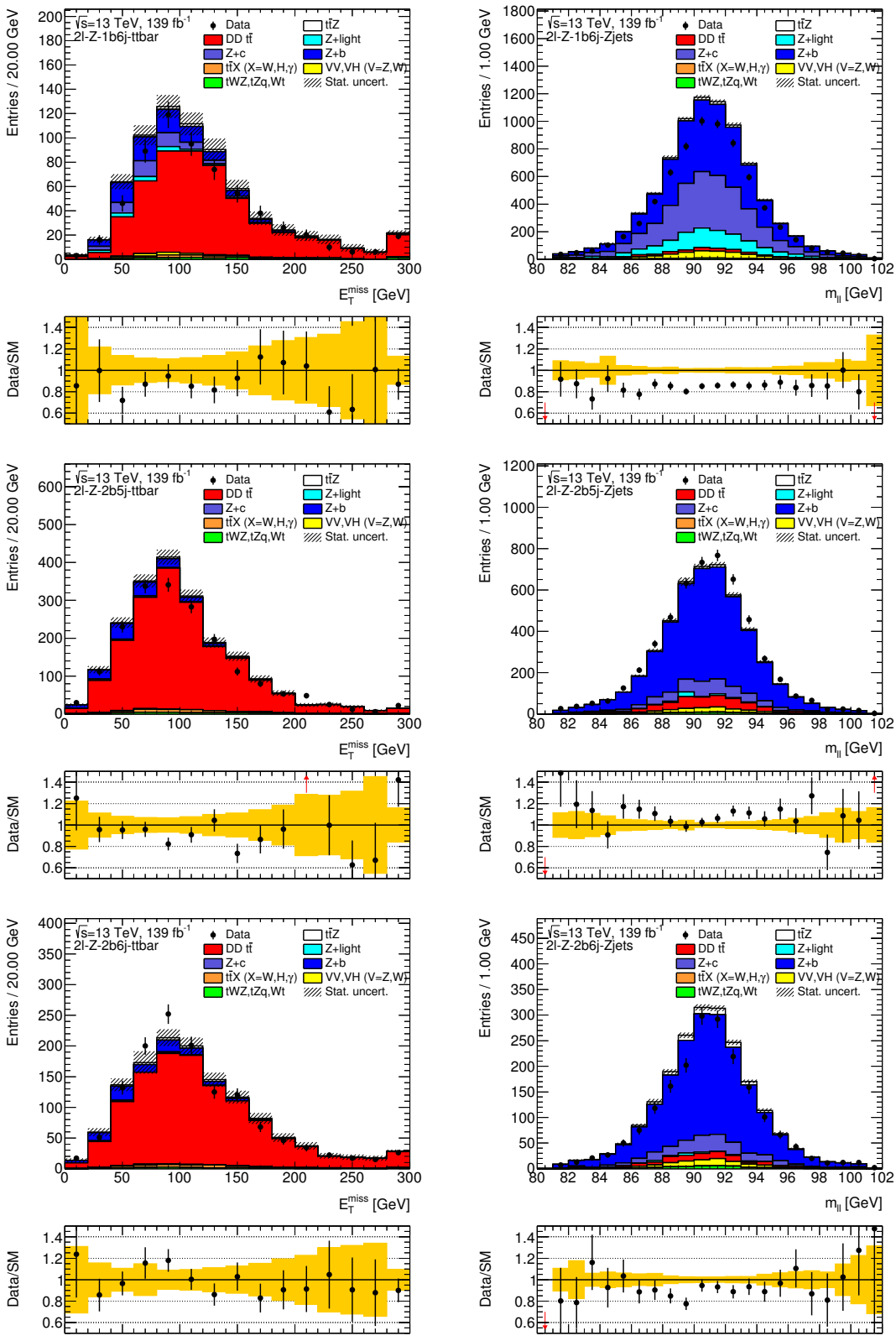


Fig. F.71.: Distributions of the missing transverse energy in the left column and of the invariant mass of the lepton pair in the right column, shown for the three $t\bar{t}Z$-enriched 2D-regions. The 2ℓ-Z-1b6j-$t\bar{t}$, 2ℓ-Z-2b5j-$t\bar{t}$ and 2ℓ-Z-2b6j-$t\bar{t}$ regions are depicted in the top, middle and bottom row of the left columns, respectively. In the right column, the three Z+jets-enriched 2D-regions are presented in the same order.

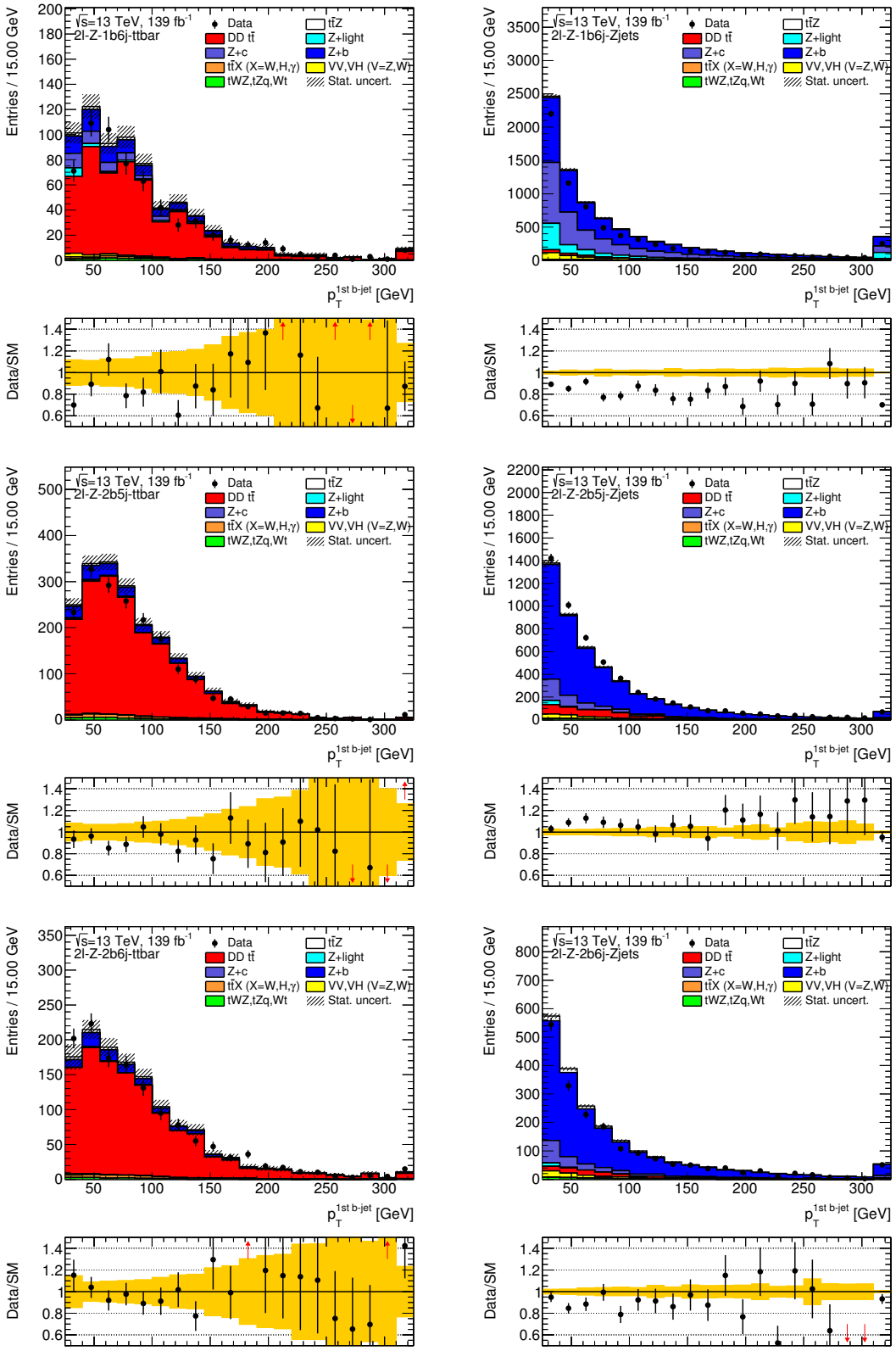
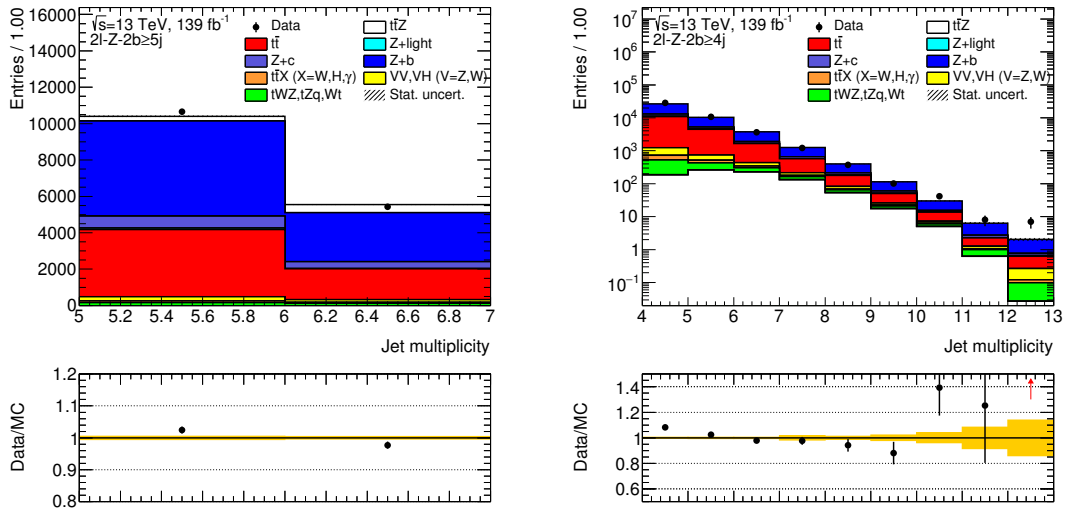


Fig. F.72.: Distributions of the the first b-jet p_{T}, shown for the three $t\bar{t}$-enriched 2D-regions in the left column and for the Z+jets-enriched region in the right column. The 2ℓ-Z-1b6j-$t\bar{t}$, 2ℓ-Z-2b5j-$t\bar{t}$ and 2ℓ-Z-2b6j-$t\bar{t}$ regions are depicted in the top, middle and bottom row of the left columns, respectively. In the right column, the three Z+jets-enriched 2D-regions are presented in the same order. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

F.14. Data-MC modelling with respect to the jet multiplicity

For the measurement of the $t\bar{t}Z$ production cross section in the 2ℓOS channel, normalisation factors for the Z+b background were accounted for in the fit by a floating parameter. Due to the different kinematics of the events with exactly five and at least six jets, a separate normalisation factor was introduced for the Z+b background contributions in the 2ℓ-Z-2b5j-Z+jets and 2ℓ-Z-2b6j-Z+jets region, respectively. In order to justify this choice, distributions of the jet multiplicity were constructed based on a looser selection requirement of the number of jets in order to compare the data with the total predictions from Monte Carlo simulation. This selection includes the requirement of at least two b-jets and is dominated by the Z+jets background. In Fig. F.73a the jet multiplicity for events with least four jets rather than at least five as in the standard selection are depicted. Fig. F.73b shows the amount of events featuring exactly five jets in the left bin and with at least six jets in the right bin. The ratio of selected data to simulated events was seen to be greater than one for the events with exactly five jets, whereas events with at least six jets yield a data-MC ratio of smaller than one. In general, the modelling of the jet multiplicity features a linearly decreasing behaviour towards higher numbers of selected jets. For all processes, including the dileptonic $t\bar{t}$ background, the expected amount of events are taken directly from the Monte Carlo simulations.



(a) (b)

Fig. F.73.: Distributions of the jet multiplicity shown for events with at least two b-jets. On the left Fig. (a) shows the amount of selected events with exactly five jets and at least six jets in the left and right bin, respectively. The second Fig. (b) shows the jet multiplicity for events with at least two selected b-jets. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped and the yellow band, respectively, represent the statistical uncertainties from Monte Carlo simulation.

F.15. Pre-fit observed and expected yields in the validation regions

The corresponding illustration of observed and expected pre-fit yields in the signal and control regions can be found in Fig. 7.14.

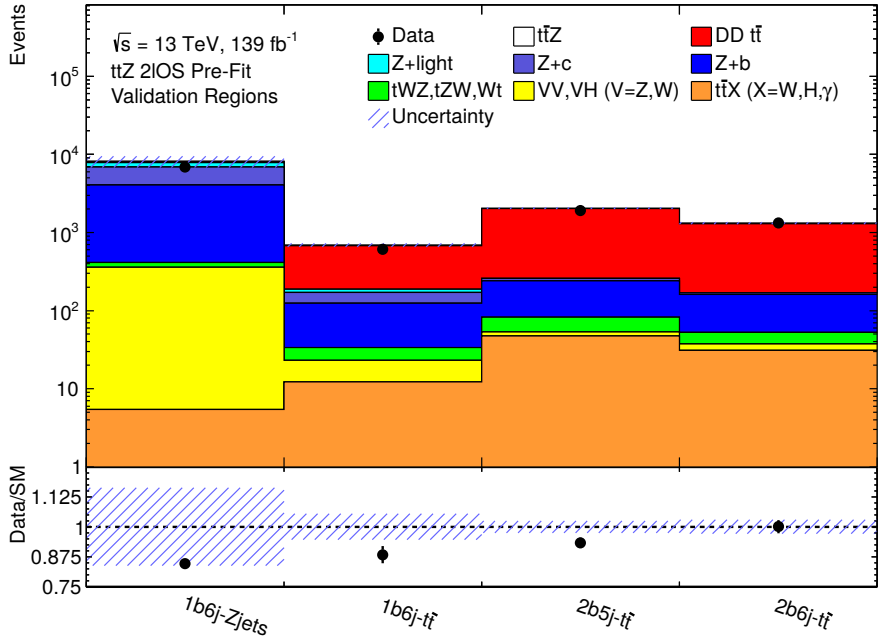


Fig. F.74.: Observed and expected event yields in the 2ℓOS validation regions prior to the fit. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped bands represent the combined statistical and systematic uncertainties of the Monte Carlo prediction.

F.16. Post-fit observed and expected yields in the validation regions

The corresponding illustration of observed and expected post-fit yields in the signal and control regions can be found in Fig. 7.15.

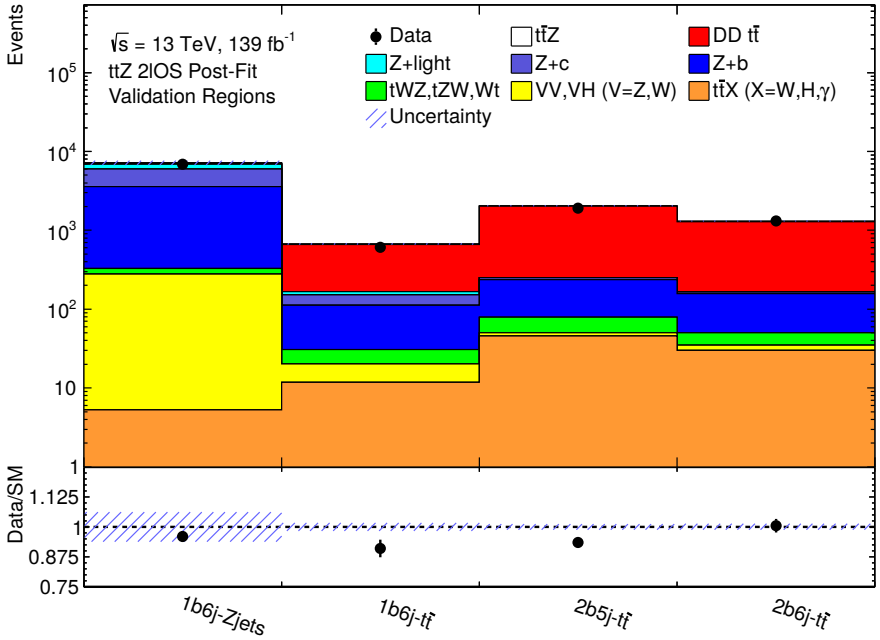


Fig. F.75.: Observed and expected event yields in the 2ℓOS validation regions after the fit. The contributions from the simulated Monte Carlo datasets were scaled to their respective theoretical cross section and to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The errors indicated by the striped bands represent the combined statistical and systematic uncertainties of the Monte Carlo prediction.

F.17. Nuisance parameters

In Fig. F.76 the pulls of the full set of nuisance parameters associated with experimental and theoretical uncertainties that were considered in the fit for the $t\bar{t}Z$ cross section measurement in the 2ℓOS channel are depicted. The green and yellow band indicate the range of one and two standard deviations of the Gaussian parametrisation of the systematic uncertainties. All nuisance parameters were sen to be constrained by at most half a standard deviation. The various sources of systematic uncertainties accounting for the normalisation of a background process and for some experimental uncertainties can be directly read off the labels on the right side of the panel. Different nuisance parameters were introduced for the identification (ID), isolation (Isol), trigger and reconstruction (Reco) efficiencies of leptons. Parameters of systematic uncertainties from flavour-tagging accounting for inefficiencies of the b-tagging, c-tagging and light-flavour tagging procedures were given separately to the fit.

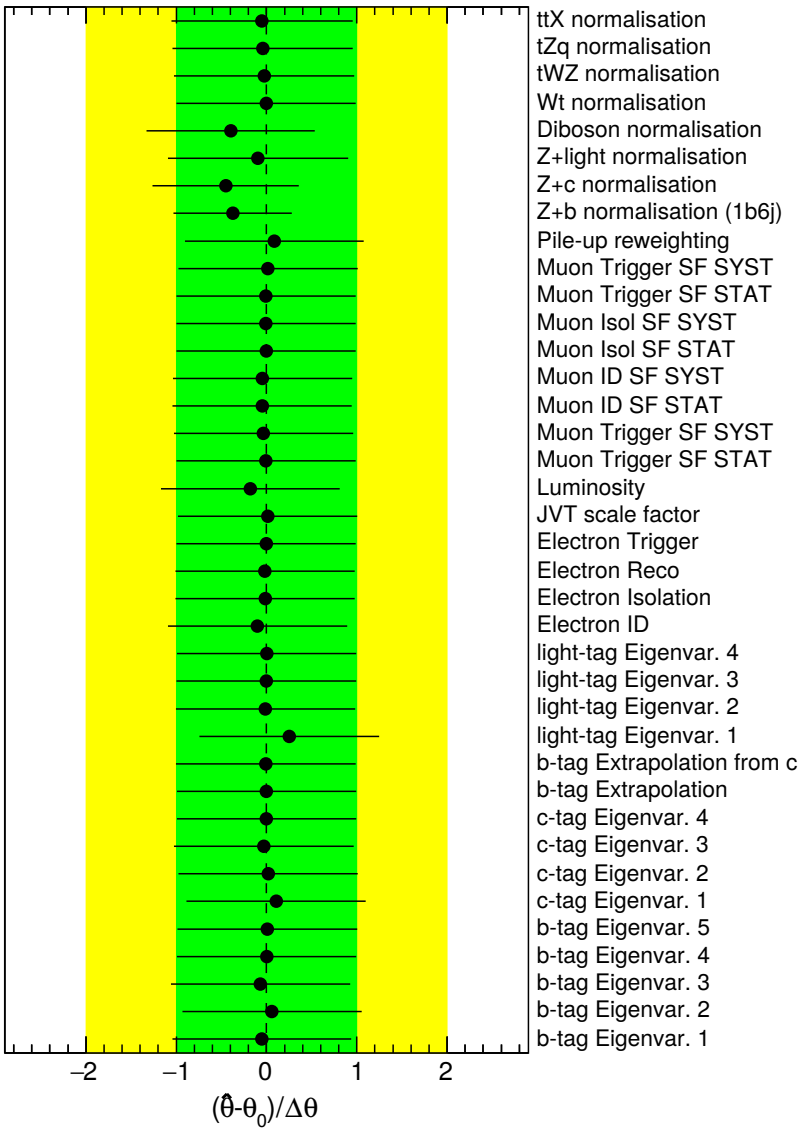


Fig. F.76.: List of the nuisance parameters and their pulls.

G. Additional tables

G.1. Variable rankings

In this section, the discriminating variables employed for the various BDT training procedures are listed, given in descending order of their separation power calculated according to Eq. (7.3). The corresponding listings presenting the variables in the order of their rankings as determined by the various training procedures can be found in Tabs. 7.9 and 7.10

Rank	2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
1	$E_{\mathrm{T}}^{\mathrm{miss}}$	$m_{b\ell}^{\mathrm{min}}$	$E_{\mathrm{T}}^{\mathrm{miss}}$
2	$p_{\mathrm{T}}^{\ell\ell}$	$E_{\mathrm{T}}^{\mathrm{miss}}$	$m_{b\ell}^{\mathrm{min}}$
3	$m_{\ell\ell}$	$w_{2\nu\mathrm{SM}}$	$w_{2\nu\mathrm{SM}}$
4	$p_{\mathrm{T}}^{\text{1st lepton}}$	$p_{\mathrm{T}}^{\ell\ell}$	$p_{\mathrm{T}}^{\ell\ell}$
5	$m_{b\ell}^{\mathrm{min}}$	$m_{\ell\ell}$	$m_{\ell\ell}$
6	$p_{\mathrm{T}}^{\text{2nd lepton}}$	$p_{\mathrm{T}}^{\text{1st lepton}}$	$p_{\mathrm{T}}^{\text{1st lepton}}$
7	$m_{b\ell}^{\text{max-min}}$	$\max m_{b\ell}^{\mathrm{min}\,\Delta R}$	$\max m_{b\ell}^{\mathrm{min}\,\Delta R}$
8	$w_{2\nu\mathrm{SM}}$	$p_{\mathrm{T}}^{\text{2nd lepton}}$	$p_{\mathrm{T}}^{\text{2nd lepton}}$
9	$\Delta R(\ell,\ell)$	$m_{b\ell}^{\text{max-min}}$	w_{2t}
10	$m_{b\ell}^{\text{max-min}}$	$H_{\mathrm{T}}^{bb\ell\ell}$	$m_{b\ell}^{\text{max-min}}$
11	w_{1t1W}	$\Delta R(\ell,\ell)$	w_{1t1W}
12	$H_{\mathrm{T}}^{bb\ell\ell}$	$\Delta R(bb,\ell\ell)$	$\Delta R(\ell,\ell)$
13	w_{2t}	w_{1t}	w_{1t}
14	w_{1t}	w_{1t1W}	$\Delta R(bb,\ell\ell)$
15	w_{2W}	w_{2W}	$H_{\mathrm{T}}^{bb\ell\ell}$
16	w_{1W}	w_{1W}	w_{2W}
17	$m_{bjj}^{\text{top-mass}}$	$m_{bjj}^{\text{top-mass}}$	$m_{bjj}^{\text{top-mass}}$
18	$N_{jj}^{W\text{-mass}}$	–	w_{1W}
19	–	–	$N_{jj}^{W\text{-mass}}$

Tab. G.1.: Ranking of input variables used to train the BDT against the $t\bar{t}$ background as a result of the training. The rankings are based on the amount of overlap of signal and background distributions, as calculated according to Eq. (7.3).

Rank	2ℓ-Z-1b6j	2ℓ-Z-2b5j	2ℓ-Z-2b6j
1	w_{1t1W}	$p_{\mathrm{T}}^{\text{4th jet}}$	w_{2t}
2	$\lvert\eta_{\ell\ell}\rvert$	$p_{\mathrm{T}}^{\text{3rd jet}}$	w_{1t1W}
3	w_{2t}	$m_{jj}^{\min \Delta R}$	w_{1t}
4	w_{1t}	$\Delta R_{jj}^{\text{avg}}$	$\lvert\eta_{\ell\ell}\rvert$
5	$p_{\mathrm{T}}^{\text{5th jet}}$	$H_{\mathrm{T}}^{bb\ell\ell}$	$p_{\mathrm{T}}^{\text{5th jet}}$
6	$p_{\mathrm{T}}^{\text{4th jet}}$	H_{T}	$\Delta R_{jj}^{\text{avg}}$
7	w_{2W}	$p_{\mathrm{T}}^{\text{5th jet}}$	$p_{\mathrm{T}}^{\text{6th jet}}$
8	$H_{\mathrm{T}}^{bb\ell\ell}$	$p_{\mathrm{T}}^{\text{2nd jet}}$	$p_{\mathrm{T}}^{\text{4th jet}}$
9	$\Delta R_{jj}^{\text{avg}}$	$m_{jj}^{\max p_{\mathrm{T}}}$	w_{2W}
10	w_{1W}	$p_{\mathrm{T}}^{b_1 jj}$	m_{bb}
11	$m_{bjj}^{\text{top-mass}}$	$p_{\mathrm{T}}^{\ell\ell}$	$m_{bjj}^{\text{top-mass}}$
12	$p_{\mathrm{T}}^{\text{1st } b\text{-jet}}$	m_{bb}	w_{1W}
13	$N_{bjj}^{\text{top-mass}}$	w_{1t}	$H_{\mathrm{T}}^{bb\ell\ell}$
14	$N_{jj}^{W\text{-mass}}$	$p_{\mathrm{T}}^{\text{1st } b\text{-jet}}$	$H_{\mathrm{T}}^{bb\ell\ell}$
15	–	w_{1t1W}	$p_{\mathrm{T}}^{\text{1st } b\text{-jet}}$
16	–	$m_{bjj}^{\text{top-mass}}$	$N_{bjj}^{\text{top-mass}}$
17	–	w_{1W}	$N_{jj}^{W\text{-mass}}$
18	–	w_{2W}	–

Tab. G.2.: Ranking of input variables used to train the BDT against the Z+jets background as a result of the training. The rankings are based on the amount of overlap of signal and background distributions, as calculated according to Eq. (7.3).

G.2. Observed and expected event yields in the 2D-regions

The expected yields of the 2D-regions enriched in $t\bar{t}$ and Z+jets events are given within this section. The numbers corresponding to simulated processes were normalised to the theoretical cross section of the respective process and were scaled to an integrated luminosity of $139\,\mathrm{fb}^{-1}$.

Process	2ℓ-Z-1b6j-$t\bar{t}$	2ℓ-Z-2b5j-$t\bar{t}$	2ℓ-Z-2b6j-$t\bar{t}$
$t\bar{t}Z$	13.77 ± 0.50	23.23 ± 0.61	24.71 ± 0.71
DD $t\bar{t}$	492 ± 23	1794 ± 45	1126 ± 34
Z+light flavour	15.8 ± 1.7	0.57 ± 0.14	–
Z+c	47 ± 25	15.7 ± 1.7	9.38 ± 1.03
Z+b	90.5 ± 2.8	158.6 ± 2.6	107.6 ± 1.9
$t\bar{t}X$ $(X = W, H, \gamma)$	12.21 ± 0.56	47.3 ± 1.2	31.04 ± 0.86
VV, VH $(V = Z, W)$	10.92 ± 0.60	6.00 ± 0.51	6.22 ± 0.47
Other (tWZ, tZq, Wt)	10.4 ± 1.0	29.2 ± 1.8	15.4 ± 1.3
Total SM	693 ± 23	2045 ± 43	1320 ± 34
Data	612	1907	1322
Data/SM	0.88	0.92	1.00
$t\bar{t}$ purity	71 %	86 %	85 %

Tab. G.3.: Estimated and measured yields from simulation as well as those from measured data of the $t\bar{t}$-enriched 2D-regions of the 2ℓOS analysis. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. Only the Monte Carlo statistical errors are shown.

Process	2ℓ-Z-1b6j-Z+jets	2ℓ-Z-2b5j-Z+jets	2ℓ-Z-2b6j-Z+jets
$t\bar{t}Z$	130.04 ± 0.93	71.63 ± 0.62	78.04 ± 0.70
DD $t\bar{t}$	186 ± 14	401 ± 21	100 ± 10
Z+light flavour	990 ± 26	60 ± 16	19.6 ± 4.0
Z+c	2814 ± 27	490 ± 19	219.8 ± 7.1
Z+b	3633 ± 14	3633 ± 19	1575.5 ± 8.7
$t\bar{t}X$ $(X = W, H, \gamma)$	5.45 ± 0.37	9.59 ± 0.50	3.68 ± 0.26
VV, VH $(V = Z, W)$	354.3 ± 3.9	149.5 ± 3.0	83.3 ± 2.1
Other (tWZ, tZq, Wt)	52.2 ± 1.2	62.6 ± 1.6	35.11 ± 0.82
Total SM	8166 ± 43	4878 ± 38	2114 ± 16
Data	6911	5186	1926
Data/SM	0.85	1.06	0.91
Z+jets purity	91 %	86 %	81 %
Z+b purity	44 %	74 %	75 %

Tab. G.4.: Estimated and measured yields from simulation as well as those from measured data of the Z+jets-enriched 2D-regions of the 2ℓOS analysis. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. Only the Monte Carlo statistical errors are shown.

G.3. Axis labels for two-dimensional linear correlation plots

As the axis labels of Figs. F.22–F.33 of Sec. F.8 would have overlapped with each other if they had been made bigger, the BDT training variables are listed in Tab. G.5 per target region and background process. Each column comprises the variables of another training procedure and is ordered from the top to the bottom as described in the following: starting in the bottom left corner of each individual correlation plot, the variables appear as one goes towards to end of the x- and y-axis, respectively.

2ℓ-Z-1b6j-$t\bar{t}Z$		2ℓ-Z-2b5j-$t\bar{t}Z$		2ℓ-Z-2b6j-$t\bar{t}Z$	
$t\bar{t}$	Z+jets	$t\bar{t}$	Z+jets	$t\bar{t}$	Z+jets
(Figs.F.22, F.23)	(Figs.F.24, F.25)	(Figs.F.26, F.27)	(Figs.F.28, F.29)	(Figs.F.30, F.31)	(Figs.F.32, F.33)
w_{1W}	w_{1W}	w_{1W}	w_{1W}	w_{1W}	w_{1W}
w_{1t}	w_{1t}	w_{1t}	w_{1t}	w_{1t}	w_{1t}
w_{2W}	w_{2W}	w_{2W}	w_{2W}	w_{2W}	w_{2W}
w_{1t1W}	w_{1t1W}	w_{1t1W}	w_{1t1W}	w_{1t1W}	w_{1t1W}
$w_{2\nu\text{SM}}$	$p_{\text{T}}^{\text{4th jet}}$	$w_{2\nu\text{SM}}$	$p_{\text{T}}^{\text{4th jet}}$	$w_{2\nu\text{SM}}$	$p_{\text{T}}^{\text{4th jet}}$
$p_{\text{T}}^{\text{1st lepton}}$	$p_{\text{T}}^{\text{5th jet}}$	$p_{\text{T}}^{\text{1st lepton}}$	$p_{\text{T}}^{\text{5th jet}}$	$p_{\text{T}}^{\text{1st lepton}}$	$p_{\text{T}}^{\text{5th jet}}$
$p_{\text{T}}^{\text{2nd lepton}}$	$p_{\text{T}}^{\text{1st }b\text{-jet}}$	$p_{\text{T}}^{\text{2nd lepton}}$	$p_{\text{T}}^{\text{1st }b\text{-jet}}$	$p_{\text{T}}^{\text{2nd lepton}}$	$p_{\text{T}}^{\text{1st }b\text{-jet}}$
$\Delta R(\ell,\ell)$	$H_{\text{T}}^{bb\ell\ell}$	$\Delta R(\ell,\ell)$	$H_{\text{T}}^{bb\ell\ell}$	$\Delta R(\ell,\ell)$	$H_{\text{T}}^{bb\ell\ell}$
$E_{\text{T}}^{\text{miss}}$	$m_{bjj}^{\text{top-mass}}$	$E_{\text{T}}^{\text{miss}}$	$m_{bjj}^{\text{top-mass}}$	$E_{\text{T}}^{\text{miss}}$	$m_{bjj}^{\text{top-mass}}$
$m_{\ell\ell}$	w_{2t}	$m_{\ell\ell}$	$p_{\text{T}}^{\text{2nd jet}}$	$m_{\ell\ell}$	w_{2t}
$p_{\text{T}}^{\ell\ell}$	$\lvert\eta_{\ell\ell}\rvert$	$p_{\text{T}}^{\ell\ell}$	$p_{\text{T}}^{\text{3rd jet}}$	$p_{\text{T}}^{\ell\ell}$	$p_{\text{T}}^{\text{6th jet}}$
$m_{b\ell}^{\text{min}}$	$\Delta R_{jj}^{\text{avg}}$	$m_{b\ell}^{\text{min}}$	m_{bb}	$m_{b\ell}^{\text{min}}$	$\lvert\eta_{\ell\ell}\rvert$
$\max m_{b\ell}^{\text{min}\,\Delta R}$	$N_{bjj}^{\text{top-mass}}$	$\max m_{b\ell}^{\text{min}\,\Delta R}$	$p_{\text{T}}^{b_1jj}$	$\max m_{b\ell}^{\text{min}\,\Delta R}$	m_{bb}
$H_{\text{T}}^{bb\ell\ell}$	$N_{jj}^{W\text{-mass}}$	$H_{\text{T}}^{bb\ell\ell}$	H_{T}	$H_{\text{T}}^{bb\ell\ell}$	$\Delta R_{jj}^{\text{avg}}$
w_{2t}		$m_{b\ell}^{\text{max-min}}$	$\Delta R_{jj}^{\text{avg}}$	w_{2t}	$N_{bjj}^{\text{top-mass}}$
$m_{b\ell}^{\text{max-min}}$		$N_{jj}^{W\text{-mass}}$	$m_{jj}^{\text{min}\,\Delta R}$	$m_{b\ell}^{\text{max-min}}$	$N_{jj}^{W\text{-mass}}$
$N_{jj}^{W\text{-mass}}$		$m_{bjj}^{\text{top-mass}}$	$m_{jj}^{\text{max}\,p_{\text{T}}}$	$\Delta R(bb,\ell\ell)$	$m_{jj}^{\text{max}\,p_{\text{T}}}$
$m_{bjj}^{\text{top-mass}}$			$p_{\text{T}}^{\ell\ell}$	$N_{jj}^{W\text{-mass}}$	
				$m_{bjj}^{\text{top-mass}}$	

Tab. G.5.: Listing of the BDT training variables per target region and background process. Each column represents the axis labels of the two-dimensional linear correlation plots shown in Sec. F.8.

G.4. Pre-fit observed and expected event yields

The observed and expected yields of all signal, control and validation regions prior to the fit are listed in Tab. G.6–G.8. The quoted errors on the numbers include both statistical and systematic uncertainties. The numbers corresponding to simulated processes were normalised to the theoretical cross section of the respective process and were scaled to an integrated luminosity of 139 fb^{-1}. The yields for the tWZ, tZq and Wt background processes were, unlike in tables in Secs. 7.1 and 7.2, split up since they are assigned different systematic uncertainties.

Process	2ℓ-Z-1b6j-$t\bar{t}Z$		2ℓ-Z-2b5j-$t\bar{t}Z$		2ℓ-Z-2b6j-$t\bar{t}Z$	
$t\bar{t}Z$	101.0	$\pm$ 3.3	96.9	$\pm$ 2.9	247.6	$\pm$ 7.3
DD $t\bar{t}$	42.6	$\pm$ 6.7	70.0	$\pm$ 8.7	114	$\pm$ 11
Z+light flavour	49	$\pm$ 17	4.1	$\pm$ 3.0	3.5	$\pm$ 2.1
Z+c	197	$\pm$ 60	28.0	$\pm$ 9.3	49	$\pm$ 16
Z+b	280	$\pm$ 71	262.5	$\pm$ 9.0	364	$\pm$ 13
$t\bar{t}X$ $(X = W, H, \gamma)$	2.0	$\pm$ 1.0	4.5	$\pm$ 2.3	7.5	$\pm$ 3.4
VV, VH $(V = Z, W)$	45	$\pm$ 23	18.5	$\pm$ 9.4	25	$\pm$ 13
tWZ	11.9	$\pm$ 6.0	7.5	$\pm$ 4.8	18.1	$\pm$ 9.1
tZq	4.9	$\pm$ 1.0	10.1	$\pm$ 2.1	16.5	$\pm$ 3.4
Wt	0.50	$\pm$ 0.26	1.54	$\pm$ 0.48	2.40	$\pm$ 0.69
Total SM	736	$\pm$ 99	504	$\pm$ 23	848	$\pm$ 37
Data	663		457		861	

Tab. G.6.: Pre-fit observed and expected event yields of the $t\bar{t}Z$-regions. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of 139 fb^{-1}. All relevant weights were applied. The quoted errors include both statistical and systematic uncertainties.

Process	2ℓ-Z-1b6j-$t\bar{t}$	2ℓ-Z-2b5j-$t\bar{t}$	2ℓ-Z-2b6j-$t\bar{t}$
$t\bar{t}Z$	13.77 ± 0.66	23.23 ± 0.95	24.7 ± 1.1
DD $t\bar{t}$	492 ± 23	1764 ± 43	1126 ± 34
Z+light flavour	15.8 ± 5.9	0.57 ± 0.36	–
Z+c	47 ± 15	15.7 ± 5.4	9.2 ± 3.3
Z+b	90 ± 23	158.6 ± 6.4	107.6 ± 5.1
$t\bar{t}X$ $(X = W, H, \gamma)$	12.2 ± 6.1	47 ± 24	31 ± 16
VV, VH $(V = Z, W)$	10.9 ± 5.5	6.0 ± 3.1	6.2 ± 3.2
tWZ	1.65 ± 0.85	1.69 ± 0.87	1.70 ± 0.88
tZq	1.24 ± 0.28	3.07 ± 0.65	2.35 ± 0.50
Wt	7.5 ± 1.6	24.5 ± 4.2	11.4 ± 2.2
Total SM	693 ± 38	2045 ± 50	1320 ± 39
Data	612	1907	1322

Tab. G.7.: Pre-fit observed and expected event yields of the $t\bar{t}$-regions. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. The quoted errors include both statistical and systematic uncertainties.

Process	2ℓ-Z-1b6j-Z+jets	2ℓ-Z-2b5j-Z+jets	2ℓ-Z-2b6j-Z+jets
$t\bar{t}Z$	130.0 ± 3.7	71.6 ± 2.4	78.0 ± 2.7
DD $t\bar{t}$	186 ± 14	401 ± 21	100 ± 10
Z+light flavour	990 ± 350	60 ± 38	20 ± 12
Z+c	2810 ± 860	490 ± 160	220 ± 73
Z+b	3630 ± 910	3630 ± 150	1575 ± 66
$t\bar{t}X$ $(X = W, H, \gamma)$	5.5 ± 2.8	9.6 ± 4.8	3.7 ± 1.9
VV, VH $(V = Z, W)$	350 ± 180	150 ± 75	83 ± 42
tWZ	21 ± 10	10.4 ± 5.2	9.3 ± 4.7
tZq	28.0 ± 5.7	38.4 ± 7.8	23.9 ± 4.9
Wt	3.43 ± 0.86	13.7 ± 2.5	1.91 ± 0.58
Total SM	8200 ± 1300	4880 ± 280	2110 ± 130
Data	6911	5186	1926

Tab. G.8.: Pre-fit observed and expected event yields of the Z+jets-regions. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. The quoted errors include both statistical and systematic uncertainties.

G.5. Post-fit observed and expected event yields

The observed and expected yields of all signal, control and validation regions after the fit was performed are listed in Tab. G.9–G.11. The quoted errors on the numbers include both statistical and systematic uncertainties. The numbers corresponding to simulated processes were normalised to the theoretical cross section of the respective process and were scaled to an integrated luminosity of $139\,\mathrm{fb}^{-1}$. The yields for the tWZ, tZq and Wt background processes were, unlike in tables in Secs. 7.1 and 7.2, split up since they are assigned different systematic uncertainties.

Process	2ℓ-Z-1b6j-$t\bar{t}Z$	2ℓ-Z-2b5j-$t\bar{t}Z$	2ℓ-Z-2b6j-$t\bar{t}Z$
$t\bar{t}Z$	105 ± 19	98 ± 16	256 ± 41
DD $t\bar{t}$	42.56 ± 0.50	69.2 ± 1.2	115.0 ± 1.6
Z+light flavour	45 ± 16	3.4 ± 2.0	3.0 ± 1.8
Z+c	167 ± 45	22.9 ± 6.7	40 ± 12
Z+b	254 ± 45	257 ± 13	362 ± 21
$t\bar{t}X$ $(X = W, H, \gamma)$	19.5 ± 09.9	4.3 ± 2.2	7.3 ± 3.7
VV, VH $(V = Z, W)$	35 ± 19	14.1 ± 7.7	19 ± 11
tWZ	11.7 ± 7.3	7.3 ± 3.7	17.7 ± 8.9
tZq	4.86 ± 0.99	9.8 ± 2.0	16.3 ± 3.3
Wt	0.501 ± 0.077	1.42 ± 0.22	2.40 ± 0.42
Total SM	666 ± 23	488 ± 19	839 ± 40
Data	663	457	861

Tab. G.9.: Post-fit observed and expected event yields of the $t\bar{t}Z$-regions. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. The quoted errors include both statistical and systematic uncertainties.

Process	2ℓ-Z-1b6j-$t\bar{t}$	2ℓ-Z-2b5j-$t\bar{t}$	2ℓ-Z-2b6j-$t\bar{t}$
$t\bar{t}Z$	14.3 ± 2.5	23.9 ± 3.9	25.4 ± 4.1
DD $t\bar{t}$	492.3 ± 4.6	1764.5 ± 6.6	1126.0 ± 5.8
Z+light flavour	14.5 ± 5.1	0.48 ± 0.29	–
Z+c	40 ± 11	13.1 ± 3.8	7.8 ± 2.3
Z+b	82 ± 15	157.4 ± 7.6	106.7 ± 6.4
$t\bar{t}X$ $(X = W, H, \gamma)$	11.9 ± 6.0	46 ± 23	30 ± 15
VV, VH $(V = Z, W)$	8.4 ± 4.6	4.6 ± 2.5	4.8 ± 2.6
tWZ	1.63 ± 0.82	1.65 ± 0.83	1.66 ± 0.84
tZq	1.23 ± 0.25	30.1 ± 06.1	2.31 ± 0.47
Wt	7.5 ± 1.2	24.3 ± 3.7	11.3 ± 1.7
Total SM	673 ± 26	2039 ± 49	1316 ± 40
Data	612	1907	1322

Tab. G.10.: Post-fit observed and expected event yields of the $t\bar{t}$-regions. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. The quoted errors include both statistical and systematic uncertainties.

Process	2ℓ-Z-1b6j-Z+jets	2ℓ-Z-2b5j-Z+jets	2ℓ-Z-2b6j-Z+jets
$t\bar{t}Z$	135 ± 24	73 ± 12	80 ± 13
DD $t\bar{t}$	186.0 ± 3.7	401.1 ± 3.1	99.73 ± 0.75
Z+light flavour	900 ± 320	50 ± 31	16.4 ± 9.9
Z+c	2280 ± 650	410 ± 120	181 ± 54
Z+b	3270 ± 590	4070 ± 160	1447 ± 78
$t\bar{t}X$ $(X = W, H, \gamma)$	5.3 ± 2.7	9.3 ± 4.7	3.6 ± 1.8
VV, VH $(V = Z, W)$	270 ± 150	115 ± 63	64 ± 35
tWZ	20 ± 10	10.1 ± 5.1	9.1 ± 4.6
tZq	27.8 ± 5.6	37.7 ± 7.7	23.4 ± 4.8
Wt	3.42 ± 0.25	13.4 ± 2.1	1.91 ± 0.30
Total SM	7200 ± 450	5186 ± 74	1926 ± 44
Data	6911	5186	1926

Tab. G.11.: Post-fit observed and expected event yields of the Z+jets-regions. The simulated processes are scaled to their respective theoretical cross section and to integrated luminosity of $139\,\mathrm{fb}^{-1}$. All relevant weights were applied. The quoted errors include both statistical and systematic uncertainties.

References

[1] N Dr. phil. Bohr, *I. On the constitution of atoms and molecules*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **26** (1913) 1–25, eprint: https://doi.org/10.1080/14786441308634955 (cit. on p. 1).

[2] E. Rutherford, *The scattering of α and β particles by matter and the structure of the atom*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **21** (1911) 669–688 (cit. on p. 1).

[3] E. Rutherford, *LIV. Collision of α particles with light atoms. IV. An anomalous effect in nitrogen*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **37** (1919) 581–587, eprint: https://doi.org/10.1080/14786440608635919, URL: https://doi.org/10.1080/14786440608635919 (cit. on p. 1).

[4] J. Chadwick, *Possible Existence of a Neutron*, Nature **129** (1932) 1476–4687, URL: https://doi.org/10.1038/129312a0 (cit. on p. 1).

[5] J. J. Thomson, *XL. Cathode Rays*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **44** (1897) 293–316, eprint: https://doi.org/10.1080/14786449708621070, URL: https://doi.org/10.1080/14786449708621070 (cit. on p. 1).

[6] Robert Hofstadter, *Electron Scattering and Nuclear Structure*, Rev. Mod. Phys. **28** (3 1956) 214–254, URL: https://link.aps.org/doi/10.1103/RevModPhys.28.214 (cit. on p. 1).

[7] Jerome I. Friedman, Henry W. Kendall, and Richard E. Taylor, *Deep inelastic scattering: Acknowledgments*, Rev. Mod. Phys. **63** (3 1991) 629–629, URL: https://link.aps.org/doi/10.1103/RevModPhys.63.629 (cit. on p. 1).

[8] Richard P. Feynman, *Very high-energy collisions of hadrons*, Phys. Rev. Lett. **23** (1969) 1415–1417 (cit. on pp. 1, 19).

[9] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964) 214–215, URL: http://www.sciencedirect.com/science/article/pii/S0031916364920013 (cit. on p. 1).

[10] G Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 1*, tech. rep. CERN-TH-401, CERN, 1964, URL: https://cds.cern.ch/record/352337 (cit. on p. 1).

[11] G Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 2*, (1964) 80 p, Version 1 is CERN preprint 8182/TH.401, Jan. 17, 1964, URL: https://cds.cern.ch/record/570209 (cit. on p. 1).

[12] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (7 1970) 1285–1292, URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285 (cit. on pp. 1, 18).

[13] J. D. Bjorken and S. L. Glashow, *Elementary Particles and SU(4)*, Phys. Lett. **11** (1964) 255–257 (cit. on p. 1).

[14] J. E. Augustin et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (23 1974) 1406–1408, URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1406 (cit. on p. 1).

[15] J. J. Aubert et al., *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (23 1974) 1404–1406, URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1404 (cit. on p. 1).

[16] Makoto Kobayashi and Toshihide Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657, URL: http://dx.doi.org/10.1143/PTP.49.652. (cit. on pp. 1, 18).

[17] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (5 1977) 252–255, URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.252 (cit. on pp. 1, 18).

[18] F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv: hep-ex/9503002 (cit. on pp. 1, 21).

[19] S. Abachi et al., *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, arXiv: hep-ex/9503003 (cit. on pp. 1, 21).

[20] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the color octet gluon picture*, Physics Letters B **47** (1973) 365 –368, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/0370269373906254 (cit. on p. 1).

[21] David J. Gross and Frank Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (26 1973) 1343–1346, URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343 (cit. on pp. 1, 14).

[22] H. David Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (26 1973) 1346–1349, URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1346 (cit. on pp. 1, 14).

[23] Sheldon L. Glashow, *The renormalizability of vector meson interactions*, Nucl. Phys. **10** (1959) 107–117 (cit. on p. 1).

[24] Abdus Salam and John Clive Ward, *Weak and electromagnetic interactions*, Nuovo Cim. **11** (1959) 568–577 (cit. on p. 1).

[25] Steven Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (21 1967) 1264–1266, URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264 (cit. on p. 1).

[26] S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*, Progress of Theoretical Physics **1** (1946) 27–42, eprint: /oup/backfile/content_public/journal/ptp/1/2/10.1143/ptp.1.27/2/1-2-27.pdf, URL: http://dx.doi.org/10.1143/PTP.1.27 (cit. on pp. 2, 7).

[27] Julian Schwinger, *On Quantum-Electrodynamics and the Magnetic Moment of the Electron*, Phys. Rev. **73** (4 1948) 416–417, URL: https://link.aps.org/doi/10.1103/PhysRev.73.416 (cit. on pp. 2, 7).

[28] J. Schwinger, *Quantum Electrodynamics. I. A Covariant Formulation*, Physical Review **74** (1948) 1439–1461 (cit. on pp. 2, 7).

[29] R. P. Feynman, *Space-Time Approach to Quantum Electrodynamics*, Phys. Rev. **76** (6 1949) 769–789, URL: https://link.aps.org/doi/10.1103/PhysRev.76.769 (cit. on pp. 2, 7).

[30] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, Physical Review **80** (1950) 440–457 (cit. on pp. 2, 7).

[31] F. J. Dyson, *The Radiation Theories of Tomonaga, Schwinger, and Feynman*, Phys. Rev. **75** (3 1949) 486–502, URL: https://link.aps.org/doi/10.1103/PhysRev.75.486 (cit. on pp. 2, 7).

[32] K. Kodama et al., *Observation of tau neutrino interactions*, Physics Letters B **504** (2001) 218 –224, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/S0370269301003070 (cit. on p. 2).

[33] Georges Aad et al., *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) 1–29, arXiv: 1207.7214 [hep-ex] (cit. on pp. 2, 5, 31).

[34] Serguei Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv: 1207.7235 [hep-ex] (cit. on pp. 2, 5, 31).

[35] Vardan Khachatryan et al., *Observation of top quark pairs produced in association with a vector boson in pp collisions at* $\sqrt{s} = 8$ *TeV*, JHEP **01** (2016) 096, arXiv: 1510.01131 [hep-ex] (cit. on pp. 2, 28, 29, 101).

[36] ATLAS Collaboration, *Search for* $t\bar{t}Z$ *production in the three lepton final state with* 4.7 fb^{-1} *of* $\sqrt{s} = 7$ *TeV pp collision data collected by the ATLAS detector*, tech. rep. ATLAS-CONF-2012-126, CERN, 2012, URL: https://cds.cern.ch/record/1474643 (cit. on pp. 2, 29).

[37] ATLAS Collaboration, *Measurement of the* $t\bar{t}W$ *and* $t\bar{t}Z$ *production cross sections in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector.* Tech. rep. ATLAS-CONF-2015-032, CERN, 2015, URL: https://cds.cern.ch/record/2038143 (cit. on pp. 2, 29).

[38] ATLAS Collaboration, *Measurement of the* $t\bar{t}Z$ *and* $t\bar{t}W$ *production cross sections in multilepton final states using 3.2* fb^{-1} *of pp collisions at 13 TeV at the LHC*, tech. rep. ATLAS-CONF-2016-003, CERN, 2016, URL: https://cds.cern.ch/record/2138947 (cit. on pp. 2, 29).

[39] Morad Aaboud et al., *Measurement of the* $t\bar{t}Z$ *and* $t\bar{t}W$ *cross sections in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: 1901.03584 [hep-ex] (cit. on pp. 2, 27, 101, 105, 112, 146, 150, 151).

[40] Vardan Khachatryan et al., *Measurement of Top Quark-Antiquark Pair Production in Association with a W or Z Boson in pp Collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C **74** (2014) 3060, arXiv: 1406.7830 [hep-ex] (cit. on pp. 2, 29).

[41] Morad Aaboud et al., *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] (cit. on p. 2).

[42] Georges Aad et al., *Search for squarks and gluinos in final states with same-sign leptons and jets using 139* fb^{-1} *of data collected with the ATLAS detector*, JHEP **06** (2020) 046, arXiv: 1909.08457 [hep-ex] (cit. on p. 2).

[43] M. Aaboud et al., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173–191, arXiv: 1806.00425 [hep-ex] (cit. on p. 2).

[44] Georges Aad et al., *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2020) 124, arXiv: 2002.07546 [hep-ex] (cit. on pp. 2, 26).

[45] Albert M Sirunyan et al., *Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at $\sqrt{s}$ =13 TeV*, Phys. Rev. Lett. **122** (2019) 132003, arXiv: 1812.05900 [hep-ex] (cit. on pp. 2, 26).

[46] ATLAS Collaboration, *Measurements of the inclusive and differential production cross sections of a top-quark-antiquark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2020), URL: https://cds.cern.ch/record/2725734 (cit. on pp. 2, 27, 79, 80, 82, 96, 98, 101, 149).

[47] Albert M Sirunyan et al., *Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2020) 056, arXiv: 1907.11270 [hep-ex] (cit. on pp. 2, 27, 101).

[48] D. Griffiths, *Introduction to Elementary Particles*, 2nd ed., WILEY-VCH Verlag GmbH & Co. KGaA, 2008, ISBN: 978-3527406012 (cit. on pp. 5, 7, 10, 11, 14, 17, 19, 21, 155).

[49] Mark Thomson, *Modern Particle Physics*, 3rd ed., Cambridge University Press, 2015, ISBN: 978-1107034266 (cit. on pp. 5, 7, 10, 11, 14, 19, 21, 36, 45).

[50] H. D. Perkins, *Introduction to High Energy Physics*, 4th ed., Cambridge University Press, 2000, ISBN: 978-0521621960 (cit. on pp. 5, 14, 17, 32, 36, 40, 45).

[51] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001, URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001 (cit. on pp. 6, 7, 9, 12, 17, 18, 24, 27, 33, 38, 46, 76, 92, 96, 156, 165).

[52] C. Berger, *Elementarteilchenphysik, Von den Grundlagen zu den modernen Experimenten*, 3rd ed., Springer Spektrum, 2014, ISBN: 978-3540231431 (cit. on p. 7).

[53] Michael E. Peskin and Daniel V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, 1995, ISBN: 978-0-201-50397-5 (cit. on pp. 7, 10, 11, 14, 17, 21).

[54] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321–323, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on pp. 9, 16).

[55] Peter W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508–509, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on pp. 9, 16).

[56] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585–587, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585 (cit. on pp. 9, 16).

[57] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics B **44** (1972) 189–213, URL: http://www.sciencedirect.com/science/article/pii/0550321372902799 (cit. on p. 9).

[58] K. Bethge and U. E. Schröder, *Elementarteilchen und ihre Wechselwirkungen*, 3rd ed., WILEY-VCH Verlag GmbH& Co. KGaA, 2006, ISBN: 978-3527405879 (cit. on pp. 11, 14, 17, 19).

[59] Francis Halzen and Alan D. Martin, *Quarks & Leptons, An Introductory in Modern Particle Physics*, 1st ed., John Wiley and Sons, Inc., 1984, ISBN: 978-0471887416 (cit. on pp. 11, 17).

[60] Michio Kaku, *Quantum Field Theory, A Modern Introduction*, Oxford University Press, 1993, ISBN: 978-0-19-509158-8 (cit. on pp. 11, 14).

[61] Salvatore Mele, *Measurements of the running of the electromagnetic coupling at LEP*, *Proceedings, 26th International Symposium on Physics in Collision (PIC 2006): Buzios, Brazil, July 6-9, 2006*, 2006, arXiv: hep-ex/0610037 [hep-ex], URL: http://www.slac.stanford.edu/econf/C060706/pdf/0610037.pdf (cit. on p. 12).

[62] Jeff Greensite, *An introduction to the Confinement Problem*, Springer-Verlag, 2011, ISBN: 978-3-642-14381-6 (cit. on p. 14).

[63] Reinhard Alkofer and J. Greensite, *Quark Confinement: The Hard Problem of Hadron Physics*, J. Phys. **G34** (2007) S3, arXiv: hep-ph/0610365 [hep-ph] (cit. on p. 14).

[64] Nathan Philip Jurik and Tomasz Skwarnicki, *Observation of $J/\psi\, p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^0 p$ decays*, Presented 01 Aug 2016, 2016, URL: https://cds.cern.ch/record/2206806 (cit. on p. 14).

[65] Peter Schmüser, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*, 2nd ed., Springer-Verlag, 1995, ISBN: 978-3540584865 (cit. on pp. 14, 17, 19).

[66] G. Dissertoi, I. Knowles, and M. Schmelling, *Quantum Chromodynamics*, 1st ed., Oxford University Press, 2003, ISBN: 978-0199566419 (cit. on p. 14).

[67] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (4 1957) 1413–1415, URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413 (cit. on p. 15).

[68] T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1 1956) 254–258, URL: https://link.aps.org/doi/10.1103/PhysRev.104.254 (cit. on p. 15).

[69] Timo van Ritbergen and Robin G. Stuart, *Complete 2-Loop Quantum Electrodynamic Contributions to the Muon Lifetime in the Fermi Model*, Phys. Rev. Lett. **82** (3 1999) 488–491, URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.488 (cit. on p. 17).

[70] M. Steinhauser and T. Seidensticker, *Second order corrections to the muon lifetime and the semileptonic B decay*, Phys. Lett. **B467** (1999) 271–278, arXiv: hep-ph/9909436 [hep-ph] (cit. on p. 17).

[71] D.M. Webber et al., *Measurement of the Positive Muon Lifetime and Determination of the Fermi Constant to Part-per-Million Precision*, Phys. Rev. Lett. **106** (2011) 041803, arXiv: 1010.0991 [hep-ex] (cit. on p. 17).

[72] Nicola Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533, [,648(1963)] (cit. on p. 18).

[73] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, arXiv: hep-ex/9807003 (cit. on p. 18).

[74] Q.R. Ahmad et al., *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 (cit. on p. 18).

[75] Ziro Maki, Masami Nakagawa, and Shoichi Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870–880 (cit. on p. 19).

[76] Richard. P. Feynman, *The behavior of hadron collisions at extreme energies*, Conf. Proc. **C690905** (1969) 237–258 (cit. on p. 19).

[77] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450, [Yad. Fiz.15,781(1972)] (cit. on p. 19).

[78] L. N. Lipatov, *The parton model and perturbation theory*, Yad. Fiz. **20** (1974) 181–198 (cit. on p. 19).

[79] Guido Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298–318 (cit. on p. 19).

[80] Yuri L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.* Sov. Phys. JETP **46** (1977) 641–653, [Zh. Eksp. Teor. Fiz.73,1216(1977)] (cit. on p. 19).

[81] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189–285, arXiv: 0901.0002 [hep-ph] (cit. on p. 20).

[82] John C. Collins, Davison E. Soper, and George F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1–91, arXiv: hep-ph/0409313 (cit. on p. 20).

[83] ATLAS, CDF, CMS, and D0 Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, tech. rep., 2014, arXiv: 1403.4427 [hep-ex] (cit. on p. 21).

[84] S. Alekhin, A. Djouadi, and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, Phys. Lett. **B716** (2012) 214–219, arXiv: 1207.0980 [hep-ph] (cit. on p. 21).

[85] Stephen P. Martin, *A Supersymmetry primer*, (1997), [Adv. Ser. Direct. High Energy Phys.18,1(1998)], arXiv: hep-ph/9709356 [hep-ph] (cit. on pp. 21, 33).

[86] Michał Czakon, Paul Fiedler, and Alexander Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_s^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph] (cit. on pp. 24, 30).

[87] Anna Kulesza, Leszek Motyka, Daniel Schwartländer, Tomasz Stebel, and Vincent Theeuwes, *Associated production of a top quark pair with a heavy electroweak gauge boson at NLO+NNLL accuracy*, Eur. Phys. J. C **79** (2019) 249, arXiv: 1812.08622 [hep-ph] (cit. on pp. 26, 82, 146).

[88] LHCTopWG, *Summary plots from the ATLAS Top physics group*, 2018, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/ (visited on 06/26/2019) (cit. on p. 30).

[89] B Stumpe, *A new Principle for X-Y Touch Screen*, tech. rep., SPS-AOP-BS-jf: CERN, 1977, URL: http://cds.cern.ch/record/1266588 (cit. on p. 31).

[90] Tim Berners-Lee, Robert Cailliau, Jean-Francois Groff, and Bernd Pollermann, *World-Wide Web: the information universe*, Electron. Netw. **2** (1992) 52–58, Reprinted in 2010, URL: https://cds.cern.ch/record/245849 (cit. on p. 31).

[91] *Convention for the establishment of a European organization for nuclear research: Paris, 1st July, 1953 : as amended. Convention pour l'établissement d'une Organisation européenne pour la Recherche nucléaire. Paris, le 1er juillet 1953 : telle qu'elle a été modifiée*, CERN, 1971, URL: https://cds.cern.ch/record/330625 (cit. on p. 31).

[92] G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. **122B** (1983) 103–116, [,611(1983)] (cit. on p. 31).

[93] M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **122B** (1983) 476–485 (cit. on p. 31).

[94] P. Bagnaia et al., *Evidence for Z0 —> e+ e- at the CERN anti-p p Collider*, Phys. Lett. **129B** (1983) 130–140 (cit. on p. 31).

[95] S. Schael et al., *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257–454, arXiv: hep-ex/0509008 (cit. on p. 31).

[96] Esma Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, (2018), General Photo, URL: https://cds.cern.ch/record/2636343 (cit. on p. 32).

[97] CERN Press Office, *LHC sets new world record*, 2009, URL: http://press.web.cern.ch/press-releases/2009/11/lhc-sets-new-world-record (visited on 08/09/2017) (cit. on p. 33).

[98] Morad Aaboud et al., *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] (cit. on pp. 35, 36, 62).

[99] Oliver Sim Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, 2004, URL: http://cds.cern.ch/record/782076 (cit. on pp. 36, 61).

[100] *LHC Machine*, JINST **3** (2008) S08001, ed. by Lyndon Evans and Philip Bryant (cit. on p. 36).

[101] G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 36, 38, 43, 45, 47, 48).

[102] A. Airapetian et al., *ATLAS: Detector and physics performance technical design report. Volume 1*, (1999) (cit. on pp. 36, 38, 43, 45, 47, 48).

[103] Joao Pequenao, *Computer generated image of the whole ATLAS detector*, 2008, URL: http://cds.cern.ch/record/1095924 (cit. on p. 37).

[104] ATLAS Collaboration, *ATLAS magnet system: Technical design report*, (1997) (cit. on p. 38).

[105] LHCTopWG, *ATLAS Experiment*, URL: http://www.lhc-facts.ch/index.php?page=atlas (visited on 07/05/2020) (cit. on p. 39).

[106] ATLAS Collaboration, *ATLAS central solenoid: Technical design report*, (1997) (cit. on p. 39).

[107] ATLAS Collaboration, *ATLAS barrel toroid: Technical design report*, (1997) (cit. on p. 39).

[108] ATLAS Collaboration, *ATLAS endcap toroids: Technical design report*, (1997) (cit. on p. 39).

[109] David J. Griffiths, *Introduction to Electrodynamics*, 4th ed., Cambridge University Press, 2017 (cit. on p. 39).

[110] R. L. Gluckstern, *Uncertainties in track momentum and direction, due to multiple scattering and measurement errors*, Nucl. Instrum. Meth. **24** (1963) 381–389 (cit. on p. 40).

[111] Joao Pequenao, *Computer generated image of the ATLAS inner detector*, 2008, URL: http://cds.cern.ch/record/1095926 (cit. on pp. 41, 42).

[112] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 1*, (1997) (cit. on pp. 40, 42, 43).

[113] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 2*, (1997) (cit. on pp. 40, 42, 43).

[114] M.S. Alam et al., *ATLAS pixel detector: Technical design report*, (1998) (cit. on p. 41).

[115] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010 (cit. on p. 41).

[116] Laura Barranco Navarro, *Alignment of the ATLAS Inner Detector in the LHC Run II*, tech. rep. ATL-PHYS-PROC-2015-190, CERN, 2015, URL: https://cds.cern.ch/record/2114708 (cit. on p. 41).

[117] Joao Pequenao, *Computer Generated image of the ATLAS calorimeter*, 2008, URL: http://cds.cern.ch/record/1095927 (cit. on p. 44).

[118] ATLAS Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, (1996) (cit. on pp. 44–46).

[119] ATLAS Collaboration, *ATLAS tile calorimeter: Technical design report*, (1996) (cit. on p. 45).

[120] Joao Pequenao, *Computer generated image of the ATLAS Muons subsystem*, 2008, URL: http://cds.cern.ch/record/1095929 (cit. on p. 47).

[121] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, Technical Design Report ATLAS (1997), URL: http://cds.cern.ch/record/331068 (cit. on pp. 47, 48).

[122] Marco Bruschi, *ATLAS muon spectrometer: Technical design report*, tech. rep., 1997 (cit. on p. 48).

[123] S. Abdel Khalek et al., *The ALFA Roman Pot Detectors of ATLAS*, JINST **11** (2016) P11013, arXiv: 1609.00249 [physics.ins-det] (cit. on p. 48).

[124] Patrick Czodrowski, *The ATLAS Trigger System: Ready for Run 2*, tech. rep. ATL-DAQ-PROC-2015-038, CERN, 2015, URL: https://cds.cern.ch/record/2058218 (cit. on p. 49).

[125] Aranzazu Ruiz-Martinez and ATLAS Collaboration, *The Run-2 ATLAS Trigger System*, tech. rep. ATL-DAQ-PROC-2016-003, CERN, 2016, URL: https://cds.cern.ch/record/2133909 (cit. on p. 49).

[126] Julian Glatzer, *Operation of the Upgraded ATLAS Level-1 Central Trigger System*, J. Phys. Conf. Ser. **664** (2015) 082013 (cit. on pp. 49, 50).

[127] S. Artz et al., *Upgrade of the ATLAS Central Trigger for LHC Run-2*, tech. rep. 02, 2015 C02030 (cit. on p. 49).

[128] H. Bertelsen et al., *Operation of the upgraded ATLAS Central Trigger Processor during the LHC Run 2*, tech. rep. 02, 2016 C02020 (cit. on p. 49).

[129] William Panduro Vazquez, *The ATLAS Data Acquisition system in LHC Run 2*, J. Phys. Conf. Ser. **898** (2017) 032017, ed. by Richard Mount and Craig Tull (cit. on pp. 49, 50).

[130] ATLAS Collaboration, *ATLAS high-level trigger, data acquisition and controls: Technical design report*, (2003) (cit. on p. 50).

[131] K. Bos et al., *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, Technical Design Report LCG (2005), URL: https://cds.cern.ch/record/840543 (cit. on p. 50).

[132] ATLAS Collaboration, *ATLAS computing: Technical design report*, ed. by G. Duckeck et al., Technical Design Report ATLAS, CERN, 2005 (cit. on pp. 50, 57).

[133] ATLAS Collaboration, *Event Displays Run 2 Physics*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Physics (visited on 07/07/2020) (cit. on p. 51).

[134] ATLAS Collaboration, *Simulation of Pile-up in the ATLAS Experiment*, J. Phys. Conf. Ser. **513** (2014) 022024 (cit. on p. 51).

[135] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, tech. rep. ATLAS-CONF-2012-042, CERN, 2012, URL: https://cds.cern.ch/record/1435196 (cit. on pp. 51, 65).

[136] Andy Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145–233, arXiv: 1101.2599 [hep-ph] (cit. on p. 51).

[137] ATLAS Collaboration, *Luminosity Public Results Run 2*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 07/07/2020) (cit. on pp. 52, 62, 63).

[138] D.J. Griffiths, *Introduction to Quantum Mechanics*, Cambridge University Press, 2017, ISBN: 9781107179868 (cit. on p. 53).

[139] Glen Cowan, *Statistical Data Analysis*, Oxford University Press, 1989, ISBN: 978-0198501558 (cit. on pp. 53, 82, 95, 129, 134–138).

[140] Stefan Höche, *Introduction to parton-shower event generators*, *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, 2015 235–295, arXiv: 1411.4085 [hep-ph] (cit. on p. 55).

[141] S. Catani, F. Krauss, R. Kuhn, and B.R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231 (cit. on p. 55).

[142] F. Krauss, *Matrix elements and parton showers in hadronic interactions*, JHEP **08** (2002) 015, arXiv: hep-ph/0205283 (cit. on p. 55).

[143] V. V. Sudakov, *Vertex Parts at Very High Energies in Quantum Electrodynamics*, Sov. Phys. JETP **3** (1956) 65 (cit. on p. 55).

[144] F. Caravaglios, Michelangelo L. Mangano, M. Moretti, and R. Pittau, *A New approach to multijet calculations in hadron collisions*, Nucl. Phys. B **539** (1999) 215–232, arXiv: hep-ph/9807570 (cit. on p. 55).

[145] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145 (cit. on p. 56).

[146] Bo Andersson, *The Lund Model*, (1998), URL: http://lccn.loc.gov/2006274063 (cit. on p. 56).

[147] B.R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238** (1984) 492–528 (cit. on p. 56).

[148] S. Agostinelli et al., *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303 (cit. on p. 57).

[149] G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823–874, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 57).

[150] I. Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, tech. rep., 2014 (cit. on p. 57).

[151] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013 (2010), URL: https://cds.cern.ch/record/1300517 (cit. on p. 57).

[152] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph] (cit. on p. 58).

[153] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele 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] (cit. on p. 58).

[154] Paolo Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on p. 58).

[155] Stefano Frixione, Paolo Nason, and Carlo Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 58).

[156] 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] (cit. on p. 58).

[157] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv: 1106.0522 [hep-ph] (cit. on p. 58).

[158] Stefano Frixione and Bryan R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv: hep-ph/0204244 (cit. on p. 59).

[159] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 59).

[160] Enrico Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 59).

[161] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 (cit. on p. 59).

[162] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv: 0710.3820 [hep-ph] (cit. on p. 59).

[163] Torbjörn Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, arXiv: 1410.3012 [hep-ph] (cit. on p. 59).

[164] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639–707, arXiv: 0803.0883 [hep-ph] (cit. on p. 59).

[165] Johannes Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 59).

[166] D.J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152–155, ed. by S. Erhan, P. Schlein, and Y. Rozen (cit. on p. 59).

[167] Richard D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 59).

[168] Stefano Frixione, Eric Laenen, Patrick Motylinski, and Bryan R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: hep-ph/0702198 (cit. on p. 59).

[169] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph] (cit. on p. 59).

[170] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV datas*, tech. rep. ATL-PHYS-PUB-2014-021, CERN, 2014, URL: http://cds.cern.ch/record/1966419 (cit. on p. 59).

[171] Richard D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244–289, arXiv: 1207.1303 [hep-ph] (cit. on p. 59).

[172] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016), arXiv: 1610.07922 [hep-ph] (cit. on pp. 59, 82, 146).

[173] 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*, tech. rep. ATL-PHYS-PUB-2016-005, CERN, 2016, URL: https://cds.cern.ch/record/2120826 (cit. on pp. 59, 82, 146).

[174] S. Catani, F. Krauss, R. Kuhn, and B.R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231 (cit. on pp. 59, 60).

[175] Stefan Höche, Frank Krauss, Steffen Schumann, and Frank Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph] (cit. on pp. 59, 60).

[176] Tanju Gleisberg and Stefan Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on p. 60).

[177] Steffen Schumann and Frank Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 60).

[178] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: 1111.1220 [hep-ph] (cit. on p. 60).

[179] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on p. 60).

[180] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on p. 60).

[181] Ansgar Denner, Stefan Dittmaier, and Lars Hofer, *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*, Comput. Phys. Commun. **212** (2017) 220–238, arXiv: 1604.06792 [hep-ph] (cit. on p. 60).

[182] Federico Demartin, Benedikt Maier, Fabio Maltoni, Kentarou Mawatari, and Marco Zaro, *tWH associated production at the LHC*, Eur. Phys. J. C **77** (2017) 34, arXiv: 1607.05862 [hep-ph] (cit. on p. 60).

[183] Stefano Frixione, Eric Laenen, Patrick Motylinski, Bryan R. Webber, and Chris D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph] (cit. on p. 60).

[184] Michal Czakon and Alexander 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] (cit. on p. 60).

[185] Joao Pequenao, *Event Cross Section in a computer generated image of the ATLAS detector.* 2008, URL: https://cds.cern.ch/record/1096081 (cit. on p. 66).

[186] M. Aaboud et al., *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex] (cit. on p. 65).

[187] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444–450 (cit. on pp. 65, 68).

[188] Thijs G. Cornelissen et al., *The global chi**2 track fitter in ATLAS*, J. Phys. Conf. Ser. **119** (2008) 032013, ed. by Randall Sobie, Reda Tafirout, and Jana Thomson (cit. on pp. 65, 68).

[189] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, and E. Moyse, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, tech. rep., 2007 (cit. on p. 66).

[190] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, J. Phys. Conf. Ser. **119** (2008) 032014, ed. by Randall Sobie, Reda Tafirout, and Jana Thomson (cit. on p. 66).

[191] G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex] (cit. on p. 66).

[192] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, *Adaptive vertex fitting*, J. Phys. G **34** (2007) N343 (cit. on p. 66).

[193] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment*, tech. rep. ATLAS-CONF-2010-069, CERN, 2010, URL: http://cds.cern.ch/record/1281344 (cit. on p. 67).

[194] Morad Aaboud et al., *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: 1611.10235 [physics.ins-det] (cit. on p. 67).

[195] Georges Aad et al., *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] (cit. on pp. 67–70, 143).

[196] Georges Aad et al., *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] (cit. on pp. 67, 73).

[197] ATLAS Collaboration, *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, 2008, URL: https://cds.cern.ch/record/1099735 (cit. on p. 67).

[198] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, tech. rep. ATLAS-CONF-2012-047, CERN, 2012, URL: https://cds.cern.ch/record/1449796 (cit. on p. 68).

[199] R. Frühwirth, *A Gaussian-mixture approximation of the Bethe-Heitler model of electron energy loss by bremsstrahlung*, Computer Physics Communications **154** (2003) 131 – 142, ISSN: 0010-4655, URL: http://www.sciencedirect.com/science/article/pii/S0010465503002923 (cit. on p. 68).

[200] Morad Aaboud et al., *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] (cit. on pp. 68, 69, 143).

[201] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, tech. rep. ATLAS-CONF-2016-024, CERN, 2016, URL: https://cds.cern.ch/record/2157687 (cit. on pp. 68–70, 143).

[202] ATLAS Collaboration, *Electron identification measurements in ATLAS using $\sqrt{s}$ = 13 TeV data with 50 ns bunch spacing*, tech. rep. ATL-PHYS-PUB-2015-041, CERN, 2015, URL: https://cds.cern.ch/record/2048202 (cit. on pp. 68, 143).

[203] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV*, tech. rep. ATL-PHYS-PUB-2016-015, CERN, 2016, URL: https://cds.cern.ch/record/2203514 (cit. on pp. 68, 69, 143).

[204] Georges Aad et al., *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] (cit. on pp. 71, 72, 143).

[205] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s}$ = 13 TeV*, tech. rep. ATLAS-CONF-2020-030, CERN, 2020, URL: https://cds.cern.ch/record/2725736 (cit. on p. 72).

[206] J. Alitti et al., *Inclusive jet cross-section and a search for quark compositeness at the CERN $\bar{p}p$ collider*, Phys. Lett. B **257** (1991) 232–240 (cit. on p. 73).

[207] F. Abe et al., *The Topology of three jet events in $\bar{p}p$ collisions at $\sqrt{s}$ = 1.8 TeV*, Phys. Rev. D **45** (1992) 1448–1458 (cit. on p. 73).

[208] S. Abachi et al., *Studies of Topological Distributions of the Three- and Four-Jet Events in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1800 GeV with the D0 Detector*, Phys. Rev. D **53** (1996) 6000–6016, arXiv: hep-ex/9509005 (cit. on p. 73).

[209] Jeff Tseng and Hannah Evans, *Sequential recombination algorithm for jet clustering and background subtraction*, Phys. Rev. D **88** (2013) 014044, arXiv: 1304.1025 [hep-ph] (cit. on p. 73).

[210] Stephen D. Ellis and Davison E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (1993) 3160–3166, arXiv: hep-ph/9305266 (cit. on p. 73).

[211] S. Catani, Yuri L. Dokshitzer, M.H. Seymour, and B.R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. B **406** (1993) 187–224 (cit. on p. 73).

[212] Yuri L. Dokshitzer, G.D. Leder, S. Moretti, and B.R. Webber, *Better jet clustering algorithms*, JHEP **08** (1997) 001, arXiv: hep-ph/9707323 (cit. on p. 73).

[213] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, (1998) 270–279, arXiv: hep-ph/9907280 (cit. on p. 73).

[214] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 73).

[215] Gavin P. Salam, *Towards Jetography*, Eur. Phys. J. C **67** (2010) 637–686, arXiv: 0906.1833 [hep-ph] (cit. on p. 73).

[216] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph] (cit. on p. 74).

[217] Matteo Cacciari and Gavin P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B **659** (2008) 119–126, arXiv: 0707.1378 [hep-ph] (cit. on p. 74).

[218] Morad Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 74).

[219] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS 2010 data*, tech. rep. ATLAS-CONF-2011-032, CERN, 2011, URL: https://cds.cern.ch/record/1337782 (cit. on p. 75).

[220] M. Aaboud et al., *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] (cit. on p. 75).

[221] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, tech. rep. ATL-PHYS-PUB-2015-015, CERN, 2015, URL: https://cds.cern.ch/record/2037613 (cit. on p. 75).

[222] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*, tech. rep. ATLAS-CONF-2015-002, CERN, 2015, URL: https://cds.cern.ch/record/2001682 (cit. on p. 75).

[223] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, tech. rep. ATLAS-CONF-2015-037, CERN, 2015, URL: https://cds.cern.ch/record/2044941 (cit. on p. 75).

[224] ATLAS Collaboration, *Determination of the jet energy scale and resolution at ATLAS using Z/γ-jet events in data at $\sqrt{s} = 8$ TeV*, tech. rep. ATL-COM-PHYS-2014-791, CERN, 2014, URL: https://cds.cern.ch/record/1741697 (cit. on p. 75).

[225] *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV*, tech. rep. ATLAS-CONF-2015-017, CERN, 2015, URL: https://cds.cern.ch/record/2008678 (cit. on p. 75).

[226] Georges Aad et al., *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: 1210.6210 [hep-ex] (cit. on p. 75).

[227] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, tech. rep. ATLAS-CONF-2014-018, CERN, 2014, URL: https://cds.cern.ch/record/1700870 (cit. on p. 75).

[228] Georges Aad et al., *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] (cit. on p. 75).

[229] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: https://cds.cern.ch/record/2160731 (cit. on p. 76).

[230] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, tech. rep. ATL-PHYS-PUB-2017-013, CERN, 2017, URL: https://cds.cern.ch/record/2273281 (cit. on p. 76).

[231] Marie Christine Lanfermann, Tobias Golling, and Andrea Coccaro, *Deep Neural Network based higher level flavour tagging algorithm at the ATLAS experiment*, tech. rep. ATL-COM-PHYS-2017-1596, CERN, 2017, URL: https://cds.cern.ch/record/2290144 (cit. on p. 76).

[232] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature (2015) 436–444 (cit. on p. 76).

[233] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, tech. rep. ATL-PHYS-PUB-2017-003, CERN, 2017, URL: https://cds.cern.ch/record/2255226 (cit. on p. 76).

[234] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, tech. rep. ATLAS-CONF-2014-004, CERN, 2014, URL: https://cds.cern.ch/record/1664335 (cit. on pp. 76, 77).

[235] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, tech. rep. ATLAS-CONF-2014-046, CERN, 2014, URL: https://cds.cern.ch/record/1741020 (cit. on p. 77).

[236] D Adams et al., *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, tech. rep. ATL-PHYS-INT-2014-018, CERN, 2014, URL: https://cds.cern.ch/record/1743654 (cit. on p. 77).

[237] D Adams et al., *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, tech. rep. ATL-COM-PHYS-2014-451, CERN, 2014, URL: https://cds.cern.ch/record/1700874 (cit. on p. 77).

[238] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*, tech. rep. ATL-PHYS-PUB-2015-027, CERN, 2015, URL: https://cds.cern.ch/record/2037904 (cit. on p. 78).

[239] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV*, tech. rep. ATL-PHYS-PUB-2015-023, CERN, 2015, URL: https://cds.cern.ch/record/2037700 (cit. on p. 78).

[240] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*, tech. rep. ATLAS-CONF-2018-023, CERN, 2018, URL: https://cds.cern.ch/record/2625233 (cit. on p. 78).

[241] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. A **362** (1995) 487–498 (cit. on p. 79).

[242] Georges Aad et al., *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **71** (2011) 1577, arXiv: 1012.1792 [hep-ex] (cit. on p. 81).

[243] Georges Aad et al., *Search for supersymmetry at $\sqrt{s}$=8 TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector*, JHEP **06** (2014) 035, arXiv: 1404.2500 [hep-ex] (cit. on p. 81).

[244] Erich W. Varnes, *A Poisson likelihood approach to fake lepton estimation with the matrix method*, (2016), arXiv: 1606.06817 [hep-ex] (cit. on p. 81).

[245] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physics Sciences*, John Wiley & Sons, Ltd., 1989 (cit. on pp. 82, 97, 129, 135, 139).

[246] Georges Aad et al., *Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Eur. Phys. J. C **79** (2019) 1028, arXiv: 1908.07305 [hep-ex] (cit. on p. 106).

[247] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, *Foundations of Machine Learning*, The MIT Press, 2012, ISBN: 9780262018258 (cit. on p. 113).

[248] Andreas Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, (2007), arXiv: physics/0703039 (cit. on pp. 113, 114, 125, 126, 160, 161).

[249] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81–86, ed. by M. Werlen and D. Perret-Gallix (cit. on p. 113).

[250] Robert E. Schapire, *The strength of weak learnability*, Machine Learning **5** (1990) 197–227, URL: https://doi.org/10.1007/BF00116037 (cit. on pp. 114, 157).

[251] Leo Breiman, *Arcing classifier (with discussion and a rejoinder by the author)*, Ann. Statist. **26** (1998) 801–849, URL: https://doi.org/10.1214/aos/1024691079 (cit. on p. 114).

[252] Morad Aaboud et al., *Probing the quantum interference between singly and doubly resonant top-quark production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. Lett. **121** (2018) 152002, arXiv: 1806.04667 [hep-ex] (cit. on p. 117).

[253] ATLAS Internal, *TRexFitter documentation*, 2020, URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs (cit. on p. 134).

[254] ATLAS Internal, *TRExFitter twiki page*, 2020, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TtHFitter (cit. on p. 134).

[255] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep. CERN-OPEN-2012-016, New York U., 2012, URL: https://cds.cern.ch/record/1456844 (cit. on pp. 134, 136).

[256] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, [Erratum: Eur.Phys.J.C 73, 2501 (2013)], arXiv: 1007.1727 [physics.data-an] (cit. on pp. 136, 138–140).

[257] Jerzy Neyman and Egon Sharpe Pearson, *On the problem of the most efficient tests of statistical hypotheses*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character (1933) 289–337, URL: https://doi.org/10.1098/rsta.1933.0009 (cit. on p. 138).

[258] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist. **9** (1938) 60–62, URL: https://doi.org/10.1214/aoms/1177732360 (cit. on p. 139).

[259] Morad Aaboud et al., *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] (cit. on p. 142).

[260] Georges Aad et al., *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on p. 143).

[261] ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*, (2018) (cit. on p. 143).

[262] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13$~TeV*, (2018) (cit. on p. 143).

[263] Morad Aaboud et al., *Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 361, arXiv: 1702.05725 [hep-ex] (cit. on p. 144).

[264] Albert M Sirunyan et al., *Measurement of differential cross sections for Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78** (2018) 965, arXiv: 1804.05252 [hep-ex] (cit. on p. 144).

[265] Albert M Sirunyan et al., *Measurement of the associated production of a Z boson with charm or bottom quark jets in proton-proton collisions at $\sqrt{s}$=13 TeV*, Phys. Rev. D **102** (2020) 032007, arXiv: 2001.06899 [hep-ex] (cit. on p. 144).

[266] Georges Aad et al., *Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2020) 044, arXiv: 2003.11960 [hep-ex] (cit. on p. 144).

[267] M. Aaboud et al., *Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at 13 TeV with the ATLAS detector*, Phys. Lett. B **780** (2018) 557–577, arXiv: 1710.03659 [hep-ex] (cit. on p. 144).

[268] Albert M Sirunyan et al., *Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **779** (2018) 358–384, arXiv: 1712.02825 [hep-ex] (cit. on p. 144).

[269] Albert M Sirunyan et al., *Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at $\sqrt{s}$ =13 TeV*, Phys. Rev. Lett. **122** (2019) 132003, arXiv: 1812.05900 [hep-ex] (cit. on p. 144).

[270] Albert M Sirunyan et al., *Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **10** (2018) 117, arXiv: 1805.07399 [hep-ex] (cit. on p. 144).

[271] Yoav Freund and Robert E Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, J. Comput. Syst. Sci. **55** (1997) 119–139, URL: http://dx.doi.org/10.1006/jcss.1997.1504 (cit. on p. 157).

[272] Robert E. Schapire and Yoram Singer, *Improved Boosting Algorithms Using Confidence-rated Predictions*, Mach. Learn. **37** (1999) 297–336, URL: http://dx.doi.org/10.1023/A:1007614523901 (cit. on p. 157).

[273] Llew Mason, Jonathan Baxter, Peter Bartlett, and Marcus Frean, *Boosting Algorithms As Gradient Descent in Funcion Space*, NIPS **7** (1999) 512–518, URL: http://dl.acm.org/citation.cfm?id=3009657.3009730 (cit. on p. 158).

[274] Jerome H. Friedman, *Greedy Function Approximation: A Gradient Boosting Machine*, The Annals of Statistics **29** (2001) 1189–1232, URL: http://www.jstor.org/stable/2699986 (cit. on p. 158).

[275] Jerome H. Friedman, *Stochastic Gradient Boosting*, Comput. Stat. Data Anal. **38** (2002) 367–378, URL: http://dx.doi.org/10.1016/S0167-9473(01)00065-2 (cit. on p. 160).

[276] Nicholas J. Higham, *Newton's Method for the Matrix Square Root*, Mathematics of Computation **46** (1986) 537–549, URL: http://www.jstor.org/stable/2007992 (cit. on p. 160).

[277] Karl Pearson F.R.S., *LIII. On lines and planes of closest fit to systems of points in space*, Philosophical Magazine **2** (1901) 559–572, eprint: http://dx.doi.org/10.1080/14786440109462720, URL: http://dx.doi.org/10.1080/14786440109462720 (cit. on p. 160).

[278] H. Hotelling, *Analysis of a complex of statistical variables into principal components.* Journal of Educational Psychology **24** (0933) 417–441 (cit. on p. 160).

[279] Morad Aaboud et al., *Measurements of top-quark pair differential cross-sections in the $e\mu$ channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Eur. Phys. J. C **77** (2017) 292, arXiv: 1612.05220 [hep-ex] (cit. on p. 167).

List of Figures

List of Tables